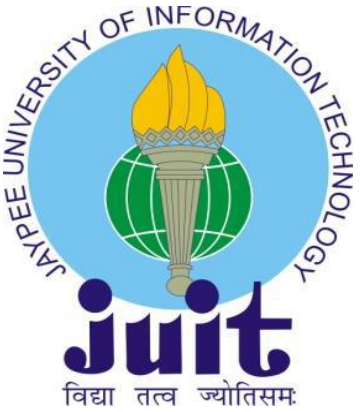
MOLECULAR DISSECTION OF LIPID BIOSYNTHETIC MACHINERY IN GREEN ALGAE, *SCENEDESMUS* SPECIES

Thesis submitted in fulfillment of the requirements for the Degree of

DOCTOR OF PHILOSOPHY

By

TAMANNA SHARMA



Department of Biotechnology and Bioinformatics
JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY
WAKNAGHAT, DISTRICT SOLAN, H.P., INDIA

JULY 2018

TABLE OF CONTENTS

DECLARATION

I hereby declare that the work reported in the Ph.D. thesis entitled "**Molecular Dissection of Lipid Biosynthetic Machinery in Green Algae, *Scenedesmus* Species**" submitted at **Jaypee University of Information Technology, Waknaghat, India,** is an authentic record of my work carried out under the supervision of **Prof. (Dr.) R. S. Chauhan**. I have not submitted this work elsewhere for any other degree or diploma. I am fully responsible for the contents of my Ph.D. thesis.

Tamanna Sharma

Department of Biotechnology & Bioinformatics

Jaypee University of Information Technology

Waknaghat, India

Date: 6th July, 2018

SUPERVISOR'S CERTIFICATE

This is to certify that the thesis entitled, **"Molecular Dissection of Lipid Biosynthetic Machinery in Green Algae, *Scenedesmus* Species"** submitted by **Tamanna Sharma** at **Jaypee University of Information Technology, Waknaghat, India**, is a bonafide record of her original work carried out under my supervision. This work has not been submitted elsewhere for any other degree or diploma.

Prof. (Dr.) Rajinder Singh Chauhan

HOD and Professor

School of Engineering and Applied Sciences

Bennett University, Greater Noida, U.P.

Ex. Dean, Department of BT and BI, JUIT, Solan, H.P.

Date: 6th July, 2018

ACKNOWLEDGEMENT

On the completion of this long and enlightening journey with the Grace of Almighty, I sincerely feel like thanking all those who guided, motivated and helped me throughout.

First of all, I convey my sincere gratitude to my supervisor ***, Prof. (Dr.) Rajinder Singh Chauhan****. His immaculate guidance, keen interest, expert evaluation, constructive criticism and extensive discussions have helped me in completing my research work. Thank you so much Sir for your support and suggestions throughout mu doctorial career.*

I am also sincerely thankful to ***Dr. Anil Kant Thakur*** *and my colleage Rakesh Singh Gour for providing Scenedesmus strains for carrying out the research work.*

I also thank the worthy administration of JUIT, Prof. (Dr.) Vinod Kumar ***(****Vice chancellor),* ***Maj Gen Rakesh Bassi (Retd.)*** *(Registrar and Dean of Students), Prof. (Dr.) Samir Dev Gupta* ***(Director & Academic Head)*** *and* ***Prof. (Dr.) Sudhir Kumar*** *(Head, Department of Biotechnology and Bioinformatics) for providing fellowship and lab infrastructure to pursue a Doctorate Degree.*

I am thankful to doctoral committee members, ***Prof. (Dr.) Satya Prakash Ghrera, Dr. Anil Kant Thakur*** *and* ***Dr. Jayashree Ramana*** *for their valuable suggestions and critical comments required to make necessary improvements in present research investigation.*

I owe my heartfelt gratitude to my parents, ***Mr. Sushil Sharma*** *and* ***Mrs. Raman Sharma*** *for their care, encouragement and understanding during my hard times. Irrespective of their health and work, they have taken care of me and my son in final stages so that I could finish my PhD. Special thanks to my brother* ***Tushar Sharma*** *whose love and silent sacrifices gave me strength to complete this study. It is my pleasant duty to place on record my thanks to* ***Mrs.Veena Sharma****, my mother-in- law for her good wishes.*

My vocabulary fails to express gratitude to my loving & understanding Husband ***Aman Sharma*** *for his constant care, personal involvement, keen interest, sympathetic attitude and persistent encouragement throughout the period of research work. His initiations, patience &*

sincere interest enabled me to achieve this arduous objective.

I am also grateful to my little son, ***Kriday Sharma*** *for his cooperation during the final stages of PhD. His cute smile and lovely acts were big support to me.*

I also extend my thankfulness to my friends and labmates for their moral support and encouragement. My special thanks to ***Dr. Archit Sood****,* ***Dr. Tamanna*** *and* ***Manali Gangwar*** *for their helpful attitude towards me.*

At last, I am also thankful to the technical staff of the department, especially ***Mrs. Somlata Sharma*** *for the assistance.*

Above all, I thank the Almighty God for everything-----He blessed me with.

TAMANNA SHARMA

ABSTRACT

Microalgal species are the most suitable feedstocks for biodiesel production. Though, to commercially utilizing the microalgae for biodiesel production, selection of strains with fast growth rate and high lipid content is required. *Scenedesmus* species are freshwater microalgae considered for obtaining biodiesel due to their high lipid content and appropriate fatty acid profile. Previous molecular studies on governing storage lipid biosynthesis in microalgae were focused on environmental stress conditions. However, there is no report on the identification of genetic factors responsible for the differential lipid accumulation among different species and strains of microalgae. The current study therefore investigated: 1) Cloning and expression analysis of genes involved in lipid biosynthesis vis-à-vis lipid content in *Scenedesmus* species 2) Deciphering molecular components channeling precursors to differential lipid accumulation and its regulation in *Scenedesmus* species.

Identification of fatty acid and triacylglycerol biosynthetic genes was done through comparative genomics and transcriptomics in *Scenedesmus dimorphus* and the gene expression investigation was performed in various developmental stages and unfavorable conditions in *Scenedesmus quadricauda* and *Scenedesmus dimorphus*. Most FA and TAG biosynthesis genes exhibited up-regulation under stress conditions and at late stationary phase, which indicates that multiple genes are involved in regulating the lipid content in *Scenedesmus* strains. Comparing the results of RT-qPCR analysis from differential conditions discovered six key genes i.e. β-ketoacyl-ACP synthase II (*KASII*), acyl carrier protein (*ACP*), fatty acyl-ACP thioestrase (*FATA*), phosphatidic acid phosphatase (*PAP*), lysophosphatidic acid acyltransferase (*LPAAT*) and diacylglycerol acyltransferase (*DGAT*) that showed correlation with lipid content in all three strains and all conditions**.** Comparative transcriptomic analysis of *Scenedesmus quadricauda* and *Scenedesmus dimorphus* having variable lipid content of 14% and 26% dry weight, respectively, although equivalent biomass uncovered the molecular mechanisms controlling differential lipid accumulation. Pathways such as carbon fixation, photosynthesis, pyruvate metabolism, citric acid cycle, glycolysis, fatty acid and triacylglycerol metabolism were explored. Genes contributing the precursors for storage lipid accumulation were ovreexpressed in high lipid

content species *S. dimorphus*. Key genes enolase (*ENO*), glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*), pyruvate dehydrogenase (*PDH*), acetyl-CoA synthetase (*ACOS*), citrate synthase (*CS*), ATP citrate lyase (*ACL*) and glycerol kinase (*GK*) were the major regulators of lipid biosynthesis in *S. dimorphus*. Molecular insight into the regulation of lipid biosynthesis was provided by the identification of WRINKLED1 transcription factor that was significantly overexpressed in *S. dimorphus*. RT-qPCR analysis for expression level estimation of WRINKLED1 transcription factor in growth stages of *S. quadricauda* and *S. dimorphus* showed a significant elevation in its expression at stationary phase of *S. dimorphus*, which revealed its association with high lipid content.

The current research provided the first report on expression analysis of lipid biosynthetic pathway genes in differential conditions of *Scenedesmus* strains, which unveils key genes that can be targeted by genetic engineering techniques for improvement of lipid profile in *Scenedesmus* species. Further, comparative transcriptome analysis and differential gene expression in *Scenedesmus* species unveiled pathways, genes and regulators responsible for cross species variation of lipid content.

LIST OF ABBREVIATIONS

ACP	Acyl carrier protein
ACAT	Acetyl- CoA acyltransferase
ACC	Acetyl- CoA carboxylase
ACL	ATP citrate lyase
ACOS	Acetyl-CoA synthetase
ACOX	Acyl-CoA oxidase
ACPD	Palmitoyl desaturase
ACSL	Long-chain acyl-CoA synthetase
ALDO	Fructose-bisphosphate aldolase
AP2	Apetala2
BLAST	Basic local alignment search tool
bZIP	Basic leucine zipper
C3H	Cysteine 3 Histidine
CBF	CCAAT box binding factors
cDNA	Complementary deoxy ribonucleic acid
CDS	Coding sequence
CoA	Coenzyme A
CPP	Cystein-rich polycomb -like protein
Cq	Quantification cycle
CS	Citrate synthase
Ct	Cycle threshold
CTAB	Cetyl trimethylammonium bromide
DGAT	Diacylglycerol acyltransferase
dNTP	Deoxynucleotide triphosphates
EDTA	Ethylenediaminetetraacetic acid
ENO	Enolase
ER	Enoyl-ACP reductase
ESP	Early stationary phase
EXP	Exponential phase

FA	Fatty acid
FAD	Omega fatty acid desaturase
FATA	Fatty acyl-ACP thioestrase
FBP	Fructose-1,6-bisphosphatase
FPKM	Fragments per kilobase of transcript per million mapped fragment
GAPA	$NADP^+$ dependent glyceraldehyde-3-phosphate dehydrogenase
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GK	Glycerol kinase
GLGC	Glucose-1-phosphate adenylyltransferase
GO	Gene ontology
GPAT	Glycerol-3-phosphate acyltransferase
HAD	β-hydroxyacyl ACP dehydratase
HAP	Heme-associated proteins
HB	Homeobox
HQ	High quality
HS	High salinity
HSF	Heat shock factor
KAAS	Kegg automatic annotation server
KAR	β-ketoacyl-ACP reductase
KASI	β -ketoacyl-ACP synthase I
KASII	β -ketoacyl-ACP synthase II
KASIII	β -ketoacyl-ACP synthase III
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	Kegg orthology
LEC	Leafy cotyledon
LFY	Leafy
LHCB2	Light-harvesting complex protein
LPAAT	Lysophosphatidic acid acyltransferase
LSP	Late stationary phase
MFP	Enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase

MISA	Microsatellite identification tool
MT	Malonyl transferase
MYB	Myeloblastosis
$NADP^+$	Nicotinamide adenine dinucleotide phosphate
NCBI	National Center for Biotechnology Information
ND	Nitrogen deficiency
NF-Y	Nuclear factor-Y
NR	Non-redundant
ORF	Open reading frame
PAP	Phosphatidic acid phosphatase
PCA	Principal component analysis
PCR	Polymerase chain reaction
PDAT	Phospholipid: diacylglycerol acyltransferase
PDH	Pyruvate dehydrogenase
PDK	Pyruvate dehydrogenase kinase
PEPC	Phosphoenolpyruvate carboxylase
PERL	Practical extraction and report language
PETF	Ferredoxin
PGM	Phosphoglucomutase
PlantTFDB	Plant transcription factor database
PPA	Pyrophosphatase
PRK	Phosphoribulokinase
PT	Palmitoyl thioestrase
PUFAs	Polyunsaturated fatty acids
QC	Quality control
QV	Quality value
RBCL	Ribulose-bisphosphate carboxylase large chain
RPIA	Ribose 5-phosphate isomerase A
rRNA	Ribosomal ribonucleic acid
RSEM	RNA-Seq by expectation maximization

RT-qPCR	Reverse transcriptase-quantitative polymerase chain reaction
SAD	Stearoyl-ACP desaturase
SBP	Sedoheptulose-bisphosphatase
SD	*Scenedesmus dimorphus*
SDHC	Succinate dehydrogenase
SQ	*Scenedesmus quadricauda*
SSR	Simple sequence repeat
ST	Stearoyl thioesterase
TAE	Tris acetate EDTA
TAG	Triacylglycerol
Taq	*Thermus aquaticus*
TCA	Tricarboxylic acid
TF	Transcription factor
TGL	TAG lipase
TKT	Transketolase
TPM	Transcripts per million
WRI1	WRINKLED1

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Research has been carried out for years on the production of biodiesel due to increasing demand of petroleum, reduction in fossil fuel resources and rise of atmospheric greenhouse gases. Biodiesel minimizes the need of depleting fossil fuel resources as well as reduces the pollution and greenhouse gas emissions from the enviornmemt [1]. Biodiesel is mainly obtained from plant oils either edible or non-edible; however utilization of edible plant oils for biodiesel production results in the struggle between fuel and food which makes the biodiesel expensive for utilization [2]. For sustainable biodiesel production, research focus should be non-edible oil or waste cooking oil. Microalgae are eukaryotic microorganisms that serve as a promising alternative to existing sources of biodiesel production due to their various advantages over other conventional sources. Microalgae accumulate significant amount of neutral lipids and have extremely fast growth rate as compared to other sources of biodiesel production. Microalgae can grow on land that is inappropriate for agriculture, hence avoids competetion with food resources for land requirement. Microalgae also utilize water sources that are inappropriate for human consumption such as waste and saline water for their growth. Also, microalgae can recycle carbon from CO_2 rich flue emissions, thereby reducing global warming. Microalgal strains can be engineered easily due to their simpler structure. All these factors enable microalgal biodiesel to compete with petroleum diesel [2-6]. However, several hurdles have to be resolved for commercial exploitation of microalgae for biodiesel, wherein foremost step is the identification and improvement of suitable strains [7].

Microalgae are aquatic photosynthetic organisms that possess the ability to convert CO2, sunlight and water into biomass. Diatoms (Bacillariophyceae), golden algae (Chrysophyceae) and green algae (Chlorophyceae) are the three most prominent classes of microalgae [8] where majority of oleaginous candidates have been identified in Chlorophyceae class (green algae) [9]. Green algae share similar metabolic pathways and photosynthetic pigments with higher plants. Moreover, oleic acid and palmitic acid are main mono and saturated fatty acids and linolenic acid and linoleic acid are main polyunsaturated fatty acids (PUFAs) in

Chlorophyceae algal class, which is prerequisite fatty acid composition for ideal biodiesel [10]. Rapid growth, comparable lipid content and suitable fatty acid composition of *Scenedesmus* genus makes its species to be considered for biodiesel production [11-14]. Girisha and coworkers observed the maximum biomass production in *Scenedesmus* species with high lipid content, when compared with *Botryococcus* and *Chlorella* oleaginous species [15]. High amount of most preferred fatty acid i.e. oleic acid has been detected in fatty acid profile of *Scenedesmus* [16].

Scenedesmus are common freshwater microalgae that belong to Phylum Chlorophyta, Class Chlorophyceae and Family Scenedesmaceae. *Scenedesmus* species are characterised by unicellular cells or coenobia of four or eight cells inside a parental mother wall. Lipid content is highly variable in *Scenedesmus* genus, of which two species under our study i.e *Scenedesmus dimorphus* (16 to 40%) and *Scenedesmus quadricauda* (1.9 to 18%) also have contrasting lipid content [2,17] (Figure 1.1).

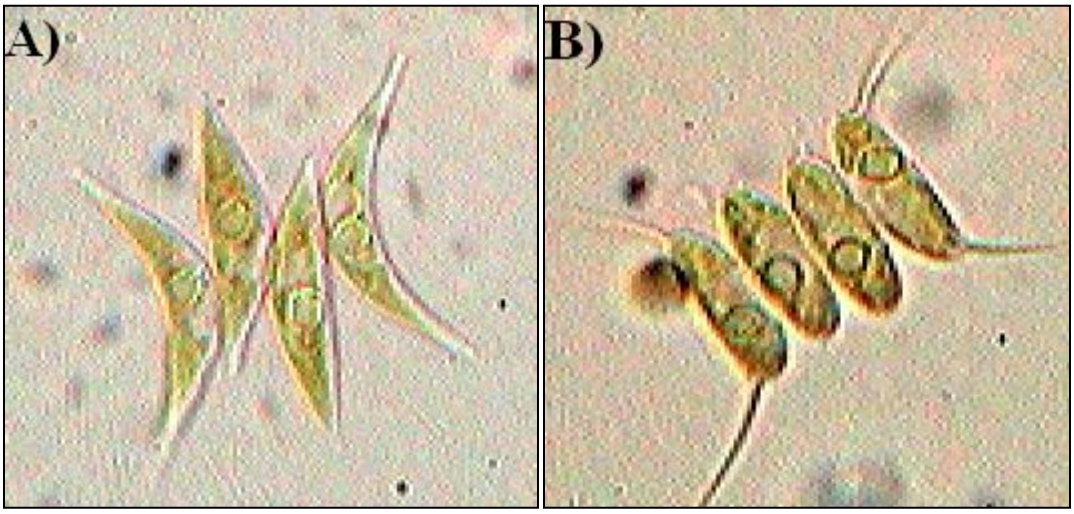


Figure 1.1 Microscopic image of **A)** *Scenedesmus dimorphus* **B)** *Scenedesmus quadricauda*

Ideal strain for overall accomplishment of biodiesel production from microalgae must have the mixture of traits i.e. high lipid content and fast growth rate. Triacylglycerol content of microalgae has shown variation among diverse strains of same genera. Usually unfavorable environmental conditions enhance the lipid content of microalgae. It has been revealed from earlier studies that nutrient deprivation is the most efficient stress condition that triggers high lipid production in microalgae [18]. However, several studies have stated that enhancement of

the lipid content under unfavorable conditions in microalgae results in low lipid productivity due to the declination of growth resuting in decreased biomass of the cells [19]. Hence, research has shifted on the identification of suitable microalgal strains and unraveling metabolic pathways for genetic manipulation of microalgae. With the advancement of molecular biology techniques and sequencing of microalgal genomes, it is possible to improve microalgal strains for enhanced lipid yields without applying stress conditions i.e. by engineering metabolic pathways involved in biodiesel production.

Lipid metabolism in algae is analogous to higher plants including *de novo* fatty acid biosynthesis and triacylglycerol biosynthesis. Fatty acid (FA) biosynthesis takes place in plastid and triacylglycerol (TAG) biosynthesis in endoplasmic reticulum. Synthesized TAG molecules are stored in densely packed lipid bodies present in the cytoplasm of the microalgal cell [9] (Figure 1.2).

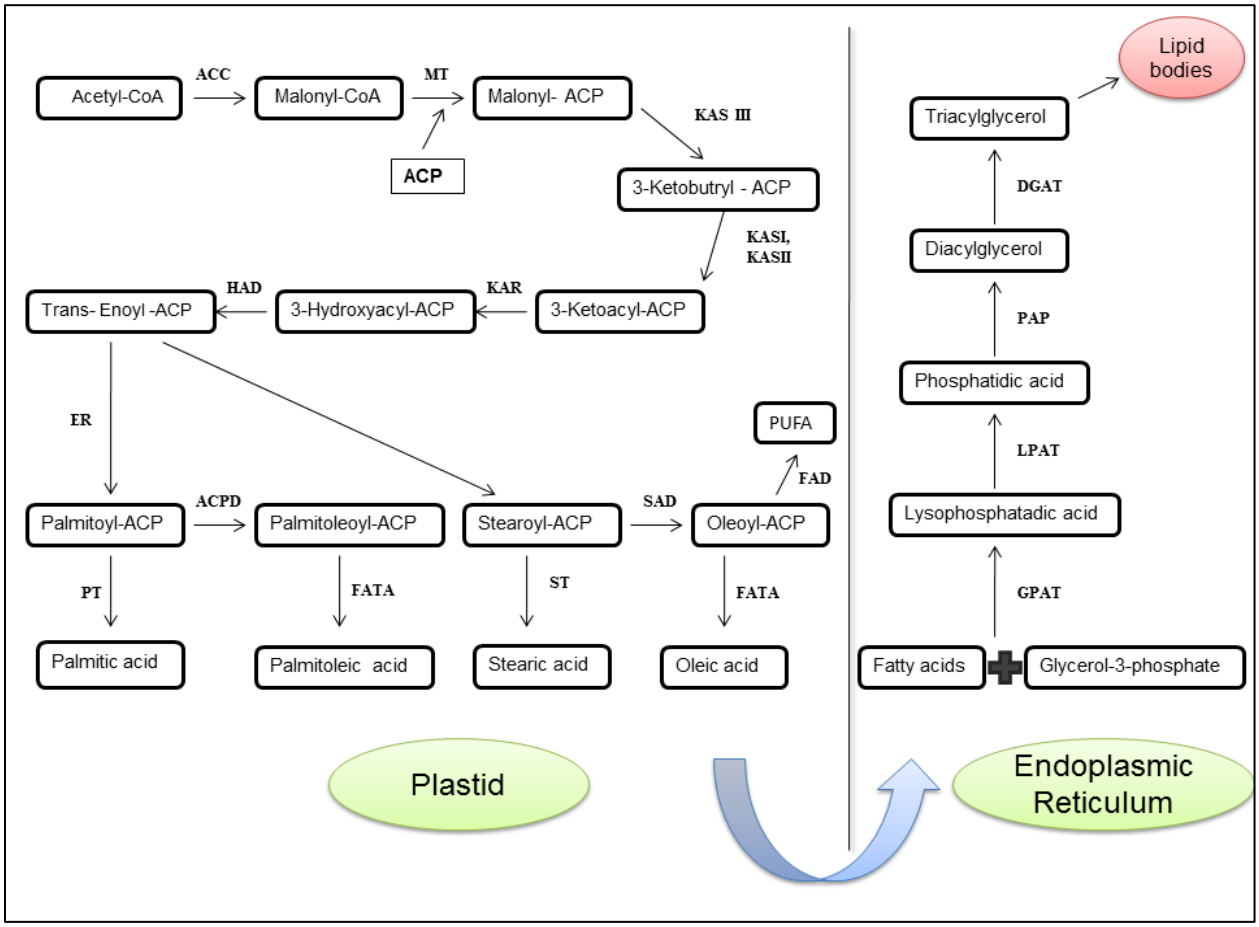


Figure 1.2 The general scheme of microalgal fatty acid and triacylglycerol biosynthetic pathway in microalgae [20, 21]. Fatty acid biosynthesis occurs in plastid wheres triacylglycerol biosynthesis in endoplasmic reticulum. Enzyme abbreviations: ACC- Acetyl-CoA carboxylase; MT- Malonyl transferase; ACP-Acyl Carrier Protein; KASI- β-ketoacyl-ACP synthase I; KASII- β-ketoacyl-ACP synthase II; KASIII- β-ketoacyl-ACP synthase III; KAR- β-ketoacyl-ACP reductase; HAD- β-hydroxyacyl ACP dehydratase; ER- Enoyl-ACP reductase; PT- Palmitoyl thioestrase; FATA- Fatty acyl-ACP thioesterase; ST- Stearoyl thioesterase; SAD- Stearoyl-ACP desaturase; ACPD- Palmitoyl desaturase; FAD- Omega fatty acid desaturase; GPAT- Glycerol-3-phosphate acyltransferase; LPAT- Lysophosphatidic acid acyltransferase; PAP- Phosphatidic acid phosphatase; DGAT- Diacylglycerol acyltransferase

Understanding of lipid metabolism, catabolism as well as pathways involved in the modification of fatty acid chain length and saturation are essential for the production of lipid rich microalgal strains. Progression of next generation sequencing techniques has resulted in the sequencing of several microalgal genomes and transcriptomes [22-30]. Various key genes of different metabolic pathways involved in lipid biosynthesis have been identified and studied for their role in increasing lipid accumulation among different species of microalgae [31-36]. Also,

several transformation methods have been developed to allow genetic engineering techniques in microalgal species [37-46]. Further, transcriptomics is an ultimate approach to gain understanding of metabolic pathways pertaining to differential lipid accumulation. Transcriptomes of various microalgal species *Chlamydomonas reinhardtii, Neochloris oleoabundans, Dunaliella tertiolecta, Chlorella protothecoides, Dunaliella parva* were sequenced and compared under contrasting conditions of nutrient replete and nutrient deplete conditions to unveil the pathways and genes linked with high lipid content [47-51]. Little knowledge was available about the transcriptional regulation of storage lipid biosynthesis in microalgae. Availability of genomic resources has directed the identification of transcription factors in microalgal species *Chlamydomonas reinhardtii* and *Nannochloropsis oceanica* which also revealed lipid-related transcription factors [52, 53].

In spite of the available information, knowledge about the inherent genetic differences that influence the production of lipids in high lipid containing strains is still lacking. All the earlier reports on unveiling the molecular basis for high lipid content in microalgae were focused on correlating the expression of genes involved in storage lipid biosynthesis with various stress conditions. But as mentioned above, lipid content enhancement in unfavourable conditions is associated with reduction in biomass productivity of microalgal cells. Enhancement of lipid accumulation with comparable growth is essential for commercial exploitation of microalgae for biodiesel production. In order to genetically modifying the components for enhancing the lipid accumulation and manipulating the lipid profile, there should be a thorough knowledge of the molecular basis of triacylglycerol biosynthesis in microalgae. The mechanism of interspecies differentiation of lipid content in closely related species and strains of microalgae was not known. Hence, to unravel the genetic basis of differences in lipid content, comparative molecular dissection of pathways and genes in high and low lipid content strains of the same genera can provide logical insights.

Thus, to ascertain the role of fatty acid and triacylglycerol biosynthesis genes for storage lipid production, cloning and relative gene expression analysis of 14 genes of fatty acid biosynthesis and 4 genes of triacylglycerol biosynthesis pathway was performed in various stress conditions and growth stages of lipid content contrasting *Scenedesmus* for the identification of genes having role in high lipid accumulation. Further, to elucidate the role of other metabolic pathways and regulatory mechanisms for differential lipid accumulation among species of same genera of

microalgae, transcriptomes of *S. quadricauda* and *S. dimorphus* at lipid accumulating phase were sequenced using Illumina NextSeq 500 platform. Comparative analysis of pathways and genes involved in storage lipid accumulation and provide precursors for storage lipid biosynthesis was performed on the basis of fragments per kilobase of transcript per million mapped fragment (FPKM). Carbon metabolism, photosynthesis, fatty acid and triacylgycerol metabolism and catabolism were the main pathways studied. Further, transcription factors were identified from the transcriptomes of both the species, of which transcript abundance of lipid biosynthesis-related transcription factors was estimated, which provided the connection of transcription factors with the high lipid content.

As understanding the molecular biology of high lipid content in contrasting strains of same genera would provide information on genetic and regulatory factors influencing the production of lipids in microalgae, the current research work was carried out with two objectives:

1) Cloning and expression analysis of genes involved in lipid biosynthesis vis-à-vis lipid content in *Scenedesmus* species.

2) Deciphering molecular components channeling precursors to differential lipid accumulation and its regulation in *Scenedesmus* species.

CHAPTER 2

REVIEW OF LITERATURE

The literature pertaining to the present study has been reviewed as under:

2.1 What is Biodiesel?

Renewable diesel and the substitute of conventional diesel based on fossil fuels, obtained from animal fat, vegetable oil or waste cooking oil is called biodiesel. Triacylglycerols are the major components of natural oil or fat, which are transesterified into alkyl esters. Triacylglycerols contain fatty acids of different lengths that are attached to one glycerol backbone. Hence, the properties of oil pertaining to biodiesel production depend on the different fatty acid composition of natural oil. In the process of transesterification, alkyl esters are formed from the reaction of triacylglycerol and alcohol carried out in the presence of catalyst [54].

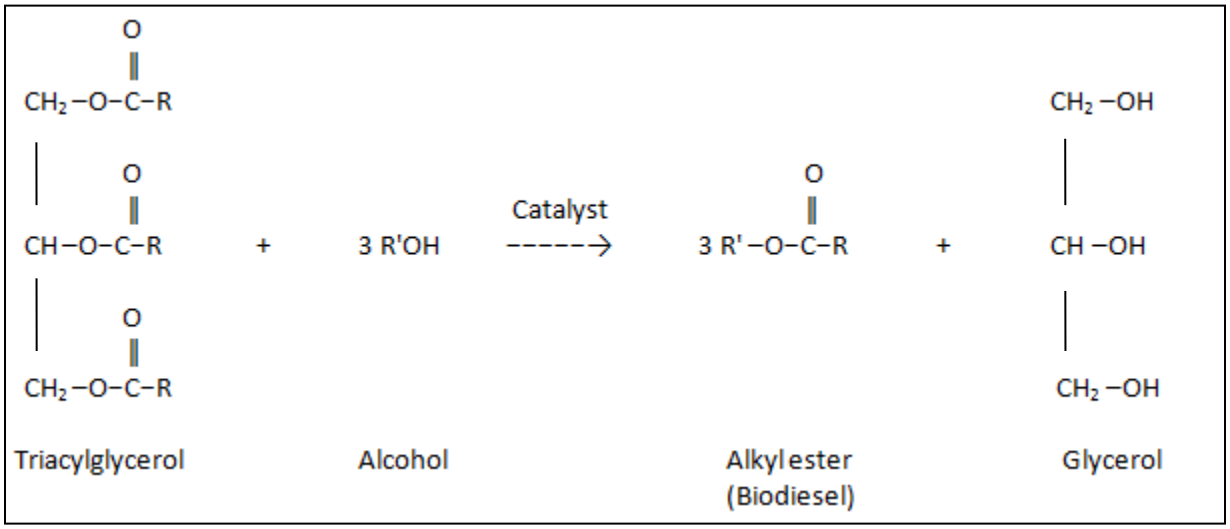


Figure 2.1 Transesterification reaction [54]. R represents fatty acid chains and R' is usually CH_3

2.2 Feedstocks of Biodiesel

Several oil crops have been identified as possible feedstocks of biodiesel production. However, rapeseed, soybean, canola, sunflower, cotton seed and palm kernels have been the main feedstocks used for commercial production of biodiesel [55]. Rapeseed oil is most commonly used world's biodiesel feedstock, followed by sunflower oil, soybean and palm oil. It depends on the availability of feedstocks in a particular geographical area. But, these edible oils compete with food security; hence non-edible oils have been the preferred sources for biodiesel production over edible oils. Inedible oil from *Jatropha curcas* has been used for the production of biodiesel in tropical areas such as India and Africa [56]. Nowadays, research has been carried out for utilizing algae based biodiesel for meeting the global petroleum needs. Comparison of microalgae with other feedstocks of biodiesel is given in Table 2.1.

Table 2.1 Comparison of different feedstocks for biodiesel [18]

Source	Oil Yield (L/ ha)	Land area required (M ha)
Corn	172	1540
Soybean	446	594
Canola	1190	223
Jatropha	1892	140
Coconut	2689	99
Oil palm	5950	45
Microalgae	58700- 136900	2- 4.5

Microalgae offer various advantages over other feedstocks for biodiesel production:

1. Have fast growth rate and high lipid content.
2. Uses non- arable land that is unsuitable for agriculture, hence does not compete with food crops.
3. Can grow in brackish, saline, and waste water that has little competing demand.

4. Biomass can be harvested batch-wise at any time in a year, unlike other sources which are harvested only once or twice in a year.
5. Recycles CO_2 from industrial waste and utilizes phosphorous and nitrogen from several wastewater sources, thereby reduces the generation of green house gases and performs wastewater bioremediation [57].
6. Biodiesel from microalgae reduces nitrous oxide release, thus have minimal environmental impact [58].
7. Produce more lipids per unit area than other sources being utilized for production of biodiesel [9].

2.3 Algal Biology

Algae are simple microorganisms that live in water and use light, water and CO_2 to produce algal biomass [8]. Algae can be classified as microalgae and macroalgae. Microalgae are unicellular photosynthetic organisms whereas macroalgae are large, multicellular organisms [60]. Microalgae can grow by sexual or asexual reproduction, but mostly microalgal reproduction occurs by asexual mode. Microalgae have diverse groups based on their biochemical constituents, pigment composition, life cycle and ultrastructure.

Mainly, microalgae are classified as: Bacillariophyta (diatoms), Chlorophyta (green algae), Prymnesiophyta or Haptophyta, Eustigmatophyta and Cyanobacteria (prokaryotic blue-green algae) [61].

2.4 Microalgal growth dynamics

The growth of microalgae has different developmental phases i.e. lag, exponential, declining growth rate, stationary and death phase (Figure 2.2).

1. **Lag phase:** Miroalgae starts adapting to the new growth conditions, when transferred to a fresh medium. This phase of physiological adaptation of microalgal culture is called lag phase.
2. **Exponential phase:** When the microalgal culture has adapted itself according to the given environmental condition, cells enter exponential phase. It is also called logarithmic phase as the cells divide at stable rate and there was increase in cell density with time.

3. **Declining growth phase:** This phase occurs when the nutrients, pH, light, carbon dioxide or other chemical and physical factors start exhausting, which results in the declining of cell division.
4. **Stationary phase:** Stationary phase is when there is no net growth of microalgal cells due to the limitation of any of the growth requirement.
5. **Death Phase:** Unfavourable environmental conditions cause the microalgal cells to collapse. Death phase of a microalgal culture is very rapid and termed as "culture crush" [17, 62, 63].

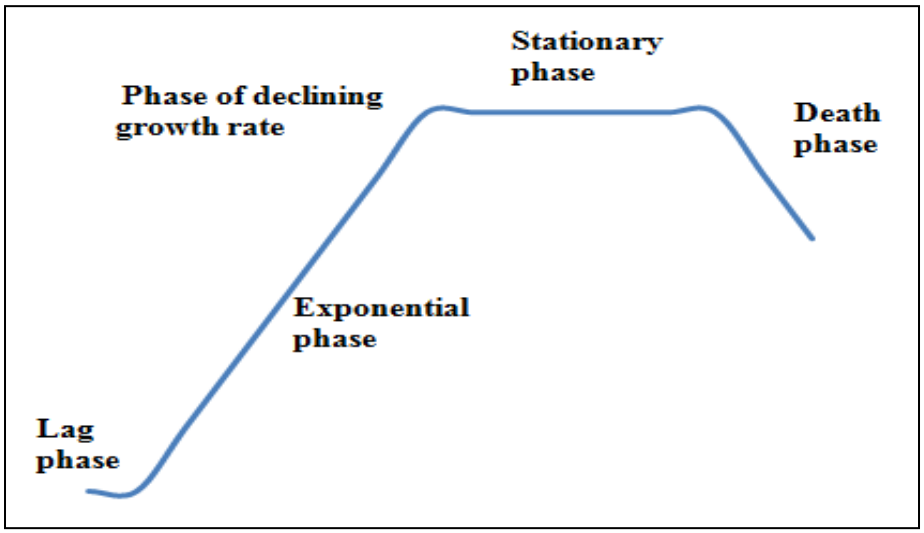


Figure 2.2 General pattern of microalgal growth

2.5 Microalgae as feedstocks for biodiesel production

Microalgal biomass includes three major components: lipids, carbohydrates and proteins [64]. Each fraction can be used for fuel generation which provides a potential alternative for products obtained from natural gas or petroleum. Of these three components, lipids are having the highest potential and energy content. Moreover, lipids from microalgae resemble vegetable oils derived from soybean, corn, canola and other feedstocks of biodiesel production [65]. Hence, microalgae are the potential sources of biodiesel production having properties similar to petroleum feedstocks.

DOE's aquatic species program, carried out from 1978 to 1996 focused its research on biodiesel from microalgal storage lipids. About 3000 different strains were collected, screening allowed selection of strains adapted to particular condition such as salinity, pH, temperature and their ability to produce storage lipids. A total of 300 strains were selected, of which green algae (Chlorophyceae) and diatoms (Bacillariophyceae) dominated. Further research of Aquatic Species program was done on the metabolic engineering of diatoms and green algae that included pathway modifications and development of genetic tools [66].

2.6 Research on microalgae for biodiesel production

There are many species of microalgae that accumulate remarkable amounts of lipids resulting in high oil yield. Lipid productivity of microalgae depends on the rate of their growth and on the level of biomass lipid content. Microalgae having high lipid productivity are desirable for biodiesel production [67]. Hence, selection of suitable microalgal strain is the first and most important step in the commercializationof microalgae-based biodiesel industry [68].

Microlage shift their metabolism towards neutral lipids biosynthesis in unfavorable environmental conditions, which results in high storage lipid content [69]. Different species of microalgae are grown under adverse conditions of nutrient deficiency, salinity stress, high temperature and pH, resulting in increased lipid production (Table 2.2). Nutrient deficiency includes limitation of major nutrients i.e. nitrogen, phosphorous, sulphur, that are required for growth of microalgae [70]. *Chlamydomonas reinhardtii*, unicellular model green algae, has shown high level of triacylglycerol production under the limitation of nitrogen and sulphur, respectively [71, 72]. Also, increase in the lipid proportion has been reported in *Scenedesmus* species subjected to nitrogen or phosphorus deficiency [73]. Other microalgal species including *Chlorella, Nannochloropsis, Phaeodactylum and Dunaliella* have shown significant increase in the level of triacylgycerol under nitrogen deficient conditions.

Variation of salinity in the growing media causes hypo or hyperosmotic stress on microalgal cells, thus shifting the pathways towards lipid production. *Dunaliella* is the most studied halophillic green microalgae found in high salinity environments [69]. Enhanced lipid production has been observed in *Dunaliella tertiolecta* grown in media containing increased NaCl concentration [74]. Other stress conditions i.e.temperature, light intensity and pH function in

altering the composition of fatty acid in microalgae [69]. Generally declination in temperature results in increased fatty acid unsaturation and elevation in temperature increases fatty acid saturation. In *Isochrysis* and *Nannochloropsis* species, increase in the temperature resulted in enhancement of lipid accumulation [75-77]. Also, alkaline pH stress causes increased triacylgycerol content and declined the proportion of membrane lipids in *Chlorella* [78]. Light/dark cycles in various growth stages have shown their impact on lipid content and composition in the diatom *Thalassiosira pseudonana* [79].

Table 2.2 Conventional approaches for increasing lipid production in microalgae

Microalgal species	Stress	Effect on lipid content	Reference
Chlamydomonas reinhardtii	Nitrogen limitation	Increase in total lipids	Dean *et al*. [72]
Chlorella vulgaris		Total lipid increased by 16.41%	Converti *et al*. [77]
Nannochloropsis oculata		Total lipid increased by 15.31%	
Phaeodactylum tricornutum		TAG levels increased from 69 to 75%	Alonso *et al*. [80]
Scenedesmus subspicatus		Increase in total lipids	Dean *et al*. [72]
Chlorella vulgaris var L3		Increased lipid (TAG) content (2.7 times)	Ikaran *et al*. [81]
Chlorella sorokiniana	Nitrogen limitation	Increase in total lipids	Li *et al*. [82]
Scenedesmus sp.	Nitrogen and phosphorus starvation	Lipids increased 30% and 53%, respectively	Xin *et al*. [73]
Chlamydomonas reinhardtii	Sulphur limitation	Increase in TAG	Matthew *et al*. [71]
Dunaliella tertiolecta	Transferred from 29 g/L to58 g/L NaCl	Increase in lipid content and TAG	Takagi *et al*. [74]
*Chlorella minutissima*UTEX 2341	NaCl, Fe^{3+} and nitrogen starvation	lipid content increase to 2.5 times	Cao *et al*.[83]
Neochloris oleoabundans	Low light, high pH and nitrogen starvation	Increase in total lipids	Santos *et al*. [84]
Nannochloropsis oceanica IMET1	High light intensity and nitrogen repletion	Higher neutral lipid and biomass	Xiao *et al*. [85]

2.7 *Scenedesmus* (Green Microalgae)

2.7.1 Taxonomic classification

Kingdom	Plantae
Subkingdom	Viridiplantae
Phylum	Chlorophyta
Class	Chlorophyceae
Order	Sphaeropleales
Family	Scenedesmaceae
Genus	*Scenedesmus*

There are about 87 taxonomically accepted species of *Scenedesmus* [86].

2.7.2 Biology

Scenedesmus is the most common freshwater green microalga having fast growth rate with simple nutritional needs. It is easy to maintain *Scenedesmus* cultures in the laboratory [87]. It exits as coenobia of 4- 32 cells arranged in 1 or 2 rows. Cells of *Scenedesmus* species differ in shape i.e. elliptical to spindle shape, but always have elongated structure containing one chloroplast and one pyrenoid. Cell wall can be smooth or can have different sculptures, with or without spines among different species [88].

2.7.3 Reproduction

Most of the *Scenedesmus* species reproduce asexually through the production of autocolonies. Parent cells divide into a number of daughter cells. The daughter cells group themselves to form a colony called autocolony. Also, sexual mode of reproduction has been reported in few *Scenedesmus* species, where *Scenedesmus* produces zoospores or gametes that fuse together to form zygote resulting in a new cell.

