

Trascriptome meta-analysis of microalga *Dunaliella tertiolecta* under stress condition

Bahman Panahi^{*1}, Mohammad Farhadian², Seyyed Abolghasem Mohammadi³, Mohammad Amin Hejazi⁴

1- Department of Genomics, Branch for Northwest & West region, Agricultural Biotechnology Research Institute of Iran (ABRII), Agricultural Research, Education and Extension Organization (AREEO), Tabriz, Iran

2-Department of Animal Science, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

3- Department of Plant Breeding and Biotechnology, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

4- Department of Food Biotechnology, Branch for Northwest & West region, Agricultural Biotechnology Research Institute of Iran (ABRII), Agricultural Research, Education and Extension Organization (AREEO), Tabriz, Iran

*Corresponding authors, Bahman Panahi, b.panahi@abrii.ac.ir

Abstract

Microalgae are photosynthetic organisms, which are considered as a potential source for sustainable metabolite production. Furthermore, stress conditions can affect metabolite production. In this study, a meta-analysis of RNA-seq experiments was performed to evaluate the response of metabolite biosynthesis pathways in *Dunaliella tertiolecta* to abiotic stress conditions, including high light, nitrogen deficiency, and high salinity. The results indicated down-regulation of light reaction, photorespiration, tetrapyrrole, and lipid-related pathways in salt stress. In comparison to salt stress, nitrogen deficiency mostly induced light reaction and photorespiration metabolisms. The up-regulation of phosphoenolpyruvate carboxylase, phosphoglucose isomerase, bisphosphoglycerate mutase, and glucose-6-phosphate-1-dehydrogenase (involved in central carbon metabolism) was observed under salt, high light, and nitrogen stress conditions. Interestingly, the results indicated that the meta-genes (i.e., modules of genes strongly correlated) tended to be located in a hub of stress-specific PPI (Protein-Protein Interaction) networks. Module enrichment of meta-genes PPI networks highlighted the cross talk between photosynthesis, fatty acids, starch, and sucrose metabolism under multiple stress conditions. Moreover, it was observed that the coordinated expression of the tetrapyrrole intermediated with meta-genes involved in starch biosynthesis. The results of the present study also showed that some pathways such as vitamin B6 metabolism, methane metabolism, ribosome biogenesis, and folate biosynthesis responded to different stress factors specifically. In conclusion, the results of this study revealed the main pathways underlying the abiotic stress responses for optimized metabolite production by the microalga *Dunaliella* in future studies. PRISMA check list was also included in the study.

40 **Keyword:** Network analysis, Central carbon metabolism, Translation-related genes, Tetrapyrrole
41 biosynthesis.

Introduction

Photosynthetic microalgae have been considered as a potential source of secondary metabolites and proposed as a promising cell factory since they have efficient photosynthesis apparatus and biomass production (Spolaore et al. 2006, Del Campo et al. 2007). Environmental stress conditions can redirect microalga energy flux towards the production and accumulation of fatty acids, starch, and carotenoids (Dellomonaco et al. 2010). This feature is exploited to produce biofuels, pigments, and starch; however, it can reduce metabolite production as well. The growth rates and productivity of microalgae are dramatically diminished under stress conditions (Adams et al. 2013). Growth media optimization is proposed as an effective method to overcome this problem (Feng et al. 2011, San Pedro et al. 2013); however, previous attempts to gain sufficient levels of productivity have failed because the underlying molecular mechanisms are not well understood (Courchesne et al. 2009, Prochnik et al. 2010).

It has been documented that abiotic stresses increase the transcription and activity of key enzymes involved in various metabolite biosynthesis pathways (Ho et al. 2017). Salt stress has also been reported to enhance the activity of glycerol metabolism enzymes such as glycerol-3-*salina* (Breuer et al. 2013). Recent study have also showed that mRNA level of zeaxanthin epoxidase, an enzyme involved in the accumulation and conversion of zeaxanthin, increases rapidly in response to the increased salinity or light/dark cycle (Kim et al. 2019). Similar results were also obtained for the enzymatic activities of fructose-bisphosphate aldolase (FBPA) involved in starch metabolism (Klok et al. 2013). The high light condition phosphate phosphatase, glycerol 2-dehydrogenase, and dihydroxyacetone kinase in *Dunaliella* leads to energy dissipation and the inhibition of photosynthetic apparatus and generates reactive oxygen species (ROS) (Jin *et al.* 2003). Cells can utilize the higher light energy for growth only after

adjusting the photosynthetic apparatus and cellular processes (Nymark et al. 2009). The expression of most photosynthesis genes decreases under high light conditions, indicating excess light capture (Binte Safie et al. 2018). Fang *et al.* (2012) demonstrated that about 15% of energy metabolism and carbon metabolism and 25% of Tricarboxylic acid (TCA) related genes were up-regulated under high light conditions. It has also been confirmed that the overexpression of genes involved in carbon metabolism (e.g., fructose -1, 6-bisphosphatase) improves photosynthetic performance as well as the total organic carbon and glycerol production of *Dunaliella bardawil* (Fang *et al.* 2012). The metabolome reprogramming effects of nitrate deficiency and/or high light in *D. tertiolecta* cells has also been examined in the literature. It is revealed that such stresses increase saturated fatty acids and decrease polyunsaturated fatty acids (Lee *et al.* 2014).

Despite the broad scientific interest in identifying the responses of *D. tertiolecta* to environmental stresses, the involved biochemical pathways and the underlying genetic mechanisms have still remained unknown (Binte Safie *et al.* 2018). Moreover, contradictory information on the transcriptional regulation of *Dunaliella* under stress conditions has been reported. For example, some studies reported that Glycerol-3-phosphate dehydrogenase *GPDH-c* transcription levels were increased under salinity stress (Kim *et al.* 2010), whereas another study showed that the *GPDH-c* expression level was decreased under salt stress (Alkayal *et al.* 2010). Additionally, various studies have reported different degrees of FBPA expression under abiotic conditions (Cai *et al.* 2013, Kim *et al.* 2010). On the other hand, not all *Dunaliella* species respond in the same manner to abiotic stress conditions (Kim et al. 2010, Pereira and Otero. 2019).

Transcriptome meta-analysis is the most promising strategies to overcome the abovementioned challenges (Zhang *et al.* 2020). It is an efficient tool for the identification of fundamental and

principal genes/pathways and often forms the basis for novel hypotheses (Zhang et al. 2018). For RNA-seq meta-analysis, the p-value combination technique is proposed based on the inverse normalization and Fisher's methods (Rau A. et al. 2014a). This approach has been successfully applied to gain insight into crop responses to environmental stresses (Rest et al. 2016). More recently, salt stress transcriptome responses in different rice genotypes were elucidated by using RNA-seq meta-analysis (Kong et al. 2019).

The microalga *Dunaliella* is frequently considered for industrial applications because of its potential to produce high value compounds such as beta-carotene, its high growth rate, tolerance under extreme physiological conditions, lack of cell wall, and low energy requirements in harvesting (Einali et al. 2017, Oren 2010, Rammuni et al. 2019).

In the present study, RNA-seq meta-analysis of *D. tertiolecta* was performed under high light, nitrogen deficiency, and high salinity stress conditions. Then, protein-protein interaction (PPI) network analysis was constructed to extend the possible functions of identified meta-genes in different stress conditions.

Materials and Methods

Dataset Collection

RNA-seq raw reads were retrieved from the Gene Expression Omnibus (GEO), Sequence Read Archive (SRA), and European Nucleotide Archive (ENA) databases. Raw reads of samples treated under high salinity/NaCl, high light intensity, and nitrogen deficiency conditions were included in the meta-analysis. The details of included datasets are presented in Table 1.

Processing RNA-Seq data

Raw FASTQ files of the above-mentioned datasets were quality controlled using FastQC software version 0.11.5. In this step, reads with quality score below 30 were excluded. Subsequently, the high-quality reads were trimmed using Trimmomatic software version 0.32 (Bolger et al. 2014) with the following parameters: LEADING: 30, TRAILING: 3, SLIDINGWINDOW: 4:20, and MINLEN: 45. We first mapped the trimmed reads to the *Dunaliella salina* genome (available at <https://www.ncbi.nlm.nih.gov/genome/10713>) using TopHat2 software (Trapnell et al. 2009). Since mapping rate was extremely low, de novo assembling with strand-specific mode (`—SS_lib_type RF`) was performed using Trinity software v2.4.0 (Haas et al., 2013). In the next step, the trimmed raw reads were aligned to *de novo* assembled transcripts using Kallisto software v0.44.0 (Bray et al. 2016) with default parameters. The expression of genes was determined as Fragments Per Kilobase Million (FPKM) values per gene, and differential expression of genes between treated and untreated samples was estimated by using the edgeR package (Robinson et al. 2010). A significant differential expression was defined as a fold change $\geq |2|$ and a False Discovery Rate (FDR) corrected p-value ≤ 0.05 (Benjamini and Hochberg 1995). Protein orthology was determined using blastx (cut off E-value of $1.0E-10$) of assembled transcript against the *Chlamydomonas reinhardtii* annotation. The top five hits were extracted using in-house python scripts.