2.7.4 Ideal for biodiesel production

Composition of fatty acids present in the lipids influences the properties of biodiesel produced. Linolenic acid (C18:3), linoleic acid (C18:2), oleic acid (C18:1), stearic acid (C18:0) and palmitic acid (C16:0) are five required fatty acids suitable to the lipid profile for biodiesel production. Oleic acid and palmitic acid are main fatty acids and linolenic and linoleic acid are

main polyunsaturated fatty acids (PUFAs) present in Chlorophyceae algal class, hence green microalgae possess the required composition of fatty acids for ideal biodiesel production [89]. *Scenedesmus* have been identified as potential lipid producing species with fast growth rate and high lipid content. High growth rate, high lipid content and suitable fatty acid profile of *Scenedesmus* species i.e. presence of oleic acid and linoleic acid in high amount have made these species to be considered for production of biodiesel [90-94]. In a recent study, different strains belonging to green microalgae were collected including *Scenedesmus* sp., *Chlorella* sp., *Desmodesmus* sp., *Tetraedron caudatum*, *Graesiella emersonii*, *Pseudomuriella* sp., *Chlamydomonas* sp., *and Mychonastes timau*ensis. *Scenedesmus dimorphus* has shown most appropriate fatty acid composition that is desired for biodiesel production [95] (Table 2.3).

Table 2.3 Fatty acids composition of different green microalgae **[95]**

Fatty acids	***Chlorella* sp.**	***Graesiella emersonii***	***Tetraedron caudatum***	***Scenedsmus dimorphus***
Palmitic acid (C16:0)	33.43	18.79	7.16	27.94
Stearic acid (C18:0)	1.03	2.04	0.46	1.91
Oleic acid (C18:1)	15.09	23.79	6.13	34.49
Linoleic acid (C18:2)	22.29	11.04	3.45	9.43
Linolenicacid (C18:3)	38.85	18.36	11.77	20.37
Lipid productivity ($\mu g\, mL^{-1} day^{-1}$)	14.61	9.99	2.71	12.39

Scenedesmus dimorphus is green microalgae, bean shaped and about 10 µm in size. Lipid content ranges from 16- 40% dry weight. This particular strain of *Scenedesmus* has the potential to be used as biodiesel feedstock as it has fast growth rate with high lipid production [96]. Comparison of different species of microalgae has revealed high fraction of lipids in the chemical composition of *Scenedesmus dimorphus* (Figure 2.3).

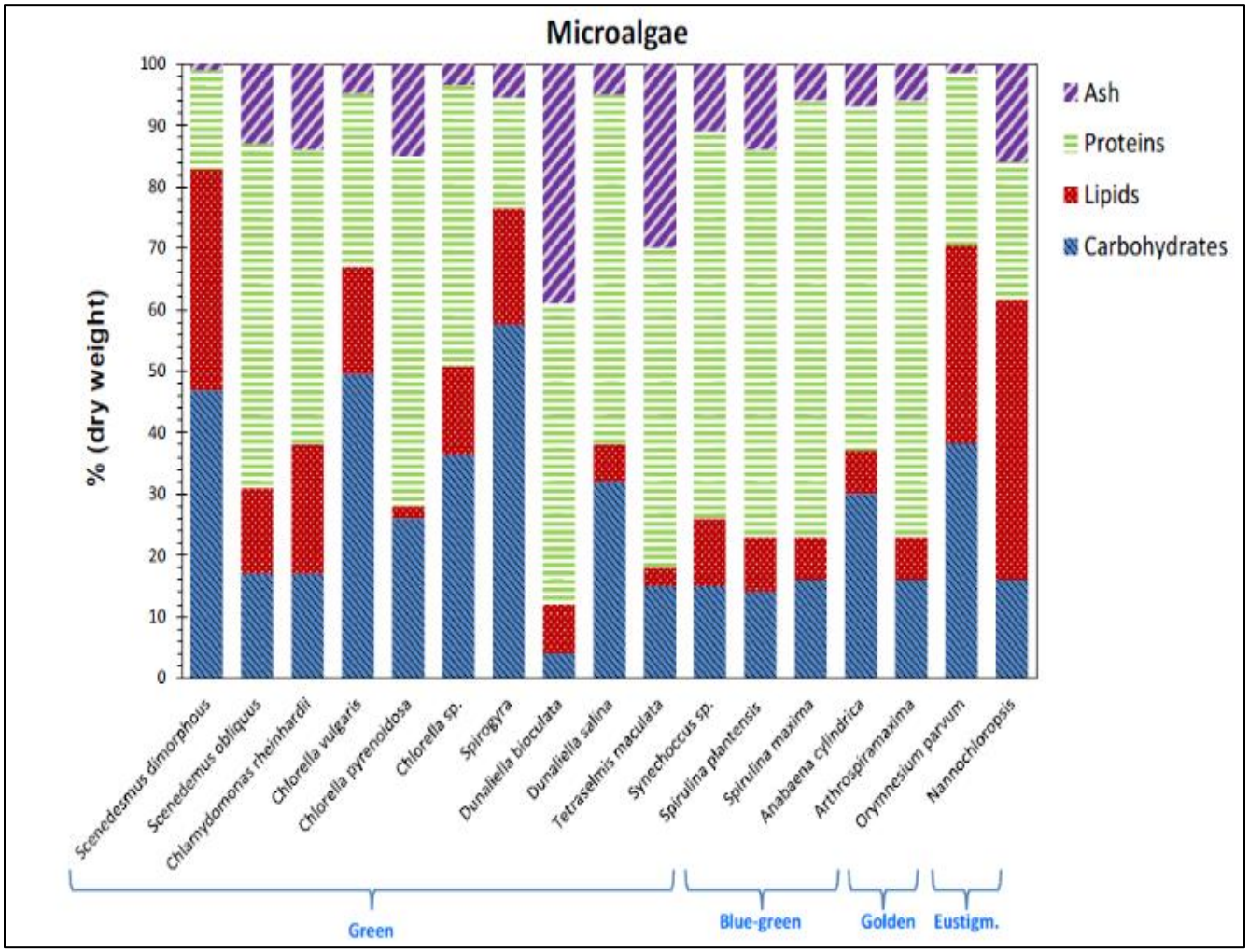


Figure 2.3 Chemical composition of microalgal species of different groups (green, blue-green, golden and eustigmatophyceae) [97].

2.7.5 Other biotechnological applications of *Scenedesmus*

Although, members of *Scenedesmus* genus are the promising species for biodiesel, many species of *Scenedesmus* are rich sources of bioactive compounds that are utilized in aquaculture, cosmetics, pharmaceuticals and human nutrition (Table 2.4). *Scenedesmus* species have high bioactivities and nutritional content, hence utilized for various biotechnological applications [98, 99]. Toxicological safety assessment of *Scenedesmus* was performed with test animals which indicated the absence of abnormalities or toxic impacts [100]. Also, features such as rapid growth, ease of cultivation and survival ability in harsh environmental conditions make *Scenedesmus* suitable to be used worldwide [101, 102].

Table 2.4 Bioactive compounds found in *Scenedesmus* species

Compound	Applications	Reference
Vitamin B	Health-Food	Becker [103]; Borowitzka [104]
Vitamin C	Health-Food additives, Pharmaceutical	Becker [103]; Borowitzka [104]
Vitamin E	Health-Food, Medicine	Becker [103]; Borowitzka [104]
Lutein	Animal nutrition, Pharmaceuticals	Tukaj *et al*. [105]; Otto and Wolfgang [102]; Ceron *et al*. [106]; Skjanes *et al*. [107]
Astaxanthin	Aquaculture, cosmetics, Human nutrition, Medicine	Otto and Wolfgang [102]; Qin *et al*. [108]; Gouveia *et al*. [109]; Jouni and Makhoul [110]
Haemagglutinin	Medicine	Chu *et al*. [111]
β- Carotene	Food colourant, Medicine	Karen *et al*. [112]; Indira and Biswajit [113]; Guedes *et al*. [114]
Mycosporine-like amino acids, sporopollenin	Cosmetics (UV - screening compounds)	Indira and Biswajit [113]; Skjanes *et al*. [107]
Chlorophyll a, b, c	Food colourants, Pharmaceuticals, cosmetics	Karen *et al*.[112]; Gouveia *et al*. [109]; Indira and Biswajit [113]; Catarina *et al*. [114]
Polysaccharides	Medicine, Bioethanol, $BioH_2$	Gouveia *et al*.[109]; Skjanes *et al*. [107]
Extracts with antimicrobial/ antifungal activities	Medicine	Abedin & Taha [116]
Amino acids	Food	Chacón-Lee and Gonzalez- Marino [97]

2.7.6 Antibacterial Activities of *Scenedesmus* species

Scenedesmus species produce antimicrobial substances and possess an antibacterial property that inhibits the growth of pathogenic strains of bacteria including *Pseudomonas aeroginosa, Salmonella* sp., *Staphylococcus aureus* and *Escherichia coli* [117]. *Scenedesmus* species are known to have chlorophyll a, b and several carotenoids. Chlorophylls and β-carotene act as microbial growth inhibitors, and some studies have reported their antioxidant as well as antimicrobial properties [118,119]. Also, antibacterial activity of *Scenedesmus costatum* has been reported by Desbois and coworkers against aquaculture bacteria [120].

2.5 Biosynthesis of fatty acids and triacylglycerol in microalgae

FA and TAG biosynthesis pathway in microalgae are similar with higher plants as discovered from sequence homology.

2.8.1 Fatty acid biosynthesis

FA biosynthesis in microalgae generally takes place in plastid. Conversion of acetyl-CoA to malonyl-CoA by irreversible carboxylation catalyzed by acetyl-CoA carboxylase (*ACC*) is the first committed step of this pathway. Next, malonyl transferase (*MT*) transfers the malonyl group from malonyl-CoA to a protein cofactor acyl carrier protein (*ACP*). This f step is catalysed by enzyme. After that, there is condensation of malonyl-ACP in a series of reactions with acyl-ACP. Condensation reaction is carried out by enzymes of fatty acid synthase complex that adds two carbon units in each elongation cycle. The first condensation reaction results in the formation of β-ketobutyl-ACP by the action of β-ketoacyl synthase III (*KAS III*). Three additional reactions takes place after each condensation i.e. reduction of β-ketoacyl-ACP into β-hydroxylacyl-ACP by β-ketoacyl-ACP reductase (*KAR*), next is dehydration to form enoyl-ACP by hydroxyacyl-ACP dehydratase (*HAD*) and further reduction to form the elongated fatty acids by enzyme enoyl-ACP reductase (*ER*). Subsequent condensation of β-ketobutyl-ACP with malonyl-ACP is carried out by β-ketoacyl synthase I (*KAS I*) enzyme, which generates fatty acids of varying carbon length (6-16). Another condensing enzyme β-ketoacyl synthase II (*KAS II*) catalyzes the elongation of 16 carbon fatty acids to 18 carbon fatty acids. Saturated fatty acids formed from the fatty acid biosynthetic pathway are further utilized by desaturases to produce unsaturated

fatty acids. Desaturation of stearic acid (18:0-ACP) is catalysed by enzyme stearoyl-ACP desaturase (*SAD*), which produces oleic acid (18:1- ACP). Further, acyl-ACP thioesterases terminate the elongation of fatty acids by hydrolyzing the acyl-ACPs into free fatty acids or by transfering these fatty acids to glycerol-3-phosphate.

2.8.2 Triacylglycerol biosynthesis

TAG biosynthesis in microalgae occurs through direct glycerol pathway (Kennedy pathway). Released fatty acids from chloroplast get attached to 1 and 2 positions of glycerol-3-phosphate by the action of enzymes glycerol-3-phosphate acyltransferase (*GPAT*) and lysophosphatidic acid acyltransferase (*LPAAT*) respectively, generating phosphatidic acid. Subsequently, phosphatidic acid is dephosphorylated to diacylglycerol by phosphatidic acid phosphatase (*PAP*). Enzyme diacylglycerol acyltransferase (*DGAT*) catalyzis the final step by transfering the third fatty acid to diacylglycerol. The acyltransferases have a significant role in the determination of the acyl composition of TAG molecules due to their preference for specific acyl-CoA.

There is also another route for TAG biosynthesis in microalgae, which is catalyzed by phospholipid: diacylglycerol acyltransferase (*PDAT*) and considered as acyl-CoA-independent synthesis of TAG. Phospholipids present in microalgae donate acyl groups to diacylglycerol for the generation triacylglycerol.

TAG biosynthesis from this pathway occurs in stress conditions in which the degradation of photosynthetic memebrane generates TAG molecules in lipid bodies of cytosol. Hence this pathway is vital for regulating the composition of memebrane lipids in various growth conditions and environmemtal circumstances [9, 121].

2.9 Available genomic resources of microalgal species

Next generation sequencing technologies have led the generation of significant genomic information of different microalgal species has to study the molecular aspects of microalgae. Whole genome information of more than ten microalgal species has been generated including *Chlamydomonas reinhardtii* [23], *Volvox carteri* [26], and *Chlorella variabilis* [28] (Table 2.5). However, the rationale of sequencing the whole genomes was not in context of lipid biosynthesis [48]. Genomes of a number of organelles including mitochondria, nucleus and plastid in microalgal species were sequenced [122, 123]. As well, a lot of EST databases are there in NCBI (www.ncbi.nlm.nih.gov/dbEST).

Table 2.5 List of ten microalgal species that have their genome sequences available

Organism	Strain	Genome size (Mb)	Reference
Chlamydomonas reinhardtii	CC-503	121	Merchant *et al*. [23]
Volvox carteri	UTEX2908	138	Prochnik *et al*. [26]
Chlorella vulgaris	NC64A	46.2	Blanc *et al*. [28]
Coccomyxa subellipsoidea	C-169	48.8	Blanc *et al*. [27]
Ostreococcus lucimarinus	CCE9901	13.2	Palenik *et al*. [24]
Ostreococcus tauri	OTH95	12.6	Derelle *et al*. [22]
Micromonas pusilla	RCC299	20.9	Worden *et al*. [25]
Cyanidioschyzon merolae	10D	16.5	Matsuzaki *et al*. [124]
Phaeodactylum tricornutum	CCP1055/1	27.4	Bowler *et al*. [125]
Thalassiosira pseudonana	CCMP1335	34.5	Armbrust *et al*. [126]

2.10 Molecular basis of fatty acid and triacylglycerol biosynthetic pathway

Growing number of sequences of microalgal species facilitated the gene identification and metabolic pathways investigation for the production of microalgal biodiesel. Also, advancement in high throughput methods to study gene expression has enhanced the understanding of molecular basis of lipid accumulation in microalgae. A number of studies have performed the identification of lipid related genes in different microalgal species and studied their gene expression in different conditions. Available genome sequences and expressed sequence tags information of *Chlamydomonas reinhardtii* was utilized by Riekhof and coworkers for *in-silico* study of fatty acid and triacylglycerol metabolism. Genes involved in these pathways were annotated based on similarity to homologs from *Arabidopsis thaliana* [127]. In 2007, genomic and biochemical investigation of unicellular rhodophyte *Cyanidioschyzon merolae* revealed significant variation of storage lipid biosynthesis in red microalgae compared to green microalgae and plants. The results suggested the absence of main desaturases of green algae *in C. merolae* viz. acyl lipid desaturases and stearoyl ACP desaturase [128]. Further, expression of fatty acid biosynthesis genes was studied under different stress conditions in *Haematococcus pluvialis* which revealed that *ACP*, *KAS* and *FATA* genes were correlated positively with fatty acid biosynthesis [32]. In another study, response of genes involved in lipid biosynthesis and accumulation was examined under three nutrition stressors in *Chlorella pyrenoidosa* where malic enzyme, ACCase and diacylglycerol acyltransferase were found to be highly related to lipid accumulation [33].

As whole genome sequencing is difficult to be performed without reference genome, transcriptome sequencing makes it possible to perform *de novo* assembly of transcriptomes from species that do not have their genomes sequenced, which aids in better understanding of lipid metabolism in unsequenced oleaginous microalgae. Transcript abundance estimation of transcriptomes calculates the relative abundance of mRNA levels in a single cell or a population of cells. It provides information about the active genes expressed in particular physiological condition in cells and thus used for the identification of putative genes that can be targeted to for lipid content enhancement in microalgae [129]. Microalgal transcriptomic investigations are focused primarily on the identification of differentially expressed transcripts under differentiating physiological stress conditions. In this regard, the first *de novo* transcriptomics analysis of microalga *D. tertiolecta* identified a range of genes from biosynthetic and

catabolism pathways of starch, fatty acids and TAG [48]. Further, transcriptome data of various microalgal species i.e. *Chlorella vulgaris, Chlamydomonas reinhardtii, Nannochloropsis oceanica, Neochloris oleoabundans, Chlorella protothecoides, Dunaliella parva* have been used to identify genes and pathways for lipid accumulation and to further identify differentially expressed genes in contrasting conditions of lipid accumulation. Various genomic and transcriptomic studies in context of lipid biosynthesis in different species of microalgae have been summarized in Table 2.6.

2.11 Carbon flux to lipid accumulation

Lipid accumulation not only depends upon fatty acid and triacylglycerol pathways, but the major control of lipid biosynthesis depends upon the carbon precursors supply. Overall lipid accumulation is a complex process and is regulated at several levels. Carbon precursor supply is the major restrictive factor for triacylglycerol biosynthesis in *C. reinhardtii* [130]. Comparative transcriptomic and proteomic analysis of *Chlorella protothecoides* under heterotrophic and autotrophic growth conditions revealed that the up-regulation most of the metabolic pathways involved in the generation of triacylglycerol and and down-regulation of fatty acid degradation under heterotrophic conditions [131]. Gene expression estimation during nutrient-depleting and lipid-accumulating conditions of *Phaeodactylum tricornutum* suggested the important role of acetyl-CoA carboxylase precursors for TAG synthesis than acetyl-CoA carboxylases [132]. Transcriptome sequencing of *Neochloris oleoabundans* under nitrogen limiting conditions unveiled the mechanism of how stress conditions causes elevation in TAG content. Activated pentose phosphate pathway, overexpressed pyruvate dehydrogenase complex, enhanced fatty acid synthesis and TAG biosynthesis and repressed β-oxidation pathway was revealed for high lipid accumulation [133]. Significant increase towards the flux of carbon precursors for storage lipid production was reported by the transcriptomic study of *Nannochloropsis* under nitrogen depletion [134]. Further, high concentration of exogenous CO_2 provided to the cultures of *Chlorella sorokiniana* with nitrogen depleting conditions revealed the up-regulation of Calvin cycle and glycolysis enzymes. This study points to the increased CO_2 assimilation rate and hence enhanced carbon flux towards fatty acid synthesis [135]. Transcriptomic analysis of green alga, *Coccomyxa subellipsoidea* C-169 with CO_2 supplementation also provided the overview of enhanced photosynthesis, Calvin cycle,

glycolysis, oxidative phosphorylation, TCA cycle, fatty acid synthesis and declined fatty acid catabolism, contributing to the tremendous lipid accumulation [136].

All these reports were focused on the lipid trigger that is environmental stress in which many microalgae appeared to produce more TAGs. The increase in lipid content does not increase the overall productivity of lipids because of reduction in overall growth rates.

The possible strategy to enhance lipid accumulation in microalgae without applying stress conditions is by appropriately modifying the genomes of suitable strains through genetic engineering [20]. Considerable research has been carried for genetically manipulating the genomes by overexpression or knock-out of genes of oleaginous microalgal species with high potential for biofuel production [9, 19, 20]. While such efforts demonstrate the feasibility of genetic engineering in improving microalgae for biodiesel, its effectiveness will be dependent on a deep understanding of the target genes and metabolic pathways responsible for lipid accumulation in microalgae. Variation of lipid content in closely related species and strains suggests that lipid metabolism is diverse in different taxa [16]. For this, it is importance to understand the genetic factors that contribute to this deviation. In spite of the available knowledge about fatty acid and triacylglycerol metabolism, mechanism behind the differential lipid accumulation in microalgal species and strains of same genera is not known.

Table 2.6 Genomic and transcriptomic studies to reveal lipid biosynthetic machinery for biodiesel production

Organism	Study	Reference
Chlamydomonas reinhardtii	Glycerolipid biosynthesis genes were annotated based on similarity to homologs from *Arabidopsis thaliana*	Riekhof *et al*. [127]
Cyanidioschyzon merolae	Fatty acid and lipid biosynthetic genes were studied by genomic and biochemical analysis	Sato and Moriyama [128]
Haematococcus pluvialis	Correlation between fatty acid synthesis and gene expression patterns under different stress conditions was investigated	Lei *et al*. [32]
Phaeodactylum tricornutum	Expression of genes during nutrient deprivation conditions and role of carbon fixation pathway was described for high lipid accumulation	Valenzuela *et al*. [132]
Chlamydomonas reinhardtii	Mechanism of oil accumulation under nitrogen deprivation was explored	Miller *et al*. [47]
Dunaliella tertiolecta	Transcriptomic investigation under nitrogen and osmotic-inducing stress to identify genes and pathways associated with biofuel precursor production	Rismani-Yazdi *et al*. [48]
Chlorella vulgaris	Comparative global transcriptomic and proteomic study under nitrogen stress for examining triacylglycerol biosynthetic pathways	Guarnieri *et al*. [21]
Neochloris oleoabundans	Transcriptomic analysis was performed in nitrogen replete and nitrogen limiting conditions for the identification of pathways linked with triacylglycerol production	Rismani-Yazdi *et al*. [133]
Chlamydomonas reinhardtii and Coccomyxa sp. C-169	Metabolic and gene expression changes subjected to nitrogen deprivation were examined under strictly photoautotrophic conditions	Msanne *et al*. [137]

Chlamydomonas reinhardtii	Genes and nitrogen-responsive regulatory components that are responsible in nitrogen starvation-induced triacylglycerol accumulation were discovered	Boyle *et al*. [138]
Chlamydomonas reinhardtii	Understanding of the molecular basis underlying increased triacylglycerol accumulation under nitrogen starvation condition in starchless mutants	Blaby *et al*. [139]
Nannochloropsis gaditana	Genome sequencing, gene annotation and transcriptional profiling to identify components for genetic intervention for improved biodiesel production	Corteggiani Carpinelli [140]
Nannochloropsis oceanica IMET1	Transcriptomic and lipidomic dynamics under nitrogen-replete and nitrogen-depleted conditions for revealing the mechanisms of oil production in microalgae	Li *et al*. [134]
Dunaliella tertiolecta	Transcriptome profiling of high lipid producing mutant to identify key pathways enhancing lipid yield	Yao *et al*. [141]
Coccomyxa subellipsoidea C-169	Global and collaborative regulation in response to CO_2 supplementation was revealed by transcriptome analysis	Peng *et al*. [136]
Dunaliella parva	Transcriptome data of nitrogen limiting and nitrogen sufficient culture conditions revealed pathways and genes important for biofuel production	Shang *et al*. [51]
Neodesmus sp. UTEX 2219-4	Transcriptome sequencing under stresse conditions elucidated the mechanism of photosynthate partitioning between fatty acid and starch biosynthesis	Chang *et al*. [142]

2.12 Transcription factors controlling lipid accumulation

Transcription factors bind to particular sequences in the promoter region of the respective genes and hence control the gene expression levels [143]. In order to enhance the productivity of target molecules, transcription factors identification is also significant as it dissecting and engineering the regulatory network. Among microalgae, computational identifications of transcriptional factors have been reported only for *Volvox carteri* and *Chlamydomonas reinhardtii*, *Galdieria sulphuraria* (red algae) [144, 145] and *Nannochloropsis* species [53, 146]. 41 differentially regulated transcription factors were identified from transcriptomic analyses of *Chlamydomonas reinhardtii* [49] and a set of transcription factors, cell-signaling proteins and cell-cycle regulators in *Chlorella vulgaris* [21] in nitrogen depleting condition. In another report, CrNRR1 transcription factor was found responsible for high lipid accumulation under nitrogen deprivation in *C. reinhardtii*. Knockout of this nitrogen responsive TF decreased the TAG content by 50 % [138]. Also, overexpression of DOF-type TF has elevated the lipid production by two-fold in *C. reinhardtii* [147]. Genomic and transcriptomic analysis of *Chlorella pyrenoidosa* identified 195 transcription factors. Seven families of lipid-related transcription factors including CAMTA, DOF, HSF, GATA, MYB-related, CO-like, Nin-like were also revealed [148].

The comprehensive review of literature therefore, highlights the following gaps in our understanding of molecular basis of lipid accumulation in microalgae:

1. No information is available on status of genes involved in fatty acid and triacylglycerol biosynthesis vis-à-vis development stages and genetically diverse species of the same genera of microalgae with varying lipid contents.
2. Fewer reports are available on transcriptional regulation of lipid biosynthesis in microalgae.

CHAPTER 3

MATERIALS AND METHODS

3.1 Culturing of *Scenedesmus* strains

Different isolates of microalgae belonging to *Scenedesmus quadricauda* and *Scenedesmus dimorphus* species were obtained from various geographical sites of Himachal Pradesh [149]. These strains were screened for their lipid content and high and low lipid content strains of *S. dimorphus* and *S. quadricauda* were maintained under optimal conditions of light and temperature and grown in BG11 medium in the glass house of Jaypee University of Information Technology, Waknaghat, H.P., India. Light intensity of 42 μmol. photons m−2 s−1 was provided to the cultures with a diurnal cycle of 16 h light and 8 h dark at temperature of 25 ± 2 °C. Cultures were continuously aerated with 0.22 μm filtered air through a mechanical pump.

3.2 Media prepration

BG11 medium consisted of four stock solutions:

Stock Solutions	**g/l**
Stock 1	
Na_2Mg EDTA	0.1
$(NH_4)_5$ Fe Citrate	0.6
Citric acid H_2O	0.6
$CaCl_2$ H_2O	3.6
Stock 2	
$MgSO_4$ $7H_2O$	7.5
Stock 3	
K_2HPO_4 $3H_2O$	4
Stock 4	
H_3BO_3	2.86
$MnCl_2$ $4H_2O$	1.81

$ZnSO_4\ 7H_2O$	0.222
$CuSO_4\ 5H_2O$	0.79
$CoCl_2\ 6H_2O$	0.05
$NaMoO_4\ 2H_2O$	0.391

These stock solutions were sterilized by autoclaving for 20 minutes and stored at 4ºC.

For preparing 1 litre of BG-11 medium:

Stock 1	10 ml
Stock 2	10 ml
Stock 3	10 ml
Stock 4	1 ml
Na_2CO_3	0.02 g
$NaNO_3$	1.5 g

The pH of BG-11 medium was 7.5 to 7.6 and sterilized by autoclaving for 20 minutes.

3.3 Establishment of axenic cultures of *Scenedesmus* species/strains

Axenic cultures of both the species were established by treating the cultures with antibiotics: penicillin, streptomycin and chloramphenicol following the protocol of Guillard [150].

100 mg of penicillin G (sodium or potassium salt), 50 mg of dihydrostreptomycin sulfate were dissolved in 10 ml of milli-Q water. 10 mg chloramphenicol was dissolved in 95% ethanol and then mixed to the solution. The antibiotic solution was sterilized by passing from 0.2 micron membrane filter and kept frozen until used in tubes.

3.3.1 Procedure

Six flasks, each for *S. dimorphus* and *S. quadricauda* strains, containing BG11 medium and a drop of bacterial test medium (i.e. Luria Broth) were taken. 1 ml algae culture from the exponential phase was inoculated to each flask. Different volumes of antibiotic solution were added to each of six flasks for providing varying concentration of antibiotics penicillin,

streptomycin and chloramphenicol. Hence, antibiotic volumes were: 0.08, 0.18, 0.35, 0.7, 1.4 and 2.1 to provide penicillin concentration in the range of 20-500 mg/L. The flasks were then placed in the glass house.

After 24 and 48 hrs, 1 ml of growing microalgal culture from flasks was transferred to flasks containing sterile BG11 medium to stop the antibiotic treatment. The flasks were then kept in glass house for microalgal growth.

3.3.2 Test for bacterial presence

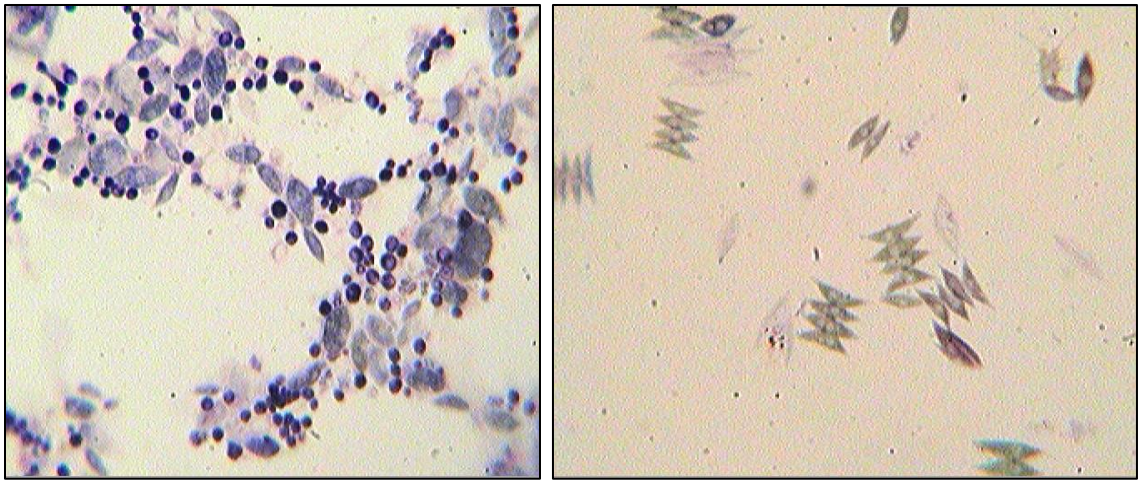
Three different media were used to test the presence of bacteria in the antibiotic treated cultures of *S. dimorphus* and *S. quadricauda*.

1. Nutrient Broth: 1g Nutrient Broth in 100 ml milli-Q water.
2. SST Medium: 1g glucose, 1g tryptone and 0.5 g yeast extract in 100 ml milli-Q water.
3. Peptone Glucose Medium: 1g glucose and 1g peptone in 1 L milli-Q water.

1.5% agar was added to each media and sterilized by autoclaving for 20 minutes. Petriplates were prepared and the antibiotic treated cultures were spreaded in each petriplate and incubated at 37 ^{0}C and examined after 1, 2, 6 days for any bacterial contamination (Figure 3.1).

Scenedesmus dimorphus

Before antibiotic treatment After antibiotic treatment



Scenedesmus quadricauda

Before antibiotic treatment After antibiotic treatment

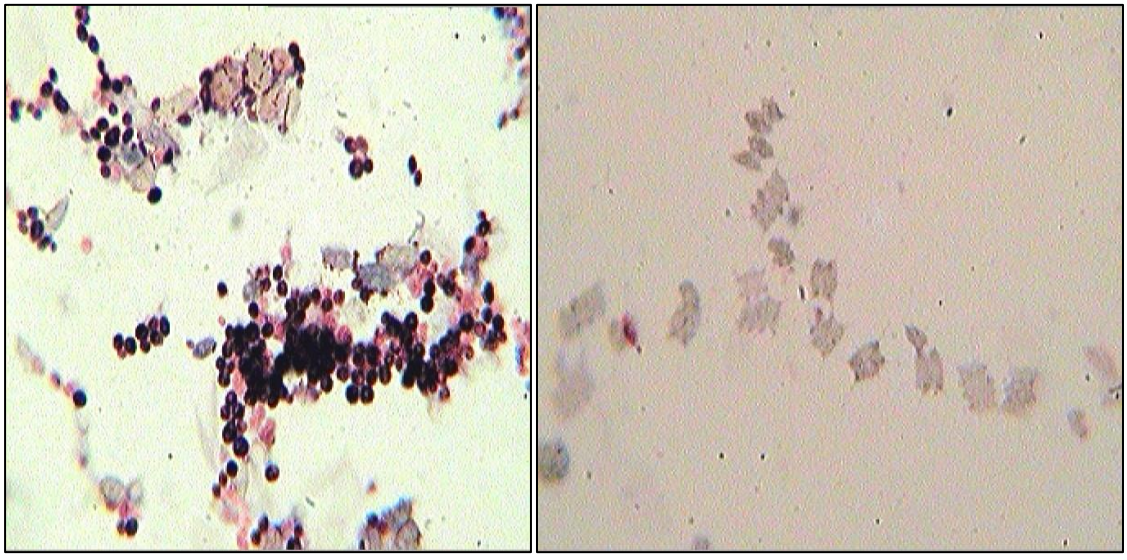


Figure 3.1 *S. dimorphus* and *S. quadricauda* cultures before and after antibiotic treatment

3.4 Growth and lipid content estimation in *Scenedesmus* strains

Bligh and Dyer method with minor modifications was utilized for the extraction of lipids from *Scenedesmus* strains [151]. 1000 ml cultures of each strain of *Scenedesmus* species were harvested at late stationary phase (LSP) and freeze-dried overnight using lyophilizer. Chloroform: methanol solution was taken in the ratio of 1:2 for suspending the lyophilized cells. The glass tubes containing the mixture of lyophilized *Scenedesmus* cells and chloroform methanol solution were immediately subjected to vortex. Then, the glass tubes were kept in sonicator for 5 min and left for shaking overnight. The next day, chloroform and distilled water were added to the glass tubes in equal amounts i.e. in ratio 1:1. The mixture was then vortexed and subjected to centrifugation for 10 min at 6000 ×g for the separation of phases. Three layers were obtained in the glass tubes after centrifugation. Top layer contains methanol and water, cell debris forms a middle layer and lipids being soluble in chloroform collected at the bottom as a dense layer. Lipids with chloroform were recovered from the centrifuge tube by inserting the micropipette by applying gentle positive pressure, so that other layers do not get into the micropipette. Whatman filter paper was put into the funnel and the obtained lipids with chloroform solution was passed through 2.5 cm thick layer of anhydrous sodium sulphate and collected into a pre-weighed container required for rotary evaporation. The chloroform was removed from the solution using a rotary evaporator under reduced pressure at 60 °C and the remaining lipids were analysed further. Weight of the lipids was recorded by placing the container with lipids on the weighing balance. Total lipid content was calculated as percentage of the total biomass (in % dry weight).

Lipid content of *Scenedesmus* strains was also estimated by fluorescence microscopy. Cells of *S. quadricauda* and *S. dimorphus* were stained with the fluorescent dye, BODIPY 505/515 (Invitrogen) in accordance with the protocol described by Cooper *et al.* [152]. Fluorescence of green colour was the measurement of the lipids present in the cells, which was determined by fluorescence microscope (Olympus BX53).

Growth of the *Scenedesmus* strains was measured by calculating the optical densities of the cultures at 730 nm using a spectrophotometer (ELICO SL- 159 UV–VIS). The biomass was calculated by weighing the lyophilized cells obtained after freeze drying.

3.5 Selection of strains

On the basis of lipid content, low lipid strain SD16 (5 %) and high lipid strain SD12 (26 %) from *S. dimorphus* and high lipid strain SQ19 (14 %) of *S. quadricauda* were considered for further experiments. For differential conditions, *Scenedesmus* cultures were harvested at day 10 for EXP, day 18 for ESP and day 24 for LSP after incubation. For stress conditions, nitrogen deficiency was induced by suspending the cultures in BG11 medium with KNO_3 nitrogen source having concentration of 2.5 mM and high salinity condition was provided by supplementing the medium with 180 mM NaCl. These cultures under stress conditions were harvested at late stationary phase i.e. after 24 days of incubation and stored at −80 °C for further use.

For transcriptome sequencing, cultures of *S. dimorphus* (26 %) and *S. quadricauda* (14 %) were harvested at stationary phase i.e. at day 20 after incubation and stored at -80°C for further use.

3.6 Genomic DNA and RNA extraction

CTAB extraction protocol was utilized for the isolation of Genomic DNA from the cultures of *Scenedesmus* strains under study [153]. RNA extraction was performed by RaFlex RNA isolation kit (GeNei™) following the manufacturer's instructions. Isolated DNA and RNA were analysed for their quality by agarose gel electrophoresis using 0.8% and 1% (w/v) ethidium bromide-stained agarose gel for DNA and RNA, respectively. Further, the absorbance spectrum DNA and RNA was checked at wavelengths 260 nm and 280 nm.

3.7 Cloning and sequencing of genes in *S. dimorphus*

The nucleotide and protein sequences of five genes from different microalgal and plant species i.e. omega fatty acid desaturase (FAD), enoyl-ACP reductase (ER), palmitoyl desaturase (ACPD), stearoyl-ACP desaturase (SAD) and fatty acyl-ACP thioesterase (FATA) of fatty acid biosynthetic pathway were retrieved from the NCBI (http://www.ncbi.nlm.nih.gov). The obtained sequences were aligned using ClustalW for the identification of conserved regions. These conserved regions were then utilized for primer designing using Primer3 software and the primers were amplified on the genomic DNA and cDNA of *S. dimorphus*. Requirements for PCR amplification were 30 ng genomic DNA and cDNA separately with varying amounts (10 mM or 25 mM) of primer pairs, dNTPs, Mg2+ and Taq DNA polymerase. Amplification programs

included initial denaturation at 94 °C for 4 min, denaturation of 35 cycles at 94 °C for 1 min, annealing at range of temperatures i.e. 50–65 °C for 55 s, extension at 72 °C for 1 min and a final extension at 72 °C for 5 min. For analyzing the results, 10 µl of each PCR product was mixed with 2 µl of 6× gel loading dye (0.2% bromophenol blue, 0.2% xylene cyanol dye and 30% glycerol) and electrophoresed in a 1.2% agarose gel prepared in 1× tris acetate-EDTA (TAE) buffer. The agarose gels with PCR products were examined in a gel documentation system (Alpha Imager EP, Alpha Innotec Corp., USA).

For cloning, the PCR products were ligated with pGEM-T vector (Promega) and then transformed into competent *E. coli* DH5α cells. The transformed cells were then spreaded into glass plates containing Luria Broth with agar and kept at 37°C overnight. Next day, the appeared colonies were observed and the white colonies were taken for further analysis. Plamid isolation of the fully grown colonies in media was performed, which were utilized for sequencing. The obtained sequences were subjected to BLAST (http://www.ncbi.nlm.nih.gov/Blast) for homology search and the positive results were then used for primer designing using Primer3. These primers were further utilized for performing gene expression analysis. Methodology adopted for cloning the genes of *S. dimorphus* is summarized in Figure 3.2.

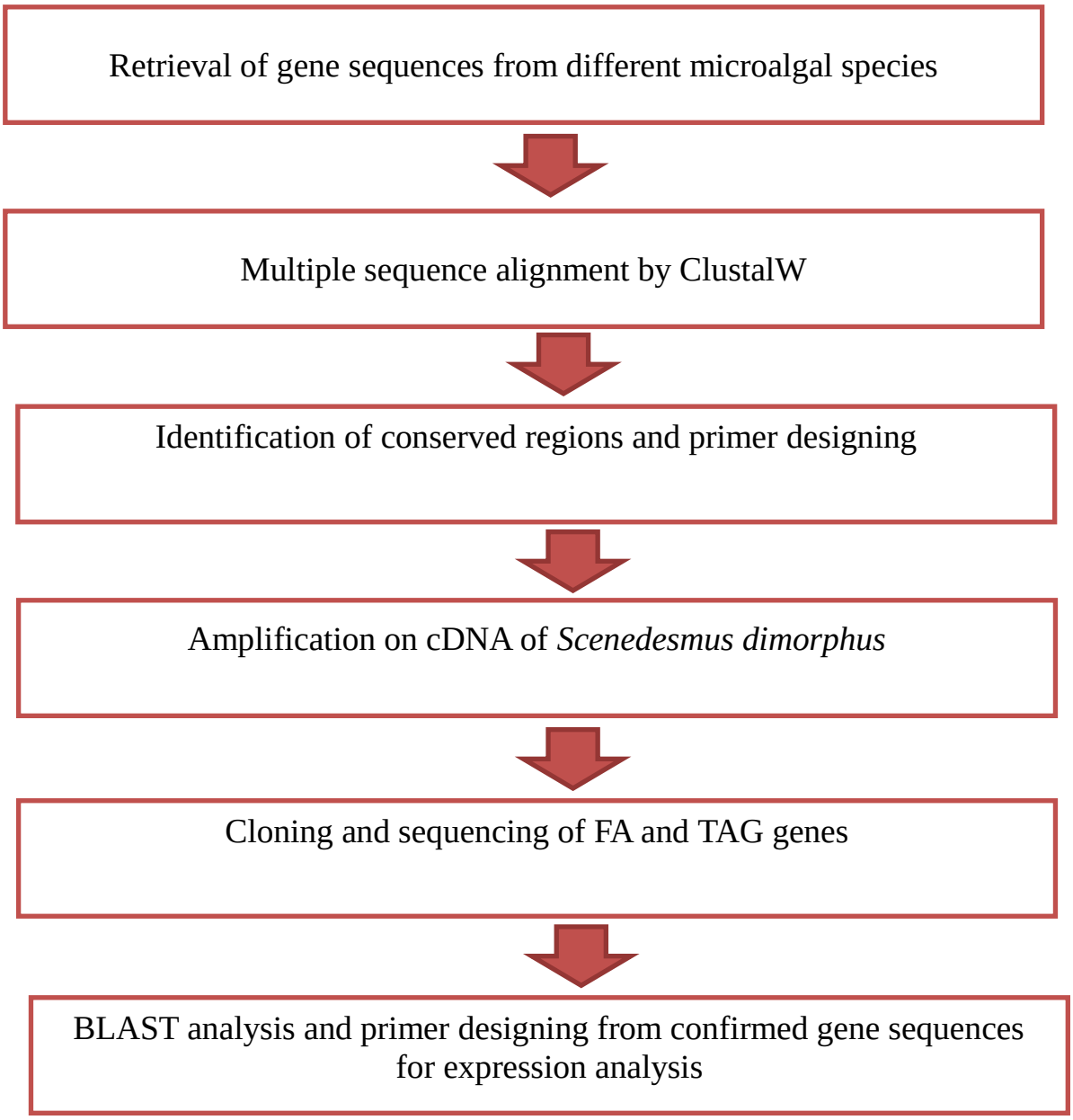


Figure 3.2 Cloning of FA and TAG pathways genes through comparative genomics

3.8 Mining of genes from transcriptome of *S. dimorphus*

The transcriptome of S. *dimorphus* was mined for the identification of nine genes of FA and TAG biosynthetic pathway. These genes include malonyl transferase (MT), β-ketoacyl-ACP reductase (KAR), acetyl-CoA carboxylase (ACC), β-ketoacyl-ACP synthase I (KAS I), β-ketoacyl-ACP synthase II (KAS II), acyl carrier protein (ACP), β-ketoacyl-ACP synthase III (KAS III), palmitoyl thioesterase (PT) and β-hydroxyacyl ACP dehydratase (HAD) of fatty acid biosynthesis and phosphatidic acid phosphatase (PAP), glycerol-3-phosphate acyltransferase

(GPAT), diacylglycerol acyltransferase (DGAT) and lysophosphatidic acid acyltransferase (LPAAT) of triacylgycerol biosynthesis. Primers for these genes were designed using Primer3 and were further amplified on the cDNA of *S. dimorphus*.

3.9 RT-qPCR analysis

First-strand cDNA synthesis from RNA (5 μg) template was performed using Verso cDNA synthesis kit (Thermo Scientific, USA) as per manufacturer's instructions. Gene expression investigation was performed by RT-qPCR using gene specific primers (Table 3.2) in triplicate on a RT-qPCR system (CFX96, Bio-Rad Laboratories; Hercules CA) with iScript RT-qPCR kit (Bio-rad). The protocol for RT-qPCR was as follows: initial denaturation at 94°C for 5 min, followed by 35 cycles each of denaturation at 94°C for 20s, annealing at 50–55°C for 30s, followed by one elongation step at 72°C for 20s.

26S rRNA and actin genes were taken as an internal reference for the normalization of the expression data as described by Vandesompele [154]. For comparing the gene expression levels between *Scenedesmus* strains, lowest lipid content strain SD16 was considered as calibrator strain. In case of growth stages, stage where minimum lipid accumulation takes place i.e. EXP was considered as calibrator condition and in stress conditions, LSP was taken as calibrator.

Relative fold expression was calculated from the Cq values using the comparative Ct (ΔΔCt) method as described by Schmittgen and Livak [155].

Table 3.1 Primer sequences of FA and TAG pathway genes used for expression analysis

S. No.	Gene	Gene Abb.	Forward Primer	Reverse Primer	Annealing temp [°C]
1	Acetyl-CoA Carboxylase	ACC	GGGTATCTTGAAGTTTGGGT	TCCGATACTTGATTTCCACC	52
2	Malonyl Transferase	MT	CTGTATTCACCCAGACTCAG	GTTACGACCTGCTGGAAC	53.2
3	Acyl Carrier Protein	ACP	CAATCCCCACCTACAGCA	CATTACAACGATAGAACACGAA	52.5
4	β-ketoacyl-ACP synthaseI	KASI	CAAGATCATGGGTAAGGCAC	GTACTTCTTGTCAATGGCTG	52.2
5	β-ketoacyl-ACP synthase II	KASII	TGACCCCGACAAATTCTATG	CCACCATGGTATACTTGAGG	52.2
6	β-ketoacyl-ACP synthase III	KASIII	CGACTTGATATTGCTGGCTA	CTTGAAGGTGCCTGTTTTG	52.1
7	β-ketoacyl-ACP reductase	KAR	CCTTAACATCTTTGAGACCA	AAAGTCCTTCTTGACCTTCT	52.3
8	β- hydroxyacyl ACP dehydratase	HAD	TCAATGACCAGTTCTTCAAT	TGTTGTACTTGGTGACCTC	52.2
9	Enoyl-ACP reductase	ER	AACATCTTTGAGACCAGCTT	GTCCTTCTTGACCTTCTCC	52

10	Palmitoyl Thioestrase	PT	CCTGAAGCAGTACATGGATT	TAGAGGGGCTGGTCTTTAAG	52.8
11	Fatty Acyl-ACP Thioestrase	FATA	TAGAGACGTATTTTGCGGAG	CATACCCATCAGAGCCAAG	52
12	Stearoyl-ACP Desaturase	SAD	GTAACCATCCAGAACCTCAT	GTCCATGATCCTGCAGTAC	52
13	Palmitoyl Desaturase	ACPD	GACAACAGCAGGAACAATTG	ATCTTTGATCCTGAAACGCT	52
14	Omega fatty acid desaturase	FAD	CTTGAGTTCCCCTTTGACC	CCAGACCACATTGATCCAG	53.2
15	Glycerol-3-phosphate acyltransferase	GPAT	CTGTTCATCCTGCTGTTCTT	TAGCTGGTCAGGATGATGTA	52.5
16	Lysophosphatidic acid acyltransferase	LPAAT	TCTTCCTTAGGCGGTATGTA	CACCTTCTCCCAGTTGTAAA	52.2
17	Phosphatidic acid phosphatase	PAP	ACGGCTATCAGATCATGTTC	CATGCAATCTTGAACTCGTG	52
18	Diacylglycerol acyltransferase	DGAT	TCGTCAACACCTTCAACAA	GTGACAAAACCTGGGATGAT	52

3.10 Statistical analysis

3.10.1 Principal component analysis (PCA)

PCA highlights the similarities and differences in the data and correlates the data on the basis of similar pattern. PCA is mainly used for data of high dimension where particular patterns are hard to find. Hence, this analysis is a powerful tool as it reduces the dimension of data and represents it in graphical form, which is easy to correlate [156]. PCA was executed to determine the correlation of expression of eighteen genes from fatty acid and triacylglycerol biosynthetic pathway in three strains of *Scenedesmus* species i.e. SD16, SD12, SQ19 and their evaluation in growth stages. Association among the genes and their involvement in experimental conditions was evaluated.

3.10.2 Heat map

The heat map was constructed from the expression values of RT-qPCR analysis of FA and TAG biosynthetic pathway genes in different stress conditions and growth stages among three lipid content contrasting strains of *Scenedesmus* using GenEx software (V 1.1).

3.11 *De novo* transcriptome assembly of *S. dimorphus* and *S. quadricauda*

The methodology followed for *de novo* whole transcriptome analysis of *S. dimorphus* and *S. quadricauda* is summarized in Figure 3.3. Sequencing of the transcriptomes was performed by NextSeq 500 platform and the data was generated in the form of raw reads. The raw reads were further filtered by Trimmomatic (v 0.30) with quality value QV > 20. Other contaminants such as adapters were also trimmed from the data. QC passed reads from the two transcriptomes were termed as high quality reads (HQ reads). HQ reads were assembled by *de novo* aaproach using Trinity resulting in the generation of transcripts. Assembled transcripts were further subjected to CD-HIT-EST run and unigenes were obtained.

ORF Predictor (http://proteomics.ysu.edu/tools/OrfPredictor.html) with default parameters was utilized for the prediction of CDSs from unigenes.

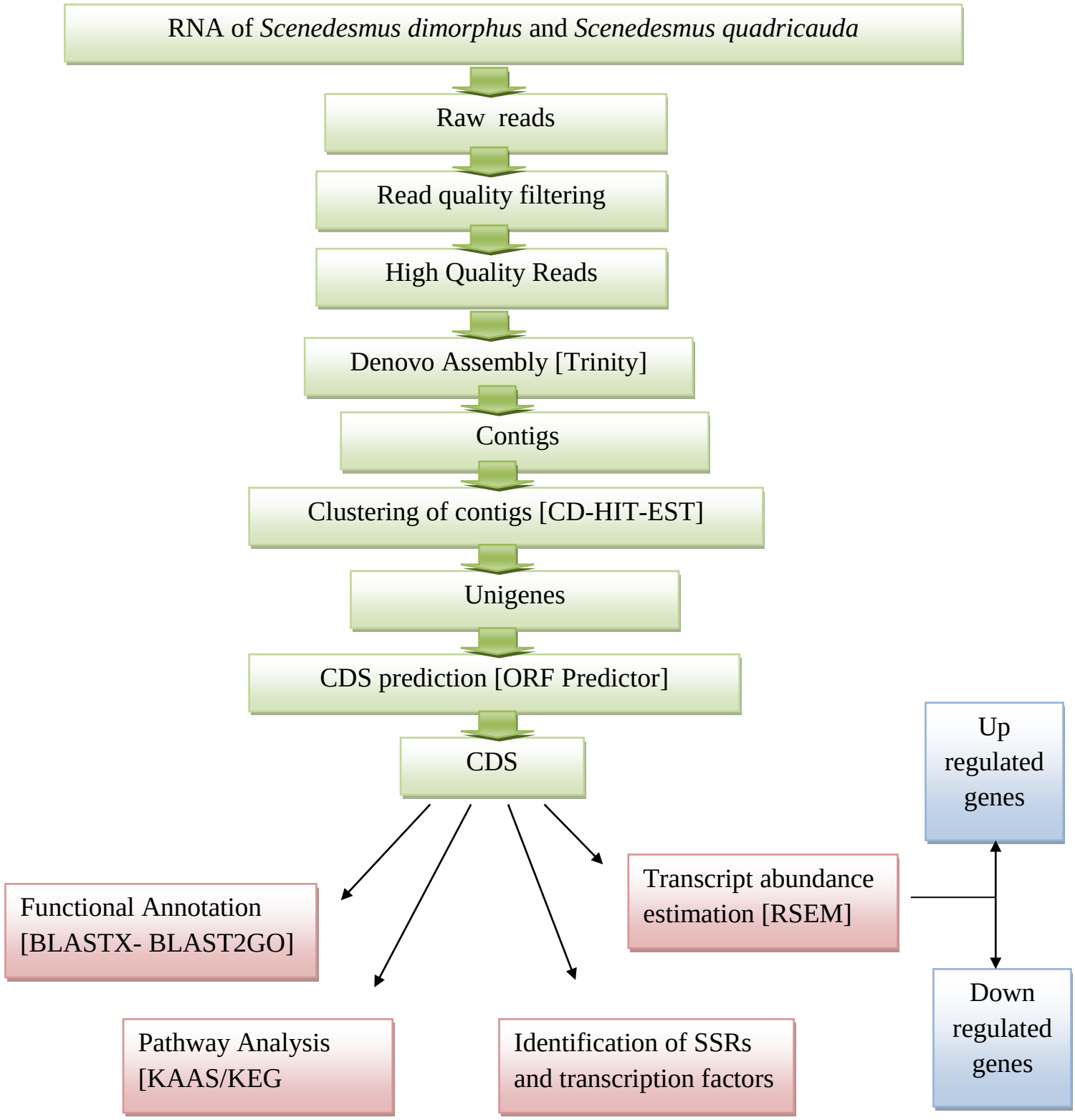


Figure 3.3 Flow chart depicting the methodology adopted for transcriptome sequencing and analysis of *S. dimorphus* and *S. quadricauda*

3.12 Functional annotation of predicted CDSs

Predicted CDSs from transcriptomes of *S. quadricauda* and *S. dimorphus* were functionally annotated by subjecting the CDSs file to BLASTx search against non-redundant (nr) database of NCBI using E-value cut-off of 10^{-6}.

3.13 GO mapping and CDS distribution

Gene ontology assignments were used to perform annotation of predicted CDSs on the basis of their function. GO mapping annotates the gene sets into three main domains: biological process, molecular function and cellular component. Accession IDs resulted from the BLASTx search of CDSs were searched directly in the gene product table of GO database, which has given the output specifying all the annotated nodes comprising GO functional groups such as cellular component, biological process and molecular function.

3.14 Pathway analysis

KAAS is a tool of KEGG database, which was utilized annotating the CDSs of *S. dimorphus* and *S. quadricauda* by performing the homology search BLAST against KEGG GENES database. The BBH (bi-directional best hit) option was used to assign gene's IDs and KO terms. These KO numbers were used for pathway mapping by KEGG Orthology database (http://www.genome.jp/kegg/ko.html).

3.15 Transcript abundance estimation

Transcript abundances of the *de novo* assembled sequences was estimated using RSEM approach, in which the RNA-Seq reads are mapped with the assembled transcriptome for the quantification of transcript abundances of sequences. RSEM calculates transcript abundance in two steps. Firstly, command rsem-prepare-reference is utilized for generating a set of reference transcript sequences and preprocessing them for the second step. Second, and rsem-calculate-expression command align the set of RNA-Seq reads to the reference transcripts. The resulting alignments from these two commands are used to calculate the abundances of sequences and their credibility intervals [157]. Transcript abundances were estimated as TPM and FPKM

values, which measure the expression of even poorly expressed transcripts using fragment count. Hence, this is a sensitive approach to detect the expression level of transcripts.

3.16 Identification of SSRs from transcriptome data

Simple sequence repeats (SSRs) were identified from the transcriptomes *S. quadricauda* and *S. dimorphus*. Perl script MISA available at "http: //pgrc.ipkgatersleben.de/misa/misa.html" was used for the identification of SSRs and determination of the frequency of SSRs. The FASTA file was allowed to search the size and type of constituting SSRs. The minimum repeat unit was set as ten for mononucleotides, six for dinucleotides and five for all the higher order motifs including tri-, tetra- and penta-nucleotides.

3.17 Identification of transcription factors from transcriptomes of *S. dimorphus* and *S. quadricauda*

Predicted CDSs from *S. quadricauda* and *S. dimorphus* were subjected to BLAST search against PlantTFDB (http://planttfdb.cbi.pku.edu.cn) with E-value cut-off of 10^{-5}. Green microalgae *Chlamydomonas reinhardtii* was taken as reference. Conserved Domain Database was utilized for the identification of conserved domains in transcription factors, which is available at NCBI with url "http: //www.ncbi.nlm.nih. gov/Structure/cdd/wrpsb.cgi". Only the top bit-scoring significant hit for each sequence was considered. Further, literature based mining of lipid-biosynthesis related transcription factors was done from the previous reports and identified transcription factors were categorized into lipid-biosynthesis related transcription factors in S. *quadricauda* and *S. dimorphus* (Figure 3.4).

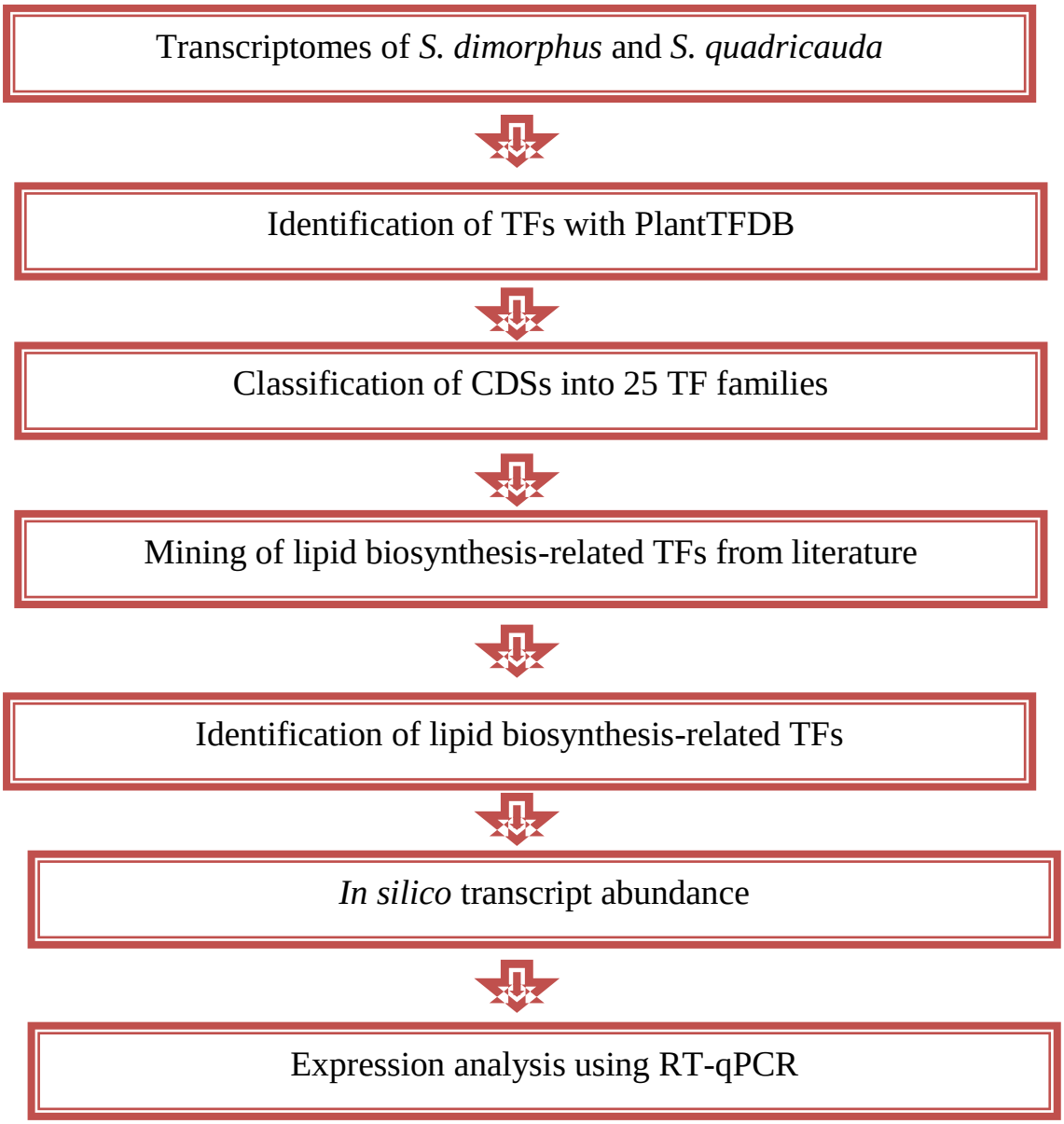


Figure 3.4 Methodology followed for the identification of lipid biosynthesis related transcription factors in *S. dimorphus* and *S. quadricauda*

3.18 Gene expression validation through RT-qPCR

Quality of isloted RNA was verified by measuring the absorbtion in spectrophotometer. First-strand cDNA synthesis from RNA (2 μg) template was performed using Verso cDNA synthesis kit (Thermo Scientific, USA) as per manufacturer's instructions. Gene expression investigation was performed by RT-qPCR using gene specific primers (Table 3.2) in triplicate on a RT-qPCR system (CFX96, Bio-Rad Laboratories; Hercules CA) with iScript RT-qPCR kit (Bio-rad). The protocol for RT-qPCR was as follows: initial denaturation at 94°C for 5 min, followed by 35 cycles each of denaturation at 94°C for 20s, annealing at 50–55°C for 30s, followed by one elongation step at 72°C for 20s.

18S rRNA and actin genes [158, 159] were taken as an internal reference for the normalization of the expression data as described by Vandesompele [154]. For comparing the gene expression levels between *Scenedesmus* species, relative fold changes were calculated from Cq values generated in *S. dimorphus* and *S. quadricauda* samples in stationary phase. While for measuring the relative fold change in expression level of CDS encoding WRI1 transcription factor in three growth stages, exponential phase (EXP), where minimum lipid accumulation takes place was kept as a calibrator. Relative fold expression in early stationary phase (ESP) and late stationary phase (LSP) were calculated with respect to EXP from obtained Cq values using the comparative Ct (ΔΔCt) method as described by Schmittgen and Livak [155]. All experiments were performed in triplicates and repeated twice.

Table 3.2 Primer sequences for expression analysis by RT-qPCR

S. No.	Gene	Gene Abb.	Forward Primer	Reverse Primer	Annealing temp [°C]
1	Enolase	ENO	AGGTGTACCACAACCTCAAG	ATACATCTTGTCCTCGGTGA	55.2
2	Glyceraldehyde3-phosphate dehydrogenase	GAPDH	CGTCAGAGACCTACCTCAAG	TCGTACCAGGCGTATATCTT	54.9
3	Acetyl-CoA synthetase	ACOS	GCTGGTGTATGAGAACACCT	TGGCTGTGTACACCATGTAG	55.3
4	Pyruvate dehydrogenase	PDH	ATCATGGAGATGGACACCTA	ACCTTCTTCTCCATGGTCTT	55.2
5	ATP citrate lyase	ACL	CTCAATGACAAAGGTGTTCA	AATCAAACAAAGCTGGTTGT	54.6
6	Citrate synthase	CS	CTGAAGGTCTCTTCTGGTTG	GCCAGTACTCCTTCTTGTTG	54.9
7	Glycerol kinase	GK	GCTGAGTCGGAGGAGATAG	GGAAGCAGATAGCTTCCAG	55

3.19 Heat map analysis

Multiple experiment viewer (MEV v4.9) was uitilized for the construction of heat map which represents the linkage hierarchical clustering of top 100 differentially expressed genes i.e. top 50 up-regulated and 50 down- regulated genes. Pearson's uncentered correlation distance was calculated and on the basis of this value log-transformed and normalized value of genes was estimated. Heat map was created using both the log-transformed and normalized value of genes as well as based on the complete linkage method.

CHAPTER 4

RESULTS

4.1 Differential conditions for gene expression analysis in *Scenedesmus* strains

For studying the gene expression of fatty acid and triacylglycerol biosynthetic pathway, three strains of *Scenedesmus* species were taken. SD16 and SD12 strains with lipid content of 5%and 26% dry weight, respectively were taken from *Scenedesmus dimorphus* species to study the intraspecies variation in gene expression. For interspecies examination, SQ19 with maximum lipid content of 14% was taken from *Scenedesmus quadricauda* and compared with SD12 of *S. dimorphus* (Figure 4.1).

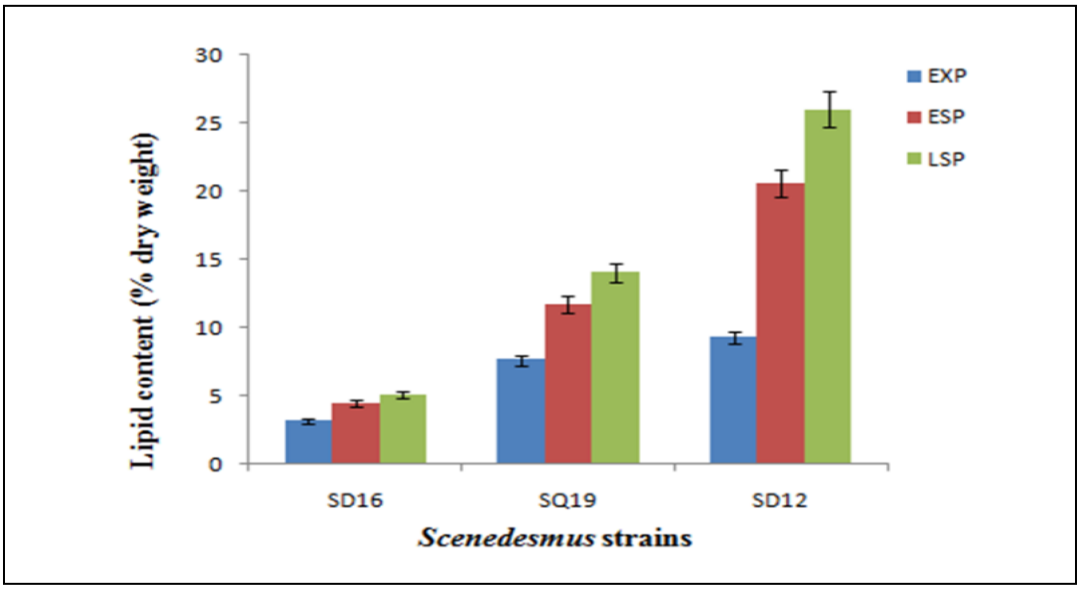


Figure 4.1 Lipid content of three contrasting strains taken for study; *S. dimorphus* (SD16 (5%), SD12 (26%)), *S. quadricauda* (SQ19 (14%))

Microalgal growth occurs in different developmental phases i.e. lag phase, exponential phase, stationary phase and death phase. Lag phase corresponds to the period when microalgae are struggling for adaptation in the new environmental conditions of fresh media. Once the microalgae get adapted to the fresh environmemt, it starts growing exponentially with time. This phase is called exponential phase. At a point of time, one or more requirements for growth starts depleting and the net growth of microalgae remains to be constant, which is known as stationary

phase. Death phase is when the cells of microalgae get damaged due to non-avalilability of nutrients and conditions in the media [160]. Growth of the three strains under study was measured by the optical density estimation of cultures at 730 nm. It was observed that *Scenedesmus* strains start growing from second day after incubation and the growth was maximum at day18, after that there was no net growth (Figure 4.2). Therefore, three growth stages were taken to study the differential gene expression i.e. exponential phase (EXP) at 10th day, early stationary phase (ESP) at 18th day and late stationary phase (LSP) at 24th day. Also two stress conditions i.e. nitrogen deficiency (ND) and high salinity (HS) at late stationary phase (LSP) were selected for differential expression analysis unser stress conditions.

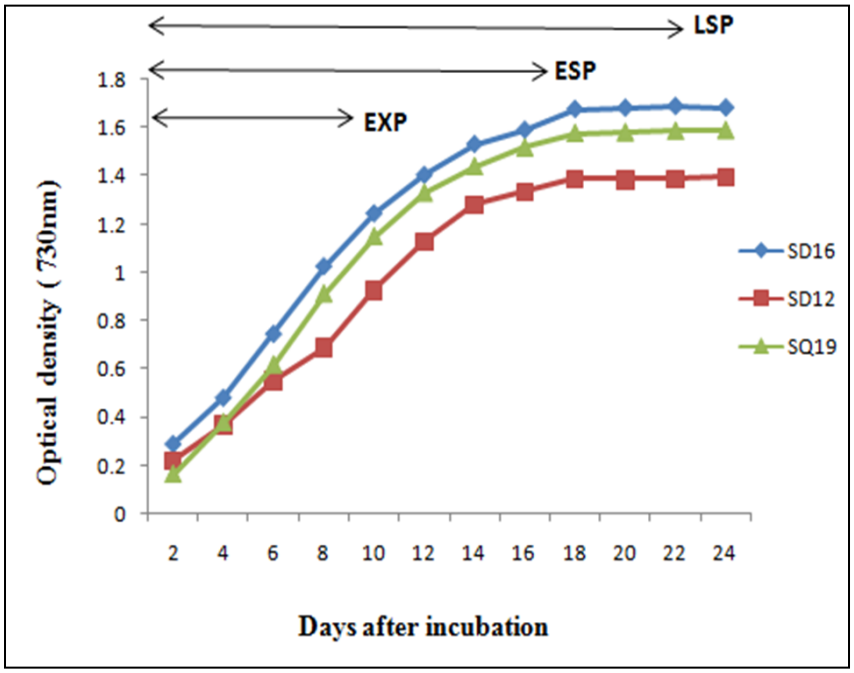


Figure 4.2 Growth of two *S. dimorphus* strains (SD16, SD12) and one *S. quadricauda* strain (SQ19) reflected through optical density at 730 nm cultured in BG11 medium, Abbreviations: EXP- Exponenetial Phase, ESP- Early Stationary Phase, LSP- Late Stationary Phase

4.2 Gene cloning and sequencing from fatty acid pathway in *S. dimorphus*

Comparative genomics approach was utilized to design degenate primers of five genes encoding enzymes omega fatty acid desaturase (FAD), enoyl-ACP reductase (ER), palmitoyl desaturase (ACPD), stearoyl-ACP desaturase (SAD) and fatty acyl-ACP thioesterase (FATA) of fatty acid biosynthetic pathway. Nucleotide sequences of the respective genes from different microalgae

and oleaginous plant species were retrieved from GenBank and the conserved regions generated from aligning the sequences were exploited for primer designing. The amplified partial sequences of the respective genes from *S. dimorphus* were subsequently cloned and sequenced. The new sequences obtained from sequencing were annotated for biological function and the positive results were further exploited for primer designing primers to study the gene expression in *Scenedesmus* strains. The size of the new partial sequences was in the range of 200-459 bp, which were submitted to GenBank under accession numbers KT203348-KT203352 (Figure 4.3). Primers of remaining genes of fatty acid and tricaylglycerol biosynthetic pathway were designed from the sequences retrieved from the transcriptome of *S. dimorphus*.

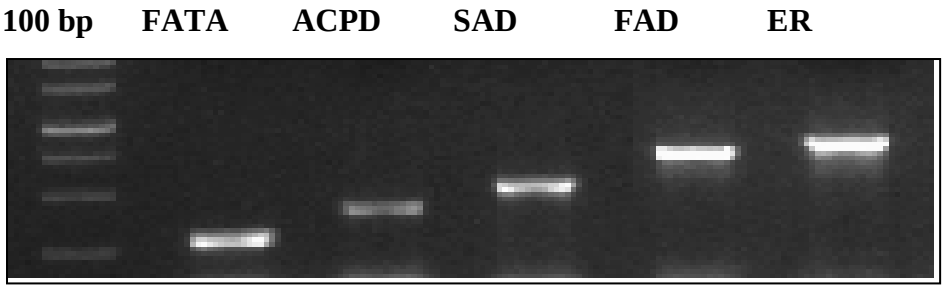


Figure 4.3 Amplication of five genes from fatty acid biosynthetic pathway on cDNA of *S. dimorphus*

4.3 Expression analysis of FA and TAG genes among three lipid content contrasting *Scenedesmus* strains

Firstly, the gene expression investigation of 14 genes from FA biosynthesis and 4 genes of TAG biosynthesis was performed among three strains SD16, SQ19 and SD12, differing in the lipid content was performed. Most of the genes showed increased expression in highest lipid strain SD12, signifying the potential role of these genes for increasing the lipid content (Figure 4.4). On comparing the expression levels of the genes among three strains, it was found that genes *ACP*, *ACC*, *KASII*, *KASI*, *FATA*, *PAP*, *LPAAT* and *DGAT* showed remarkably high fold difference (up to 50 folds) in SD12 as compared to SD16 whereas it was up to 25 folds high as compared to SQ19, which is consistent with the lipid content of *Scenedesmus* species (Table 4.1).

Table 4.1 Relative ratios of expression of lipid biosynthetic pathway genes across three lipid content contrasting strains of *Scenedesmus*; SQ19: *S. quadricauda* 19; SD12: *S. dimorphus* 12; SD16: *S. dimorphus* 16 at late stationary phase (LSP)

Genes	**Ratio of transcripts of SD12 to that of SQ19**	**Ratio of transcripts of SD12 to that of SD16**
ACC	1.03	1.17
MT	1.21	0.63
ACP	15.66	28.18
ACPD	5.00	2.15
KASI	1.55	2.24
KASII	13.29	26.97
KASIII	13.43	1.76
ER	4.66	3.64
KAR	13.41	3.15
HAD	3.05	1.01
PT	0.63	0.69
SD	81.40	13.02
FAD	0.64	0.09
FATA	5.45	6.70
GPAT	32.70	4.31
LPAAT	3.61	32.11
PAP	25.97	49.21
DGAT	20.34	25.63

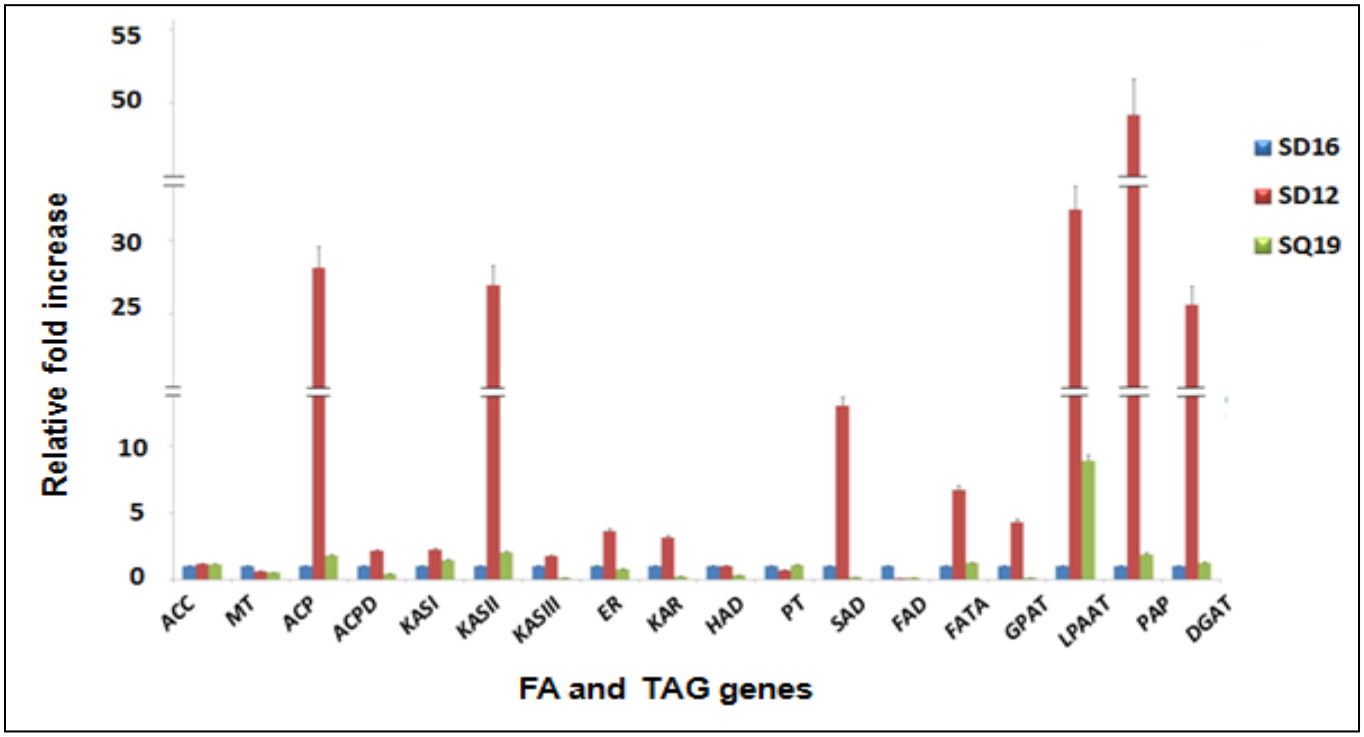


Figure 4.4 Gene expression analysis of pathway genes across three lipid contrasting strains of *Scenedesmus*; SQ19: *S.quadricauda*19; SD12: *S.dimorphus*12; SD16: *S.dimorphus*16 at late stationary phase (LSP). Enzyme abbreviations: ACC- Acetyl-CoA Carboxylase; MT- Malonyl Transferase; ACP- Acyl Carrier Protein; KASI- β-ketoacyl-ACP synthase I; KASII- β-ketoacyl-ACP synthase II; KASIII- β-ketoacyl-ACP synthase III; KAR- β-ketoacyl-ACP reductase; HAD- β-hydroxyacyl ACP dehydratase; ER- Enoyl-ACP reductase; PT- Palmitoyl Thioestrase; FATA- Fatty Acyl-ACP Thioestrase; SAD- Stearoyl-ACP Desaturase; ACPD- Palmitoyl Desaturase; FAD- Omega fatty acid desaturase; GPAT- Glycerol-3-phosphate acyltransferase; LPAAT- Lyso-phosphatidic acid acyltransferase; PAP- Phosphatidic acid phosphatase; DGAT- Diacylglycerol acyltransferase

4.4 Gene expression vis-à-vis variation in lipid accumulation

Further, the study was extended to examine the changes in the expression levels of genes from FA and TAG biosynthesis. Temporal expression analysis of *Scenedesmus* strains was carried out in three stages of growth i.e. EXP, ESP and LSP and exponential phase was taken as control. Gene expression patten of majority of FA and TAG genes was consistent among three strains and in the three stages (Figure 4.5). Genes *KASII, ACP, KASIII, KAR, ER, HAD, SAD, FATA, PT, PAP, LPAAT, DGAT* exhibited maximum expression in LSP (where maximum lipid accumulation takes place). This revealed the association of these genes with high lipid content. Expression of genes *GPAT* and *ACPD* was maximum at EXP and ESP growth stages,

respectively. Expression of some genes i.e. *KASI, ACC, FAD, MT* was not consistent among the three strains. It was observed from the temporal gene expression that majority of genes showed consistent expression pattern, however the level of expression was different in the three strains, which is in accordance with the lipid content of the strains. Genes *ACP, ACC, KASI, KASII, KASIII, ACPD, FATA, ER, PAP,LPAAT, DGAT* showed high fold change in expression between SD16 and SD12 than SQ19 and SD12 in both stages of growth i.e. ESP and LSP. This demonstrates their role in enhancement of lipid accumulation (Table 4.2).

Table 4.2 Relative ratios of expression of FA and TAG biosynthetic genes in growth stages, EXP, ESP and LSP for three lipid contrasting strains of *Scenedesmus*

ESP			LSP		
Genes	**Ratio of transcripts of SD12 to that of SQ19**	**Ratio of transcripts of SD12 to that of SD16**	**Genes**	**Ratio of transcripts of SD12 to that of SQ19**	**Ratio of transcripts of SD12 to that of SD16**
ACC	1.165847	2.34145	ACC	0.455831	2.259899
MT	3.158016	1.979535	MT	0.918829	0.152777
ACP	2.119837	4.12574	ACP	4.113866	5.249069
ACPD	1.9873	3.038796	ACPD	1.792391	3.758286
KASI	1.452041	3.691868	KASI	1.510773	3.362407
KASII	1.361414	2.176958	KASII	1.571733	2.292889
KASIII	0.722705	0.927209	KASIII	0.623812	1.315229
ER	0.652883	1.165892	ER	1.830617	4.85202
KAR	0.97465	0.870652	KAR	1.963867	0.769137
HAD	1.51275	0.860551	HAD	4.995878	2.129089
PT	9.279949	18.102	PT	2.357592	1.545177
SD	0.746197	0.620681	SD	10.97932	7.133411
FAD	1.141002	0.063765	FAD	0.663335	0.012639
FATA	2.197343	2.559641	FATA	2.59701	2.959178
GPAT	1.561404	1.1125	GPAT	2.045455	1.403688
LPAAT	2.387501	4.196189	LPAAT	2.089685	2.36318
PAP	1.17529	1.202885	PAP	1.583323	2.133236
DGAT	1.121884	1.394998	DGAT	1.357692	1.524607

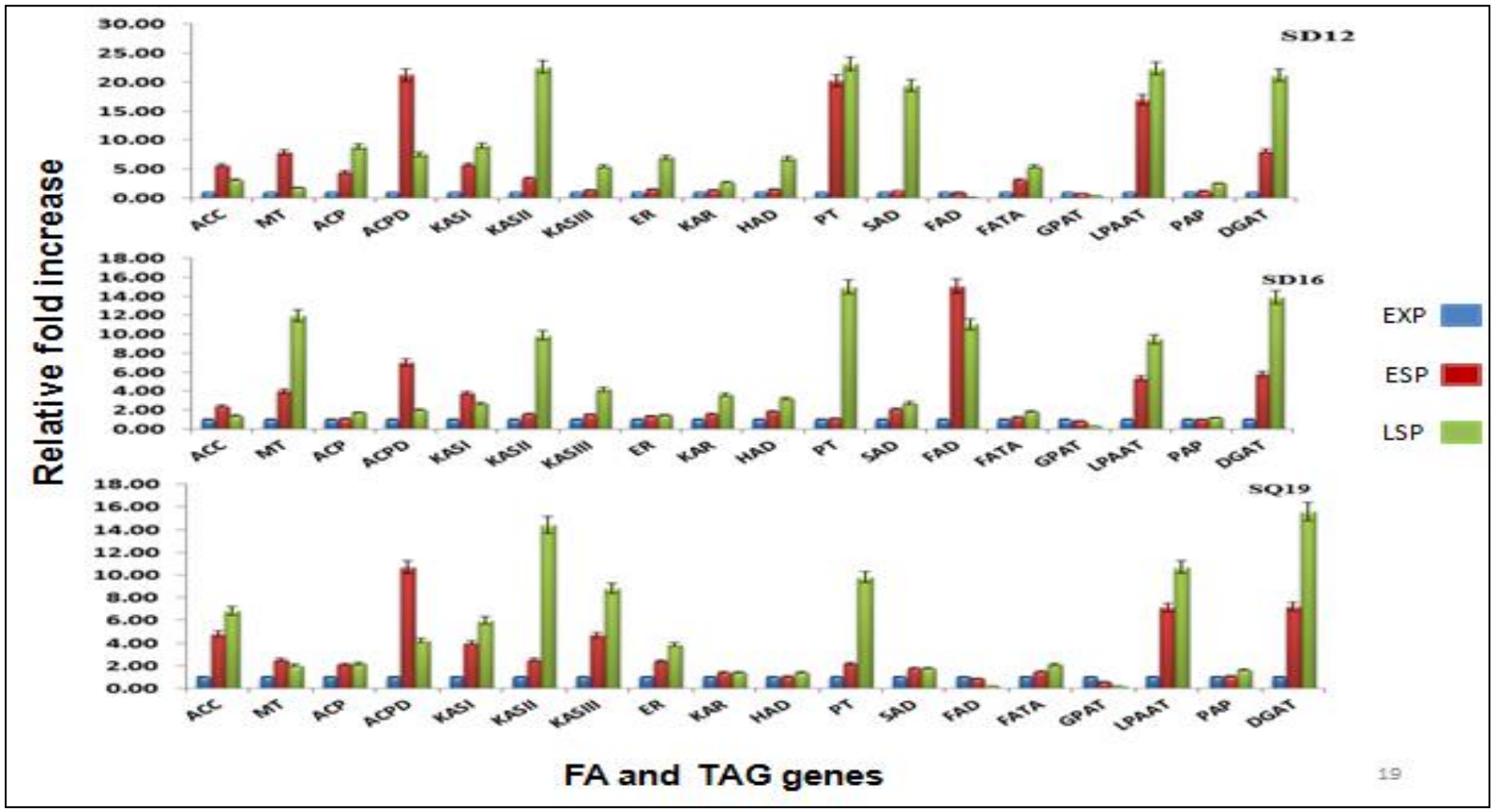

Relative fold increase
SD12
SD16
SQ19
EXP
ESP
LSP
ACC
MT
ACP
ACPD
KASI
KASII
KASIII
ER
KAR
HAD
PT
SAD
FAD
FATA
GPAT
LPAAT
PAP
DGAT
FA and TAG genes
19

Figure 4.5 Relative expression of FA and TAG biosynthetic genes in growth stages, ESP and LSP w.r.t EXP for three lipid contrasting strains of *Scenedesmus*; SD12: *S.dimorphus*12; SQ19: *S.quadricauda*19; SD16: *S.dimorphus*16. Abbreviations: EXP- Exponenetial Phase, ESP- Early Stationary Phase, LSP- Late Stationary Phase, ACC- Acetyl-CoA Carboxylase; MT- MalonylTransferase; ACP-Acyl Carrier Protein; KASI- β-ketoacyl-ACP synthase I; ; KASII- β-ketoacyl-ACP synthase II; KASIII- β-ketoacyl-ACP synthase III; KAR- β-ketoacyl-ACP reductase; HAD- β-hydroxyacyl ACP dehydratase; ER- Enoyl-ACP reductase; PT- Palmitoyl Thioestrase; FATA- Fatty Acyl-ACP Thioestrase; SAD- Stearoyl-ACP Desaturase; ACPD- Palmitoyl Desaturase; FAD- Omega fatty acid desaturase; GPAT- Glycerol-3-phosphate acyltransferase; LPAAT- Lysophosphatidic acid acyltransferase; PAP- Phosphatidic acid phosphatase; DGAT- Diacylglycerol acyltransferase

4.5 Variation of gene expression in strains under stress conditions

Microalgae grown under stress conditions have higher lipid content than microalage grown in optimal conditions. Hence, gene expression analysis was further extended in stress conditions. *Scenedesmus* strains were grown in two stress conditions i.e. nitrogen deficiency and high salinity and harvested at LSP. LSP in optimal growth condition was taken as a control for the relative gene expression estimation. Both stress conditions have considerably increased the expression of most of the genes i.e. *ACP, ACC, KASII, KASIII, ACPD, HAD, KAR, ER, FATA, SAD, PAP, LPAAT, DGAT* in all three strains. However, high salinity condition exerted more impact on gene expression than nitrogen deficiency. KASII exhibited maximum fold expression in SD12 and SQ19 with 27 folds and and 21 folds, respectively and SAD showed highest up-regulation in SD16 with 19 folds (Table 4.3) (Figure 4.6).