Meta-analysis of differential gene expression

To run the meta-analysis, p-values of differentially expressed genes were combined with two “Fisher’s and inverse” normalization methods complimented in metaRNASeq package (Rau Andrea et al. 2014b). To this end, each stress type dataset was analysed separately using the GLM (Generalized Linear Models). The raw *P*-values for each gene from each of the datasets

were combined using Fisher and Stouffer's methods (inverse normal method). Fisher's method combines the P -values from each experiment with one test statistic defined as:

$$X^2 = -2 \sum_{s=1}^S \ln(p_{gs})$$

where, p_i denotes the raw P -value obtained from gene g in experiment s . S is the number of the combined experiments. Under the null hypothesis, the test statistic χ follows a χ^2 distribution with $2S$ degrees of freedom. This test provides a meta P -value, and the classical procedures for multiple testing correction can be applied to obtain the adjusted P -values to control the false discovery rate. In the Stouffer's method, standard normal deviate, the effect size, and its standard are often directly calculated as:

$$Z = \frac{\sum_{i=1}^k Z_i}{\sqrt{k}}$$

where, Z_i denotes the one-side P -value obtained from gene i in the experiment K under the null hypothesis indicating no effect in each study. The Benjamini-Hochberg method (Benjamini and Hochberg 1995) was used to correct for multiple testing. Genes with the adjusted meta P -value ≤ 0.05 were considered to be statistically significant.

Functional analysis of DEGs (Differentially Expressed Genes) and meta-genes

GO enrichment analysis was performed in three categories of biological processes (BP), molecular functions (MF), and cellular components (CC) using the Algal Functional Annotation tool (Lopez *et al.* 2011). $P \leq 0.05$ was set as the cut-off threshold. To visualize DEGs

(Differentially Expressed Genes) and meta-genes within the enriched pathways, MapMan software (Thimm *et al.* 2004) was used. The PPI analysis was also performed to extend the possible functions of meta-genes in the system level. The STRING database (<https://string-db.org/>) (Szklarczyk *et al.* 2015) was employed for PPI network construction based on the previous knowledge of *Chlamydomonas reinhardtii* at high level of confidence (score ≥ 0.70). Only experimentally validated interactions were included for the network construction. Then the K-means algorithm was used to statistically distinguish functional modules. Finally, the biological functions of each module were enriched using the KEGG database (Kanehisa and Goto 2000). *P*-values were considered to be statistically significant at Benjamini-Hochberg adjusted $P \leq 0.05$.

Results and Discussion

Metabolism overview of differentially expressed genes

The results demonstrated down-regulation of light reactions, photorespiration, tetrapyrrole and lipid-related pathways under salt stress. In comparison to salt stress, nitrogen deficiency mostly induced light reaction and photorespiration metabolisms. Down-regulation of the TCA, starch, and sucrose metabolism-related genes was also revealed under high light stress condition (Fig. 1). It was proposed that the TCA cycle, glycolysis, oxidative pentose phosphate pathway (OPPP), and the carbon concentrating mechanism (CCM) were the main determinants in the carbon flux towards different forms of stored carbon (Tan *et al.* 2016). Conserved up-regulation of phosphoenolpyruvate carboxylase, phosphoglucose isomerase, bisphosphoglycerate mutase, and glucose-6-phosphate-1-dehydrogenase involved in central carbon metabolism was also observed across high salt, high light, and nitrogen stress conditions (Fig. 1 and Supplementary Table S1). Despite down-regulation of the genes directly involved in the Triacylglycerol (TAG)

and Fatty acid (FA) synthesis, up-regulation of glycolytic-related transcripts such as pyruvate kinase and pyruvate dehydrogenase are observed (Fig. 1 and Supplementary table S1).

The findings also demonstrated that nitrogen deficiency down-regulated the ACP, whilst salt stress up-regulated this gene. In response to light stress, the majority of genes involved in lipid metabolism were up-regulated. Ketoacyl ACP synthase and Omega-6 desaturase were two genes with a striking increase in expression (fold change >3) under the same condition. Moreover, in line with the result of this study, it has been reported that genes related to photosynthetic antenna proteins were down-regulated in high light stress condition (Binte Safie et al., 2018). As nitrogen is an important component for the synthesis of chlorophyll and photosystem proteins, the reduction in nitrogen availability could hinder the expression of genes related to photosynthesis (Yao et al., 2017). The results also indicated that nitrogen deficiency up-regulated biotin carboxylase and down-regulated the diacylglycerol kinase. Interestingly, a similar expression pattern of diacylglycerol kinase and biotin carboxylase were observed under both salt and nitrogen stress conditions. Down-regulation of starch degrading enzymes were observed in *D. tertiolecta* cells under high light and nitrogen deficiency conditions (Fig. 1). These results support the idea that *D. tertiolecta* cells tend to accumulate more starch than lipids under high light and nitrogen deficiency conditions (Tan et al. 2016).

Regulation overview of differentially expressed genes

The effects of different stress conditions on regulatory elements are presented in Fig. 2 and Supplementary Figs S4-7. Transcription factors, protein modification, and degradation-related genes, and G-proteins have been frequently identified in the differentially expressed genes. Interestingly, most of the regulatory elements were expressed with stress specific manners. In the

category of transcription factors, Cytokinesis related protein (*MOBI*), Histone deacetylase (*HDA*), and C3H zing finger family were down-regulated, whereas Histone acetyltransferase (*HAM*) and Chromatin remodeling factor (*CHR*) were dramatically up-regulated under salt stress condition (Fig. 2). On the other hand, the up-regulation of the *MOBI* and *HDA* was observed under nitrogen deficiency conditions. Moreover, the *CHR* was moderately down-regulated under high light condition. Peptide methionine sulfoxide reductase (*MSRA3*), involved in post-translational modification, was significantly down-regulated under salt stress, whereas it was not involved in the differentially expressed gene list under high light and nitrogen deficiency conditions (Fig. 2). Regarding the protein degradation-related genes, the down-regulation of 26S proteasome regulatory subunit *RPN7*, ubiquitin-activating enzyme E1 (*UBA1*), and active subunits of chloroplast ClpP complex (*CLPP5*) and the up-regulation of ClpD chaperon (*CLPD*) were noticed under the salt stress condition. On the contrary, the *CLPD* was down-regulated under light and nitrogen stresses.

bZIP (basic leucine zipper) which trigger the cascade of signaling pathways (Zheng et al. 2016), is other key regulatory element of stress responses in *D. tertiolecta*. Results of our study showed that there were similar expression patterns of *bZIP* and *bHLH* (basic helix-loop-helix) (Supplementary Figs. S5-7) with 6-phosphogluconate dehydrogenase (*6PGD*) and glucose-6-phosphate dehydrogenase (*G6PD*), which are involved in oxidative pentose phosphate (OPPP) pathway. The results of our analysis also revealed that *PHD*- and *SET*- transcriptional regulator family is another important regulatory factor in multiple stress responses. As reported, these TFs are regulatory factors in CCM metabolism (in the TCA pathway) and chromatin structure (Bienz 2006). Similar expression patterns of PHD-finger family and some TCA pathway genes (Fig. 2) also highlight this hypothesis.

Meta-analysis of transcriptome data

Although the transcriptome analysis simplifies some complexities of stress response pathways, it poses some questions in this regard. What are the principal pathways mediating the responses to each type of stress? Are there any putative and unique pathways between different stresses? To answer these questions, we performed the statistical meta-analysis along with the system-level survey of meta-genes. Fig. 3 presents the statistical results of the transcriptome meta-analysis of multiple stresses. Under salt stress condition, 477 (317 down-regulated and 144 up-regulated genes) and 144 (101 down-regulated and 43 up-regulated genes) genes were defined as meta-genes based on the Fisher's and reverse normalization methods, respectively (Fig. 3A). Under nitrogen deficiency condition, 50 (39 up-regulated and 11 down-regulated genes) and 15 (10 genes up-regulated and 5 genes down-regulated) genes were identified as meta-genes based on the Fisher's and reverse normalization methods, respectively (Fig. 3B). Moreover, under high light stress condition, 108 (50 up-regulated and 58 down-regulated genes) and 41 (22 up-regulated and 19 down-regulated genes) genes were defined using Fisher's and reverse normalization methods, respectively (Fig. 3C). Supplementary Table S2 lists the meta-genes for each abiotic stress. Oxygen-evolving enhancer protein (*PSBQ*), superoxide dismutase (*FSDI*) and ascorbate peroxidase (ASP) are examples of the key genes under nitrogen deficiency condition. *FATI*, encoding the acyl carrier protein thioesterase, and *PPXI*, encoding the protoporphyrinogen oxidase specifically, were determined as the key genes under salinity stress condition (Supplementary Table S2).

Interestingly, the results indicate that the key meta-analysis genes tend to be located in a hub situation of the stress-specific PPI networks (indicated by a circle, Fig. 4). Surprisingly,

some of these hubs such as serine hydroxyl methyltransferase (*SHMT1*) and *PSBO* are the key regulators of CCM pathways (Fang *et al.* 2017). Not unexpectedly, the overproduction of storage metabolites seems to be interconnected with defense-related pathways. It is also reported that the functions of *SHMT1* influence resistance to different stress, and the mutation of *SHMT1* leads to increased cell damage due to accumulation of H₂O₂ (Moreno *et al.* 2005). *STAI*, which is the key enzyme in starch granulation under stress condition, was defined as a core responsive gene (Supplementary Table S2). It has been proposed that the *STAI* mutants significantly reduce granular starch deposition (Wattebled *et al.* 2003).