Table 4.3 Relative ratios of expression of FA and TAG biosynthetic genes under two stress conditions, ND and HS (at LSP) among three *Scenedesmus* strains; SD12: *S. dimorphus* 12; SQ19: *S. quadricauda* 19; SD16: *S. dimorphus* 16

ND			HS		
Genes	**Ratio of transcripts of SD12 to that of SQ19**	**Ratio of transcripts of SD12 to that of SD16**	**Genes**	**Ratio of transcripts of SD12 to that of SQ19**	**Ratio of transcripts of SD12 to that of SD16**
ACC	4.185841	1.58194	ACC	0.5325	2.098522
MT	14.25	10.96154	MT	1.257353	19.73077
ACP	0.418842	2.016393	ACP	2.103627	3.052632
ACPD	0.714286	0.47541	ACPD	0.539823	2.08547
KASI	0.39726	0.783784	KASI	0.044444	0.5
KASII	2.676259	1.61039	KASII	1.258197	13.2201
KASIII	0.261053	0.270742	KASIII	0.173724	0.392157
ER	0.878049	0.232258	ER	0.394161	0.837209
KAR	0.765152	0.181329	KAR	0.62973	2.137615
HAD	2.575342	0.876457	HAD	0.638356	0.920949
PT	1.69223	20.20482	PT	24.96875	79.9
SD	0.956204	0.100924	SD	1.947674	0.173575
FAD	15.11628	86.66667	FAD	3.315985	16.83019
FATA	0.265997	1.277108	FATA	1.35023	4.650794
GPAT	0.434211	0.458333	GPAT	1.230769	1.333333
LPAAT	2.955357	0.547107	LPAAT	2.588652	10
PAP	3.549296	3.876923	PAP	3.438871	4.12406
DGAT	4.70892	6.388535	DGAT	1	1.168224

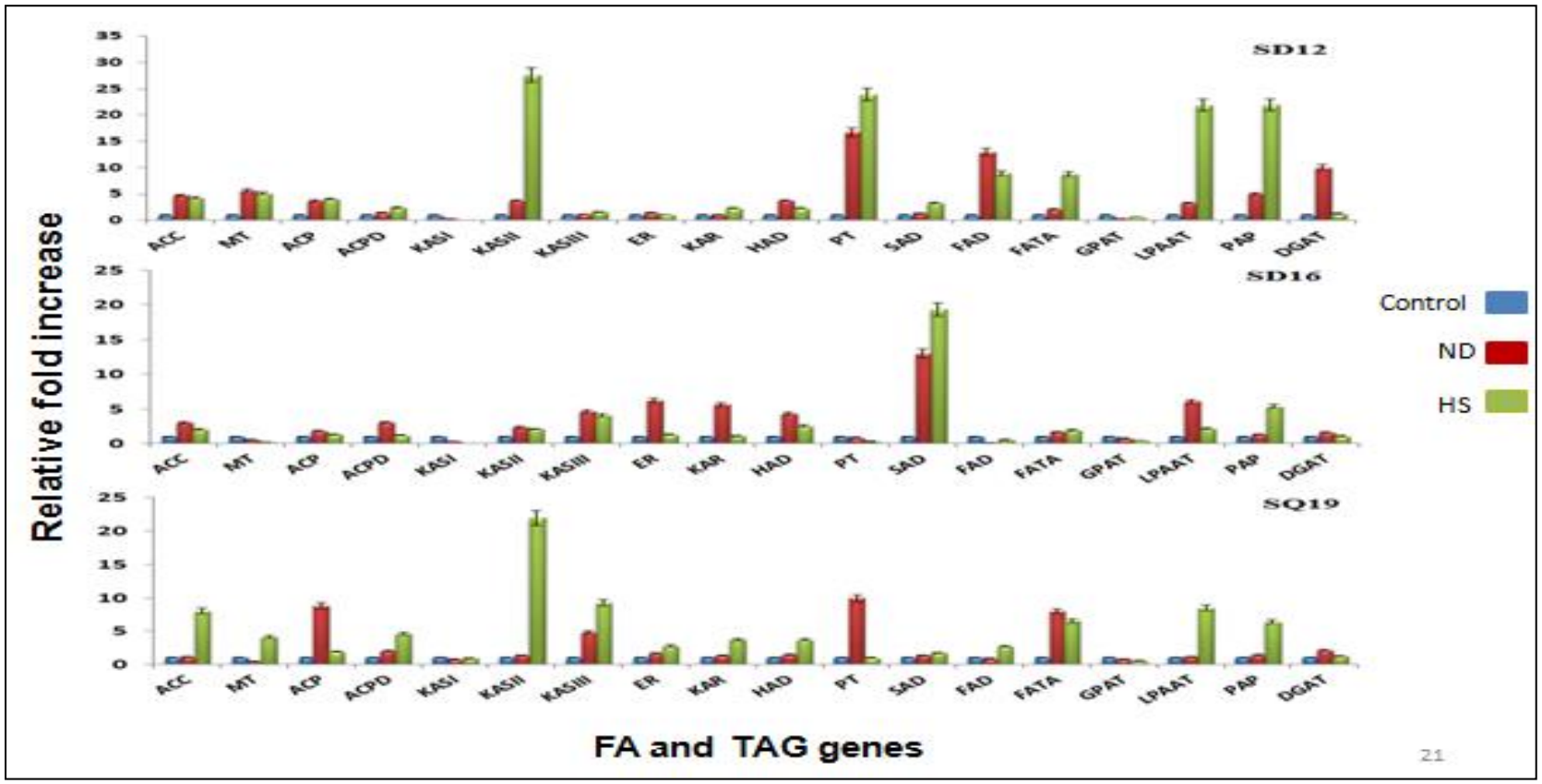


Figure 4.6 Relative expression of FA and TAG biosynthetic genes under two stress conditions, ND and HS (at LSP) w.r.t LSP for three lipid contrasting strains of *Scenedesmus;* SD12: *S. imorphus*12; SQ19: *S. quadricauda*19; SD16: *S. dimorphus*16, Abbreviations: ND- Nitrogen Deficiency, HS- High Salinity, ACC- Acetyl-CoA Carboxylase; MT- Malonyl Transferase; ACP-Acyl Carrier Protein; KASI- β-ketoacyl-ACP synthase I; KASII- β-ketoacyl-ACP synthase II; KASIII- β-ketoacyl-ACP synthase III; KAR- β-ketoacyl-ACP reductase; HAD- β-hydroxyacyl ACP dehydratase; ER- Enoyl-ACP reductase; PT- Palmitoyl thioestrase; FATA- Fatty Acyl-ACP thioestrase; SAD- Stearoyl-ACP desaturase; ACPD- Palmitoyl desaturase; FAD- Omega fatty acid desaturase; GPAT- Glycerol-3-phosphate acyltransferase; LPAAT- Lysophosphatidic acid acyltransferase; PAP- Phosphatidic acid phosphatase; DGAT- Diacylglycerol acyltransferase

4.6 Categorization of genes showing similar expression pattern

Combining of results from three experiments i.e. general gene expression, temporal expression in growth stages and gene expression under stress conditions carried out in three lipid content contrasting strains of *Scenedesmus* has categorised the FA and TAG biosynthesis genes into five clusters. The clusters were summarized in Table 4.4 containing the respective genes in each strain. Cluster I contains genes exhibiting increase in expression with lipid accumulation and majority of genes belonged to this category. Cluster II includes genes having maximum expression in early stationary phase and genes that exhibited down-regulated expression in LSP belonged to Cluster III. Genes that were up-regulated and down-regulated under both the stress conditions were assembled into Cluster IV and Cluster V, respectively.

Table 4.4 Clustering of genes on the basis of their expression pattern in growth stages and stress conditions for *Scenedesmus* species with different lipid content

Cluster	Gene Expression	***S.dimorphus*16 [5%]**		***S.dimorphus*12[26%]**		***S.quadricauda*19 [14%]**	
		FA genes	**TAG genes**	**FA genes**	**TAG genes**	**FA genes**	**TAG genes**
Cluster I	Increased with lipid accumulation	ACP,MT, KASII, KASIII, FATA, PT, SAD, KAR, HAD , ER	LPAAT, PAP, DGAT	ACP, ER, KASI, KASII, KASIII, FATA, PT, SAD, KAR, HAD	LPAAT, PAP, DGAT	ACC, ACP, KASI, KASII, KASIII,FATA, PT, SAD, KAR, HAD, ER	LPAAT, PAP, DGAT,
Cluster II	Increased at early stationary phase and declined later	ACC, ACPD, KASI, FAD,		ACC, ACPD, MT		ACPD, ER, MT , KAR, HAD, KASI	
Cluster III	Decreased with lipid accumulation		GPAT	FAD	GPAT	FAD	GPAT
Cluster IV	Up-regulated under stress conditions	ACC, ACPD, ACP, ER, KASII,KASIII, FATA,SAD, KAR, HAD	LPAAT, PAP, DGAT	ACC, MT, ACPD, ACP, ER, KASII, KASIII, FAD, FATA, PT, SAD, KAR, HAD	LPAAT, PAP, DGAT	ACC, ACPD, ACP, ER, KASII, KASIII, FATA, SAD, KAR, HAD	LPAAT, PAP, DGAT
ClusterV	Down-regulated under stress conditions	MT, KASI, PT, FAD	GPAT	KASI	GPAT	KASI	GPAT

4.7 Statistical Analysis

For the establishment of correlation between fatty acid and triacylglyercol biosynthesis genes and the lipid content, principle component analysis (PCA) was done on the data obtained from gene expression analysis in differential conditions among the three strains of *Scenedesmus*. Arrangement of the whole dataset was done into six components with first two components i.e. F1 and F2 contributed 70% variability in data. F1 component exhibited a critical role in storage lipid biosynthesis as it covered both the growth stages, ESP and LSP of SQ19 and SD12 (Figure 4.7). Biplot was created from the expression values of the eighteen genes in differential conditions (Figure 4.8). Squared cosine values were calculated, which revealed the significance of genes *KASII, HAD, ACP, KAR, PT, ER, FAT*A from FA biosynthesis and *PAP, LPAAT, DGAT* from TAG biosynthesis for lipid content enhancement (Table 4.5).

Heat map was also constructed from the expression values of FA and TAG biosynthesis genes in stress conditions and growth stages of three *Scenedesmus* strains using GenEx (Figure 4.9).

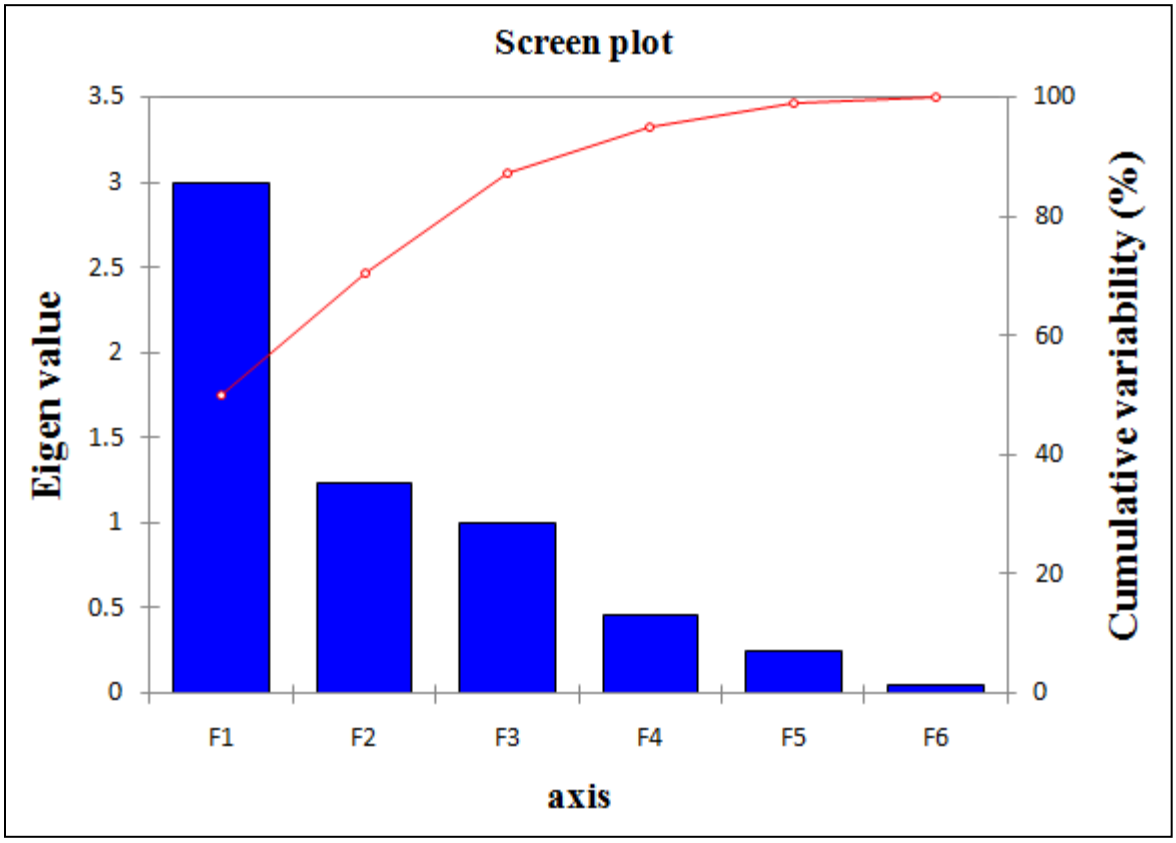


Figure 4.7 Screen plot for principal components (F1 – F6), their respective Eigen values, and cumulative variability. Major variance was contributed by component F1 and F2 where F1 played significant role covering most of the genes

Table 4.5 Squared cosine values for the respective PCA analysis of 18 lipid biosynthetic genes in growth stages among two strains of *S. dimorphus* (SD16, SD12) and one *S. quadricauda* strain (SQ19)

Genes	F1	F2	F3	F4	F5	F6
ACC	0.068	0.000	**0.571**	0.197	0.123	0.041
MT	0.005	0.261	0.161	0.218	**0.289**	0.067
ACP	**0.637**	0.173	0.078	0.055	0.057	0.000
ACPD	0.226	0.241	**0.522**	0.007	0.003	0.000
KASI	0.005	0.011	**0.486**	0.309	0.076	0.113
KASII	**0.376**	0.348	0.189	0.080	0.001	0.007
KASIII	0.007	0.111	0.076	**0.564**	0.237	0.006
ER	**0.739**	0.165	0.047	0.045	0.005	0.000
KAR	**0.942**	0.006	0.005	0.009	0.028	0.009
HAD	**0.892**	0.042	0.026	0.022	0.013	0.005
PT	**0.596**	0.073	0.061	0.259	0.005	0.006
SAD	0.116	0.227	0.032	0.004	**0.573**	0.049
FAD	0.032	**0.688**	0.263	0.006	0.008	0.003
FATA	**0.888**	0.070	0.015	0.023	0.000	0.004
GPAT	0.015	**0.958**	0.009	0.009	0.005	0.003
LPAAT	**0.951**	0.003	0.022	0.004	0.020	0.000
PAP	**0.947**	0.036	0.006	0.002	0.006	0.003
DGAT	**0.850**	0.000	0.036	0.103	0.004	0.007

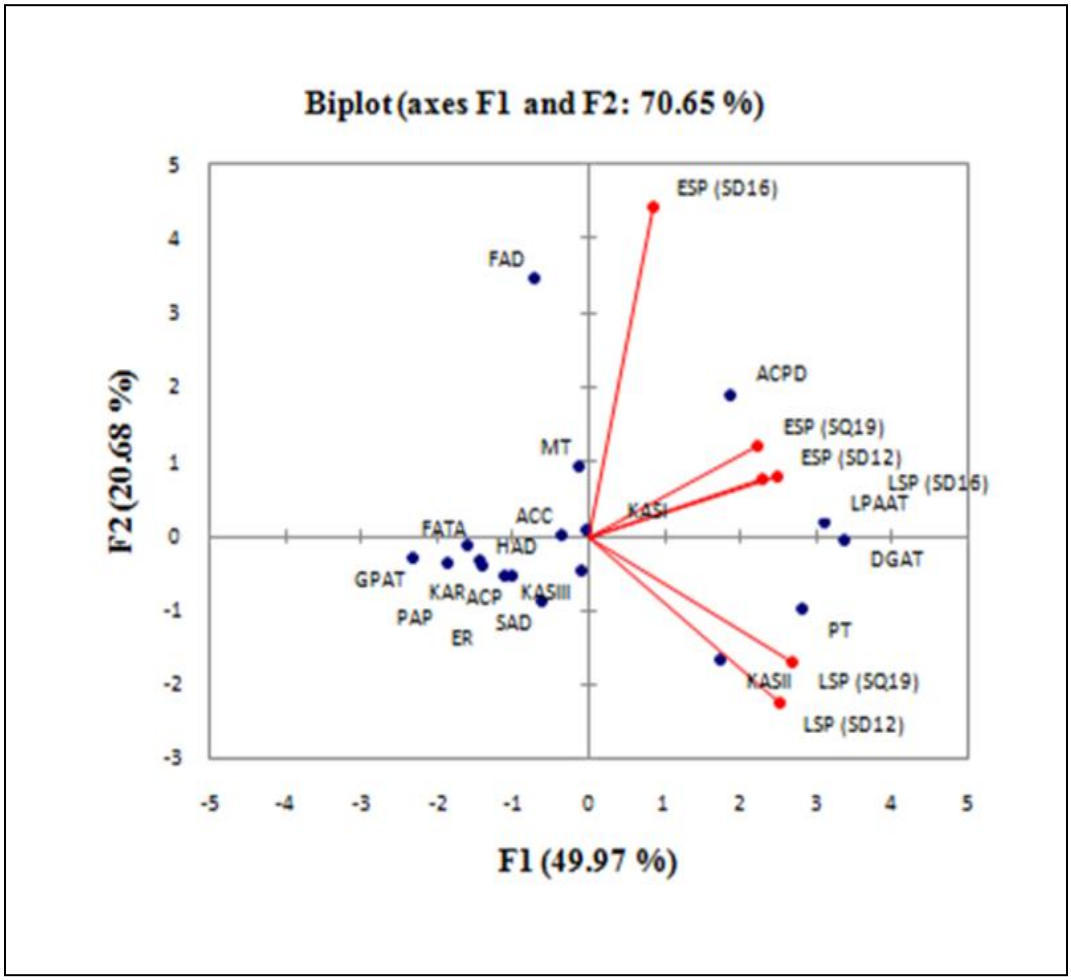


Figure 4.8 Biplot for the respective PCA analysis of 18 lipid biosynthetic genes in growth stages among two strains of *S.dimorphus* (SD16, SD12) and one *S.quadricauda* strain (SQ19), where variables are shown in red and observations for 18 genes in blue

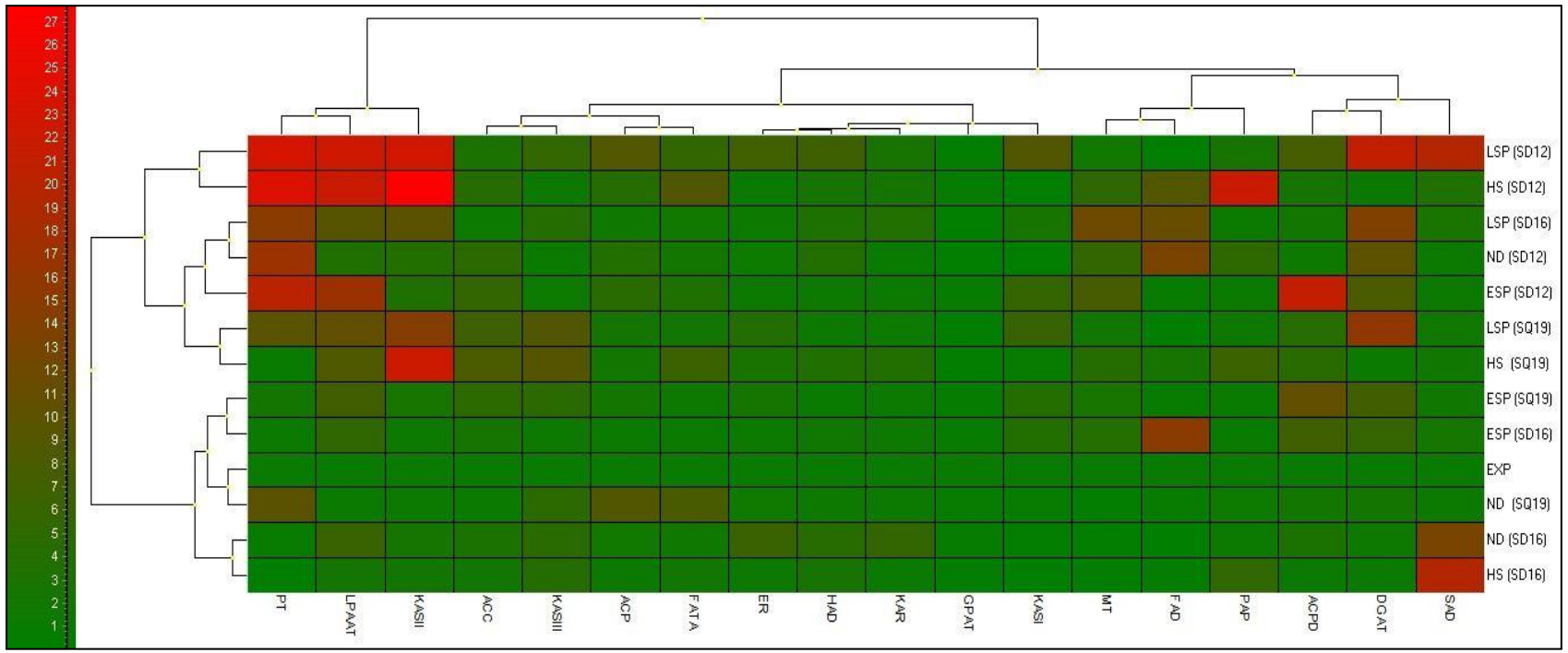


Figure 4.9 A representative heat map demonstrating the expression profiling of fatty acid and triacylglycerol biosynthetic pathway genes in various growth stages and stress conditions among three lipid contrasting *Scenedesmus* strains, Abbreviations: EXP- Exponenetial Phase, ESP- Early Stationary Phase, LSP- Late Stationary Phase, ND- Nitrogen Deficiency, HS- High Salinity, ACC- Acetyl-CoA Carboxylase;MT- Malonyl Transferase; ACP-Acyl Carrier Protein; KASI- β-ketoacyl-ACP synthase I; ; KASII- β-ketoacyl-ACP synthase II; KASIII- β-ketoacyl-ACP synthase III; KAR- β-ketoacyl-ACP reductase; HAD- β-hydroxyacyl ACP dehydratase; ER- enoyl-ACP reductase; PT- Palmitoyl Thioestrase; FATA- Fatty Acyl-ACP Thioestrase; SAD- Stearoyl-ACP Desaturase; ACPD- Palmitoyl Desaturase; FAD- Omega fatty acid desaturase; GPAT- Glycerol-3-phosphate acyltransferase; LPAAT- Lyso-phosphatidic acid acyltransferase; PAP- Phosphatidic acid phosphatase; DGAT- Diacylglycerol acyltransferase

4.8 Transcriptome sequencing of *Scenedesmus* species

Comparative transcriptomic analysis was performed between *Scenedesmus quadricauda* and *Scenedesmus dimorphus* species of *Scenedesmus*, having maximum lipid content of 14 % and 26% dry weight, respectively with equivalent biomass production. The focus of transcriptome sequencing was to uncover the molecular mechanism behind the cross species variation in lipid content (Figure 4.10).

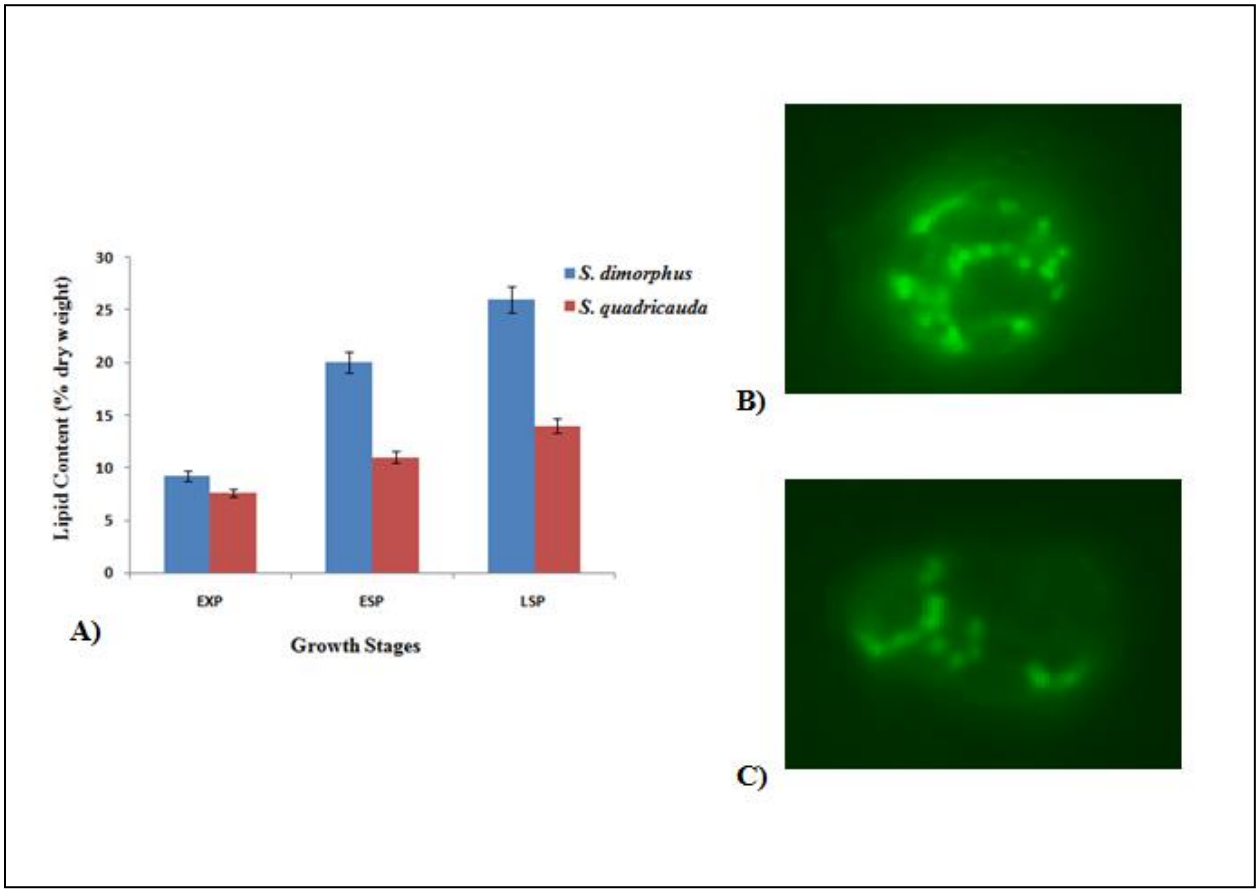


Figure 4.10 A) Lipid Content of *S. quadricauda* and *S. dimorphus* in three Growth Stages i.e. EXP, ESP, LSP [31]. Error bars indicate 5% standard error.N=3 microalgal cultures. Abbreviations: EXP-Exponential Phase; ESP-Early Stationary Phase; LSP- Late Stationary Phase **B)** Fluorescence microscopy image of *S. dimorphus* **C)** Fluorescence microscopy image of *S. quadricauda*. The brightness indicate amount of triacylglycerol content

4.9 *De novo* sequence assembly

Illumina sequencing was performed for the generation of the 7 paired-end cDNA sequencing libraries from both species. The mean size of the fragment distribution ranged from 550-700 bp. Library quantification and validation process resulted in the production of raw reads, which were further subjected to filteration with quality value QV > 20. Other contaminants present such as adapters were also trimmed. Parameters for performing the filtration were as follows:

- Trimming of adapters
- Sliding window trimming was performed upto 20 bp and cutting was done when the average quality within the window became lower than the threshold of 20 (Sliding Window)
- If the value was lower than the threshold quality of 30, then the bases were cut off at the start of a read (Leading)
- Bases were cut off at the end of a read, if value was below a threshold quality of 30 (Trailing)
- If it was below 40 bp length then the read was dropped (Minlength)

Quality filtering of raw reads has generated 11,290,041 and 17,907,034 high quality (HQ) reads from *S. quadricauda* and *S. dimorphus*, respectively. *De novo* assembly of these HQ reads was performed with Trinity producing 145,112 transcripts for *S. quadricauda* and 60,826 transcripts for *S. dimorphus* (Table 4.6). Further, CD-HIT-EST program was run on the assembled transcripts to get 78,386 unigenes for *S. quadricauda*, generating 110 Mb data (N50= 1736) and 42,979 unigenes in *S. dimorphus* constituting 38 Mb data (N50= 1148).

These unigenes were further utilized by ORF Predictor for prediction of CDSs. Parameters of ORF Predictor was set as default. 76,969 and 40,979 CDSs were successfully predicted in *S. quadricauda* and *S. dimorphus*, respectively. Predicted CDSs were of the size ranging from 200 to 6000 bp in both S*cenedesmus* species. 400-500 bp was the average length of CDSs in *S. quadricauda* and 300-400 bp was in *S. dimorphus*. The size distribution of CDSs in *S. quadricauda* and *S. dimorphus* has been shown in Figure 4.11.

Table 4.6 *De novo* assembly statistics

Parameter	*S. dimorphus*	*S. quadricauda*
Number of reads	17,907,034	11,290,041
Number of assembled transcripts	60,826	145,112
Total number of bases	45.5 Mb	140.7 Mb
N50	1,099	1,513

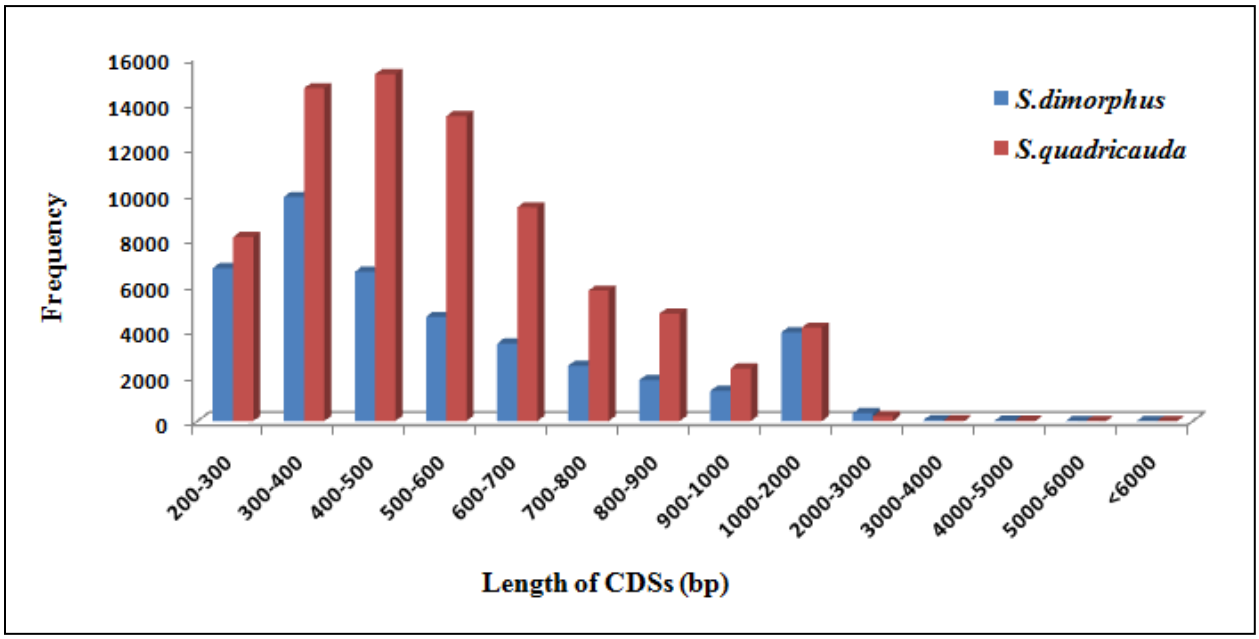


Figure 4.11 Length distribution of CDSs for *S. dimorphus* and *S. quadricauda*

4.10 Homology search and functional annotation of predicted CDSs

Predicted CDSs from both transcriptomes were subjected to BLASTx search against non-redundant (nr) protein database of NCBI with an E-value cut-off of 10^{-6}. Best hits obtained from the database were used as annotations of particular sequences. 16,678 CDSs were annotated in *S. dimorphus* and 11,917 CDSs were annotated in *S. quadricauda*. From the annotated sequences, 10,194 CDSs were having significant hits from green algal species in *S. dimorphus* and 11,917 CDSs were having best hits from green algal species in *S. quadricauda*. Majority of significant

hits were found to be from *Volvox carteri* after that *Chlamydomonas reinhardtii* and *Chlorella variabilis* in both the species (Figure 4.12).

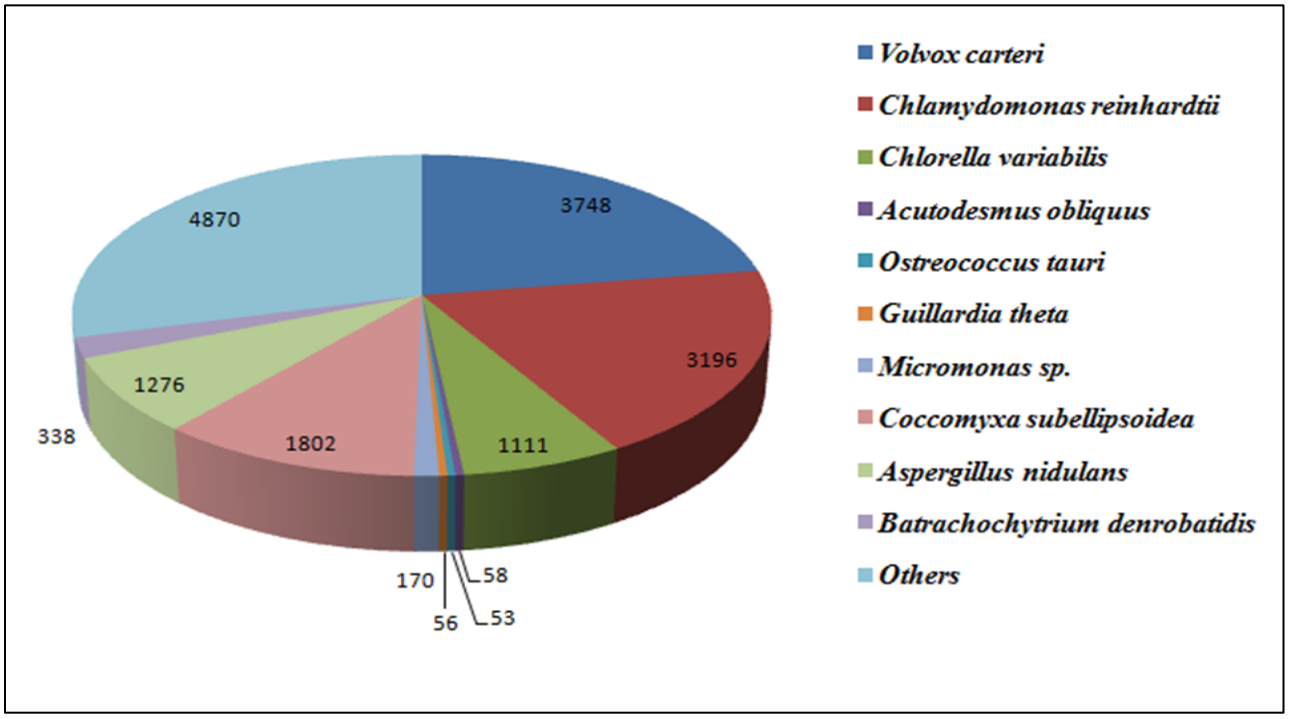


Figure 4.12.1 Top Blast species distribution for BLASTx matches in *S. dimorphus*

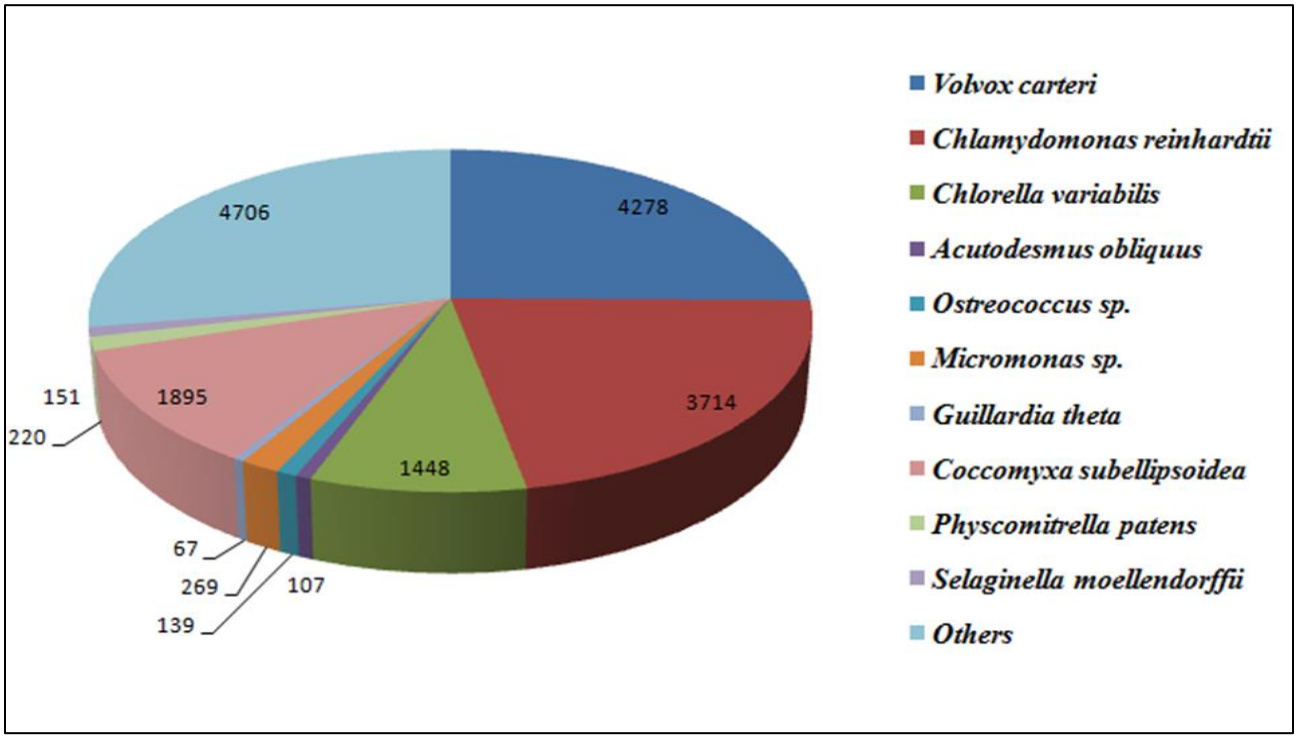
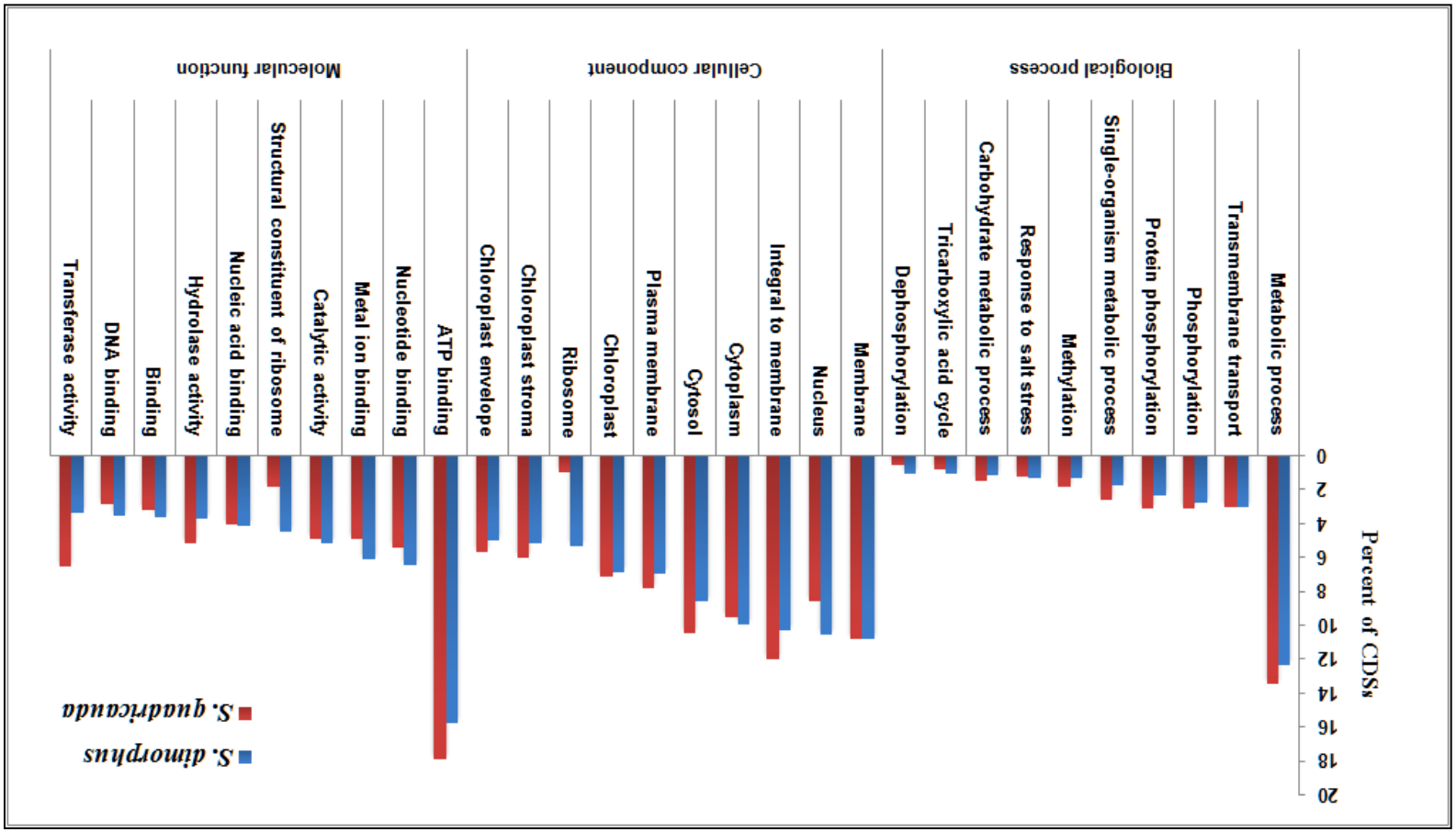


Figure 4.12.2 Top Blast species distribution for BLASTx matches in *S. quadricauda*

Further, classification of predicted CDSs was performed by GO mapping on the basis of their function. Accession IDs obtained from functional annotation using BLASTx were explored in the gene product table of GO database. Majority of assignments of predicted CDSs were found to be in molecular function category in both the species (7,591 in *S. quadricauda*; 6,551 in *S. dimorphus*) followed by biological process (6,090 in *S. quadricauda*; 6,123 in *S. dimorphus*) and cellular component (4,430 in *S. quadricauda*; 3,839 in *S. dimorphus*) (Figure 4.13).

Figure 4.13 GO categories in *S. dimorphus* and *S. quadricauda*



4.11 SSRs identification from transcriptomes of *S. quadricauda* and *S. dimorphus*

Mining of SSRs was performed from the transcriptomes of *S. quadricauda* and *S. dimorphus* having di- to penta-nucleotide motifs. Minimum repeat unit was set as ten for mononucleotides, six for dinucleotides and five for all higher order motifs including tri-, tetra-, and penta-nucleotides. Total of 12,169 and 6,800 SSRs were identified from *S. quadricauda* and *S. dimorphus*, respectively. 1,173 and 785 sequences were found with more than one SSR in *S. quadricauda* and *S. dimorphus*, respectively (Table 4.7).

Table 4.7 SSRs identified in transcriptomes of *S. dimorphus* and *S. quadricauda*

Description	*S. dimorphus*	*S. quadricauda*
Total number of sequences examined	40,979	76,969
Total size of examined sequences (bp)	23,585,019	42,992,676
Total number of identified SSRs	6,800	12,169
Number of SSR containing sequences	5,877	10,866
Number of sequences containing more than one SSR	785	1,173
Number of SSRs present in compound formation	9	8
Di-nucleotide	154	1,587
Tri-nucleotide	6,621	10,395
Tetra-nucleotide	23	181
Penta-nucleotide	2	6

Most abundant repeat patterns were tri-nucleotide repeats in both species (85% in *S. quadricauda*, 97% in *S. dimorphus*) followed by di-, tetra- and penta-nucleotides. GCA, CAG, CTG, GCT and AGC were the most frequent repeats in both species (Table 4.8, 4.9). Frequency of repeats was found to be inversely proportional to the repeat pattern length. Most of the potential SSRs were having the repeat unit of five, after that six and then others. The pattern of

the identified SSRs in both the species was almost similar; however the variation occurs in terms of abundance. SSRs are essential for studying the polymorphism in contrasting species.

Table 4.8 Repeat unit distribution of SSRs in transcriptome of *S. dimorphus*

Repeat Units	Penta	Hexa	Hepta	Octa	Nona	Deca	Hendeca	Total
AC	-	14	7	4	1	2	2	30
AG	-	3					2	5
CA	-	25	8	7	1	1	2	44
CT	-	1	1					2
GA	-	1						1
GC	-	4		1				5
GT	-	11	6	4	1	2	1	25
TC	-	1						1
TG	-	25	8	3	1	4		41
AAC	1							1
AAG	13	1						14
ACA	2							2
ACC	31	8	2					41
AGA	2							2
AGC	571	222	29					822
AGG	60	11	2					73
ATC	1			1				2
ATG	1	1	1					3
ATT	1		1					2
CAA	3	1						4
CAC	31	12	2					45
CAG	1557	603	51					2211
CAT	1	2	1	1				5
CCA	62	8	3					73
CCG	15	2	1					18
CCT	25	13	5					43
CGA	1							1
CGC	19	2						21
CGG	11	3						14
CTC	29	5	2					36
CTG	396	169	22	1				588

CTT	4							4
GAA	5		1					6
GAC	1							1
GAG	41	15	6	1				63
GAT	2	1						3
GCA	1025	395	20					1440
GCC	24	1	1					26
GCG	34	4	1					39
GCT	331	148	11					490
GGA	37	10	3					50
GGC	28	2	2					32
GGT	30	16	2	1				49
GTC			1					1
GTG	15	3	1					19
GTT			3					3
TCA	1							1
TCC	13	6	3					22
TCT	3							3
TGA	1							1
TGC	180	106	8					294
TGG	34	15	1					50
TGT	1	1						2
TTC	1							1
ATGC	1							1
ATGT	1							1
ATTC	4							4
CAGG	1	1						2
CCTG	4							4
CTGT	1							1
GCAG	1							1
GCAT	1							1
GCCA		1						1
GCCT	1	1						2
GGAT	2							2
TGCA	1							1
TGCT	1							1
TTGG		1						1
TATGG	1							1
TCCAG	1							1

Table 4.9 Repeat unit distribution of SSRs in transcriptome of *S. quadricauda*

Repeat Units	Penta	Hexa	Hepta	Octa	Nona	Deca	Hendeca	Total
AC	-	121	72	50	29	17	5	294
AG	-	6	10			3	1	20
AT	-	3			1		1	5
CA	-	257	142	69	38	6	2	514
CG	-	3						3
CT	-	18	11	4	3	2		38
GA	-	9	9	4	4	2		28
GC	-	9	3					12
GT	-	128	103	50	30	8	4	323
TA	-	3	1					4
TC	-	7		4	1			12
TG	-	145	94	56	27	11	1	334
AAC	1		1					2
AAG	6	2						8
AAT		2						2
ACA	9	4	1					14
ACC	42	15	2					59
ACG	1							1
ACT	1		2					3
AGA	1		2					3
AGC	958	231	18					1207
AGG	55	8	2					65
ATC	2							2
ATG	2		3					5
CAA	6	1	3					10
CAC	68	27	3					98
CAG	2390	472	43					2905
CAT	4	3	1					8
CCA	75	23	8					106
CCG	14	6	4					24
CCT	49	19	4					72
CGA	1			1				2
CGC	19	2	1					22
CGG	14	3	1					18
CGT	1							1

CTC	30	6	2					38
CTG	1038	165	10					1213
CTT	1							1
GAA	3	1						4
GAC	1			2				3
GAG	61	17	2					80
GAT	2	3						5
GCA	1763	366	25					2154
GCC	48	5	1					54
GCG	27	10	1					38
GCT	812	150	10					972
GGA	40	20	2					62
GGC	43	2						45
GGT	48	17	4					69
GTG	36	19	3					58
GTT	2	1	1					4
TAA		1						1
TAT		1						1
TCA	1		1					2
TCC	16	7	1					24
TCT	4	1						5
TGA	2	3	1					6
TGC	683	139	4					826
TGG	50	18	2					70
TGT	6		1					7
TTA	1		2					3
TTC			2					2
TTG	5	3	3					11
AACC		1						1
ACAG	1							1
ACAT	1							1
ACCA	1							1
AGAC	1							1
AGCA	2							2
AGGC	5							5

4.12 Pathway classification by KEGG

Pathway mapping was performed through KEGG Automatic Annotation Server (KAAS). Predicted CDSs files from both species were subjected into KAAS which computed the BLAST scores by aligning the query sequence file with KEGG GENES database. On the basis of BLAST scores, homologs were identified and bi-directional best hit method was utilized for assigning the KO identifiers to the predicted CDSs. Distribution of the orthologs into KO groups and assignment of K numbers to CDSs generates assignment score. Further this assignment resulted in the reconstruction of KEGG pathways, as KO identifiers are associated with the gene product in the KEGG pathway [161]. Genes encoding enzymes involved in major metabolic pathways, environmental information processing, genetic information processing, organism systems and cellular processes were identified. Distribution of CDS from *S. quadricauda* and *S. dimorphus* transcriptomes in KEGG pathways has been shown in Figure 4.14.

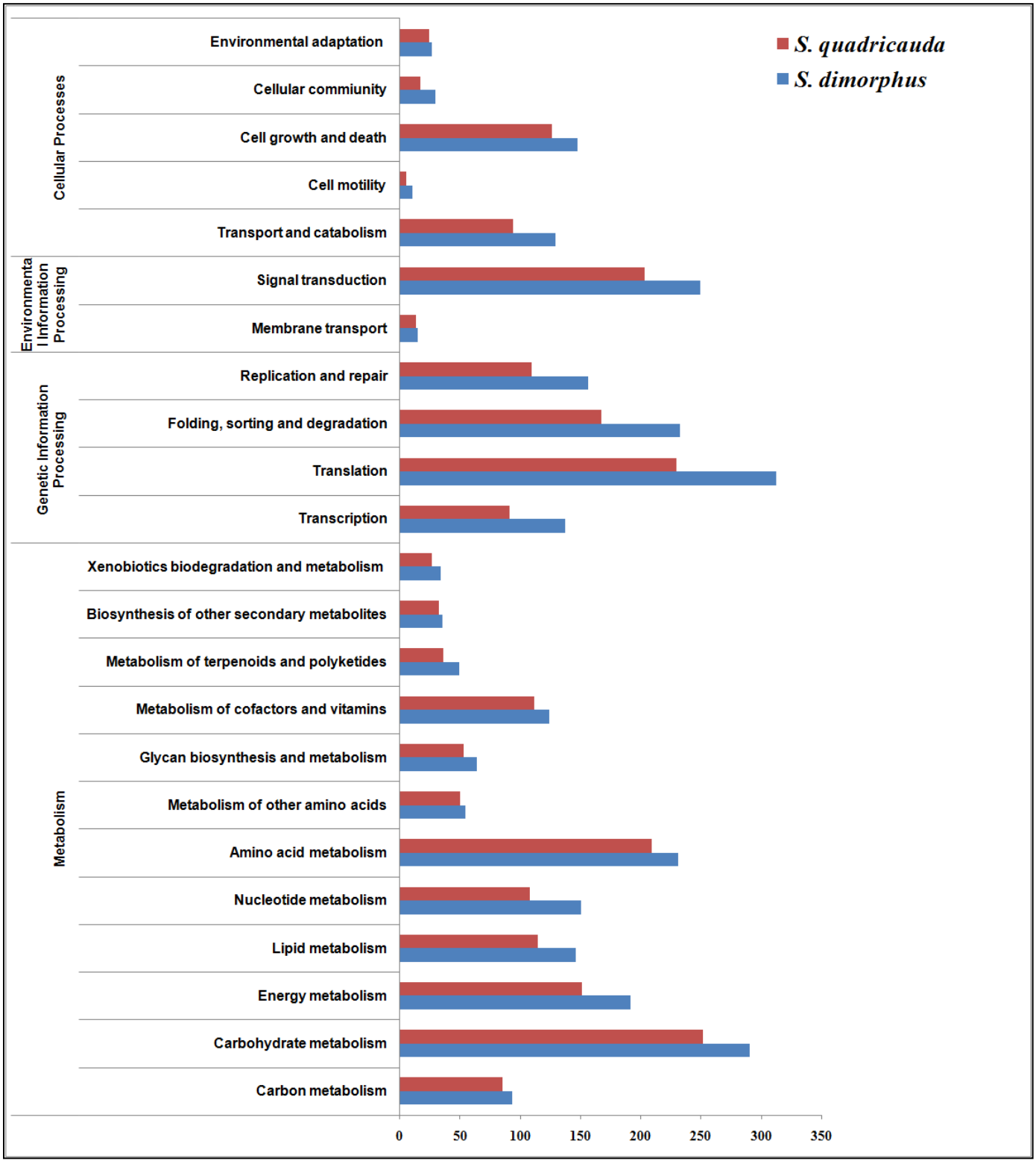


Figure 4.14 Functional characterization and abundance of KEGG pathways in *S. dimorphus* and *S. quadricauda*

4.13 Overview of differentially expressed genes in both transcriptomes

Variation in lipid content of closely related species of microalgae may be contributed from the differential gene expression. Comparison of the transcriptomes of *S. quadricauda* and *S. dimorphus* at stationary phase was performed for the identification of differentially expressed genes, playing crucial role in cross species lipid content variation in microalgae. Distribution of genes expressed in both the species has been shown in Figure 4.15. About 3,893 genes were found to be expressed in both the species, whereas 412 genes were expressed exclusively in *S. quadricauda* and 1,530 genes were expressed exclusively in *S. dimorphus*. Out of 3,893 genes expressed in both species, 2,537 genes were overexpressed in *S. dimorphus* (high lipid content containing strain) whereas 1, 356 genes were overexpressed in *S. quadricauda*. On the basis of differential gene expression, the important metabolic pathways having role in lipid biosynthesis were studied in both species. These pathways include carbon metabolism, photosynthesis, energy metabolism and fatty acid and triacylglycerol metabolism (Table 4.10).

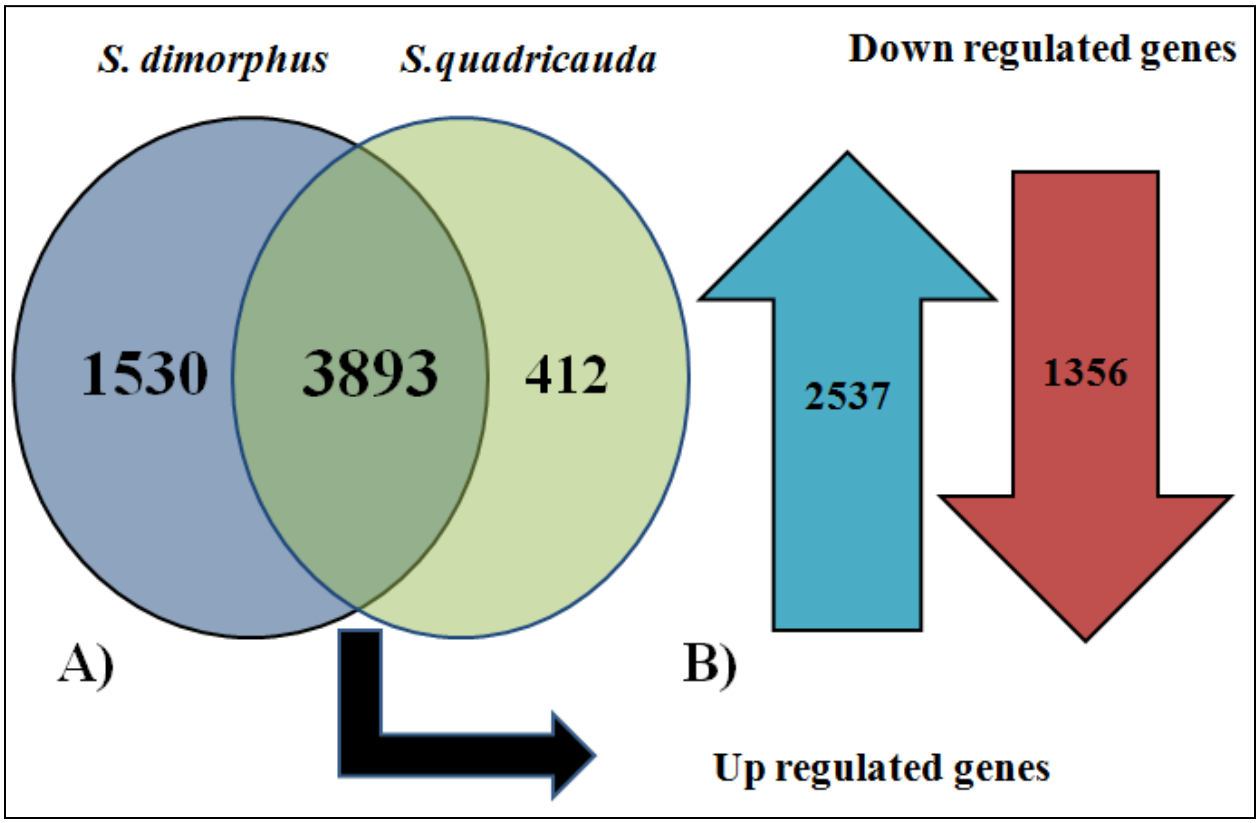


Figure 4.15 A) Distribution of genes expressed in *S. dimorphus* and *S. quadricauda* **B)** Up-regulated and down-regulated genes among the common genes of *S. dimorphus* and *S. quadricauda*

4.14 Pathways showing up-regulation of genes in *S. dimorphus*

4.14.1 Photosynthesis

Twenty-six photosynthetic genes were identified in *S. quadricauda* and *S. dimorphus*. Transcript abundance of these genes revealed enhanced expression of majority of the genes in *S. dimorphus* compared to *S. quadricauda*. However, certain genes i.e. photosystem I (*PSAO*), photosystem II (*PSBP*), ferredoxin (*PETF*) and light-harvesting complex protein (*LHCB2*) showed considerable increase of expression in *S. dimorphus.*

4.14.2 Carbon fixation

Enhancement of carbon fixation pathway was observed in *S. dimorphus* as almost all genes exhibited increased expression in *S. dimorphus* compared to *S. quadricauda*. Genes included phosphoglycerate kinase (*PGK*), ribulose-bisphosphate carboxylase large chain (*RBCL*), glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*), sedoheptulose-bisphosphatase (*SBP*), fructose-bisphosphate aldolase (*ALDO*), fructose-1,6-bisphosphatase (*FBP*), transketolase (*TKT*), phosphoribulokinase (*PRK*) and ribose 5-phosphate isomerase A (*RPIA*).

4.14.3 Oxidative phosphorylation

Thirty one genes of oxidative phosphorylation were found to be overexpressed in *S. dimorphus* compared to *S. quadricauda*. Genes such as succinate dehydrogenase (*SDHC*), ATPase (*ATPF1G, ATPF0B, ATPeF1G, ATPeV1D*) and pyrophosphatase (*PPA*) were significantly up-regulated in *S. dimorphus*.

4.14.4 Carbon metabolism

Carbon metabolism pathways i.e. glycolysis, TCA cycle, pyruvate metabolism and pentose phosphate pathway were found to be overexpressed in *S. dimorphus*. Total 25 genes were identified in glycolysis pathway from both *S. dimorphus* and *S. quadricauda*. Genes encoding enzymes catalyzing the later steps of glycolysis showed enhanced expression while only three genes belonging to early glycolysis were found to be down-regulated in *S. dimorphus*. Genes glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) and enolase (*ENO*) had a considerable increase of expression in *S. dimorphus* compared to *S. quadricauda*.

Citric acid cycle was also upregulated in *S. dimorphus*, as all genes except citrate synthase (*CS*) showed increase in transcript abundance in *S. dimorphus* than *S. quadricauda*.

Also overexpression of genes involved in pentose phosphate pathway in *S. dimorphus* was observed. Elevation of this pathway may provide energy for fatty acid biosynthesis and fatty acid chain elongation.

Also, genes encoding enzyme pyruvate dehydrogenase complex (*PDH*) including E1 and E2 components were up-regulated in *S. dimorphus* while gene encoding pyruvate decarboxylase (*PDC*) showed higher transcript abundance in low lipid isolate *S. quadricauda.*

4.14.5 Fatty acid and triacylglycerol biosynthetic pathway

Elevation of fatty acid biosynthesis and triacylglyerol biosynthesis pathway was observed in *S. dimorphus*. Majority of the genes from fatty acid biosynthesis, such as β-ketoacyl -ACP synthase II (*KASII*), fatty acyl-ACP thioesterase A (*FATA*), stearoyl-ACP desaturase (*SAD*), β-ketoacyl -ACP reductase (*KAR*) and hydroxyacyl-ACP dehydratase (*HAD*) showed significant increase of expression in *S. dimorphus.*

Triacylglycerol biosynthetic genes i.e. lysophosphatidic acid acyltransferase (*LPAAT*), glycerol kinase (*GK*), phosphatidic acid phosphatase (*PAP*) and diacylglycerol acyltransferase (*DGAT*) except glycerol-3-phosphate acyltransferase (*GPAT*) were up-regulated in *S. dimorphus* compared to *S. quadricauda*. Further, gene utilizing membrane lipids for triacylglycerol production i.e. phospholipid: diacylglycerol acyltransferase (*PDAT*) has been found with a remarkable expression in *S. dimorphus*, while it has not been identified in *S. quadricauda*.

4.15 Pathways showing down-regulation of genes in *S. dimorphus*

4.15.1 Catabolism of fatty acids and triacylglycerol

TAG lipase (*TGL*) catalyzing the degradation of triacylglycerol in lipid bodies and genes acyl-CoA oxidase (*ACOX*), long-chain acyl-CoA synthetase (*ACSL*) and enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase (*MFP*) belonging to fatty caid catabolism have shown decreased expression in *S. dimorphus*. Only one gene i.e. acetyl-CoA acyltransferase (*ACAT*) was found to be overexpressed in *S. dimorphus*. Down-regulated fatty acid and triacylglycerol catabolism *S. dimorphus* indicated its role in high lipid accumulation.

4.15.2 Starch metabolism

Transcript abundance data of 15 common genes identified in both the species revealed that most of the genes involved in starch biosynthesis pathway had decreased expression in *S. dimorphus*. However, genes encoding phosphoglucomutase (*PGM*) and glucose-1-phosphate adenylyltransferase (*GLGC*) were found to be overexpressed in *S. dimorphus*.

Table 4.10 Differentially expressed genes of carbon metabolism and fatty acid and triacylglycerol metabolism in *S. dimorphus* and *S. quadricauda*

Gene symbol	Gene name	EC number	Fold change
	Carbon metabolism		
HK	Hexokinase	2.7.1.1	- 0.25
GPI	Glucose-6-Phosphate isomerase	5.3.1.9	+ 0.62
PFK	Phosphofructokinase 1	2.7.1.11	- 0.50
ALDO	Fructose-bisphosphate Aldolase	4.1.2.13	+ 13.65
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase	1.2.1.12	+ 15.67
PGK	Phosphoglycerate kinase	2.7.2.3	+ 107.67
GPMI	Bisphosphoglycerate-independent phosphoglycerate mutase	5.4.2.12	+ 1.08
ENO	Enolase	4.2.1.11	+ 20.75
PK	Pyruvate kinase	2.7.1.40	- 1.88
FBP	Fructose-1,6-bisphosphatase	3.1.3.11	- 4.57
RBCL	Ribulose-bisphosphate carboxylase large chain	4.1.1.39	+ 1.48
TKT	Transketolase	2.2.1.1	+ 9.30
TAL	Transaldolase	2.2.1.2	+ 6.10
PPDK	Pyruvate, orthophosphate dikinase	2.7.9.1	+ 2.06
PEPC	Phosphoenolpyruvate carboxylase	4.1.1.31	- 0.14
PEPCK	Phosphoenolpyruvate carboxykinase (ATP)	4.1.1.49	+ 7.04
MDH	Malate dehydrogenase	1.1.1.37	+ 8.27
ACOS	Acetyl-CoA synthetase	6.2.1.1	+ 1.86
CS	Citrate synthase	2.3.3.1	- 0.63
ACL	ATP citrate lyase	2.3.3.8	+ 1.01
PDH	Pyruvate dehydrogenase E1 component alpha subunit	1.2.4.1	+ 2.79
DLD	Dihydrolipoamide dehydrogenase	1.8.1.4	+ 1.53
G6PD	Glucose-6-phosphate 1-dehydrogenase	1.1.1.49	+ 41.96
PGD	Phosphogluconate dehydrogenase	1.1.1.343	+ 67.81
GAPN	Glyceraldehyde-3-phosphate dehydrogenase	1.2.1.9	+ 1.56

	(NA DP+)		
	Glyceraldehyde-3-phosphate dehydrogenase		
GAPA	(NADP+) (phosphorylating)	1.2.1.13	+ 19.30
	Fatty acid biosynthesis		
	Acetyl-CoA carboxylase carboxyl transferase subunit		
ACCA	alpha	6.4.1.2	- 0.20
ACCC	Acetyl-CoA carboxylase, biotin carboxylase subunit	6.3.4.14	+ 2.77
	Acetyl-CoA carboxylase carboxyl transferase subunit		
ACCD	beta	6.4.1.2	- 0.21
KAS III	3-oxoacyl-ACP synthase III	2.3.1.180	- 0.16
KAS II	3-oxoacyl-ACP synthase II	2.3.1.179	-1.14
KAR	3-oxoacyl-ACP reductase	1.1.1.100	+ 1.14
HAD	3-hydroxyacyl-ACP dehydratase	4.2.1.59	+ 4.00
FATA	Fatty acyl-ACP thioesterase	3.1.2.14	+ 5.45
EAR	Enoyl-ACP reductase	1.3.1.10	- 0.53
SAD	Stearoyl-ACP desaturase	1.14.19.2	+ 5.05
FAD	Omega-6 fatty acid desaturase	1.14.19.	- 0.64
	Triacylglycerol biosynthesis		
GPAT	Glycerol-3-phosphate acyltransferase	2.3.1.15	- 0.35
PAP	Phosphatidate phosphatase	3.1.3.4	+ 41.94
DGAT	Diacylglycerol acyltransferase	2.3.1.20	+ 14.07
GK	Glycerol kinase	2.7.1.30	+1.70
	Fatty acid and triacylglycerol degradation		
ACSL	Long-chain acyl-CoA synthetase	6.2.1.3	- 0.47
	Enoyl-CoA hydratase/3-hydroxyacyl-CoA		
MFP	dehydrogenase	4.2.1.17	- 0.61
ACOX	Acyl-CoA oxidase	1.3.3.6	- 0.50
ACD	Acyl-CoA dehydrogenase	1.3.8.7	- 0.22
ACAT	Acetyl-CoA acetyltransferase	2.3.1.9	+ 2.93
TGL	TAG lipase	3.1.1.3	- 0.27

4.16 Heat map construction for differentially expressed genes in both species

Heat map was constructed for the differentially expressed genes in *S. quadricauda* and *S. dimorphus* on the basis of hierarchical clustering. Log-transformed and normalized value of genes was calculated from the Pearson's uncentered correlation distance, which was utilized for heat map generation (Figure 4.16).

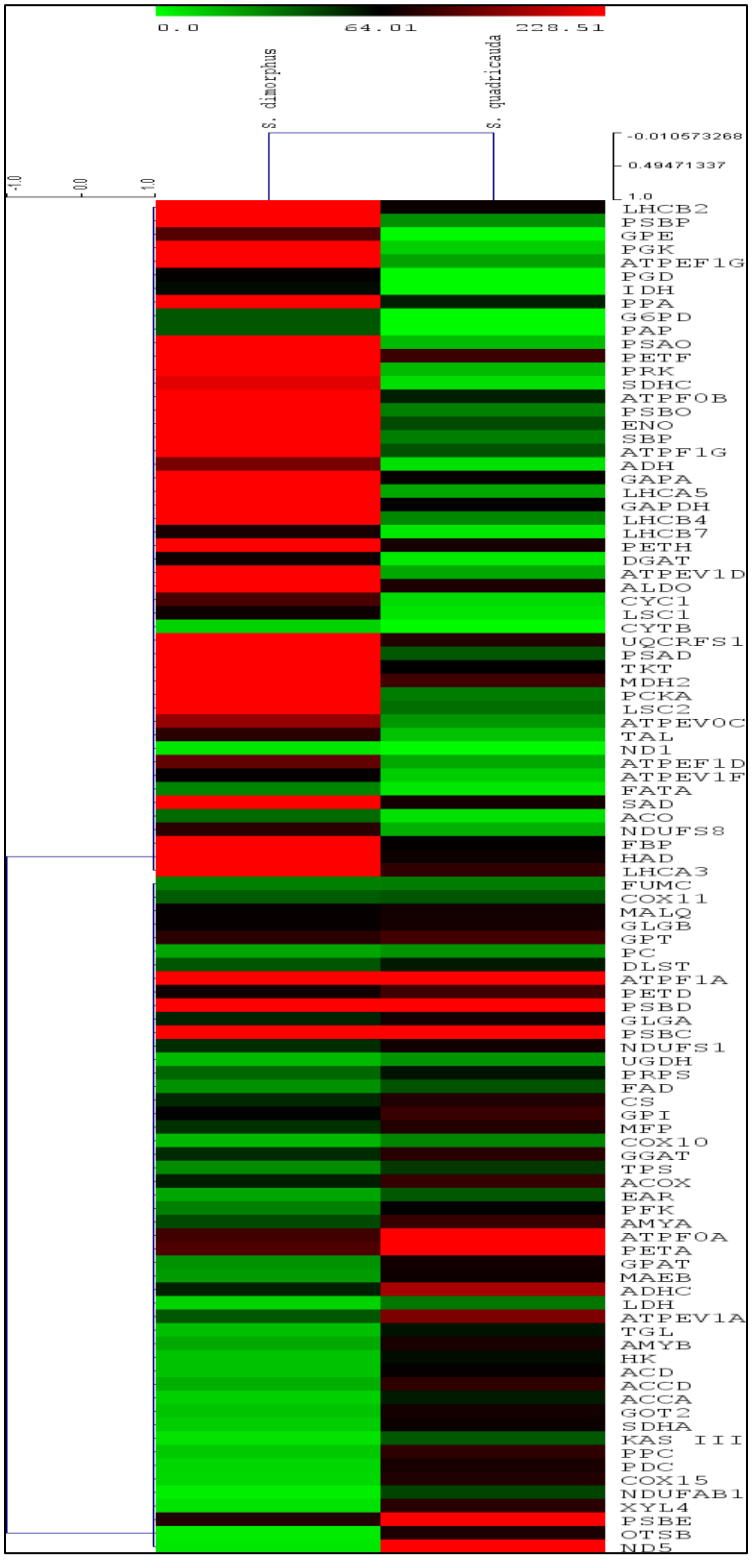


Figure 4.16 Heat map representing top 50 up-regulated and top 50 down-regulated genes in *S. dimorphus*

4.17 Identification of lipid biosynthesis related transcription factors from the transcriptomes of *S. dimorphus* and *S. quadricauda*

For the identification of transcription factors (TFs), CDSs of *S. quadricauda* and *S. dimorphus* were aligned to PlantTFDB using BLAST algorithm with an E-value cut-off of 10^{-5}. Total of 454 and 613 CDSs (encoding transcription factors) were identified in *S. quadricauda* and *S. dimorphus*, respectively which were categorized into 25 TF families. MYB TF family contained the highest number of TFs in both the species, which also included MYB and MYB-related families. MYB TF family represents 27% and 32% of total CDSs encoding transcription factors in *S. quadricauda* and *S. dimorphus* respectively. Nin-like TF family also contained large number of TFs in *S. dimorphus* as comparison to *S. quadricauda*.

Further, to identify transcription factors associated with lipid metabolism, literature based mining was performed. Hu *et al.* have reported thirty transcription factors playing role in lipid metabolism by *in-silico* approach. These thirty transcription factors belonged to TF families, viz. HB-other, AP2, C3H, E2F/DP, MYB, MYB_related, CPP, NF-YC, bZIP, HSF, LFY [53]. Searching these TFs in both the transcriptomes resulted in the identification of twenty transcription factors in *S. dimorphus* and only eight in *S. quadricauda*. Transcript abundance of the seven common transcription factors of *S. quadricauda* and *S. dimorphus* revealed the up-regulation of five transcription factors. These TFs were belonged to MYB, AP2, LFY and NF-YC TF families. One transcription factor of AP2 family was found to be expressed significantly in *S. dimorphus* than in *S. quadricauda* (Table 4.11). When subjected to BLAST search, it was found to be WRINKLED1 (*WRI1*) homolog. One more CDS encoding transcription factor of NF-YB family was found with considerable enhanced expression in *S. dimorphus*.

Table 4.11 Lipid related transcription factor families in *S. dimorphus* and *S. quadricauda*

TF family	*S. dimorphus*	*S. quadricauda*	Up-regulation(↑) / Down-regulation (↓)
NF-YC	+	+	↑
C3H	+	-	
C3H	+	-	
MYB_related	+	-	
NF-YC	+	+	↑
MYB_related	+	-	
E2F/DP	+	-	
C3H	+	+	↓
AP2	+	-	
MYB	+	+	↑
NF-YC	+	+	↓
MYB_related	+	-	
MYB_related	+	-	
E2F/DP	+	-	
NF-YC	+	-	
LFY	+	+	↑↑
CPP	+	-	
AP2	+	-	
C3H	+	-	
AP2	+	+	↑↑
E2F/DP	-	+	

(+: Present; - : Absent; ↑↑: Significantly up-regulated)

4.18 RT-qPCR analysis of the identified key genes from *in-silico* approach

Transcript abundance data from the computational anlaysis of both transcriptomes identified seven key genes that were found to be involved in regulationg the lipid content of *Scenedesmus* species. These genes were found to be the control points of lipid biosynthesis as they provide precursors for fatty acid and triacylglycerol biosynthesis in microalgae. Expression of the identified genes was further examined in both species by RT-qPCR. The results obtained from RT-qPCR analysis were in agreement with the computational analysis, i.e. enolase (*ENO*), genes glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*), pyruvate dehydrogenase (*PDH*), acetyl-CoA synthetase (*ACOS*), glycerol kinase (*GK*), and ATP citrate lyase (*ACL*) were overexpressed in *S. dimorphus* whereas citrate synthase (*CS*) have shown down-regulation in *S. dimorphus* (Figure 4.17A).

Also, a CDS encoding transcription factor *WRI1* was found with significant expression in *S. dimorphus.* RT-qPCR analysis of *WRI1* transcription factor was performed in three growth stages of both the species (Figure 4.17B). The result was in agreement with *in-silico* analysis. *WRI1* transcription factor exhibited a considerable increase in expression at the start of stationary phase in *S. dimorphus* whereas there was a slight increase in the expression in *S. quadricauda*. RT-qPCR analysis validated the computational analysis of transcriptomes and identified the genes responsible for increasing the lipid of *S. dimorphus.*

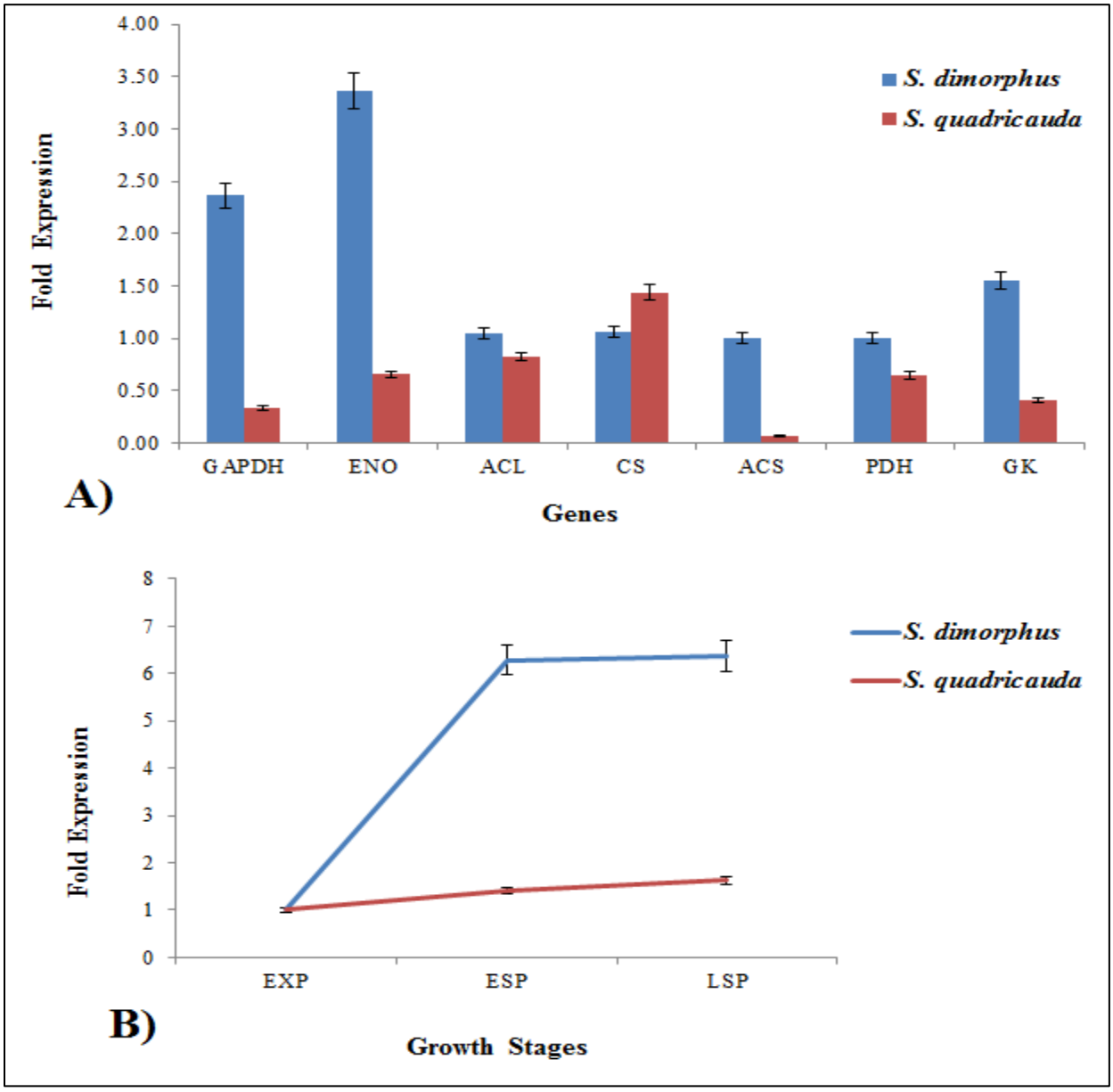


Figure 4.17 A) Expression analysis of seven identified key genes in *S. dimorphus* and *S. quadricauda* at stationary phase by RT-qPCR. Abbreviations: GAPDH -Glyceraldehyde 3-phosphate dehydrogenase; ENO- Enolase; ACOS- Acetyl-CoA synthetase; PDH - Pyruvate dehydrogenase; ACL- ATP citrate lyase; GK- Glycerol kinase; CS- Citrate synthase **B)** Relative expression analysis of CDS encoding transcription factor WRINKLED1in *S. dimorphus* and *S. quadricauda* in three growth stages

CHAPTER 5
DISCUSSION

Microalgae have recently emerged as a most appropriate renewable energy source for biodiesel production due to their benefits over other biodiesel sources. However, for commercial production of biodiesel from microalgae, there are few challenges including nutrient sourcing, strain selection and production management. One of the potential strategies to overcome these hurdles is to enhance the lipid production of microalgae by providing environmental stress conditions including nutrient stress. However, previous studies have reported that providing stress does not increase the overall lipid productivity due to the reduction in the growth of microalgae. Hence, there is a need to identify gene targets associated with lipid production in microalgae, in order to develop engineered strains with high lipid content. In case of *Scenedesmus*, there was no information available on the lipid biosynthetic pathways. Therefore, the current research was performed with the aim of elucidating the fatty acid and triacylglycerol biosynthetic pathway and to gain thorough understanding of molecular basis and transcriptional regulation of lipid biosynthesis in contrasting strains of *Scenedesmus* in terms of lipid content. This research work has provided initial leads which can be taken forward to carry out any genetic improvement strategy for enhancing lipid production in *Scenedesmus* species. The significant outcomes have been discussed below.