The module enrichment of the PPI networks also highlighted the cross talk between photosynthesis, fatty acids, starch, and sucrose metabolism under multiple stress conditions (Fig. 4). The KEGG enrichment of the identified meta-genes under the above-mentioned stress conditions suggested that tetrapyrrole biosynthesis was a key stress responsive pathway in *D. tertiolecta*. Among different genes involved in the tetrapyrrole biosynthetic pathways, glutamyl-tRNA reductase (*GluTR*), uroporphyrinogen decarboxylase (*UROD*), porphobilinogen synthase (*ALAD*), chlorophyllide a oxygenase (*CAO*), Mg²⁺-protoporphyrin IX monomethyl ester cyclase (*CHLE*), glutamyl-tRNA synthetase (*EARS*), and protochlorophyllide reductase (*POR*) were defined as meta-genes under high light condition (Supplementary Table S2); however, protoporphyrin/coproporphyrin ferrochelatase (hemH), geranylgeranyl diphosphate (chlP), hydroxymethylbilane synthase (*HMBS*), and *GluTR* were detected as meta-genes under high salt stress condition. Under nitrogen deficiency condition, coproporphyrinogen III oxidase (*CPOX*), geranylgeranyl diphosphate (chlP), and hydroxymethylbilane synthase (*HMBS*) were identified as meta-genes (Supplementary Table S2). Two main functions of the tetrapyrrole biosynthesis pathway are the chlorophyll (Chl) biosynthesis and plastid to nucleus retrograde signaling

(Kobayashi and Masuda 2016). In contrast with previous findings in plants, our analysis indicated that a majority of meta-genes involved in the tetrapyrrole pathway genes are dominantly up-regulated by high light (Cihlář *et al.* 2016). In line with the results of our analysis, the up-regulation of tetrapyrrole biosynthesis genes under photo-oxidative stress condition has been documented, as reported in *Chlamydomonas reinhardtii* (Nogaj *et al.* 2005). It has been also reported that oxidative stress triggers chloroplast to nucleus communication (Wang *et al.* 2018). Similar expression patterns of the identified meta-genes were noticed under nitrogen deficiency condition. The coordinated expression of the tetrapyrrole intermediated with starch biosynthesis involved meta-genes was observed in this study. Moreover, the co-occurrence and physical interaction of the starch biosynthesis with tetrapyrrole biosynthesis involved meta-genes under high salt, high light, and nitrogen deficiency conditions confirm the hypothesis indicating that plastid to nucleus retrograde signals are essential for the expression of nuclear starch biosynthesis genes (Wang *et al.* 2018).

The meta-genes of the three abiotic stresses were also enriched in photosynthesis processes. The functional module enrichment of PPI highlighted this finding. Photosynthetic-reaction center complexes, light-harvesting complexes, and photosynthetic antennae were the most abundant functional groups of photosynthesis-involved meta-genes. Transcripts of light-harvesting chlorophyll (Chl) a/b-binding proteins B2 (Lhcb2) and *LHCB4* were constitutively decreased under the three abiotic stress conditions (Supplementary Table S2). *LHCs* allows the reaction centers of photosystem (PS) I and II to initiate charge separation and fuel the photosynthetic electron transport chain in the thylakoid membranes (Ishizaki *et al.*, 2012; Perrine *et al.*, 2012). It has been proposed that feedback regulation of Lhcb is essential to reduce the effects of reactive oxygen species on Chls and to accumulate metabolic intermediates under oxidative stress

conditions (Wang *et al.* 2018). Hence, the reduced accumulation of the Mg-Proto, corresponded with decreased expression of *LHCB2* and *LHCB4*, proposed the regulatory effects of tetrapyrrole biosynthesis intermediates on Chl and carotenoid pigment biosynthesis under multiple stress conditions.

Furthermore, the KEGG enrichment analysis of the meta-genes and functional modules of PPI networks indicated that vitamin B6 metabolism, methane metabolism, ribosome biogenesis, folate biosynthesis, and tetracycline biosynthesis responded to the different stresses (Fig. 4 and Supplementary Table S2). Vitamin B6 metabolism and methane metabolism were specifically enriched in light stress meta-genes (Supplementary Table S2). Indeed, the accumulation of the vitamin B6 confers oxidative stress tolerance (Raschke *et al.* 2011). It has been also reported that the vitamin B6 alleviates oxidative damage of Chl and carotenoid pigments intermediates (Havaux *et al.* 2009).

The findings also indicated that ribosome biogenesis specifically enriched under salt stress (Fig. 4). Among protein synthesis apparatus involved genes in *D. tertiolecta*, nine large (*RPL15*, *RPL37*, *RPL27A*, *RPL8*, *RPL23A*, *MRPL9*, *MRPL17*, *MRPL27*, and *RPL7*) and 3 small (*RPSA*, *RPS19*, *RPS27*) subunits of ribosomal proteins were defined as core responsive-genes under salt stress condition. Most interestingly the cytosolic ribosomal proteins, including the *RPL15*, *RPL23A*, *RPL27A*, *RPL37*, *RPL7*, *RPL8*, *RPS19*, and *RPS27* were found to be up-regulated under salt stress condition, while mitochondrial ribosomal proteins such as *MRPL9*, *MRPL17*, *MRPL27*, and *RPSA* were down-regulated under salt stress (Supplementary Table S2). The salt stress-dependent down-regulation of *MRPL9*, *MRPL17*, and *MRPL27* may contribute to mitochondrial functionality through maintaining redox status. It was proposed that the transcriptional control of the ribosome synthesis is an important regulatory point of signal and

regulon generation under salt stress condition (Omidbakhshfard *et al.* 2012). Although, microalgae ribosomal proteins have not yet been studied in detail, preliminary studies showed that a mutation in the RP genes (*RPS18A*, *RPL24B*, *RPS5B*, *RPS13B*, and *RPL27A*) leads to decreased cell division and growth rate (Liang *et al.* 2017).

Conclusion

The present study employed a meta-analysis of RNAseq studies under the abiotic stress conditions, including high light, high salt, and nitrogen deficiency on *D. tertiolecta*. This approach magnifies the efficiencies in the identification of key and core stress response genes/pathways, which might be overlooked by previous studies. The comparative analysis of differential expression analysis and KEGG pathway enrichment of the three stress conditions indicated the existence of shared and unique pathways under high light, high salt, and nitrogen deficient conditions. The shared stress-responsive genes were found to be involved in the tetrapyrrole biosynthesis pathway, starch, and sucrose metabolism. These changes in metabolism allow *D. tertiolecta* to maintain its homeostatic state and redirect the carbon flux to over-accumulate multiple secondary metabolites.

Author contributions:

Conceived and designed the experiments: BP. Analyse the data: BP, MF, writing the paper: BP, SAM, and MAH.

Acknowledgements

The authors are grateful to Dr Anusuya Willis (CSIRO, Australia) for critical reviewing of manuscript. The financial support of the Agricultural Biotechnology Research Institute of Iran (ABRII) is gratefully acknowledged.

Competing interests

The authors declare that they have no competing interests.

References

Adams C, Godfrey V, Wahlen B, Seefeldt L, Bugbee B. 2013. Understanding precision nitrogen stress to optimize the growth and lipid content tradeoff in oleaginous green microalgae. *Bioresour Technol* 131: 188-194.

Alkayal F, Albion RL, Tillett RL, Hathwaik LT, Lemos MS, Cushman JC. 2010. Expressed sequence tag (EST) profiling in hyper saline shocked *Dunaliella salina* reveals high expression of protein synthetic apparatus components. *Plant Sci* 179: 437-449.

Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57: 289-300.

Bienz M. 2006. The PHD finger, a nuclear protein-interaction domain. *Trends Biochem Sci* 31: 35-40.

Binte Safie SR, Ng YK, Yao L, Lee YK. 2018. Growth bottlenecks of microalga *Dunaliella tertiolecta* in response to an up-shift in light intensity. *European Journal of Phycology* 53: 509-519.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)* 30: 2114-2120.

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* 34: 525.

Breuer G, Evers WA, de Vree JH, Kleinegris DM, Martens DE, Wijffels RH, Lamers PP. 2013. Analysis of fatty acid content and composition in microalgae. *J Vis Exp*.

Cai T, Park SY, Li Y. 2013. Nutrient recovery from wastewater streams by microalgae: Status and prospects. *Renewable and Sustainable Energy Reviews* 19: 360-369.

Cihlář J, Füßy Z, Horák A, Oborník M. 2016. Evolution of the Tetrapyrrole Biosynthetic Pathway in Secondary Algae: Conservation, Redundancy and Replacement. *PLOS ONE* 11: e0166338.

Courchesne NM, Parisien A, Wang B, Lan CQ. 2009. Enhancement of lipid production using biochemical, genetic and transcription factor engineering approaches. *J Biotechnol* 141: 31-41.

370 Del Campo JA, Garcia-Gonzalez M, Guerrero MG. 2007. Outdoor cultivation of microalgae for
371 carotenoid production: current state and perspectives. *Appl Microbiol Biotechnol* 74: 1163-1174.

372 Dellomonaco C, Fava F, Gonzalez R. 2010. The path to next generation biofuels: successes and
373 challenges in the era of synthetic biology. *Microbial Cell Factories* 9: 3.