5.1 FA and TAG genes expression analysis among *Scenedesmus* strains

To study the interspecies variation in gene expression, high lipid content strains of *Scenedesmus quadricauda* (SQ19) and *Scenedesmus dimorphus* (SD12) having lipid contents of 14 % and 26 % respectively were taken. Further, for intraspecies variation, low lipid content (SD16) and high lipid content (SD12) strains of *S. dimorphus* with lipid contents of 5% and 26 %, respectively were taken.

Most of the genes from FA and TAG biosynthesis have shown overexpression in high lipid content strain SD12 as observed from the gene expression examination among three strains of *Scenedesmus*. The expression pattern suggested the major control of lipid biosynthesis was exerted by genes of tricaylglycerol biosynthesis i.e. genes *PAP*, *LPAAT* and *DGAT* for regulating

the lipid content of *Scenedesmus* species. This study is in accord with the previous reports on olive and oil palm [162]. Further, high fold change in gene expression was observed in genes *ACP*, *ACC*, *FATA*, *KASI*, *LPAAT*, *KASII*, *PAP* and *DGAT* in SD12 versus SD16 than SQ19, which indicated their association with high lipid content.

5.2 Molecular basis of lipid accumulation vis-à-vis growth stages

Microalgal growth occurs in different developmental stages as described in results. We have taken three phases of growth viz. exponential phase (EXP), early stationary phase (ESP) and late stationary phase (LSP) for expression studies. Exponential phase was taken as control for calculating relative expression in ESP and LSP among all three strains. Maximum lipid accumulation takes place in late stationary phase (LSP) of *Scenedesmus* species and the gene expression investigation of FA and TAG biosynthesis genes have shown the enhanced expression of most of the genes in late stationary phase in all three strains of *Scenedesmus,* which is in accordance with the lipid accumulation. On the basis of expression pattern of genes in the three lipid content contrasting strains, FA and TAG biosynthesis genes were categorized into different clusters. Cluster I included genes that exhibited increase in the expression with growth of *Scenedesmus* culture. These genes belonged to *de novo* fatty acid biosynthesis (*KAR*, *ACP*, *ER*, *HAD*), condensation reactions in FA biosynthesis (*KASIII*, *KASII*), thioesterification (*FATA*, *PT*)**,** deasturation (*SAD*) and TAG biosynthesis (*LPAAT*, *PAP*, *DGAT*). Genes that showed high transcript abundance in early stationary phase (ESP) and decline of expression in late stationary phase were included in Cluster II. *FAD* and *ACPD*, both are desaturases were grouped into Cluster II. *ACPD* adds a double bond to an acyl group which is further esterifies to *ACP* [163] and *FAD* is omega fatty acid desaturase that performs the desaturation of n-6 fatty acids into n-3 fatty acids (PUFAs) [164]. It has been reported that in stationary phase of microalgae, saturated and monounsaturated fatty acids have the tendency to increase and polyunsaturated fatty acids (PUFAs) tend to decrease [165]. *FAD* and *ACPD* are not directly involved in the lipid production as their role is to alter the lipid composition. As demonstrated by the expression pattern of these genes, fatty acids desaturation for generating palmitoleic acid and other PUFAs occurs primarily in early stationary phase. Cluster III includes only one gene i.e. *GPAT* that showed down regulated pattern of expression in LSP and ESP among all the three strains. This result is in agreement with the earlier report of *Chlamydomonas*, where down-

regulated expression patten of *GPAT* was observed in stationary phase where maximum lipid accumulation takes place [49]. Remaining genes exhibited expression variation in three strains, which indicated the functional differentiation of these genes during lipid biosynthesis.

5.3 Molecular basis of enhanced lipid accumulation under stress conditions

Further, the study was extended for performing gene expression analysis under two stress conditions viz. high salinity and nitrogen deficiency. Late stationary phase (LSP) was considered for calculating the relative expression fold in both the stress conditions. The majority of the genes exhibited enhanced expression under both stress conditions, however high salinity exerted more impact in enhancement of gene expression in all three strains. Some genes were up-regulated in one stress condition and other genes showed enhanced expression in other stress condition, suggesting the involvement of multiple genes in lipid content regulation among three different strains of *Scenedesmus*.

Cluster IV and V includes genes exhibiting up-regulation and down-regulation under stress conditions. *ACPD, ACC, ER, ACP, FATA, KASII, KASIII, KAR, SAD, HAD* (Cluster IV) are the genes that were found to be overexpressed under both stress conditions whereas *KASI* and *GPAT* (Cluster V) have shown down-regulation under both nitrogen deficiency and high salinity condition in all three strains. Down-regulation of KASI was also reported previously in *Chlamydomonas reinhardtii* under high salt and stress nitrogen deficiency [137]. However, the reason behind the down regulation was unknown. Remaining genes viz. *MT, PT* and *FAD* have shown inconsistency in expression pattern among three *Scenedesmus* strains which suggests the differential regulation for fatty acid and triacylglycerol biosynthetic pathway in *Scenedesmus*. Also, differences in regulatory elements present in the promoter regions of genes might be influencing the inconsistent expression patterns in strains.

Further, principle component analysis (PCA) reduced the data obtained from expression analysis of FA and TAG biosynthesis genes performed in differential conditions and established the correlation among genes with lipid accumulation. Genes *FATA, ACP, KASII, PAP, LPAAT* and *DGAT* were found to be completely associated with the storage lipid production as revealed by the comparative analysis and principle component analysis. *ACP, KASII, FATA* belong to *de novo* fatty acid biosynthesis and *PAP, LPAAT* and *DGAT* were from triacylglycerol biosynthesis

pathway. *ACP* cofactor is involved in the transfer of acyl chains in condensation, dehydration and reduction steps of fatty acid biosynthesis [166]. Elongation of fatty acid chain is terminated by thioesterase *FATA* that generates free fatty acids and *ACP* by hydrolyzing the newly formed acyl-ACP [167] and *KASII* component belongs to *KAS* (β-ketoacyl-ACP-synthase) complex which performs condensation reaction and elongation of 16:0-ACP to 18:0-ACP during fatty acid synthesis [168]. Monounsaturated acyl-ACPs are the preferred substrates of *FATA* [169]**.** These three genes viz. *ACP*, *FATA*, *KASII* have shown a remarkable rise of expression in highest lipid content strain and also in lipid accumulating late stationary phase. As discussed above, *FATA* and *KASII* are associated with the regulation of oleic acid content in storage lipid of microalgae and enhanced expression of these genes signifies favourable fatty acid profile for ideal biodiesel. Expression of the above mentioned genes was also found to be significantly enhanced in stress conditions. These three genes were reported as rate limiting genes in previous study by Lei *et al*. that were involved in fatty acid biosynthesis of green algae, *Haematococcus pluvialis* [32].

Other genes belonged to triacylglycerol biosynthesis where the steps i.e. phosphatidic acid production from lysophosphatidic acid acylation [170], diacylglycerol synthesis from phosphatidic acid by dephosphorylation process and production of triacylglycerol from diacylglycerol were catalyzed by *LPAAT*, *PAP*, *DGAT*, respectively [9]. Our results were supported by earlier reports where increased expression of *LPAAT* gene resulted in enhanced lipid content of *Chlamydomonas* sp. [49]. Also, *PAP* gene was found to be important in regulating the lipid biosynthesis of microalgal cells as indicated by its overexpression in *Chlamydomonas reinhardtii* [171]. *DGAT* and *PAP* activities were found to be associated with elevated lipid content in *Brassica napus* lines [172]. *DGAT* was reported to be the key gene for performing genetic engineering to enhace lipid content in *Chlorella pyrenoidosa* and *Chlamydomonas reinhardtii*, as it was found to be in correlation with high lipid accumulation [33, 49]. *DGAT* was also found to be involved in the modulation of fatty acid composition of TAG in marine diatom *Thalassiosira pseudonana* [173]. From this, it was concluded that expression of multiple genes from fatty acid and triacylglycerol biosynthesis controls the lipid accumulation in *Scenedesmus*. Genes *KASII*, *ACP*, *FATA* of FA biosynthesis and *PAP*, *LPAAT*

and *DGAT* from TAG biosynthetic pathway can be taken as the key genes for genetic engineering strategies targeted to increase storage lipid accumulation.

5.4 Comparative transcriptome sequencing to reveal molecular components associated with lipid production

Variation of lipid content in closely related species and strains suggests that lipid content in microalgae is influenced not only by fatty acid and triacylglycerol biosynthesis but by several other metabolic pathways. Hence, there should be a comprehensive understanding of molecular pathways and components responsible for high lipid accumulation in high lipid content strains. Whole genome and transcriptome sequencing provides complete information of genes expressed in a particular condition. High lipid strains of *S. quadricauda* and *S. dimorphus* with lipid contents of 14 % and 26 % dry weight respectively were taken for transcriptome sequencing. Comparative analysis and differential expression investigation of down-regulated and up-regulated pathways as well as genes was performed (Figure 5.1).

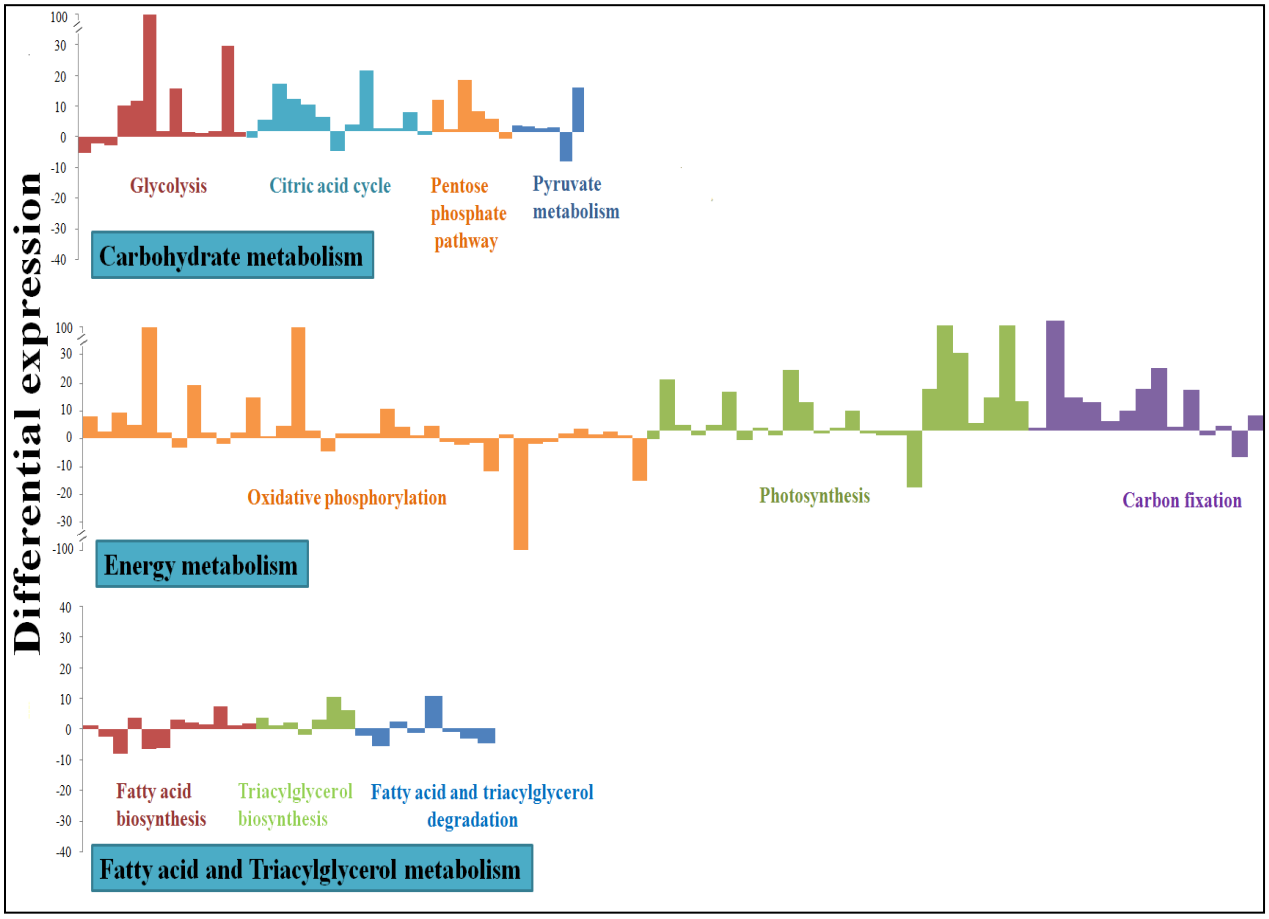


Figure 5.1 Pathways showing up-regulation and down-regulation based on differential gene expression in *S. dimorphus* versus *S. quadricauda*

5.5 Photosynthesis genes with enhanced expression in *S. dimorphus*

Photosynthesis utilizes the captured light energy for the generation of ATP and NADPH (reducing power) [174]. The energy from photosynthesis is further exploited by carbon fixation pathway, which is transferred to other metabolic pathways. Hence, photosynthesis is the vital process for other metabolic reactions and cellular processes. Out of the 26 commonly identified genes from photosynthetic pathway in species, *S. quadricauda* and *S. dimorphus*, most of the genes showed enhanced expression in *S. dimorphus*. Main photosynthetic genes include photosystem II (*PSBP*), photosystem I (*PSAO*), light-harvesting complex protein (*LHCB2*) and ferredoxin (*PETF*). Gene catalyzing the last electron transfer from photosystem I i.e. Ferredoxin-$NADP^+$ reductase [175] was also significantly expressed in *S. dimorphus*. Upregulation of photosynthetic genes in high lipid strain *S. dimorphus* is consistent with the

previous report, enhancement of photosynthetic light reactions resulted in high lipid accumulation in *Chlorella vulgaris* [176].

5.6 Increased supply of precursors for lipid biosynthesis in *S. dimorphus*

ATP and NADPH from photosynthesis are utilized by carbon fixation pathway for the generation of ATP, carbon skeletons and reducing power, which are further utilized by other metabolic pathways [174]. Glycolysis, pyruvate metabolism, TCA cycle are important metabolic pathways for regulating the lipid content in microalgae as they provide substrates for FA and TAG biosynthesis. NADPH is also produced from pentose phosphate pathway by the utilization of glucose, which is fed into FA biosynthesis [177].

Genes fructose-bisphosphate aldolase (*ALDO*), glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*), sedoheptulose-bisphosphatase (*SBP*), transketolase (*TKT*), fructose-1,6-bisphosphatase (*FBP*), phosphoglycerate kinase (*PGK*), ribose 5-phosphate isomerase A (*RPIA*), phosphoribulokinase (*PRK*) and ribulose-bisphosphate carboxylase large chain (*RBCL*) of carbon fixation pathways have shown enhanced expression in *S. dimorphus*. This increase of expression signifies the up-regulaed carbon fixation in *S. dimorphus* and hence, excess carbon supply from this pathway may have contributed for increasing the storage lipid production process (Figure 5.2).

End product from gycolytic pathway is being fed into fatty acid biosynthesis process. Up-regulation of late glycolysis in *S. dimorphus* symbolizes the high flux of substrates for fatty acid biosynthesis. Pyruvate dehydrogenase (*PDH*) gene is involved in the formation of acetyl-CoA from pyruvate, which is the end product of glycolysis. Simultaneosly another enzyme, pyruvate decarboxylase (*PDC*) converts pyruvate into acetaldehyde and subsequently to alcohol. Overexpression of *PDH* and down-regulated *PDC* in *S. dimorphus* indicates the basis for its high lipid content. This can be taken as contol point for the high lipid accumulation in *Scenedemsus* species. Also, phosphoenolpyruvate carboxylase (*PEPC*) enzyme converts phosphoenolpyruvate into oxaloacetate and inorganic phosphate by irreversible β-carboxylation process [178]. *PEPC* gene was down-regulated in *S. dimorphus* as revealed from transcript abundance data, which is in accord with previous study where enhancement of lipid content was observed through decreased expression of *PEPC* in *Chlamydomonas reinhartii* [179, 180]. Hence, *PEPC* can also

be taken as key gene due to its critical function for high lipid accumulation in *S. dimorphus*. One more important gene enolase (*ENO*) has shown significant overexpression in *S. dimorphus*. *ENO* is the crutial enzyme for lipid enhancement as it catalyzes the generation of phosphoenolpyruvate, which is converted into pyruvate [181]. Previous studies which support this observation are high lipid accumulation in cultures of *Isochrysis galbana* with nitrogen deficiency [182] and sunflower seeds [183] due to rise in the transcript abundance of genes enolase (*ENO*) and glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*).

GAPA i.e. $NADP^+$ dependent glyceraldehyde-3-phosphate dehydrogenase is another enzyme that provides precursors for triacylglycerol biosynthesis. It catalyzis the production of 3-phosphoglycerate and NADPH by performing the irreversible oxidation of glyceraldehyde-3-phosphate and $NADP^+$. NADPH molecules are required during fatty acids elongation for the incorporation of carbon unit and 3-phosphoglycerate is the precursor for TAG biosynthesis [184]. Gene encoding enzyme *GAPA* has shown a dramatic rise of expression in *S. dimorphus* coinciding with the earlier report which states that enhanced expression of glyceraldehyde-3-phosphate dehydrogenase (*GAPA*) enhances the biofuels productivity derived from fatty acids with promoting the NADPH supply [185].

One more gene found to be associated with high lipid content was citrate synthase (*CS*). It utilizes acetyl-CoA for condensation with oxaloacetate to form the product citrate and considered as rate-limiting gene of citric acid cycle. CS was down-regulated in *S. dimorphus* which indicates the extra flow of acetyl-CoAtowards fatty acid biosynthesis in *S. dimorphus.* This observation is supported by previous study where inhibition of citrate synthase in *Chlamydomonas reinhardtii* resulted in overall increase of lipid content [186]. This step was also taken as another regulatory step for lipid content improvement in microalgae.

Acetyl-CoA is the main substrate for lipid production in microalgae. Three altenative routes are there for the synthesis of this substrate catalyzed by enzymes viz. pyruvate dehydrogenase (*PDH*), acetyl-CoA synthetase (*ACOS*) and ATP citrate lyase (*ACL*) [187]. *PDH* enzyme catalyzes the direct conversion of pyruvate (end product of glycolysis) into acetyl-CoA. The gene encoding *PDH* was overexpressed in *S. dimorphus* supported by the previous study, where rise in gene expression of *PDH* was observed in *Nannochloropsis oceanica*, oleaginous microalga grown under nitrogen depletion i.e. lipid accumulating conditions [134]. Also,

suppression of pyruvate dehydrogenase kinase (*PDK*), enzyme that deactivates pyruvate dehydrogenase, had significantly enhanced the lipid content of *Phaeodactylum tricornutum* [188]. Now, *ACOS* enzyme converts acetate into acetyl-CoA and was overexpressed in *S. dimorphus* coinciding with the report where increase in its expression has revealed improved fatty acid proportion in *Schizochytrium* [189]. Third gene of acetyl-CoA biosynthesis i.e. *ACL* is the rate limiting gene of triacylglycerol production, catalyzes the cleavage of citrate to form acetyl-CoA [190]. In a previous report, in which three microalgal species, viz. *Dunaliella tertiolecta*, *Chlorella desiccata* and *Chlamydomonas reinhardtii* have shown differential triacylglycerol accumulation, highest expression of *ACL* was observed in *Chlorella desiccata* with maximum TAG accumulation [187].

Hence, enhanced expression of genes that provide substrates for carrying the fatty acid and triacylglycerol biosynthesis have revealed their putative roles for increasing the lipid content of *S. dimorphus* (Figure 5.3). Importance of carbon precursors supply was demonstrated by the present study, in order to increase the lipid accumulation in microalgae.

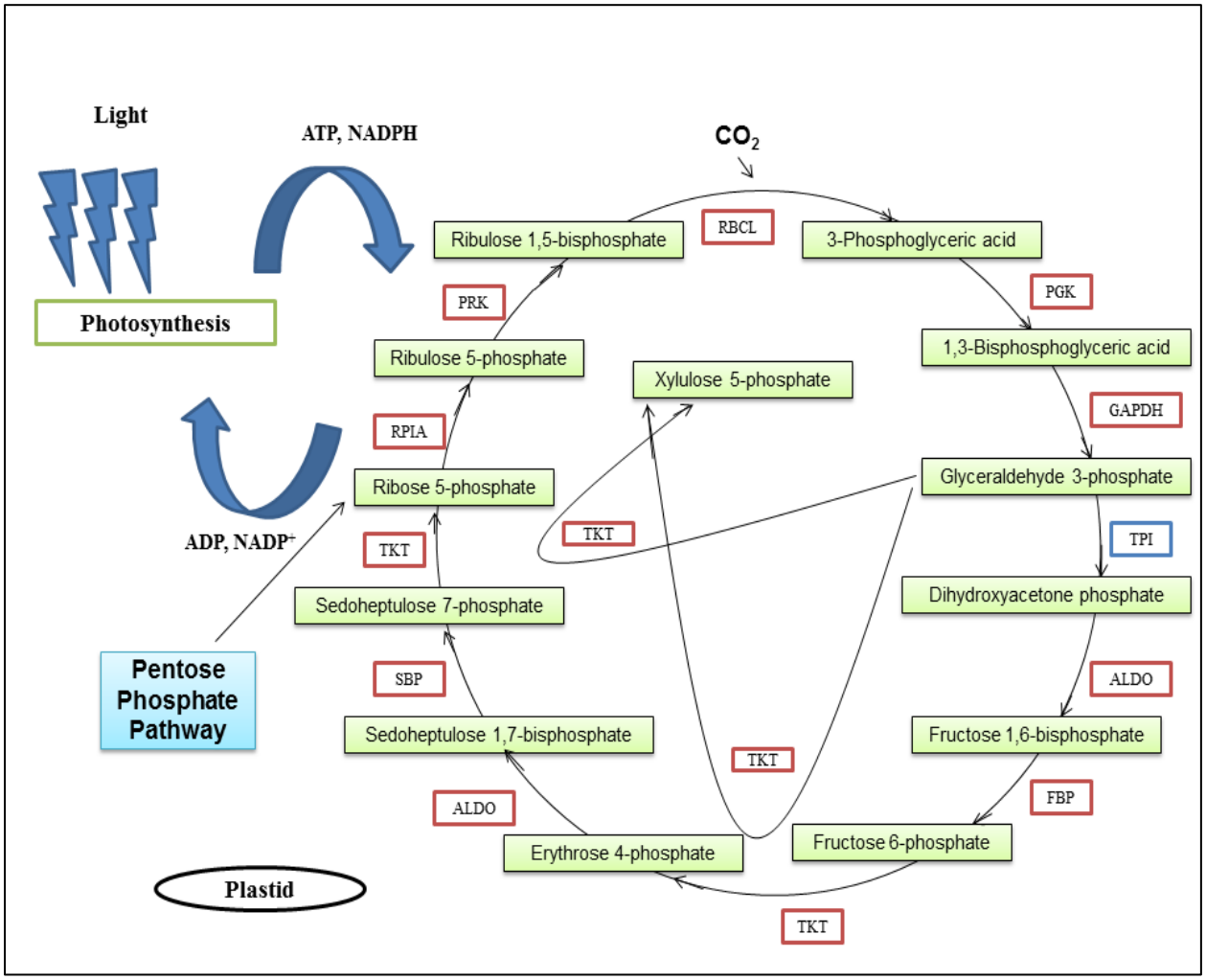


Figure 5.2 Carbon fixation pathway, enzymes are shown in red and blue colour boxes with red colour signifies up-regulation and blue colour signifies down- regulation in *S. dimorphus*. Enzyme abbreviations: RBCL- Ribulose-bisphosphate carboxylase large chain; PGK- Phosphoglycerate kinase; GAPDH- Glyceraldehyde 3-phosphate dehydrogenase; ALDO- Fructose-bisphosphatealdolase; FBP- Fructose-1,6-bisphosphatase; TK- Transketolase; SBP- Sedoheptulose-bisphosphatase; PRK- Phosphoribulokinase; RPIA- Ribose 5-phosphate isomerase A

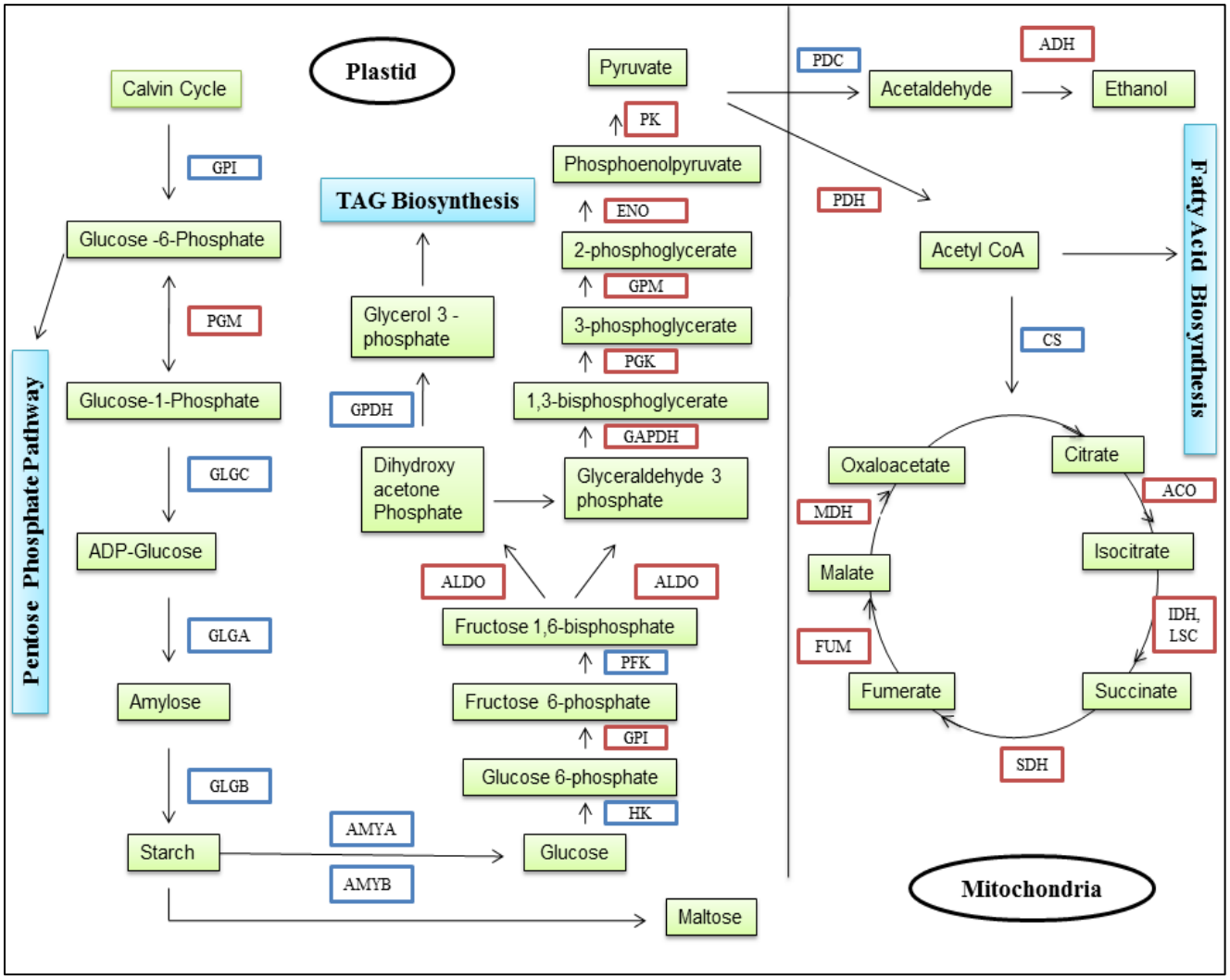


Figure 5.3 Carbon metabolism, enzymes are shown in red and blue colour boxes with red colour signifies up-regulation and blue colour signifies down-regulation in *S. dimorphus*. Enzyme abbreviations: PGK- Phosphoglycerate kinase; GAPDH- Glyceraldehyde 3-phosphate dehydrogenase; ALDO- Fructose-bisphosphatealdolase; PRK- Phosphoribulokinase; RPIA- Ribose 5-phosphate isomerase A; GPI- Glucose-6-phosphate isomerase; PGM- Phosphoglucomutase; GLGC- Glucose-1-phosphate adenylyltransferase; GLGA- Starch synthase; GLGB- 1,4-alpha-glucan branching enzyme; AMYA- Alpha-amylase; AMYB- Beta-amylase; HK- Hexokinase; GPI- Glucose-6-phosphate isomerase; PFK- Phosphofructokinase; ENO- Enolase; PK- Pyruvate kinase; PDH- Pyruvate dehydrogenase; PDC- Pyruvate decarboxylase; ADH- Alcohol dehydrogenase; CS- Citrate synthase; ACO- Aconitate hydratase; IDH- Isocitrate dehydrogenase; LSC- Succinyl-CoA synthetase; SDH- Succinate dehydrogenase; FUM- Fumarate hydratase; MDH- Malate dehydrogenase

5.7 Comparison of FA and TAG biosynthesis in *S. dimorphus* and *S. quadricauda*

Now the main pathway for storage lipid production in microalgae consists of two steps viz. FA biosynthesis and TAG biosynthesis. FA biosynthesis occurs in the chloroplast and TAG biosynthesis in endoplasmic reticulum. Acetyl-CoA is the major substrate of FA biosynthesis, produced from the above described pathways. Enzyme acetyl-CoA carboxylase (*ACC*) converts acetyl-CoA into malonyl-CoA. Fatty acids are produced by a series of reactions catalyzed by different enzymes of this pathway and then transferred into the cytosol, where these FAs are exploited by the enzymes of TAG biosynthesis in endoplasmic reticulum [9].

Out of the common fatty acid biosynthetic genes in *S. quadricauda* and *S. dimorphus*, most of the genes showed increase of expression in *S. dimorphus*. Up-regulated genes include β-ketoacyl -ACP synthase II (*KASII*), fatty acyl-ACP thioesterase A (*FATA*), β-ketoacyl-ACP reductase (*KAR*), hydroxyacyl-ACP dehydratase (*HAD*) and stearoyl-ACP desaturase (*SAD*) . High ratio of monounsaturated fatty acids in the fatty acid profile of microalgae is prerequisite for the production of biodiesel. Gene stearoyl-ACP desaturase (*SAD*) is involved in the production of monounsaturated fatty acids as it introduces the double bond into stearic acid to form the most desired fatty acid for biodiesel production i.e. oleic acid [191]. Also, fatty acyl-ACP thioesterase A (*FATA*) has important role as it terminates the fatty acid chain elongation and performs hydrolyzation of monounsaturated acyl-ACP to generate free fatty acids and ACP [192]. Free fatty acids from FA biosynthetic pathway are recruited by enzymes of TAG biosynthesis for sequential acylation of glycerol-3-phosphate [193]. Glycerol-3-phosphate is the main substrate of TAG biosynthesis and its production is catalysed by enzyme glycerol kinase (*GK*) [194]. *GK* gene has shown elevated expression in *S. dimorphus*, supported by previous study which has reported increase in lipid content and biomass productivity by overexpressing the enzyme glycerol kinase [195]. Further, esterification reactions for the formation of triacyglycerol are catalysed by enzymes lysophosphatidic acid acyltransferase (*LPAAT*), phosphatidic acid phosphatase (*PAP*), glycerol-3-phosphate acyltransferase (*GPAT*) and diacylglycerol acyltransferase (*DGAT*) [48]. Genes encoding glycerol-3-phosphate acyltransferase (*GPAT*) has shown down-regulation in *S. dimorphus* and other genes *LPAAT*, *PAP* and *DGAT* were significantly expressed in *S. dimorphus*.

Further, triacylglycerol formation can also takes place through extracting fatty acids from membrane lipids and consequently catalyzing reacylation by enzyme phospholipid: diacylglycerol acyltransferase (*PDAT*). In a previous report on *C. reinhardtii*, it has been observed that recycling of fatty acids from the membrane lipids has resulted in increase of lipid accumulation [196]. Gene *PDAT* was found to be expressed in *S. dimorphus*, however it was not expressed in *S. quadricauda*, suggesting important step of consideration for enhancement of lipid accumulation *Scenedesmus*.

5.8 Fatty acids and triacylglycerol degradation

Enzyme that catalyzes the degradation of triacylglycerol by hydrolyzing the ester bonds between fatty acids and glycerol backbone is triacylglycerol lipase (*TGL*). Free fatty acids are then released from triacylglycerol and diacylglycerol for further breakdown [197]. TAG lipase (*TGL*) has shown down-regulation in *S. dimorphus* signifying the increase in lipid accumulation. Further, degradation of free fatty acids by β-oxidation generates acetyl-CoA for TCA cycle [198]. Genes acyl-CoA oxidase (*ACOX*), enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase (*MFP*) and long-chain acyl-CoA synthetase (*ACSL*) have shown down-regulation whereas gene acetyl-CoA acyltransferase (*ACAT*) was found to be overexpressed in *S. dimorphus*. Hence, active catabolism of triacylglycerol and fatty acids may also be responsible for low lipid content in *S. quadricauda* (Figure 5.4).

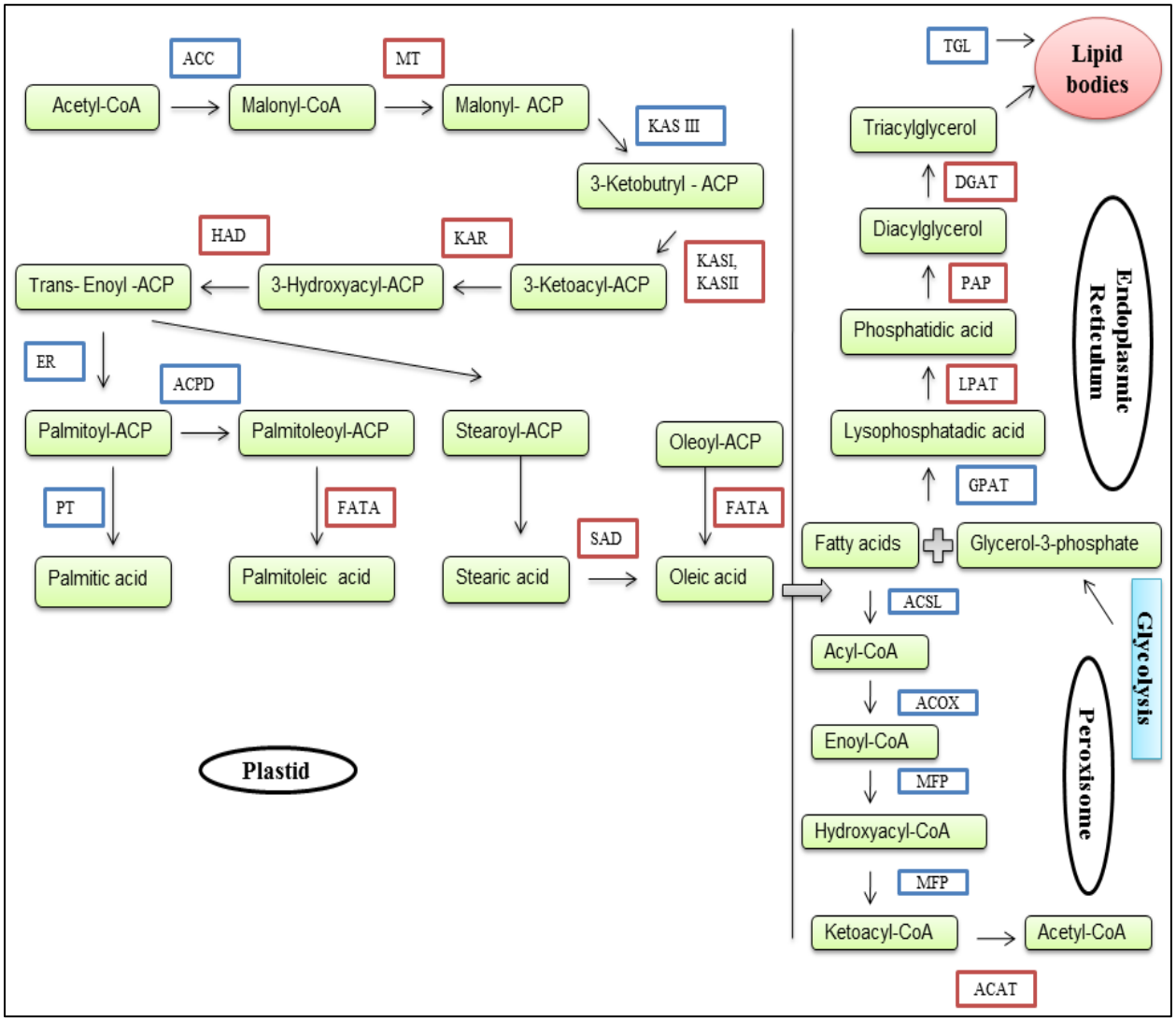


Figure 5.4 Fatty acid and triacylglyceol metabolism, enzymes are shown in red and blue colour boxes with red colour signifies up-regulation and blue colour signifies down-regulation in *S. dimorphus*. Enzyme abbreviations: ACC- Acetyl-CoA Carboxylase; MT- Malonyl transferase; KASI- β Ketoacyl-ACP synthase I; KASII- β-Ketoacyl-ACP synthase II; KASIII- β-Ketoacyl-ACP synthase III; KAR- β-Ketoacyl-ACP reductase; HAD- β-Hydroxyacyl ACP dehydratase; ER- Enoyl-ACP reductase; PT- Palmitoylthioestrase; FATA- Fatty Acyl-ACP thioestrase; SAD- Stearoyl CoA desaturase; ACPD- Palmitoyl desaturase; GPAT- Glycerol-3-phosphate acyltransferase; LPAAT- Lysophosphatidic acid acyltransferase; PAP- Phosphatidic acid phosphatase; DGAT- Diacylglycerol acyltransferase; ACSL- Long-chain acyl-CoA synthetase; ACOX- Acyl-CoA oxidase; MFP- Enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase; ACAT- Acetyl-CoA acyltransferase

5.9 SSRs from transcriptome data of *S. quadricauda* and *S. dimorphus*

SSRs are generally identified in both genomes and mRNAs and are the array of repeats of one to six nucleotides. SSRs are considered as tremendously variable entities across species due to the difference in numbers of repeat units in array of different species. Due to their co-dominant, locus-specific, highly polymorphic and PCR-based behavior, they are dominant genetic markers [199]. SSRs identification has revealed the prominence of tri-nucleotide repeats (85% in *S. quadricauda*, 97% in *S. dimorphus*,), where GCA, CAG, CTG, AGC and GCT were the most numerous repeats after that di-, tetra- and penta-nucleotides. Frequency of repeats was found to be inversely proportional to the repeat pattern length. Most SSRs have the repeat unit represention as five, followed by six and then others. It was observed that there was difference in the abundance of SSRs whereas the repeat pattern was identical in *S. quadricauda* and *S. dimorphus*. SSRs identification is vital to unveil the polymorphism in the contrasting species.