374 Einali A, Mazang-Ghasemi S, Valizadeh J, Noorozifar M. 2017. Metabolic responses and ²-
375 carotene production by the unicellular green alga *Dunaliella salina* exposed to leaf extracts. *Acta*
376 *Botanica Brasilica* 31: 180-190.

377 Fang L, Lin HX, Low CS, Wu MH, Chow Y, Lee YK. 2012. Expression of the *Chlamydomonas*
378 *reinhardtii* sedoheptulose-1,7-bisphosphatase in *Dunaliella bardawil* leads to enhanced
379 photosynthesis and increased glycerol production. *Plant Biotechnol J* 10: 1129-1135.

380 Fang L, Qi S, Xu Z, Wang W, He J, Chen X, Liu J. 2017. De novo transcriptomic profiling of
381 *Dunaliella salina* reveals concordant flows of glycerol metabolic pathways upon reciprocal
382 salinity changes. *Algal Research* 23: 135-149.

383 Feng D, Chen Z, Xue S, Zhang W. 2011. Increased lipid production of the marine oleaginous
384 microalgae *Isochrysis zhangjiangensis* (Chrysophyta) by nitrogen supplement. *Bioresour*
385 *Technol* 102: 6710-6716.

386 Havaux M, Ksas B, Szewczyk A, Rumeau D, Franck F, Caffarri S, Triantaphylides C. 2009.
387 Vitamin B6 deficient plants display increased sensitivity to high light and photo-oxidative stress.
388 *BMC Plant Biol* 9: 130.

389 Ho S-H, Nakanishi A, Kato Y, Yamasaki H, Chang J-S, Misawa N, Hirose Y, Minagawa J,
390 Hasunuma T, Kondo A. 2017. Dynamic metabolic profiling together with transcription analysis
391 reveals salinity-induced starch-to-lipid biosynthesis in alga *Chlamydomonas* sp. JSC4. *Scientific*
392 *reports* 7: 45471-45471.

393 Ishizaki A, and Fleming GR, 2012. Quantum coherence in photosynthetic light harvesting. *Annu.*
394 *Rev. Condens. Matter Phys.*, 3(1), pp.333-361.

395 Jin E, Yokthongwattana K, Polle JEW, Melis A. 2003. Role of the Reversible Xanthophyll Cycle
396 in the Photosystem II Damage and Repair Cycle in *Dunaliella salina*. *Plant Physiology* 132: 352-
397 364.

398 Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids*
399 *Res* 28: 27-30.

400 Kim M, Kang Y, Jin E. 2019. Gene Expression Analysis of Zeaxanthin Epoxidase from the
401 Marine Microalga *Dunaliella tertiolecta* in Response to Light/Dark Cycle and Salinity. *Journal*
402 *of microbiology and biotechnology* 29: 1453-1459.

403 Kim M, Park S, Polle JEW, Jin E. 2010. Gene expression profiling of *Dunaliella* sp. acclimated
404 to different salinities. *Phycological Research* 58: 17-28.

405 Klok AJ, Martens DE, Wijffels RH, Lamers PP. 2013. Simultaneous growth and neutral lipid
406 accumulation in microalgae. *Bioresource Technology* 134: 233-243.

407 Kobayashi K, Masuda T. 2016. Transcriptional Regulation of Tetrapyrrole Biosynthesis in
408 *Arabidopsis thaliana*. *Frontiers in plant science* 7: 1811-1811.

409 Kong W, Zhong H, Gong Z, Fang X, Sun T, Deng X, Li Y. 2019. Meta-Analysis of Salt Stress
410 Transcriptome Responses in Different Rice Genotypes at the Seedling Stage. *Plants (Basel)* 8.

411 Lee S-Y, Kim S-H, Hyun S-H, Suh HW, Hong S-J, Cho B-K, Lee C-G, Lee H, Choi H-K. 2014.
412 Fatty acids and global metabolites profiling of *Dunaliella tertiolecta* by shifting culture
413 conditions to nitrate deficiency and high light at different growth phases. *Process Biochemistry*
414 49: 996-1004.

415 Liang S, Qi Y, Zhao J, Li Y, Wang R, Shao J, Liu X, An L, Yu F. 2017. Mutations in the
416 *Arabidopsis* AtMRS2-11/AtMGT10/VAR5 Gene Cause Leaf Reticulation. *Front Plant Sci* 8:
417 2007.

418 Lopez D, Casero D, Cokus SJ, Merchant SS, Pellegrini M. 2011. Algal Functional Annotation
419 Tool: a web-based analysis suite to functionally interpret large gene lists using integrated
420 annotation and expression data. *BMC Bioinformatics* 12: 282.

421 Moreno JI, Martín R, Castresana C. 2005. *Arabidopsis* SHMT1, a serine
422 hydroxymethyltransferase that functions in the photorespiratory pathway influences resistance to
423 biotic and abiotic stress. *The Plant Journal* 41: 451-463.

424 Nogaj LA, Srivastava A, van Lis R, Beale SI. 2005. Cellular levels of glutamyl-tRNA reductase
425 and glutamate-1-semialdehyde aminotransferase do not control chlorophyll synthesis in
426 *Chlamydomonas reinhardtii*. *Plant Physiol* 139: 389-396.

427 Nymark M, Valle KC, Brembu T, Hancke K, Winge P, Andresen K, Johnsen G, Bones AM.
428 2009. An Integrated Analysis of Molecular Acclimation to High Light in the Marine Diatom
429 *Phaeodactylum tricornutum*. *PLOS ONE* 4: e7743.

430 Omidbakhshfard MA, Omranian N, Ahmadi FS, Nikoloski Z, Mueller-Roeber B. 2012. Effect of
431 salt stress on genes encoding translation-associated proteins in *Arabidopsis thaliana*. *Plant*
432 *signaling & behavior* 7: 1095-1102.

433 Oren A. 2010. Industrial and environmental applications of halophilic microorganisms.
434 *Environmental Technology* 31: 825-834.

435 Perrine Z, Negi S, and Sayre RT. 2012. Optimization of photosynthetic light energy utilization
436 by microalgae. *Algal Research*, 1(2), pp.134-142.

437 Prochnik SE, et al. 2010. Genomic analysis of organismal complexity in the multicellular green
438 alga *Volvox carteri*. *Science* 329: 223-226.

439 Rammuni MN, Ariyadasa TU, Nimarshana PHV, Attalage RA. 2019. Comparative assessment
440 on the extraction of carotenoids from microalgal sources: Astaxanthin from *H. pluvialis* and β -
441 carotene from *Dunaliella salina*. *Food Chemistry* 277: 128-134.

442 Raschke M, Boycheva S, Crevecœur M, Nunes-Nesi A, Witt S, Fernie AR, Amrhein N,
443 Fitzpatrick TB. 2011. Enhanced levels of vitamin B(6) increase aerial organ size and positively
444 affect stress tolerance in Arabidopsis. Plant J 66: 414-432.

445 Rau A, Marot G, Jaffrezic F. 2014a. Differential meta-analysis of RNA-seq data from multiple
446 studies. BMC Bioinformatics 15: 91.

447 Rau A, Marot G, Jaffrézic F. 2014b. Differential meta-analysis of RNA-seq data from multiple
448 studies. BMC Bioinformatics 15: 91.

449 Rest JS, Wilkins O, Yuan W, Purugganan MD, Gurevitch J. 2016. Meta-analysis and meta-
450 regression of transcriptomic responses to water stress in Arabidopsis. The Plant Journal 85: 548-
451 560.

452 Rismani-Yazdi H, Haznedaroglu BZ, Bibby K, and Peccia J. 2011. Transcriptome sequencing
453 and annotation of the microalgae *Dunaliella tertiolecta*: pathway description and gene discovery
454 for production of next-generation biofuels. BMC genomics, 12(1), p.148.

455 Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential
456 expression analysis of digital gene expression data. Bioinformatics 26: 139-140.

457 San Pedro A, González-López CV, Acién FG, Molina-Grima E. 2013. Marine microalgae
458 selection and culture conditions optimization for biodiesel production. Bioresource Technology
459 134: 353-361.

460 Shin H, Hong SJ, Kim H, Yoo C, Lee H, Choi HK, Lee CG, Cho BK. 2015. Elucidation of the
461 growth delimitation of *Dunaliella tertiolecta* under nitrogen stress by integrating transcriptome
462 and peptidome analysis. Bioresource technology, 194: 57-66.

463 Spolaore P, Joannis-Cassan C, Duran E, Isambert A. 2006. Commercial applications of
464 microalgae. Journal of Bioscience and Bioengineering 101: 87-96.

465 Szklarczyk D, et al. 2015. STRING v10: protein-protein interaction networks, integrated over the
466 tree of life. Nucleic acids research 43: D447-D452.

467 Tan KWM, Lin H, Shen H, Lee YK. 2016. Nitrogen-induced metabolic changes and molecular
468 determinants of carbon allocation in *Dunaliella tertiolecta*. Scientific Reports 6: 37235.

469 Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, Kruger P, Selbig J, Muller LA, Rhee SY,
470 Stitt M. 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of
471 metabolic pathways and other biological processes. Plant J 37: 914-939.

472 Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq.
473 Bioinformatics 25: 1105-1111.