5.10 Transcription factor *WRI1* regulated the lipid production in *S. dimorphus*

Till now, various genes associated with lipid accumulation have been identified. However, applying genetic engineering techniques to multiple genes is quite difficult. Hence, the study was further extended towards the identification of transcription factors related to lipid production in *Scenedesmus*. Transcription factors regulate the expression of specific genes.

Transcription factors were predicted in *S. quadricauda* and *S. dimorphus* and it was observed that the most abundant TFs were found in MYB TF family in both the species. MYB TF family also included MYB-related families. A conserved DNA binding domain is present at N-terminus in MYB TFs and these transcription factors can be found in all eukaryotic organisms [200]. These transcription factors are involved in secondary metabolism, gene expression regulation and responses to environmental stresses [201]. In a report, soybean MYB was overexpressed in transgenic *Arabidopsis* plants, which resulted in the enhancement of the lipid content [2]. Nin-like transcription factor family was also found to be having numerous TFs in *S. dimorphus* than *S. quadricauda*. These transcription factors are involved in nitrate signaling [203] and have crucial role for high lipid content as nitrogen is the main requirement for the growth and biomass production in microalgae [204].

Further, lipid biosynthesis related transcription factors were identified and transcript abundance data revealed the enhanced expression of majority of the transcription factors in *S. dimorphus.* A transcription factor belonging to AP2 family was significantly enhanced in *S. dimorphus*, which was identified to be Wrinkled1 homolog. *WRI1* transcription factor controls genes involved in fatty acid synthesis and late glycolysis and hence have a direct role in TAG accumulation in seed and non-seed tissues [205]. Expression studies have revealed that *WRI1* mainly regulates genes such as enoyl-ACP reductase, BCCP (a subunit of ACCase), fatty acid desaturase, β-ketoacyl-ACP reductase, pyruvate dehydrogenase, sucrose synthase, plastidial pyruvate kinase and acyl-carrier protein [206, 207]. WRINKLED1 (*WRI1*) transcription factor controls the TAG production and storage in Arabidopsis seeds [208, 209]. *WRI1*-like genes have also reported to have related function in regulation of lipid accumulation in other plants i.e. maize and *Brassica napus* [210, 211]. In a study on closely related species oil palm (oleaginous) and date palm (non-oleaginous), *WRI1* homolog was found to be overexpressed in oil palm, which increased the overall TAG level [212]. RT-qPCR analysis of *WRI1* in three growth stages of *S. quadricauda* and *S. dimorphus* revealed its significant up-regulation in *S. dimorphus*, which validated the involvement of *WRI1* transcription factor in enhanced lipid biosynthesis.

One more transcription factor that belongs to NF-YB family was significantly overexpressed in *S. dimorphus*. NF-YA (CBF-B/HAP2), NF-YB (CBF-A/HAP3) and NF-YC (CBF-C/HAP5) are three subunits of Nuclear Factor Y and hence known as a multimeric transcription factor family. These transcription factors regulate several stress-induced responses and plant developmental processes and also known as CBFs (CCAAT box binding factors) and Heme-associated proteins (HAPs) [213]. NF-YB family includes *LEC* gene that encodes a HAP3 subunit of the CCAAT-binding TFs [214] and elevation in the seed oil content by overexpressing this gene has been reported in various plant species [215]. Elevated expression of NF-YB in *S. dimorphus*, suggests its association with high lipid accumulation in *S. dimorphus*.

SUMMARY

Understanding molecular mechanism of lipid content variation in different species and strains of microalgae is important for implementing any genetic intervention strategy to increase lipid content. Current study is the first report on the gene identification and expression analysis of fatty acid and triacylglycerol biosynthetic pathway in relation to lipid content in *Scenedesmus* species. Also, comparative transcriptome analysis was performed to unveil the molecular components and transcription factors involved in differential lipid accumulation of *Scenedesmu*s species.

Eighteen genes involved in FA and TAG biosynthetic pathway were identified by comparative genomics and transcriptome sequencing in *Scenedesmus dimorphus.* Investigation of expression level of the respective genes was carried out in three lipid content contrasting strains of *Scenedesmus* viz. SD12 (26%) and SD16 (5%) belonging to *Scenedesmus dimorphus* and SQ19 (14%) belonged to *Scenedesmus quadricauda*. Further, relative gene expression analysis was performed in growth stages i.e. late stationary phase, early stationary phase, exponential phase and two stress conditions i.e. high salinity and nitrogen deficiency. Six and three genes from FA and TAG biosynthesis, respectively, were identified for their association with lipid content among all three strains as inferred from combining and analyzing the results of three different expression studies. PCA analysis was also perfomed to establish the correlation of genes with lipid accumulation, on the basis of expression in stress conditions and growth stages. Six key genes were identified including β-ketoacyl-ACP synthase II (*KASII*), acyl carrier protein (*ACP*), fatty acyl-ACP thioestrase (*FATA*), lysophosphatidic acid acyltransferase (*LPAAT*), diacylglycerol acyl transferase (*DGAT*) and phosphatidic acid phosphatase (*PAP*), having correlation with lipid biosynthesis in *Scenedesmus*.

Transcriptomes of highest lipid containing strains of *S. quadricauda* (14 %) and *S. dimorphus* (26 %) were sequenced, assembled and then analyzed for performing comparative analysis. Metabolic pathways and genes were studied in both transcriptomes that were associated with lipid biosynthesis and transcriptional regulation was also analysed to reveal cross species

variation of lipid content. Similar core functional pathways were shared by two species and homology analysis showed high level of similarity with *Volvox carteri* after that *Chlamydomonas reinhardtii*. Pathway analysis and differential gene expression was performed to unveil molecular mechanisms playing significant role for lipid content variation in *Scenedesmus* species. Metabolic pathways viz. carbon fixation, photosynthesis, glycolysis, TCA cycle, pyruvate metabolism, FA and TAG biosynthesis and catabolism were examined for variation in gene expression levels. Pathways contributing the precursors and energy for fatty acid and triacylglycerol biosynthesis were enhanced in *S. dimorphus* whereas degradation of fatty acids and triacylglycerol was downregulated in *S. dimorphus*. Certain genes of carbon metabolism viz. acetyl-CoA synthetase (*ACOS*), enolase (*ENO*), ATP citrate lyase (*ACL*), pyruvate dehydrogenase (*PDH*), citrate synthase (*CS*), glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) and glycerol kinase (*GK*) were predicted as key regulators for increased lipid content of *S. dimorphus* as they divert the carbon and energy flux towards storage lipid production. Above study revealed that instead of a single gene, multiple genes are supposed to be targeted for genetic intervention in order to increase the lipid content of microalgal species.

Moreover, identification of transcription factors related to lipid biosynthesis and their transcript abundance provided insights towards molecular dynamics of differential accumulationof lipids in *S. quadricauda* and *S. dimorphus*. Homolog of Wrinkled1 transcription factor (*WRI1*) was identified and its expression analysis in growth stages of *S. quadricauda* and *S. dimorphus* revealed its enhanced expression in *S. dimorphus*. Correlated expression of WRI1 TF with lipid content also provided the clue for differential lipid accumulation in *Scenedesmus* species. This research work provided the understanding of molecular mechanism of lipid biosynthesis in *Scenedesmus* and identified molecular components to be utilized for genetic intervention, in order to develop suitable strains of microalgae to be used commercially for biodiesel production.

APPENDIX

Table A1 Fold change in expression of genes of various pathways in *S.dimorphus* versus *S. quadricauda*

KO ID	Symbol	Gene Name	EC Number	Fold Change
Photosynthesis				
K02704	PSBB	photosystem II CP47 chlorophyll apoprotein		1.17
K02705	PSBC	photosystem II CP43 chlorophyll apoprotein		0.70
K02706	PSBD	photosystem II P680 reaction center D2 protein		0.72
K02707	PSBE	photosystem II cytochrome b559 subunit alpha		0.06
K02716	PSBO	photosystem II oxygen-evolving enhancer protein 1		20.82
K02717	PSBP	photosystem II oxygen-evolving enhancer protein 2		120.67
K02689	PSAA	photosystem I P700 chlorophyll a apoprotein A1		1.31
K02690	PSAB	photosystem I P700 chlorophyll a apoprotein A2		1.40
K02692	PSAD	photosystem I subunit II		10.05
K14332	PSAO	photosystem I subunit PsaO		38.95
K02634	PETA	apocytochrome f		0.38
K02635	PETB	cytochrome b6		1.38
K02637	PETD	cytochrome b6-f complex subunit 4		0.72
K02639	PETF	ferredoxin		30.42
K02641	PETH	ferredoxin--NADP+ reductase	1.18.1.2	14.20
K02112	ATPF1B	F-type H+-transporting ATPase subunit beta	3.6.3.14	2.75
K02115	ATPF1G	F-type H+-transporting ATPase subunit gamma		19.68
K02111	ATPF1A	F-type H+-transporting ATPase subunit alpha	3.6.3.14	0.73
K02109	ATPF0B	F-type H+-transporting ATPase subunit b		25.53
K02110	ATPF0C	F-type H+-transporting ATPase subunit c		2.72
K02108	ATPF0A	F-type H+-transporting ATPase subunit a		0.39
K08909	LHCA3	light-harvesting complex I chlorophyll a/b binding protein 3		3.90
K08911	LHCA5	light-harvesting complex I chlorophyll a/b binding protein 5		16.88

K08913	LHCB2	light-harvesting complex II chlorophyll a/b binding protein 2		134.23
K08915	LHCB4	light-harvesting complex II chlorophyll a/b binding protein 4		14.89
K14172	LHCB7	light-harvesting complex II chlorophyll a/b binding protein 7		14.35
CARBON FIXATION				
K01601	RBCL	ribulose-bisphosphate carboxylase large chain	4.1.1.39	1.48
K00927	PGK,PGK	phosphoglycerate kinase	2.7.2.3	107.67
K00134	GAPDH	glyceraldehyde 3-phosphate dehydrogenase	1.2.1.12	15.67
K05298	GAPA	glyceraldehyde-3-phosphate dehydrogenase (NADP+) (phosphorylating)	1.2.1.13	19.30
K01623	ALDO	fructose-bisphosphate aldolase, class I	4.1.2.13	13.65
K03841	FBP,FBP	fructose-1,6-bisphosphatase I	3.1.3.11	4.57
K00615	TKT	transketolase	2.2.1.1	9.30
K01100	SBP	sedoheptulose-bisphosphatase	3.1.3.37	19.94
K01807	RPIA	ribose 5-phosphate isomerase A	5.3.1.6	1.84
K00855	PRK	phosphoribulokinase	2.7.1.19	29.67
K01595	PPC	phosphoenolpyruvate carboxylase	4.1.1.31	0.14
K01610	PCKA	phosphoenolpyruvate carboxykinase (ATP)	4.1.1.49	7.04
K00814	GPT	alanine transaminase	2.6.1.2	0.84
K14272	GGAT	glutamate--glyoxylate aminotransferase	2.6.1.4	0.58
K01006	PPDK	pyruvate, orthophosphate dikinase	2.7.9.1	2.06
K00029	MAEB	malate dehydrogenase (oxaloacetate-decarboxylating)(NADP+)	1.1.1.40	0.34
K00051	MDH	malate dehydrogenase (NADP+)	1.1.1.82	3.06
K14455	GOT2	aspartate aminotransferase, mitochondrial	2.6.1.1	0.19
K00026	MDH2	malate dehydrogenase	1.1.1.37	8.27
OXIDATIVE PHOSPHORYLATION				
K03878	ND1	NADH-ubiquinone oxidoreductase chain 1	1.6.5.3	6.02
K03881	ND4	NADH-ubiquinone oxidoreductase chain 4	1.6.5.3	1.00
K03883	ND5	NADH-ubiquinone oxidoreductase chain 5	1.6.5.3	0.03
K03934	NDUFS1	NADH dehydrogenase (ubiquinone) Fe-S protein 1	1.6.5.3	0.70
K03935	NDUFS2	NADH dehydrogenase (ubiquinone) Fe-S protein 2	1.6.5.3	1.02
K03936	NDUFS3	NADH dehydrogenase (ubiquinone) Fe-S protein 3	1.6.5.3	2.32
K03941	NDUFS8	NADH dehydrogenase (ubiquinone) Fe-S protein 8	1.6.5.3	4.70

K03942	NDUFV1	NADH dehydrogenase (ubiquinone) flavoprotein 1	1.6.5.3	2.00
K03943	NDUFV2	NADH dehydrogenase (ubiquinone) flavoprotein 2	1.6.5.3	3.34
K03953	NDUFA9	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex subunit 9		1.73
K03955	NDUFAB1	NADH dehydrogenase (ubiquinone) 1 alpha/beta subcomplex 1		0.09
K00234	SDHA	succinate dehydrogenase (ubiquinone) flavoprotein subunit	1.3.5.1	0.17
K00235	SDHB	succinate dehydrogenase (ubiquinone) iron-sulfur subunit	1.3.5.1	3.02
K00236	SDHC	succinate dehydrogenase (ubiquinone) cytochrome b560 subunit		26.34
K00411	UQCRFS1	ubiquinol-cytochrome c reductase iron-sulfur subunit	1.10.2.2	10.38
K00412	CYTB	ubiquinol-cytochrome c reductase cytochrome b subunit		10.75
K00413	CYC1	ubiquinol-cytochrome c reductase cytochrome c1 subunit		12.31
K02257	COX10	protoheme IX farnesyltransferase	2.5.1.-	0.58
K02256	COX1	cytochrome c oxidase subunit 1	1.9.3.1	1.25
K02258	COX11	cytochrome c oxidase assembly protein subunit 11		0.96
K02259	COX15	cytochrome c oxidase assembly protein subunit 15		0.11
K02111	ATPF1A	F-type H+-transporting ATPase subunit alpha	3.6.3.14	0.73
K02112	ATPF1B	F-type H+-transporting ATPase subunit beta	3.6.3.14	2.75
K02115	ATPF1G	F-type H+-transporting ATPase subunit gamma		19.68
K02108	ATPF0A	F-type H+-transporting ATPase subunit a		0.39
K02109	ATPF0B	F-type H+-transporting ATPase subunit b		25.53
K02110	ATPF0C	F-type H+-transporting ATPase subunit c		2.72
K02132	ATPEF1A	F-type H+-transporting ATPase subunit alpha		1.34
K02136	ATPEF1G	F-type H+-transporting ATPase subunit gamma		74.60
K02134	ATPEF1D	F-type H+-transporting ATPase subunit delta		5.98
K02145	ATPEV1A	V-type H+-transporting ATPase subunit A	3.6.3.14	0.29
K02147	ATPEV1B	V-type H+-transporting ATPase subunit B		2.68
K02148	ATPEV1C	V-type H+-transporting ATPase subunit C		2.66
K02149	ATPEV1D	V-type H+-transporting ATPase subunit D		13.96
K02151	ATPEV1F	V-type H+-transporting ATPase subunit F		5.51
K02144	ATPEV1H	V-type H+-transporting ATPase subunit H		3.80
K02154	ATPEV0A	V-type H+-transporting ATPase subunit a		1.31
K02155	ATPEV0C	V-type H+-transporting ATPase 16kDa proteolipid subunit		6.18
K02146	ATPEV0D,AT	V-type H+-transporting ATPase subunit d		2.46

	P6D			
K01535	HTA	H+-transporting ATPase	3.6.3.6	2.85
K01507	PPA	inorganic pyrophosphatase	3.6.1.1	57.76
GLYCOLYSIS				
K00844	HK	hexokinase	2.7.1.1	0.25
K01810	GPI	glucose-6-phosphate isomerase	5.3.1.9	0.62
K00850	PFK	6-phosphofructokinase 1	2.7.1.11	0.50
K03841	FBP	fructose-1,6-bisphosphatase I	3.1.3.11	4.57
K01623	ALDO	fructose-bisphosphate aldolase, class I	4.1.2.13	13.65
K00134	GAPDH	glyceraldehyde 3-phosphate dehydrogenase	1.2.1.12	15.67
K00927	PGK	phosphoglycerate kinase	2.7.2.3	107.67
K15633	GPMI	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	5.4.2.12	1.08
K01689	ENO	enolase	4.2.1.11	20.75
K00873	PK	pyruvate kinase	2.7.1.40	1.88
K00161	PDHA	pyruvate dehydrogenase E1 component alpha subunit	1.2.4.1	2.79
K00162	PDHB	pyruvate dehydrogenase E1 component beta subunit	1.2.4.1	1.10
K00627	PDHC	pyruvate dehydrogenase E2 component	2.3.1.12	1.93
K00382	DLD	dihydrolipoamide dehydrogenase	1.8.1.4	1.53
K00016	LDH	L-lactate dehydrogenase	1.1.1.27	0.31
K01568	PDC	pyruvate decarboxylase	4.1.1.1	0.12
K00121	ADHC	S-(hydroxymethyl)glutathione dehydrogenase	1.1.1.284	0.32
K00002	ADH	alcohol dehydrogenase (NADP+)	1.1.1.2	19.33
K00128	ADHN	aldehyde dehydrogenase (NAD+)	1.2.1.3	1.01
K01895	ACOS	acetyl-CoA synthetase	6.2.1.1	1.86
K01785	GALM	aldose 1-epimerase	5.1.3.3	2.61
K01835	PGM	phosphoglucomutase	5.4.2.2	2.52
K01792	GPE	glucose-6-phosphate 1-epimerase	5.1.3.15	117.92
K00131	GAPN	glyceraldehyde-3-phosphate dehydrogenase (NADP+)	1.2.1.9	1.56
K01610	PCKA	phosphoenolpyruvate carboxykinase (ATP)	4.1.1.49	7.04

TCA CYCLE				
K01647	CS	citrate synthase	2.3.3.1	0.63
K01648	ACL	ATP citrate lyase	2.3.3.8	1.01
K01681	ACO	aconitate hydratase	4.2.1.3	5.03
K00031	IDH	isocitrate dehydrogenase	1.1.1.42	61.22
K00030	IDH3	isocitrate dehydrogenase (NAD+)	1.1.1.41	1.07
K00164	OGDH	2-oxoglutarate dehydrogenase E1 component	1.2.4.2	1.01
K00658	DLST	2-oxoglutarate dehydrogenase E2 component	2.3.1.61	0.74
K00382	DLD	dihydrolipoamide dehydrogenase	1.8.1.4	1.53
K01899	LSC1	succinyl-CoA synthetase alpha subunit	6.2.1.4	11.60
K01900	LSC2	succinyl-CoA synthetase beta subunit	6.2.1.5	6.33
K00234	SDHA	succinate dehydrogenase (ubiquinone) flavoprotein subunit	1.3.5.1	0.17
K00235	SDHB	succinate dehydrogenase (ubiquinone) iron-sulfur subunit	1.3.5.1	3.02
K00236	SDHC	succinate dehydrogenase (ubiquinone) cytochrome b560 subunit		26.34
K01679	FUMC	fumarate hydratase, class II	4.2.1.2	0.97
K00026	MDH2	malate dehydrogenase	1.1.1.37	8.27
K01958	PC	pyruvate carboxylase	6.4.1.1	0.83
K01610	PCKA	phosphoenolpyruvate carboxykinase (ATP)	4.1.1.49	7.04
PYRUVATE METABOLISM				
K00161	PDHA	pyruvate dehydrogenase E1 component alpha subunit	1.2.4.1	2.79
K00162	PDHB	pyruvate dehydrogenase E1 component beta subunit	1.2.4.1	1.10
K00627	DLAT	pyruvate dehydrogenase E2 component	2.3.1.12	1.93
PENTOSE PHOSPHATE PATHWAY				
K01810	GPI	glucose-6-phosphate isomerase	5.3.1.9	0.62
K00036	G6PD	glucose-6-phosphate 1-dehydrogenase	1.1.1.49	41.96
K01057	PGLS	6-phosphogluconolactonase	3.1.1.31	1.62
K00033	PGD	6-phosphogluconate dehydrogenase	1.1.1.44	67.81
K00615	TKT	transketolase	2.2.1.1	9.30
K00616	TAL	transaldolase	2.2.1.2	6.10
K01807	RPIA	ribose 5-phosphate isomerase A	5.3.1.6	1.84
K01835	PGM	phosphoglucomutase	5.4.2.2	2.52
K00948	PRPS	ribose-phosphate pyrophosphokinase	2.7.6.1	0.65

K00131	GAPN	glyceraldehyde-3-phosphate dehydrogenase (NADP+)	1.2.1.9	1.56
K01623	ALDO	fructose-bisphosphate aldolase, class I	4.1.2.13	13.65
K03841	FBP,FBP	fructose-1,6-bisphosphatase I	3.1.3.11	4.57
K00850	PFK	6-phosphofructokinase 1	2.7.1.11	0.50
STARCH METABOLISM				
K01087	OTSB	trehalose 6-phosphate phosphatase	3.1.3.12	0.06
K16055	TPS	trehalose 6-phosphate synthase/phosphatase	2.4.1.15	0.58
K00012	UGDH	UDPglucose 6-dehydrogenase	1.1.1.22	0.65
K08678	UXS1	UDP-glucuronate decarboxylase	4.1.1.35	1.13
K15920	XYL4	beta-D-xylosidase 4	3.2.1.37	0.07
K01835	PGM	phosphoglucomutase	5.4.2.2	2.52
K00844	HK	hexokinase	2.7.1.1	0.25
K01810	GPI	glucose-6-phosphate isomerase	5.3.1.9	0.62
K00975	GLGC	glucose-1-phosphate adenylyltransferase	2.7.7.27	1.23
K00703	GLGA	starch synthase	2.4.1.21	0.71
K00700	GLGB	1,4-alpha-glucan branching enzyme	2.4.1.18	0.88
K00688	GLGP	starch phosphorylase	2.4.1.1	1.29
K01176	AMYA	alpha-amylase	3.2.1.1	0.46
K01177	AMYB	beta-amylase	3.2.1.2	0.26
K00705	MALQ	4-alpha-glucanotransferase	2.4.1.25	0.90
FATTY ACID BIOSYNTHESIS				
K01962	ACCA	acetyl-CoA carboxylase carboxyl transferase subunit alpha	6.4.1.2	0.20
K01961	ACCC	acetyl-CoA carboxylase, biotin carboxylase subunit	6.3.4.14	2.77
K01963	ACCD	acetyl-CoA carboxylase carboxyl transferase subunit beta	6.4.1.2	0.21
K00648	KAS III	3-oxoacyl-ACP synthase III	2.3.1.180	0.16
K09458	KAS II	3-oxoacyl-ACP synthase II	2.3.1.179	1.14
K00059	KAR	3-oxoacyl-ACP reductase	1.1.1.100	1.14
K02372	HAD	3-hydroxyacyl-ACP dehydratase	4.2.1.59	4.00
K00208	EAR	enoyl-ACP reductase	1.3.1.9	0.53
K10782	FATA	fatty acyl-ACP thioesterase A	3.1.2.14	5.45
K03921	SAD	stearoyl-ACP desaturase	1.14.19.2	5.05
K10255	FAD	omega-6 fatty acid desaturase	1.14.19.	0.64

TRICAYLGLYCEROL BIOSYNTHESIS				
K00864	GK	glycerol kinase	2.7.1.30	1.70
K00630	GPAT	glycerol-3-phosphate O-acyltransferase	2.3.1.15	0.35
K15728	PAP	phosphatidate phosphatase	3.1.3.4	41.94
K11155	DGAT	diacylglycerol O-acyltransferase	2.3.1.20	14.07
FATTY ACID AND TRICAYLGLYCEROL CATABOLISM				
K07513	ACAT	acetyl-CoA acyltransferase 1	2.3.1.16	1.06
K10527	MFP	enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase	4.2.1.17	0.61
K00232	ACOX	acyl-CoA oxidase	1.3.3.6	0.57
K00249	ACD	acyl-CoA dehydrogenase	1.3.8.7	0.22
K01897	LACS	long-chain acyl-CoA synthetase	6.2.1.3	2.37
K14674	TGL	TAG lipase	3.1.1.3	0.27

REFERENCES

[1] C. Wan, Y. Li, "Effect of hot water extraction and liquid hot water pretreatment on the fungal degradation of biomass feedstocks", *Bioresource Technology*, vol. 102 (20), pp. 9788-9793 2011.

[2] T. M. Mata, A. A. Martins, N. S. Caetano, "Microalgae for biodiesel production and other applications: A review", *Renewable and Sustainable Energy Reviews*, vol. 14, pp. 217-232, 2010.

[3] Y. Chisti, "Biodiesel from microalgae beats bioethanol", *Trends In Biotechnology*, vol. 26, pp. 126–13, 2008.

[4] P. Schenk, S. Thomas-Hall, E. Stephens, U. Marx, J. Mussgnug, C. Posten, O. Kruse, B. Hankamer, "Second generation biofuels: high-efficiency microalgae for biodiesel production", *BioEnergy Research*, vol. 1, pp. 20-43, 2008.

[5] F. X. Malcata, "Microalgae and biofuels: a promising partnership", *Trends In Biotechnology*, vol. 29 (11), pp. 542-549, 2011.

[6] M. K. Lam, K. T. Lee, "Microalgae biofuels: a critical review of issues, problems and the way forward", *Biotechnology Advances*, vol. 30 (3), pp. 673-690, 2012.

[7] R. H. Wijffels, M. J. Barbosa, M. H. Eppink, "Microalgae for the production of bulk chemicals and biofuels", *Biofuels, Bioproducts and Biorefining*, vol. 4 (3), pp. 287-295, 2010.

[8] A. Demirbas, M. F. Demirbas, "Algae energy: algae as a new source of biodiesel", *Springer Science & Business Media*, 2010.

[9] Q. Hu, M. Sommerfeld, E. Jarvis, M. Ghirardi, M. Posewitz, "Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances", *The Plant Journol*, vol. 54, pp. 621-639, 2008.

[10] P.I. Leonardi, C. Popovich, C. Damiani, "Feedstock for second- generation biodiesel: microalgae's biology and oil composition", *Economic Effects of Biofuel Production*, InTech Publisher, Rijeka, pp. 318-346, 2011.

[11] G. Ahlgren, L. Lundstedt, M. Brett, C. Forsberg, "Lipid composition and food quality of some freshwater phytoplankton for cladoceran zooplankters", *Journol of Plankton Research*, vol.(4), pp. 809–818, 1990.

[12] L. Rodolfi, G. ChiniZittelli, N. Bassi, "Microalgae for oil: strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor", *Biotechnology and Bioengineering*, vol. 102 (1), pp. 100–112, 2009.

[13] L. Xin, H. Hong-Ying, Y. Jia, "Lipid accumulation and nutrient removal properties of a newly isolated freshwater microalga, *Scenedesmus sp.* LX1, growing in secondary effluent", *New Biotechnology*, vol. 27 (1), pp. 59–63, 2009.

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

[15] S. T. Girisha, K. Ravikumar, V. Girish , B. R. Mrunalini, "Growing of *Chlorella*, *Scenedesmus* and *Botryococus* in sewage water for biodiesel production", *Archives of Applied Science Research*, vol. 6 (1), pp. 131-138, 2014.

[16] B. Liu, C. Benning, "Lipid metabolism in microalgae distinguishes itself", Current Opinion Biotechnology, vol. 14 (2), pp. 300–309, 2013.

[17] E. W. Becker, "Microalgae: Biotechnology and Microbiolog", *Cambridge University Press*, Cambridge, UK, pp. 350-382, 1994.

[18] Y. Chisti, "Biodiesel from microalgae", *Biotechnology Advances*, vol. 25, pp. 294-306, 2007.

[19] N. M. D. Courchesne, A. Parisien, B.Wang, C. Q. Lan, "Enhancement of lipid production using biochemical, genetic and transcription factor engineering approaches", *Journal of Biotechnology*, vol. 141(1), pp. 31-41, 2009.

[20] R. Radakovits, R. E. Jinkerson, A. Darzins, M. C. Posewitz, "Genetic engineering of algae for enhanced biofuel production", *Eukaryotic cell*, vol. 9(4), pp. 486-501, 2010.

[21] M.T. Guarnieri, A. Nag, S.L. Smolinski, A. Darzins, M. Seibert, P.T. Pienkos, "Examination of triacylglycerol biosynthetic pathways via *de novo* transcriptomic and proteomic analyses in an unsequenced microalga", *PLoS ONE*, vol. 6, e25851, 2011.

[22] E. Derelle, C. Ferraz, S. Rombaut, P. Rouzé, A. Z. Worden, S. Robbens, F. Partensky, S. Degroeve, S. Echeynié, R. Cooke, Y. Saeys, "Genome analysis of the smallest free-living

eukaryote *Ostreococcus tauri* unveils many unique features", *Proceedings of the National Academy of Sciences,* USA, vol. 103, pp. 11647–11652, 2006.

[23] S. S. Merchant, S. E. Prochnik, O. Vallon, E.H. Harris, S. J. Karpowicz, G.B. Witman, A. Terry, A. Salamov, L. K. Fritz-Laylin, L. Maréchal-Drouard, W. F. Marshall, "The *Chlamydomonas* genome reveals the evolution of key animal and plant functions", *Science*, vol. 318, pp. 245–250, 2007.

[24] B. Palenik, J. Grimwood, A. Aerts, P. Rouzé, A. Salamov, N. Putnam, C. Dupont, R. Jorgensen, E. Derelle, S. Rombauts, K. Zhou,"The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation", *Proceedings of the National Academy of Sciences*, USA , vol. 104, pp. 7705–7710, 2007.

[25] A.Z. Worden, J.H. Lee, T. Mock, P. Rouzé, M. P. Simmons, A. L. Aerts, A. E. Allen, M. L. Cuvelier, E. Derelle, M.V. Everett, E. Foulon, "Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*", *Science*, vol. 324 (5924), pp.268-272, 2009.

[26] S. E. Prochnik, J. Umen, A. M. Nedelcu, A. Hallmann, S. M. Miller, I. Nishii, P. Ferris, A. Kuo, T. Mitros, L. K. Fritz-Laylin, U. Hellsten, "Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri*", *Science*, vol. 329 (5988), pp.223-226, 2010.

[27] G. Blanc, I. Agarkova, J. Grimwood, A. Kuo, A. Brueggeman, D. D. Dunigan, J. Gurnon, I. Ladunga, E. Lindquist, S. Lucas, J. Pangilinan, "The genome of the polar eukaryotic microalga *Coccomyxa subellipsoidea* reveals traits of cold adaptation", *Genome biology*, vol. 13(5), pp.1, 2012.

[28] G. Blanc, G. Duncan, I. Agarkova, M. Borodovsky, J. Gurnon, A. Kuo, E. Lindquist, S. Lucas, J. Pangilinan, J. Polle, A. Salamov, "The Chlorella variabilis NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex", *The Plant Cell*, vol. 22 (9), pp.2943-2955, 2010

[29] L. Fang, D. Sun, Z. Xu, J. He, S. Qi, X. Chen, W. Chew, J. Liu, "Transcriptomic analysis of a moderately growing subisolate *Botryococcus braunii* 779 (Chlorophyta) in response to nitrogen deprivation", *Biotechnology for Biofuels*, 8 (1), p.1, 2015.

[30] Q. Wang, Y. Lu, Y. Xin, L. Wei, S. Huang, J. Xu, "Genome editing of model oleaginous microalgae *Nannochloropsis* spp. by CRISPR/Cas9", *The Plant Journal*, 2016.

[31] M. Wan, P. Liu, J. Xia, J. N. Rosenberg, G. A. Oyler, M. J. Betenbaugh, Z. Nie, G. Qiu, "The effect of mixotrophy on microalgal growth, lipid content, and expression levels of three pathway genes in *Chlorella sorokiniana*", *Applied Microbiology and Biotechnology*, vol. 91(3), pp. 835-844, 2011.

[32] A. Lei, H. Chen, G. Shen, Z. Hu, L. Chen, J. Wang, "Expression of fatty acid synthesis genes and fatty acid accumulation in *Haematococcus pluvialis* under different stressors", *Biotechnology for Biofuels*, vol. 5(1), pp. 1, 2012.

[33] J. Fan, Y. Cui, M. Wan, W. Wang, Y. Li, "Lipid accumulation and biosynthesis genes response of the oleaginous *Chlorella pyrenoidosa* under three nutrition stressors", *Biotechnology for Biofuels*, vol. 7 (17) 2014.

[34] R. Huerlimann, E. J. Steinig, H. Loxton, K.R. Zenger, D. R. Jerry, K. Heimann, "Effects of growth phase and nitrogen starvation on expression of fatty acid desaturases and fatty acid composition of *Isochrysis aff. galbana* (TISO)", *Gene*, vol. 545(1), pp. 36-44, 2014.

[35] O. Avidan, U. Pick, "Acetyl-CoA synthetase is activated as part of the PDH-bypass in the oleaginous green alga *Chlorella desiccate*", *Journol of Experimental Botany*, vol. 66, pp. 7287-7298, 2015.

[36] J. Liu, D. Han, K. Yoon, Q. Hu, Y. Li, "Characterization of type 2 diacylglycerol acyltransferases in *Chlamydomonas reinhardtii* reveals their distinct substrate specificities and functions in triacylglycerol biosynthesis", *The Plant Journal*, 2016.

[37] R. Radakovits, R. E. Jinkerson, S. I. Fuerstenberg, H. Tae, R. E. Settlage, J. L. Boore, M. C. Posewitz, "Draft genome sequence and genetic transformation of the oleaginous alga *Nannochloropsis gaditana*", *Nature communications*, vol. 3 (686), 2012.

[38] S. L. Guo, X. Q. Zhao, Y. Tang, C. Wan, M. A. Alam, S. H. Ho, J. S. Chang, "Establishment of an efficient genetic transformation system in *Scenedesmus obliquus*", *Journal of biotechnology*, vol. 163 (1), pp. 61-68, 2013.

[39] M. Muto, Y. Fukuda, M. Nemoto, T. Yoshino, T. Matsunaga, T. Tanaka, "Establishment of a genetic transformation system for the marine pennate diatom *Fistulifera* sp. strain JPCC DA0580—A high triglyceride producer", *Marine biotechnology*, vol. 15 (1), pp. 48-55, 2013.

[40] B. Zorin, O. Grundman, I. Khozin-Goldberg, S. Leu, M. Shapira, Y. Kaye, S. Boussiba, "Development of a nuclear transformation system for oleaginous green alga *Lobosphaera*

(*Parietochloris*) *incisa* and genetic complementation of a mutant strain, deficient in arachidonic acid biosynthesis", *PloS ONE,* vol. 9 (8), e105223, 2014.

[41] S. Kim, Y. C. Lee, D. H. Cho, H. U. Lee, Y. S. Huh, G. J. Kim, H. S. Kim, "A simple and non-invasive method for nuclear transformation of intact-walled *Chlamydomonas reinhardtii*", *PloS ONE,* vol. 9 (7), e101018, 2014.

[42] B. Prasad, N. Vadakedath, H. J. Jeong, T. General, M. G. Cho, W. Lein, "*Agrobacterium tumefaciens*-mediated genetic transformation of haptophytes (*Isochrysis* species)", *Applied Microbiology and Biotechnology*, vol. 98 (20), pp. 8629-8639, 2014.

[43] W. H. Xie, C. C. Zhu, N. S. Zhang, D. W. Li, W. D. Yang, J. S. Liu, H. Y. Li, "Construction of novel chloroplast expression vector and development of an efficient transformation system for the diatom *Phaeodactylum tricornutum*", *Marine Biotechnology*, vol. 16 (5), pp. 538-546, 2014.

[44] P. T. Pratheesh, M. Vineetha, G. M. Kurup, "An efficient protocol for the *Agrobacteriu*m-mediated genetic transformation of microalga *Chlamydomonas reinhardtii*" *Molecular Biotechnology*, vol. 56 (6), pp. 507-515, 2014.

[45] S. Bae, S. Park, J. Kim, J. S. Choi, K. H. Kim, D. Kwon, T. S. Seo, "Exogenous gene integration for microalgal cell transformation using a nanowire-incorporated microdevice", *ACS Applied Materials & Interfaces*, vol. 7 (49), pp. 7554-27561, 2015.

[46] Y. Yamaoka, D. Achard, S. Jang, B. Legéret, S. Kamisuki, D. Ko, Y. Li-Beisson, "Identification of a *Chlamydomonas* plastidial 2-lysophosphatidic acid acyltransferase and its use to engineer microalgae with increased oil content", *Plant biotechnology journal*, 2016.

[47] R. Miller, G. Wu, R.R. Deshpande, A. Vieler, K. Gartner, X. Li, E.R. Moellering, S. Zauner, A.J. Cornish, B. Liu, "Changes in transcript abundance in *Chlamydomonas reinhardtii* following nitrogen deprivation predict diversion of metabolism" *Plant Physiology*, vol. 154 pp. 1737-1752, 2010.

[48] H. Rismani-Yazdi, B.Z. Haznedaroglu, K. Bibby, J. Peccia, "Transcriptome sequencing and annotation of the microalgae *Dunaliella tertiolecta*: pathway description and gene discovery for production of next-generation biofuels", *BMC Genomics*, vol. 12 (148) 2011.

[49] H. Lv, G. Qu, X. Qi, L. Lu, C. Tian, Y. Ma, "Transcriptome analysis of *Chlamydomonas reinhardtii* during the process of lipid accumulation", *Genomics*, vol. 101, pp. 229-237, 2013.

[50] J.O. Davidson, J. Overton, R. Waikel, "Transcriptome analysis of *Chlorella protothecoides* to identify novel pro-lipid genes for biofuel production", *FASEB Journol*, vol. 26, pp. 790-12, 2012.

[51] C. Shang, G. Bi, Z. Yuan, Z. Wang, M. A. Alam, J. Xie, "Discovery of genes for production of biofuels through transcriptome sequencing of *Dunaliella parva*", *Algal Research*, vol. 13, pp. 318–326, 2016.

[52] D. M. Riaño-Pachón, L. G. G. Corrêa, R. Trejos-Espinosa, B. Mueller-Roeber, "Green transcription factors: a *Chlamydomona*s overview", *Genetics*, vol. 179 (1), pp. 31-39, 2008.

[53] J. Hu, D. Wang, J. Li, G. Jing, K. Ning, J. Xu, "Genome-wide identification of transcription factors and transcription-factor binding sites in oleaginous microalgae *Nannochloropsis*", *Scientific Reports*, vol. 4, 2014.

[54] G. Knothe, J. Krahl, J. Van Gerpen (Eds.), The biodiesel handbook. Elsevier, 2015.

[55] S. Al-Zuhair, "Production of biodiesel: possibilities and challenges", *Biofuels, Bioproducts and Biorefining*, vol. 1(1), pp. 57-66, 2007.

[56] C. E. Goering, A. W. Schwab, M. J. Daugherty, E. H. Pryde, A. J. Heakin, "Fuel properties of eleven vegetable oils", *Transactions of the ASAE*, vol. 25(6), pp. 1472-1477, 1982.

[57] A. Darzins, P. Pienkos, L. Edye, "Current status and potential for algal biofuels production", *A report to IEA Bioenergy Task*, vol. 39, 2010.

[58] Y. Li, M. Horsman, N. Wu, C. Q. Lan, N. Dubois-Calero, "Biofuels from microalgae", *Biotechnology progress*, vol. 24(4), pp. 815-820, 2008.

[59] A. Demirbas, "Progress and recent trends in biodiesel fuels", *Energy conversion and management*, vol. 50(1), pp. 14-34, 2009.

[60] Z. Wen, M. B. Johnson, "Microalgae as a feedstock for biofuel production", 2009.

[61] R. Garofalo, "Algae and aquatic biomass for a sustainable production of 2nd generation biofuels", *AquaFUELs-Taxonomy, Biology and Biotechnology*, pp.1-258, 2009.

[62] P. Lavens, P. Sorgeloos, "Manual on the production and use of live food for aquaculture", *Food and Agriculture Organization*, 1996.

[63] G. E. Fogg, B. Thake, "Algal cultures and phytoplankton ecology", *University of Wisconsin Press*, 1987.

[64] P. T. Pienkos, Project Peer Review Algal Biomass Conversion, National Renewable Energy Laboratory, 2015.

[65] A. Darzins, P. Pienkos, L. Edye, "Current status and potential for algal biofuels production", *A report to IEA Bioenergy Task,* vol. 39, 2010.

[66] T. G. Dunahay, E. E. Jarvis, S. S. Dais, P. G. Roessler, "Manipulation of microalgal lipid production using genetic engineering", In *Seventeenth Symposium on Biotechnology for Fuels and Chemicals*, Humana Press, pp. 223-231, 1996.

[67] M. Frac, S. Jezierska-Tys, J. Tys, "Microalgae for biofuels production and environmental applications: A review", *African Journal of Biotechnology*, vol. 9(54), pp. 9227-9236, 2010.

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

[69] K. K. Sharma, H. Schuhmann, P. M. Schenk, "High lipid induction in microalgae for biodiesel production", *Energies*, vol. 5(5), pp. 1532-1553, 2012.

[70] K. Skjanes, C. Rebours, P. Lindblad, "Potential for green microalgae to produce hydrogen, pharmaceuticals and other high value products in a combined process", *Critical reviews in biotechnology*, vol. 33(2), pp. 172-215, 2013.

[71] T. Matthew, W. Zhou, J. Rupprecht, L. Lim, S. R. Thomas-Hall, A. Doebbe, P. M. Schenk, "The metabolome of *Chlamydomonas reinhardtii* following induction of anaerobic H2 production by sulfur depletion", *Journal of Biological Chemistry*, vol. 284(35), pp. 23415-23425, 2009.

[72] A. P. Dean, D. C. Sigee, B. Estrada, J. K. Pittman, "Using FTIR spectroscopy for rapid determination of lipid accumulation in response to nitrogen limitation in freshwater microalgae", *Bioresource Technology*, vol. 101(12), pp. 4499-4507, 2010.

[73] L. Xin, H. Hong-ying, G. Ke, S. Ying-xue, "Effects of different nitrogen and phosphorus concentrations on the growth, nutrient uptake, and lipid accumulation of a freshwater microalga *Scenedesmus* sp.", *Bioresource technology*, vol. 101(14), pp. 5494-5500, 2010.

[74] M. Takagi, T. Yoshida, "Effect of salt concentration on intracellular accumulation of lipids and triacylglyceride in marine microalgae *Dunaliella* cells", *Journal of Bioscience And Bioengineering*, vol. 101(3), pp. 223-226, 2006.

[75] C. J. Zhu, Y. K. Lee, T. M. Chao, "Effects of temperature and growth phase on lipid and biochemical composition of *Isochrysis galbana* TK1", *Journal of Applied Phycology*, vol. 9(5), pp. 451-457, 1997.

[76] S. M. Renaud, L. V. Thinh, G. Lambrinidis, D. L. Parry, "Effect of temperature on growth, chemical composition and fatty acid composition of tropical Australian microalgae grown in batch cultures", *Aquaculture*, vol. 211(1), pp. 195-214, 2002.

[77] A. Converti, A. A. Casazza, E. Y. Ortiz, P. Perego, M. Del Borghi, "Effect of temperature and nitrogen concentration on the growth and lipid content of *Nannochloropsis oculata* and *Chlorella vulgaris* for biodiesel production", *Chemical Engineering and Processing: Process Intensification*, vol. 48(6), pp. 1146-1151, 2009.

[78] J. B. Guckert, K. E. Cooksey, "Triglyceride accumulation and fatty acid profile changes in *Chlorella* (Chlorophyta) during high Ph-induced cell cycle Inhibition", *Journal of Phycology*, vol. 26(1), pp. 72-79, 1990.

[79] M. R. Brown, G. A. Dunstan, S. Norwood, K. A. Miller, "Effects of harvest stage and light on the biochemical composition of the diatom *Thalassiosira pseudonana", Journal of phycology*, vol. 32(1), pp. 64-73, 1996.

[80] D. L. Alonso, E. H. Belarbi, J. M. Fernández-Sevilla, J. Rodríguez-Ruiz, E. M. Grima, "Acyl lipid composition variation related to culture age and nitrogen concentration in continuous culture of the microalga *Phaeodactylum tricornutum*", *Phytochemistry*, vol. 54(5), pp. 461-471, 2000.

[81] Z. Ikaran, S. Suárez-Alvarez, I. Urreta, S. Castañón, "The effect of nitrogen limitation on the physiology and metabolism of *Chlorella vulgaris* var L3", *Algal Research*, vol. 10, pp. 134-144, 2015.

[82] T. Li, M. Gargouri, J. Feng, J. J. Park, D. Gao, C. Miao, S. Chen, "Regulation of starch and lipid accumulation in a microalga *Chlorella sorokiniana*", *Bioresource technology*, vol. 180, pp. 250-257, 2015.

[83] J. Cao, H. Yuan, B. Li, J. Yang, "Significance evaluation of the effects of environmental factors on the lipid accumulation of *Chlorella minutissima* UTEX 2341 under low-nutrition heterotrophic condition", *Bioresource technology*, vol. 152, pp. 177-184, 2014.

[84] A. M. Santos, R. H. Wijffels, P. P. Lamers, "pH-upshock yields more lipids in nitrogen-starved *Neochloris oleoabundans*", *Bioresource technology*, vol. 152, pp. 299-306, 2014.

[85] Y. Xiao, J. Zhang, J. Cui, Y. Feng, Q. Cui, "Metabolic profiles of *Nannochloropsis oceanica* IMET1 under nitrogen-deficiency stress", *Bioresource Technology*, vol. 130, pp. 731-738, (2013).

[86] http://sussle.org

[87] F. R. Trainor, J. R. Cain, L. E. Shubert, "Morphology and nutrition of the colonial green alga *Scenedesmus*: 80 years later" *The Botanical Review*, vol. 42(1), pp. 5-25, 1976.

[88] E. Hegewald, "Taxonomy and phylogeny of *Scenedesmus*", *Algae*, vol. 12(4), pp. 235-246, 1997.

[89] P. I. Leonardi, C. Popovich, C. Damiani, "Feedstock for second- generation biodiesel: microalgae's biology and oil composition", *Economic effects of biofuel production*, InTech Publisher, Rijeka, pp. 318-346, 2011.

[90] G. Ahlgren, L. Lundstedt, M. Brett, C. Forsberg, "Lipid composition and food quality of some freshwater phytoplankton for cladoceran zooplankters", *Journol of Plankton Research*, vol. 12(4), pp. 809–818, 1990.

[91] L. Xin, H. Hong-Ying, Y. Jia, "Lipid accumulation and nutrient removal properties of a newly isolated freshwater microalga, *Scenedesmus sp.* LX1, growing in secondary effluent", *New Biotechnology*, vol. 27(1), pp. 59–63, 2009.

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

[93] P. Prabakaran, A.D. Ravindran, "*Scenedesmus* as a potential source of biodiesel among selected microalgae", *Current Science*, vol. 102, pp. 616-620, 2012.

[94] J. Jena, M. Nayak, H.S. Panda, N. Pradhan, C. Sarika, P.K. Panda, B.V.K. Rao, R.B. Prasad, L.B. Sukla, "Microalgae of Odisha coast as a potential source for biodiesel production", *World Environment*, vol. 2(1), pp.11-16, 2012.

[95] V. T. Duong, F. Ahmed, S. R. Thomas-Hall, S. Quigley, E. Nowak, P. M. Schenk, "High protein-and high lipid-producing microalgae from northern Australia as potential feedstock for animal feed and biodiesel", *Frontiers in Bioengineering and Biotechnology*, vol. 3(53) 2015.

[96] H. V. Rani, G. S. Geeta, "Isolation of microalgae with biodiesel productivity prospects", *Karnataka Journal of Agricultural Sciences*, vol. 24(4), 2012.

[97] S. Roccaa, A. Agostini, J. Giuntoli, L. Marelli, "Biofuels from algae: technology options, energy balance and GHG emissions: Insights from a literature review", *Scientific and Technical Research Reports*, Publications Office of the European Union, 2015.

[98] T. L. Chacón-Lee, G. E. González-Marino, "Microalgae for healthy foods – Possibilities and challenges", *Comprehensive reviews in food science and food safety*, vol. 9(6), pp. 655-675, 2010.

[99] A. C. Guedes, H. M. Amaro, F. X. Malcata, "Microalgae as sources of high added value compounds—a brief review of recent work", *Biotechnology Progress*, vol. 27(3), pp. 597-613, 2011b.

[100] E. W. Becker, "Biotechnology and exploitation of the green alga *Scenedesmus obliquus* in India:, *Biomass*, vol. 4(1), pp. 1-19, 1984.

[101] M. Lürling, "Phenotypic plasticity in the green algae *Desmodesmu*s and *Scenedesmus* with special reference to the induction of defensive morphology." *International Journal of Limnology*, vol. 39(2), pp. 85-101, 2003.

[102] O. Pultz, W. Gross, "Valuable products from biotechnology of microalgae", *Applied Microbiology Biotechnology*, vol. 65(6), pp. 635-648, 2004.

[103] W. Becker, "Microalgae in human and animal nutrition", In A. Richmond (Ed.), Handbook of microalgal culture: Biotechnology and applied phycology, pp. 312-351, 2004

[104] M. A. Borowitzka, "Vitamins and fine chemicals from micro-algae", In M. A. Borowitzka, L. J. Borowitzka (Eds.), *Micro-algal biotechnology*, Cambridge University Press, Cambridge, pp. 153-196, 1988.

[105] Z. Tukaj, M. K. Matusiak, J. Lewandowska, J. Szurkowski, "Changes in the pigment pattern and the photosynthetic activity during a light-induced cell circle of the green alga *Scenedesmus armatus*" *Plant Physiology Biochemistry*, vol. 41(4), pp. 337-344, 2003.

[106] M. C. Ceron, I. Campos, J. F. Sanchez, F. G. Acien, E. Molina, J. M. Fernandez-Sevilla, "Recovery of lutein from Microalgae biomass: Development of a process for *Scenedesmus almeriensis* biomass", *Journal of Agriculture and Food Chemistry*, vol. 56(24), pp. 11761-11766, 2008.

[107] K. Skjanes, R. Celine, L. Peter, "Potential for green microalgae to produce hydrogen, pharmaceuticals and other high value products in a combined process", *Critical reviews in Biotechnology*, vol. 33(2), pp. 172-215, 2013.

[108] S. Qin, G. X. Liu, Z. Y. Hu, "The accumulation and metabolism of astaxanthin in *Scenedesmus obliquus* (Chlorophyceae)", *Process Biochemistry,* vol. 43(8), pp. 795-802, 2008.

[109] L. Gouveia,, A. P. Batista, I. Sousa, A. Raymundo, N. M. Bandarra, "Microalgae in novel food products In *K. Papadoupoulos*", *Food chemistry research developments*, Nova Science Publishers, 2008.

[110] Z. Jouni, Z. Makhoul, "Patent US 20120238522, 2012, Sep 20". *Carotenoid containing compositions and methods*, 2012

[111] C. Y. Chu, W. R. Liao, R. Huang, L. P. Lin, "Haemagglutinating and antibiotic activities of freshwater microalgae", *World Journal of Microbiology & Biotechnology,* vol. 20(8), pp. 817-825, 2004.

[112] H. W. Karen, B. Maarten, M. Anita, B. Heinke, "Extraction of pigments and fatty acids from the green alga *Scenedesmus Obliquus* (Chlorophyceae)", *Aquatic Ecology,* vol. 34(2), pp. 119- 126, 2000.

[113] P. Indira, R. Biswajit, "Commercial and industrial applications of micro-algae – A review", *Journal of Algal Biomass Utilization,* vol. 3(4), pp. 89-100, 2012.

[114] A. C. Guedes, M. S. Gião, A. Matias, A. V. M. Nunes, C. M. M. Duarte, M. E. Pintado, F. X. Malcata, "Supercritical fluid extraction of carotenoids – including lutein and chlorophylls a, b and c from a wild strain of *Scenedesmus obliquus* for use in food processing", *Journal of Food Engineering,* vol. 116(2), pp. 478-482, 2013.

[115] G. A. Catarina, S. G. Maria, A. M. Ana, V. M. N. Ana, E. P. Manuela, M. M. D. Catarina, M. F. Xavier, "Supercritical fluid extraction of carotenoids and chlorophylls a, b, c, from a wild strain of *Scenedesmus obliquus* for use in food processing", *Journal of Food Engineering,* vol. 116(2), pp. 478-482, 2013.

[116] R. M. A. Abedin, H. M. Taha, "Antibacterial and antifungal activity of cyanobacteria and green microalgae. Evaluation of medium components by Plackett-Burman design for antimicrobial activity of *Spirulina Platensis*", *Global Journal of Biotechnology and Biochemistry,* vol. 3(1), pp. 22-31, 2008.

[117] A. C. Guedes, R. B. Catarina, M. A. Helena, I. P. Claudia, X. M. Francisco, "Microalgal and cyanobacterial cell extracts for use as natural antibacterial additives against

food pathogen", *International Journal of Food Science and Technology*, vol. 46(4), pp. 862-870, 2011a.

[118] A. M. Humphrey, "Chlorophyll as a color and functional ingredient", *Journal of Food Science*, vol. 69(5), pp. 422-425, 2004.

[119] S. Bhagavathy, P. Sumathi, S. B. I. Jancy, "Green algae *Chlorococcum humicola* - A new source of bioactive compounds with antimicrobial activity", *Asian Journal of Tropical Biomedicine*, vol. 1(1), pp. S1-S7, 2011.

[120] A. P. Desbois, A. Mearns-Spragg, V. J. Smith, "A fatty acid from the diatom *Phaeodactylum tricornutum* is antibacterial agent diverse bacteria including multi resistant *Staphylococcus aureus* (MRSA)", *Marine Biotechnology*, vol. 11(1), pp. 45-52, 2009.

[121] A. Cagliari, R. Margis, F. dos Santos Maraschin, A. C. Turchetto-Zolet, G. Loss, M. Margis-Pinheiro, "Biosynthesis of triacylglycerols in plants and algae", *International Journol of Plant Biology*, vol. 2(1) pp. 10, 2011.

[122] A. R. Grossman, M. Croft, V. N. Gladyshev, S. S. Merchant, M. C. Posewitz, S. Prochnik, M. H. Spalding, "Novel metabolism in *Chlamydomonas* through the lens of genomics", *Current Opinion in Plant Biology*, vol. 10(2), pp. 190-198, 2007.

[123] S. S. Merchant, S. E. Prochnik, O. Vallon, E. H. Harris, S. J. Karpowicz, G. B. Witman, W. F. Marshall, "The Chlamydomonas genome reveals the evolution of key animal and plant functions", *Science*, vol. 318(5848), pp. 245-250, 2007.

[124] M. Matsuzaki, O. Misumi, T. Shin-i, S. Maruyama, M. Takahara, S.Y. Miyagishima, T. Mori, K. Nishida, F., Nishida, K. Yagisawa, Y. Yoshida, "Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D", *Nature*, vol. 428 (6983), pp. 653-657, 2004.

[125] C. Bowler, A.E. Allen, J.H. Badger, J. Grimwood, K. Jabbari, A. Kuo, U. Maheswari, C. Martens, F. Maumus, R.P. Otillar, E. Rayko, "The Phaeodactylum genome reveals the evolutionary history of diatom genomes". *Nature*, vol. 456 (7219), pp.239-244, 2008.

[126] E.V. Armbrust, J.A. Berges, C. Bowler, B.R. Green, D. Martinez, N.H. Putnam, S. Zhou, A.E. Allen, K.E. Apt, M. Bechner, M.A. Brzezinski, "The genome of the diatom Thalassiosira pseudonana: ecology, evolution, and metabolism" *Science*, vol. 306 (5693), pp.79-86, 2004.

[127] W. R. Riekhof, B. B. Sears, C. Benning, "Annotation of genes involved in glycerolipid biosynthesis in *Chlamydomonas reinhardtii*: discovery of the betaine lipid synthase BTA1Cr", *Eukaryotic Cell*, vol. 4(2), pp. 242-252, 2005.

[128] N. Sato, T. Moriyama, "Genomic and biochemical analysis of lipid biosynthesis in the unicellular rhodophyte *Cyanidioschyzon merolae*: lack of a plastidic desaturation pathway results in the coupled pathway of galactolipid synthesis", *Eukaryotic cell*, vol. 6(6), pp. 1006-1017, 2007.

[129] G. Zhang, G. Guo, X. Hu, Y. Zhang, Q. Li, R. Li, L. Chen, "Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome" *Genome research*, vol. 20(5), pp. 646-654, 2010.

[130] J. Fan, C. Yan, C. Andre, J. Shanklin, J. Schwender, C. Xu, "Oil accumulation is controlled by carbon precursor supply for fatty acid synthesis in *Chlamydomonas reinhardtii*", *Plant and Cell Physiology*, vol. 53(8), pp. 1380-1390, 2012.

[131] C. Gao, Y. Wang, Y. Shen, D. Yan, X. He, J. Dai, Q. Wu, "Oil accumulation mechanisms of the oleaginous microalga *Chlorella protothecoides* revealed through its genome, transcriptomes, and proteomes", *BMC genomics*, vol. 15(1), pp. 1, 2014.

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

[133] H. Rismani-Yazdi, B. Z. Haznedaroglu, C. Hsin, J. Peccia, "Transcriptomic analysis of the oleaginous microalga *Neochloris oleoabundans* reveals metabolic insights into triacylglyceride accumulation", *Biotechnology for Biofuels*, vol. 5(1), pp.1, 2012.

[134] J. Li, D. Han, D. Wang, K. Ning, J. Jia, L. Wei, Q. Hu, "Choreography of transcriptomes and lipidomes of *Nannochloropsis* reveals the mechanisms of oil synthesis in microalgae", *The Plant Cell*, vol. 26(4), pp. 1645-1665, 2014.

[135] Z. Sun, Z. G. Zhou, H. Gerken, F. Chen, J. Liu, "Screening and characterization of oleaginous *Chlorella* strains and exploration of photoautotrophic *Chlorella protothecoides* for oil production", *Bioresource technology*, vol. 184, pp. 53-62, 2015.

[136] H. Peng, D. Wei, G. Chen, F. Chen, "Transcriptome analysis reveals global regulation in response to CO 2 supplementation in oleaginous microalga *Coccomyxa subellipsoidea* C-169", *Biotechnology for Biofuels*, vol. 9(1), pp. 151, 2016.

[137] J. Msanne, D. Xu, A.R. Konda, J.A. Casas-Mollano, T. Awada, E.B. Cahoon, "Metabolic and gene expression changes triggered by nitrogen deprivation in the photoautotrophically grown microalgae *Chlamydomonas reinhardtii* and *Coccomyxa* sp. C-169", *Phytochemistry*, vol. 75, pp. 50–59, 2012.

[138] N. R. Boyle, M. D. Page, B. Liu, I. K. Blaby, D. Casero, J. Kropat, S. D. Gallaher, "Three acyltransferases and nitrogen-responsive regulator are implicated in nitrogen starvation-induced triacylglycerol accumulation in *Chlamydomonas*", *Journal of Biological Chemistry*, vol. 287(19), pp. 15811-15825, 2012.

[139] I. K. Blaby, A. G. Glaesener, T. Mettler, S. T. Fitz-Gibbon, S. D. Gallaher, B. Liu, C. Benning, "Systems-level analysis of nitrogen starvation–induced modifications of carbon metabolism in a *Chlamydomonas reinhardtii* starchless mutant", *The Plant Cell*, vol. 25(11), pp. 4305-4323, 2013.

[140] E. C. Carpinelli, A. Telatin, N. Vitulo, C. Forcato, M. D'Angelo, R. Schiavon, G. Valle, "Chromosome scale genome assembly and transcriptome profiling of *Nannochloropsis gaditana* in nitrogen depletion" *Molecular plant*, vol. 7(2), pp. 323-335, 2014.

[141] L. Yao, T. W. Tan, Y. K. Ng, K. H. K. Ban, H. Shen, H. Lin, Y. K. Lee, "RNA-Seq transcriptomic analysis with Bag2D software identifies key pathways enhancing lipid yield in a high lipid-producing mutant of the non-model green alga *Dunaliella tertiolecta*" *Biotechnology for Biofuels*, vol. 8(1), pp. 1, (2015).

[142] W. C. Chang, H. Q. Zheng, C. N. N. Chen, "Comparative transcriptome analysis reveals a potential photosynthate partitioning mechanism between lipid and starch biosynthetic pathways in green microalgae", *Algal Research*, vol .16, pp. 54-62, 2016.

[143] D. M. Riaño-Pachón, L. G. G. Corrêa, R. Trejos-Espinosa, B. Mueller-Roeber, "Green transcription factors: a *Chlamydomonas* overview", *Genetics*, vol. 179(1), pp. 31-39, 2008.

[144] P. Pérez-Rodríguez, D.M. Riaño-Pachón, L.G.G. Corrêa, S.A. Rensing, B. Kersten, B. Mueller-Roeber, "PlnTFDB: updated content and new features of the plant transcription factor database", *Nucleic Acids Research*, vol. 38(Database issue) pp. D822-D827, 2010.

[145] H. Zhang, J. Jin, L. Tang, Y. Zhao, X. Gu, G. Gao, J. Luo, "PlantTFDB 2.0: update and improvement of the comprehensive plant transcription factor database", *Nucleic acids research*, vol. 39(suppl 1), pp. D1114-D1117, 2011.

[146] A. Vieler, G. Wu, C. H. Tsai, B. Bullard, A. J. Cornish, C. Harvey, M. S. Campbell, "Genome, functional gene annotation, and nuclear transformation of the heterokont oleaginous alga *Nannochloropsis oceanica* CCMP1779", *PLoS Genetics*, vol. 8(11), e1003064, 2012.

[147] A. Ibánez-Salazar, S. Rosales-Mendoza, A. Rocha-Uribe, J. I. Ramírez-Alonso, I. Lara-Hernández, A. Hernández-Torres, R. E. Soria-Guerra, "Over-expression of Dof-type transcription factor increases lipid production in *Chlamydomonas reinhardtii*", *Journal of Biotechnology*, vol. 184, pp. 27-38, 2014.

[148] J. Fan, H. Xu, Y. Luo, M. Wan, J. Huang, W. Wang, Y. Li, "Impacts of CO2 concentration on growth, lipid accumulation, and carbon-concentrating-mechanism-related gene expression in oleaginous *Chlorella*", *Applied microbiology and biotechnology*, vol. *99*(5), pp. 2451-2462, 2015.

[149] R.S. Gour , A. Kant, R.S. Chauhan, "Screening of micro algae for Growth and lipid accumulation properties", *Journol of Algal Biomass Utilization*, vol. 5 (1), pp. 38–46, 2014.

[150] R.R.L. Guillard, "Methods for microflagellates and nannoplankton", In Handbook of Phycological Methods: Culture Methods and Growth Measurements, Cambridge University Press, New York, pp. 69-85, 1973.

[151] E.G. Bligh, W.J. Dyer, "A rapid method for total lipid extraction and purification", *Canadian Journol of Biochemistry and Physiology*, vol. 37, pp. 911-917, 1959.

[152] M.S. Cooper, W. R. Hardin, T.W. Peterse , R.A. Cattolico, "Visualizing "green oil" in live algal cells", *Journol of Biosciences and Bioengineering*, vol. 109, pp.198–201, 2010.

[153] J.J. Doyle, J.L. Doyle, "Isolation of plant DNA from fresh tissue", *Focus*, vol. 12 (1), pp. 13-15, 1990.

[154] J. Vandesompele, K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, F. Speleman, "Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes", *Genome Biology*, vol. 3(7), pp. research0034.1–research0034.11, 2002.

[155] T.D. Schmittgen, K.J. Livak, "Analyzing real-time PCR data by the comparative C (T) method", *Nature Protocols*, vol. 3 (6), pp. 1101–1108, 2008.

[156] L. I. Smith, "A tutorial on Principal Components Analysis", 2002.

[157] B. Li, C.N. Dewey, "RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome", *BMC Bioinformatics*, vol. 12, pp. 323, 2011.

[158] N. Ji, L. Li, L. Lin, S. Lin, "Screening for suitable reference genes for quantitative real-time PCR in *Heterosigma akashiwo* (Raphidophyceae)", *PloS ONE*, vol. 10(7), pp. e0132183, 2015.

[159] S.Wu, J. Zhou, X. Cao, S. Xue, "Determination of internal controls for quantitative gene expression of *Isochrysis zhangjiangensis* at nitrogen stress condition", *Journol of Ocean University China*, vol.15 (1), pp. 137-144, 2016.

[160] K. Hwangbo, Joon-Woo Ahn, Jong-Min Lim, Youn-Il Park, Jang Ryol Liu, Won-Joong Jeong, "Overexpression of stearoyl-ACP desaturase enhances accumulations of oleic acid in the green alga *Chlamydomonas reinhardtii*", *Plant Biotechnology Rep*orts, vol. 8 (2), pp. 135-142, 2014.

[161] Y. Moriya, M. Itoh, S. Okuda, A.C. Yoshizawa, M. Kanehisa, "KAAS: an automatic genome annotation and pathway reconstruction server", *Nucleic Acids Research*, vol. 35, pp. W182-W185, 2007.

[162] U.S. Ramli, D.S. Baker, P.A. Quant, J.L. Harwood, "Control analysis of lipid biosynthesis in tissue cultures from oil crops shows that flux control is shared between fatty acid synthesis and lipid assembly", *Biochemical Journol*, vol. 364, pp. 393–401, 2002.

[163] D. J. Schultz, M. C. Suh, J. B. Ohlrogge, "Stearoyl-Acyl Carrier Protein and Unusual Acyl-Acyl Carrier Protein Desaturase Activities Are Differentially Influenced by Ferredoxin", *Plant Physiology*, vol. 124 (2), pp. 681-692, 2000.

[164] R. Vaezi, J. A. Napier, O. Sayanova, "Identification and Functional Characterization of Genes Encoding Omega-3 Polyunsaturated Fatty Acid Biosynthetic Activities from Unicellular Microalgae", *Marine Drugs*, vol. 11(12), pp. 5116–5129, 2013.

[165] J. Liu , Z. Sun , Y. Zhong, J. Huang , Q. Hu , F. Chen, "Stearoyl-acyl carrier protein desaturase gene from the oleaginous microalga *Chlorella zofingiensis*: cloning, characterization and transcriptional analysis", *Planta*, vol. 236(6), pp.1665-76, 2012.

[166] J.K. Branen, Tzyy-Jen Chiou, N.J. Engeseth, "Overexpression of Acyl Carrier Protein-1 Alters Fatty Acid Composition of Leaf Tissue in Arabidopsis", *Plant Physiology*, vol. 127 (1), pp. 222-229, 2001.

[167] Othman, Ashida, "Gene expression and promoter characterisation of stearoyl-ACP desaturase and acyl carrier protein genes from oil palm fruits", Masters Thesis, Universiti Putra Malaysia, 2010.

[168] J.G. Jaworski, E.E. Goldschmidt, P.K. Stumpf, "Fat metabolism in higher plants. Properties of the palmityl acyl carrier protein: Stearyl acyl carrier protein elongation system in maturing safflower seed extracts", Archives of Biochemistry and Biophysics, vol. 163, pp. 769–776, 1974.

[169] A. Jones, H.M. Davies, T.A. Voelker, "Palmitoyl-Acyl Carrier Protein (ACP) Thioesterase and the Evolutionary-Origin of Plant Acyl-ACP Thioesterases", *The Plant Cell*, vol. 7, pp. 359-371, 1995.

[170] R. Sakthivel, S. Elumalai, M.M. Arif, "Microalgae lipid research, past, present: A critical review for biodiesel production in the future", *Journol of Experimental Sciences*, vol. 2 (10), pp. 29-49, 2011.

[171] Xiao-dong Deng, Jia-jiaCai, Xiao-wen Fei, "Involvement of phosphatidate phosphatase in the biosynthesis of triacylglycerols in *Chlamydomonas reinhardtii*", *Journal of Zhejiang University SCIENCE B*, vol. 14 (12), pp. 1121–1131, 2013.

[172] J. M. Chen, W. C. Qi, S. Y. Wang, R. Z. Guan, H. S. Zhang, "Correlation of Kennedy pathway efficiency with seed oil content of canola (*Brassica napus* L.) lines", Canadian Journol of Plant Pathology, vol. 91(2), pp. 251-259, 2011.

[173] J. Xu, M. Kazachkov, Y. Jia, Z. Zheng, J. Zou, "Expression of a type 2 diacylglycerol acyltransferase from *Thalassiosira pseudonana* in yeast leads to incorporation of docosahexaenoic acid b-oxidation intermediates into triacylglycerol", *FEBS Journol*, vol. 280, pp. 6162–6172, 2013.

[174] X. Johnson, J. Alric, "Central carbon metabolism and electron transport in *Chlamydomonas reinhardtii*: metabolic constraints for carbon partitioning between oil and starch", *Eukaryotic Cell*, vol.12, pp. 776-793, 2013.

[175] Z. K. Yang, Y. H. Ma, J.W. Zheng, W.D. Yang, J. S. Liu, H.Y. Li, "Proteomics to reveal metabolic network shifts towards lipid accumulation following nitrogen deprivation in the

diatom *Phaeodactylum tricornutum*", *Journol of Applied Phycology*, vol. 26, pp. 73-82, 2014.

[176] B.D. Woodworth, R.L. Mead, C.N. Nichols, D.R.J. Kolling, "Photosynthetic light reactions increase total lipid accumulation in carbon-supplemented batch cultures of *Chlorella vulgaris*", *Bioresource Technology*, vol. 179, pp.159-164, 2015.

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

[178] Z. K.Yang, Y.F. Niu, Y.H. Ma, J. Xue, M. H. Zhang, W.D. Yang, H.Y. Li, "Molecular and cellular mechanisms of neutral lipid accumulation in diatom following nitrogen deprivation", *Biotechnology for Biofuels*, vol.6 (1), pp. 1, 2013.

[179] X. Deng, J. Cai, Y. Li, X. Fei, "Expression and knockdown of the PEPC1 gene affect carbon flux in the biosynthesis of triacylglycerols by the green alga *Chlamydomonas reinhardtii*", *Biotechnology Letters*, vol. 36 (11), pp. 2199-2208, 2014.

[180] Q. L. Tian, D. J. Shi, X. H. Jia, H. L. Mi, X. W. Huang, P. M. He, "Recombinant expression and functional analysis of a *Chlamydomonas reinhardtii* bacterial-type phosphoenolpyruvate carboxylase gene fragment", *Biotechnology Letters*, vol. 36 (4), pp. 821-827, 2014.

[181] J.E. Polle, P. Neofotis, A. Huang, W. Chang, K. Sury, E. M. Wiech, "Carbon partitioning in green algae (chlorophyta) and the enolase enzyme", *Metabolites*, vol. 4 (3), pp. 612-628, 2014.

[182] P. Song, L. Li, J. Liu, "Proteomic analysis in nitrogen-deprived *Isochrysis galbana* during lipid accumulation", *PLoS ONE*, vol. 8, pp. e82188, 2013.

[183] M. A. Troncoso-Ponce, R. Garces, E. Martinez-Force, "Glycolytic enzymatic activities in developing seeds involved in the differences between standard and low oil content sunflowers (*Helianthus annuus* L.)", *Plant Physiology and Biochemistry*, vol. 48, pp. 961–965, 2010.

[184] X. Tang, J. Lee, W. N. Zhen, "Engineering the fatty acid metabolic pathway in *Saccharomyces cerevisiae* for advanced biofuel production", *Metabolic Engineering Comunications*, vol. 2, pp. 58-66, 2015.

[185] J. Sheng, X. Feng, "Metabolic engineering of yeast to produce fatty acid-derived biofuels: bottlenecks and solutions", *Frontiors in Microbiology*,vol. 6, 2015.

[186] X. Deng, J. Cai, X. Fei, "Effect of the expression and knockdown of citrate synthase gene on carbon flux during triacylglycerol biosynthesis by green algae *Chlamydomonas reinhardtii*", *BMC Biochemistry*, vol. 14, pp. 38, 2013.

[187] O. Avidan, U. Pick, "Acetyl-CoA synthetase is activated as part of the PDH-bypass in the oleaginous green alga *Chlorella desiccate*", *Journol of Experimental Botany*, vol. 66, pp. 7287-7298, 2015.

[188] H. Ma, X. Wang, Y. F. Niu, Z.-K. Yang, M. H. Zhang, Z. M. Wang, W. D. Yang, J. S. Liu, H.Y. Li, "Antisense knockdown of pyruvate dehydrogenase kinase promotes the neutral lipid accumulation in the diatom *Phaeodactylum tricornutum*", *Microbial Cell Factories*, vol. 13, pp. 100, 2014.

[189] J. Yan, R. Cheng, X. Lin, S. You, K. Li, H. Rong, Y. Ma, "Overexpression of acetyl-CoA synthetase increased the biomass and fatty acid proportion in microalga *Schizochytrium*", *Applied Microbialogy and Biotechnology*, vol. 97, pp. 1933-1939, 2013.

[190] Z. Ma, C.H. Chu, D. Cheng, "A novel direct homogeneous assay for ATP citrate lyase", *Journol of Lipid Research*, vol. 50, pp. 2131-2135, 2009.

[191] J.L. Watts, J. Browse, "A palmitoyl-CoA- specific delta9 fatty acid desaturase from *Caenorhabditis elegans*", *Biochemical and Biophysical Research Communications*, vol. 272 (1), pp. 263-269, 2000.

[192] A. Jones, H.M. Davies, T.A. Voelker, "Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases", *Plant Cell*, vol. 7 (3), pp. 359-371, 1995.

[193] J. Fan, C. Andre, C. Xu, "A chloroplast pathway for the de novo biosynthesis of triacylglycerol in *Chlamydomonas reinhardtii*", *FEBS letters*, vol. 585 (12), pp. 1985-1991, 2011.

[194] A. Cagliari, R. Margis, F. dos Santos Maraschin, A. C. Turchetto-Zolet, G. Loss, M. Margis-Pinheiro, "Biosynthesis of triacylglycerols in plants and algae", International Journol of Plant Biology, vol. 2(1), pp. 10, 2011.

[195] M. Muto, Y. Fukuda, M. Nemoto, T. Yoshino, T. Matsunaga, T. Tanaka, "Establishment of a genetic transformation system for the marine pennate diatom *Fistulifera sp.* strain JPCC DA0580 - a high triglyceride producer", *Marine Biotechnology*, vol.15, 2014.

[196] K. Yoon, D. Han, Y. Li, M. Sommerfeld, Q. Hu, "Phospholipid: diacylglycerol acyltransferase is a multifunctional enzyme involved in membrane lipid turnover and degradation while synthesizing triacylglycerol in the unicellular green microalga *Chlamydomonas reinhardtii*", *Plant Cell*, vol. 24, pp. 3708-3724, 2012.

[197] J. Chen, R.K. Tan, X.J. Guo, Z. L. Fu, Z. Wang, Z.-Y. Zhang, X.-L. Tan, "Transcriptome analysis comparison of lipid biosynthesis in the leaves and developing seeds of *Brassica napus*", *PLoS ONE*, vol. 10, 2015.

[198] A. Mühlroth, K. Li, G. Røkke, P. Winge, Y. Olsen, M.F. Hohmann-Marriott, O. Vadstein, A.M. Bones, "Pathways of lipid metabolism in marine algae, co-expression network, bottlenecks and candidate genes for enhanced production of EPA and DHA in species of *Chromista*", *Marine Drugs*, vol. 11 (11), pp. 4662-4697, 2013.

[199] H. Kuntal, V. Sharma, H. Daniell, "Microsatellite analysis in organelle genomes of Chlorophyta", *Bioinformation*, vol. 8, pp. 255-259, 2012.

[200] H. Du, B. R. Feng, S. S. Yang, Y. B. Huang, Y. X. Tang, "The R2R3-MYB transcription factor gene family in maize", *PloS ONE*, vol. 7 (6), pp. e37463, 2012.

[201] S. Ambawat, P. Sharma, N. R. Yadav, R. C. Yadav, "MYB transcription factor genes as regulators for plant responses: an overview", *Physiology and Molecular Biology of Plants*, vol. 19(3), pp. 307-321, 2013.

[202] Y.F. Liu, Q.T. X. Li, Q.X LuSong, S.M. Lam, W.K. Zhang, B. Ma, Q. Lin, W.Q. Man, W.G. Du, G.H. Shui, "Soybean GmMYB73 promotes lipid accumulation in transgenic plants", *BMC Plant Biology*, vol. 14 (1), pp. 1, 2014.

[203] M. Konishi, S. Yanagisawa, "*Arabidopsis* NIN-like transcription factors have a central role in nitrate signaling", *Nature Communications*, vol. 4, pp. 1617, 2013.

[204] C. Remacle, G. Eppe, N. Coosemans, E. Fernandez, H. Vigeolas, "Combined intracellular nitrate and NIT2 effects on storage carbohydrate metabolism in *Chlamydomonas*", *Journol of Experimental Botany*, vol. 65 (1), pp. 23-33, 2014.

[205] S. Baud, S. Wuillème, A. To, C. Rochat, L. Lepiniec, "Role of WRINKLED1 in the transcriptional regulation of glycolytic and fatty acid biosynthetic genes in Arabidopsis", *The Plant Journal*, vol. 60(6), pp. 933-947, 2009.

[206] S. A. Ruuska, T. Girke, C. Benning, J. B. Ohlrogge, "Contrapuntal networks of gene expression during Arabidopsis seed filling", *The Plant Cell*, vol. *14*(6), pp. 1191-1206, 2002.

[207] S. Baud, M. S. Mendoza, A. To, E. Harscoët, L. Lepiniec, B. Dubreucq, "WRINKLED1 specifies the regulatory action of LEAFY COTYLEDON2 towards fatty acid metabolism during seed maturation in Arabidopsis", *The Plant Journal*, vol. 50(5), pp. 825-838, 2007.

[208] N. Focks, C. Benning, "Wrinkled1: a novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism", *Plant Physiology*, vol. 118 (1), pp. 91-101, 1998.

[209] A. Cernac, C. Benning, "WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in Arabidopsis", *The Plant Journal*, vol. 40 (4), pp. 575-585, 2004.

[210] J. Liu, W. Hua, G. Zhan, F. Wei, X. Wang, G. Liu, H. Wang, "Increasing seed mass and oil content in transgenic Arabidopsis by the overexpression of wri1-like gene from *Brassica napus*", *Plant Physiology and Biochemistry*, vol. 48 (1), pp. 9-15, 2010.

[211] B. Pouvreau, S. Baud, V. Vernoud, V. Morin, C. Py, G. Gendrot, P. M. Rogowsky, "Duplicate maize Wrinkled1 transcription factors activate target genes involved in seed oil biosynthesis", *Plant Physiology*, vol. 156 (2), pp. 674-686, 2011.

[212] F. Bourgis, A. Kilaru, X. Cao, G. F. Ngando-Ebongue, N. Drira, J. B. Ohlrogge, V. Arondel, "Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning", *Proceedings of the National Academy of Sciences*, vol. 108 (30), pp. 12527-12532, 2011.

[213] F. Zhang, M. Han, Q. Lv, F. Bao, Y. He, "Identification and expression profile analysis of NUCLEAR FACTOR-Y families in *Physcomitrella patens*", *Frontiors of Plant Science*, vol. 6, pp. 642, 2015.

[214] M. Suzuki, D.R. McCarty, "Functional symmetry of the B3 network controlling seed development", *Current Opinion in Plant Biology*, vol. 11, pp. 548-553, 2008.

[215] B. Shen, W.B. Allen, P. Zheng, C. Li, K. Glassman, J. Ranch, D. Nubel, M.C. Tarczynski, "Expression of ZmLEC1 and ZmWRI1 increases seed oil production in maize", *Plant Physiology*, vol. 153, pp. 980-987, 2010.

PUBLICATIONS

Research publications:

- **Tamanna Sharma**, Rakesh Singh Gaur, Anil Kant Thakur, Rajinder S. Chauhan (2015). Lipid content in *Scenedesmus* species correlates with multiple genes of fatty acid and triacylglycerol biosynthetic pathways. *Algal Research*, 12, 341-349. [ISSN 2211-9264, IF: 5.014].

- **Tamanna Sharma**, Rajinder S. Chauhan (2016). Comparative Transcriptomics Reveals Molecular Components Associated with Differential Lipid Accumulation between Microalgal sp., *Scenedesmus dimorphus* and *Scenedesmus quadricauda*. *Algal Research*, 19, 109-122. [ISSN 2211-9264, IF: 4.694].

Conference publication:

- **Tamanna Sharma**, Rajinder Singh Chauhan (2015). Transcriptome Wide Computational Mining and Differential Expression Of Genes Involved in Fatty Acid and Triacylglycerol Metabolism in *Scenedesmus* sp. *Proceedings of the International Conference on Advances in Biomedical Engineering, Cancer Biology, Bioinformatics and Applied Biotechnology*, Jawaharlal Nehru University, New Delhi, India, 25- 26 July 2015.

NCBI Genbank sequence submissions

- Sharma T., Sood A. and Chauhan R.S. (2015) *Scenedesmus dimorphus* enoyl-ACP reductase mRNA, partial cds. **Accession No.** KT203348

- Sharma T., Sood A. and Chauhan R.S. (2015) *Scenedesmus dimorphus* stearoyl-ACP desaturase mRNA, partial cds. **Accession No.** KT203349

- Sharma T., Sood A. and Chauhan R.S. (2015) *Scenedesmus dimorphus* fatty acyl-ACP thioesterase mRNA, partial cds. **Accession No.** KT203350

- Sharma T., Sood A. and Chauhan R.S. (2015) *Scenedesmus dimorphus* palmitoyl desaturase mRNA, partial cds. **Accession No.** KT203351

- Sharma T., Sood A. and Chauhan R.S. (2015) *Scenedesmus dimorphus* omega-3 fatty acid desaturase mRNA, partial cds. **Accession No.** KT203352