474 Wang P, Liang FC, Wittmann D, Siegel A, Shan SO. 2018. Chloroplast SRP43 acts as a
475 chaperone for glutamyl-tRNA reductase, the rate-limiting enzyme in tetrapyrrole biosynthesis.
476 115: E3588-e3596.

Wattebled F, Ral JP, Dauvillee D, Myers AM, James MG, Schlichting R, Giersch C, Ball SG, D'Hulst C. 2003. STA11, a *Chlamydomonas reinhardtii* locus required for normal starch granule biogenesis, encodes disproportionating enzyme. Further evidence for a function of alpha-1,4 glucanotransferases during starch granule biosynthesis in green algae. *Plant Physiol* 132: 137-145.

Yao L, Tan KWM, Tan TW, and Lee YK. 2017. Exploring the transcriptome of non-model oleaginous microalga *Dunaliella tertiolecta* through high-throughput sequencing and high performance computing. *BMC bioinformatics*, 18(1), p.122.

Zhang L, Zhang Z, Zhang X, Yao Y, Wang R, Duan B, Fan S. 2018. Comprehensive meta-analysis and co-expression network analysis identify candidate genes for salt stress response in *Arabidopsis*. *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology*: 1-11.

Zhang X, Yao Y, Li X, Zhang L, Fan S. 2020. Transcriptomic analysis identifies novel genes and pathways for salt stress responses in *Suaeda salsa* leaves. *Scientific Reports* 10: 4236.

Zheng Y, et al. 2016. iTAK: A Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators, and Protein Kinases. *Mol Plant* 9: 1667-1670.

Table 1 . Raw data set ID, run accession and read count (before and after trimming).

Fig. 1: Metabolic overview of differentially expressed genes in Light stress conditions in *D. tertiolecta*. Green and red box represented the down and up-regulation in corresponding stress, respectively. Up-regulation of myo-inositol-phosphatase synthase, phospholipid synthase PGP1b and fatty acid desaturase are indicated. Most of the genes involved in photorespiration were down-regulated in response to light stress.

Fig. 2: Regulation overview of differentially expressed genes under salt stress condition. Most of the genes involved in protein modification and degradation were down regulated in response to salt stress.

508

509 **Fig. 3: Venn diagram for DEGs of Salt (A), Light (B) and Nitrogen (C) stress (A) and meta-**
 510 **analysis results for each corresponding stress. In total 621, 65 and 149 meta-genes were**
 511 **defined for salt, nitrogen deficiency and high light stress conditions, respectively. Most of**
 512 **the identified meta-genes were down regulated at all of the mentioned stresses.**

513 **Fig. 4: Protein- Protein interaction network of meta genes for light stress (A), Salt stress**
 514 **(B) and nitrogen deficiency (C) conditions in *D. tertiolecta* based on the *chlamydomonas***
 515 ***reinhardtii* knowledge. The results indicate that the key meta-analysis genes (signed with red**
 516 **circle) tend to be located in a hub situation of the stress-specific PPI networks**

517

518

519

bioRxiv preprint doi: <https://doi.org/10.1101/2022.04.11.487921>; this version posted April 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

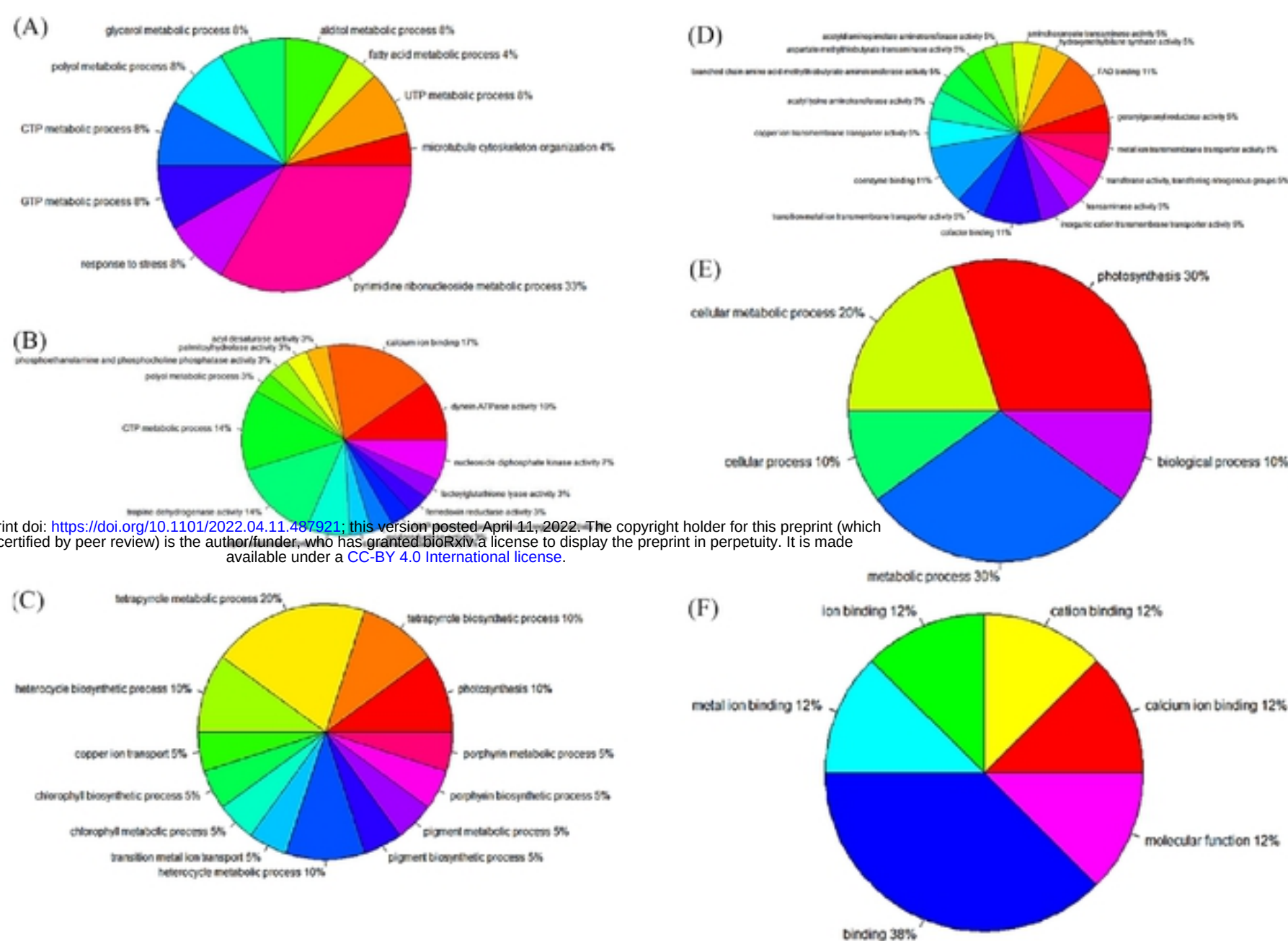


Fig. 1: Gen Ontology of DEGs of three Nitrogen stress related experiments. A, C and E, represented the Biological Process, and B, D and F represented the Molecular functions in the SRP003930, SRP052730 and SRP081094 experiments, respectively

bioRxiv preprint doi: <https://doi.org/10.1101/2022.04.11.487921>; this version posted April 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

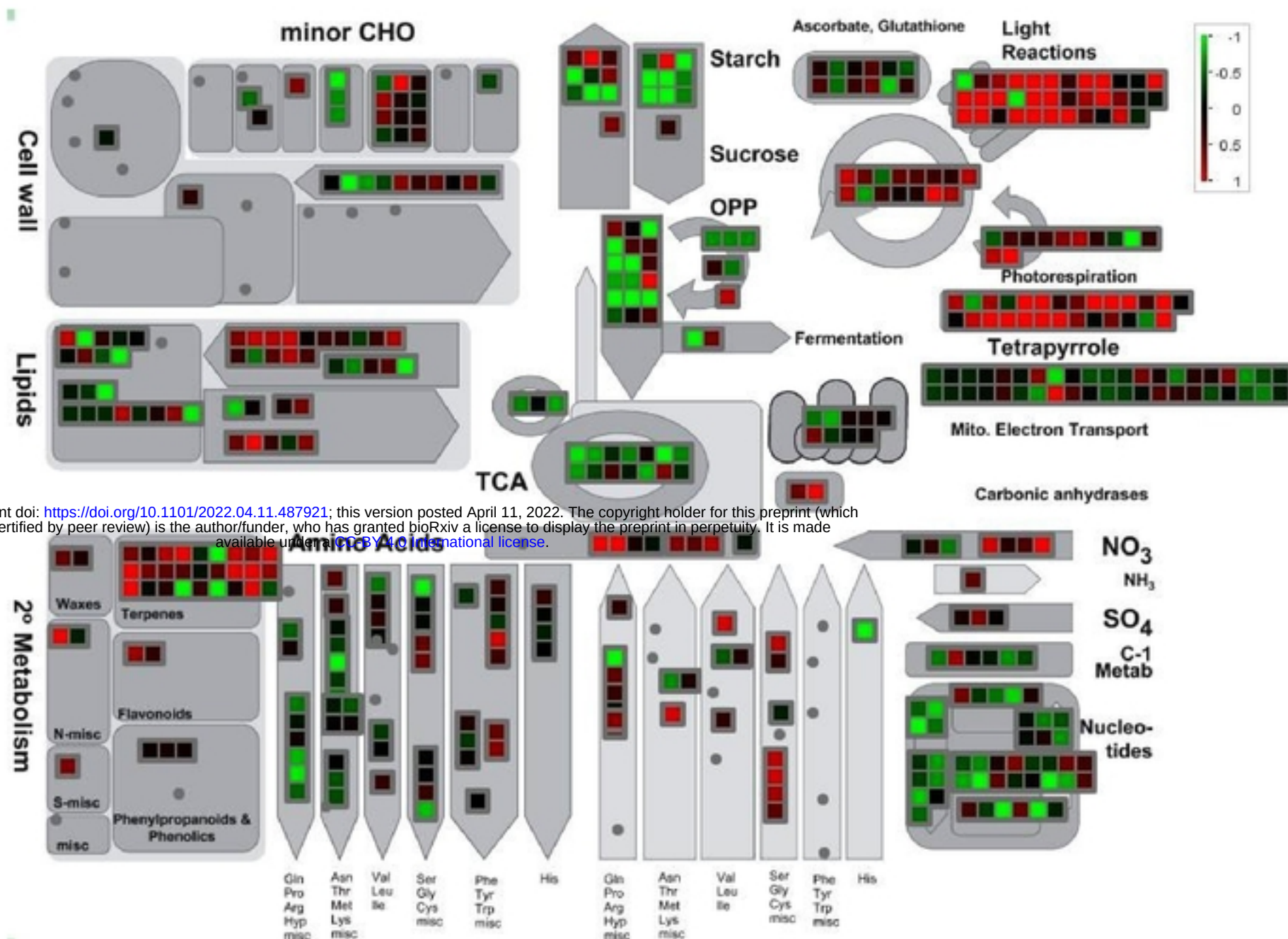
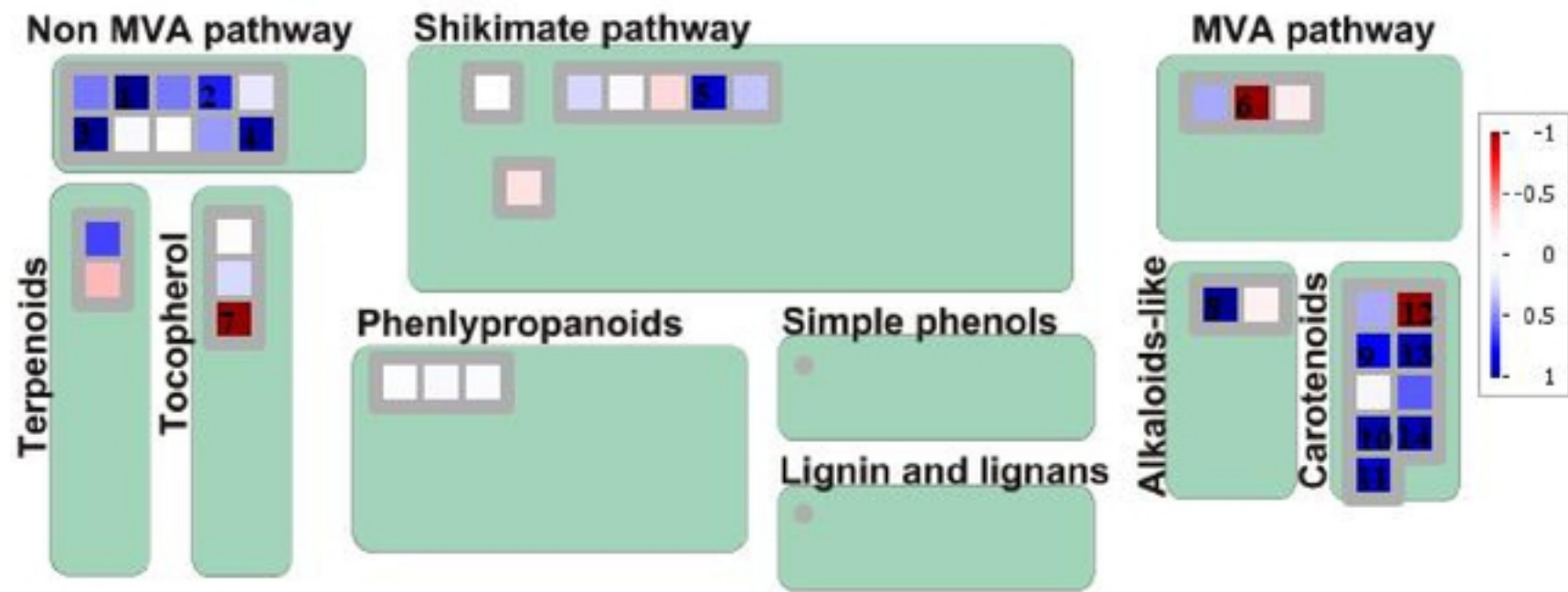


Fig. 2: metabolic overview of differentially expressed genes in Light stress conditions in *D. tertiolecta*. Green and red box represented the down and up-regulation in corresponding stress, respectively.



bioRxiv preprint doi: <https://doi.org/10.1101/2022.04.11.487971>; this version posted April 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

1-PDS: plastidial 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase
2-MCS: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase
3-GGR: geranylgeranyl reductase
4-GGPS: geranylgeranyl diphosphate synthase
5-SHKG1: 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase
6-GPPS: geranyl diphosphate synthase
7-VTE4
8-STRL: putative strictosidine synthase
9-no description
10-LCYE: lycopene epsilon cyclase
11-VDR1: violaxanthin de-epoxidase related
12-CHYB: beta-carotene hydroxylase
13-CRTISO: carotenoid isomerase
14-LCYE: lycopene epsilon cyclase

Fig. 3: Secondary metabolism genes affected by high light stress. The blue and red boxes represented the up and down regulation severity, respectively.

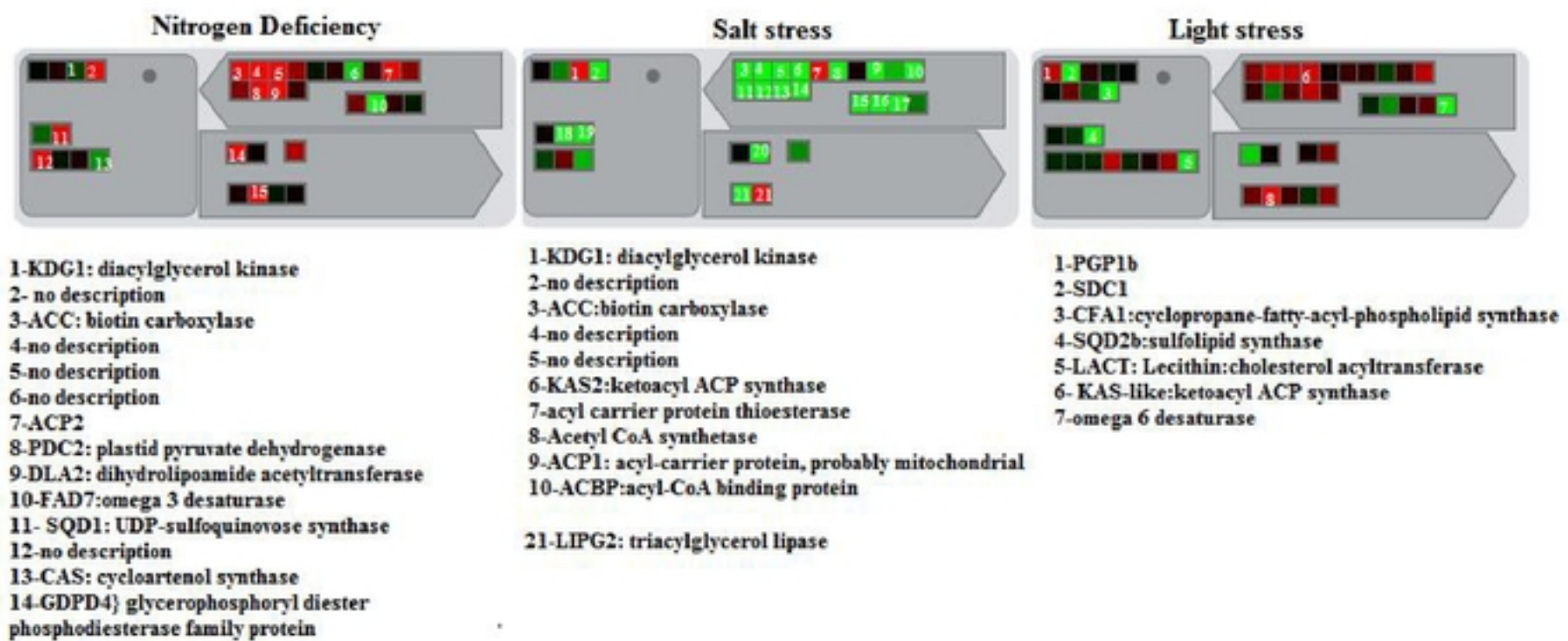


Fig. 4: transcriptional changes of lipid metabolism related genes in responses to nitrogen deficiency, salt and high light stress conditions.

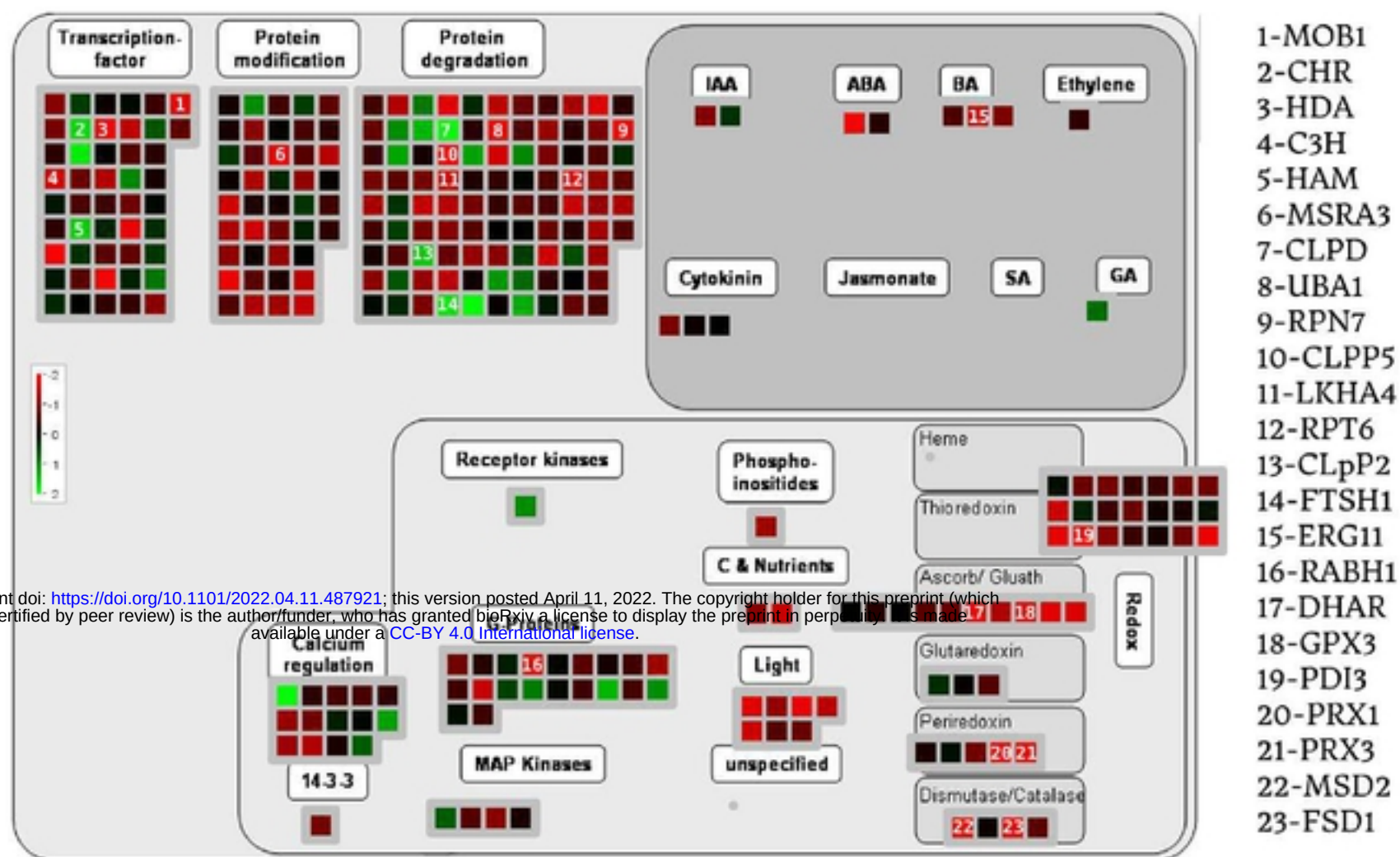
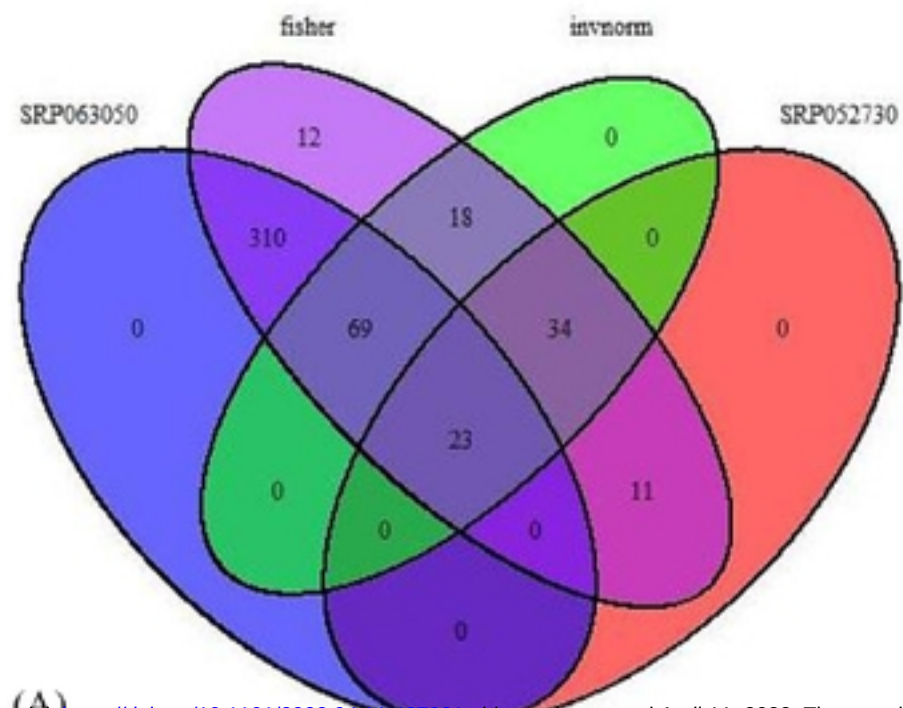


Fig. 5: Regulation overview of differentially expressed genes under salt stress condition.



bioRxiv preprint doi: <https://doi.org/10.1101/2022.04.11.487821>; this version posted April 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](#).

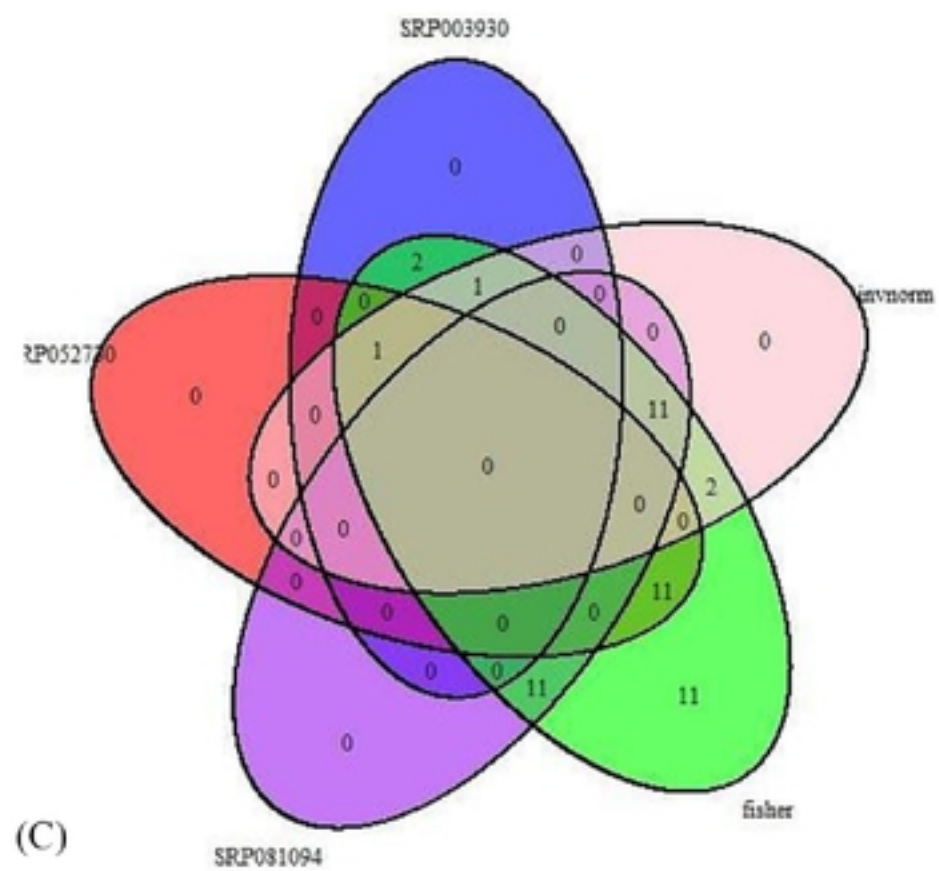
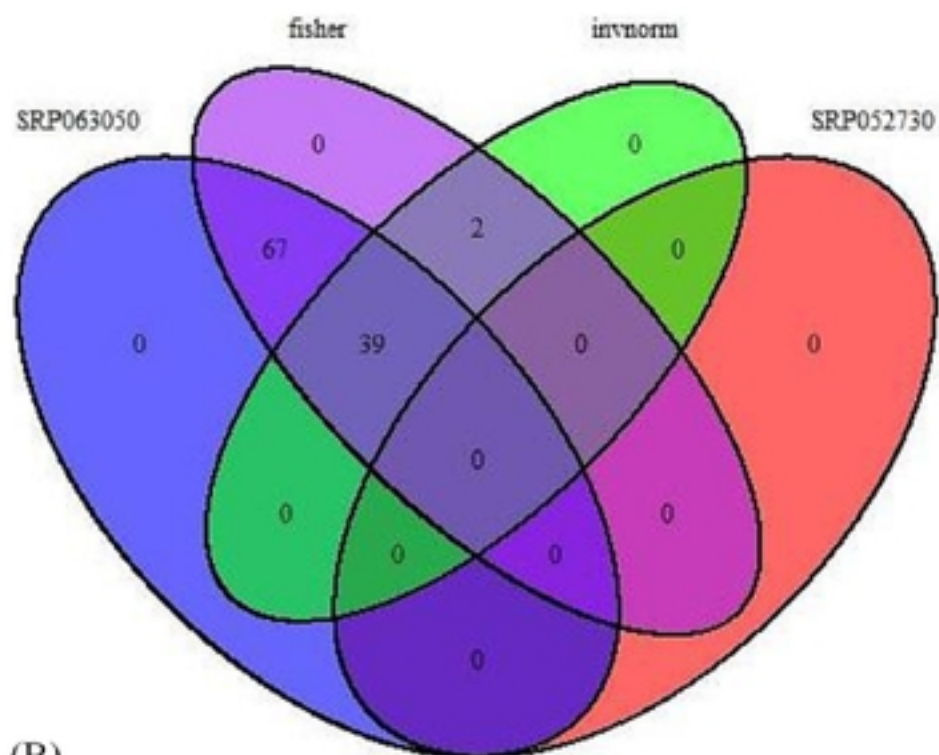


Fig. 6: Venn diagram for DEGs of Salt (A), Light (B) and Nitrogen (C) stress (A) and meta results for each corresponding stress.

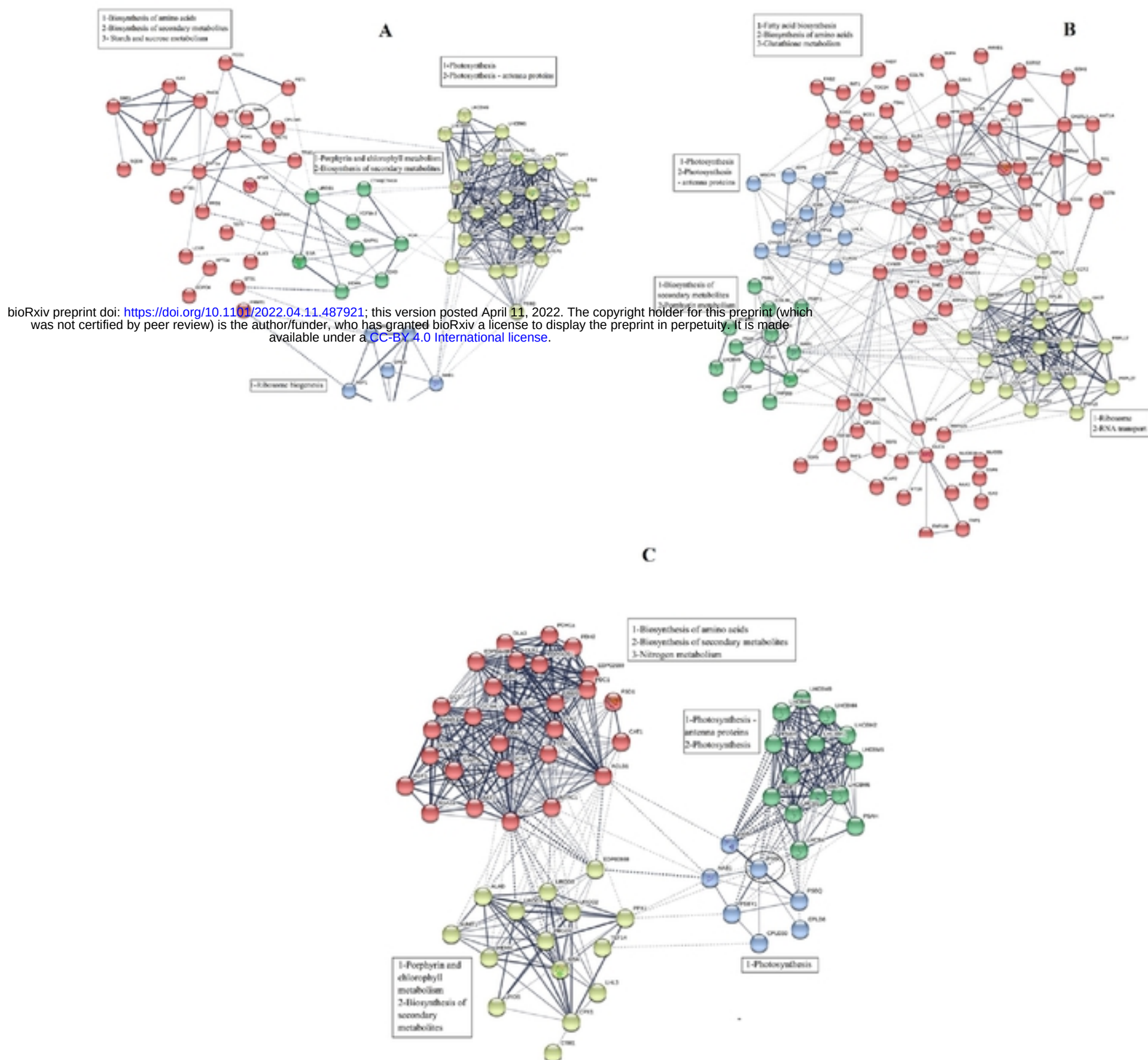


Fig. 7: Protein- Protein interaction network of meta genes for light stress (A), Salt stress (B) and nitrogen deficiency (C) conditions in *D. tertiolecta* based on the *chlamydomonas reinhardtii* knowledge.

Table 1 . Raw data set ID, run accession and read count (before and after trimming).

Data set ID	stress type	Reference	Data base	Run accession	Library layout	Read Count (before trimming)	Read Count (after trimming)
SRP052730	Light and Nitrogen	Shin et al., 2015	SRA	SRR2120818	SINGLE	43,014,956	38,345,362
				SRR2120819	SINGLE	43,342,701	39,921,921
				SRR2120820	SINGLE	42,219,714	40,632,920
				SRR2120821	SINGLE	49,394,621	47,982,792
SRP081094	Nitrogen	Not available	ENA	SRR4011621	PAIRED	3,765,251	2,946,132
				SRR4011622	PAIRED	3,990,351	3,011,629
				SRR4011623	PAIRED	3,501,613	2,720,419
				SRR4011624	PAIRED	3,852,588	2,940,240
				SRR4011625	PAIRED	3,203,153	2,708,372
				SRR4011626	PAIRED	3,303,731	2,699,457
				SRR4011627	PAIRED	3,536,396	2,947,635
				SRR4011628	PAIRED	4,079,043	3,745,201
SRP106620	Salinity	Safie et al., 2017	ENA	SRR5520490	PAIRED	1,370,286	1,056,393
				SRR5520491	PAIRED	1,498,758	1,139,345
				SRR5520492	PAIRED	1,395,338	1,183,222
				SRR5520493	PAIRED	1,421,445	1,103,289
				SRR5520494	PAIRED	944,136	826,323
				SRR5520495	PAIRED	2,044,161	1,838,644
SRP005987	Salinity	Yao et al., 2017	GEO	SRR124253	SINGLE	21,798,484	19,972,634
				SRR124254	SINGLE	19,855,406	18,928,637
SRP003930	Nitrogen and Salinity	Rismani-Yazdi et al., 2011	ENA	SRR070439	SINGLE	305,673	299,374
				SRR070440	SINGLE	333,880	300,341
				SRR070441	SINGLE	183,731	173,922
				SRR070442	SINGLE	229,538	200,231
SRP063050	Light	Safie et al., 2017	ENA	SRR6517590	PAIRED	1,399,689	1,200,384
				SRR6517591	PAIRED	1,522,878	1,392,774
				SRR6517592	PAIRED	1,572,853	1,234,093
				SRR6517593	PAIRED	1,961,343	1,893,674
				SRR6517594	PAIRED	1,839,092	1,735,468
				SRR6517595	PAIRED	3,441,148	2,992,346
				SRR5514896	PAIRED	3,653,262	3,481,926
				SRR5514897	PAIRED	2,410,762	2,001,018
				SRR5514898	PAIRED	2,199,299	1,972,364
				SRR5514899	PAIRED	2,238,539	1,999,233
				SRR5514900	PAIRED	3,017,386	2,873,087
				SRR5514901	PAIRED	3,721,648	3,200,132
				SRR5514902	PAIRED	3,071,374	2,963,692
				SRR5514903	PAIRED	2,768,536	2,517,291
				SRR5514904	PAIRED	3,017,778	2,904,828
				SRR5514905	PAIRED	2,483,867	2,018,310
				SRR5514906	PAIRED	2,184,064	1,927,438

bioRxiv preprint doi: <https://doi.org/10.1101/2022.04.11.487921>; this version posted April 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](#).