International Journal on Agriculture Research and Environmental Sciences

Transcriptome of Laser Micro Dissected Tissue from SAM of Tomato Reveals Genes Responsible for Chloroplast Biogenesis

Research Article

Volume 3 Issue 1- 2022

Author Details

*V K Dalal**

Department of Botany, Dayalbagh Educational Institute, India

*Corresponding author

Vijay Kumar Dalal, Department of Botany, Dayalbagh Educational Institute, Agra, Uttar Pradesh, India

Article History

Received: March 08, 2022 Accepted: April 05, 2022 Published: April 30, 2022

Abstract

Chloroplast biogenesis was studied in the shoot Apical Meristem of tomato through employing mRNA sequencing of Laser micronized samples in different chloroplast developmental stages, by [1]. A stringent criterion was used for the determination of Differentially Expressed Genes. However this strategy masked most of the important genes that are involved in chloroplast biogenesis and thus undermined the importance of the study. Here, the same transcriptome data is being analyzed using a less stringent criteria of 1.5 fold expression between samples and minimum 2 mapping reads per gene to get 1266 DEGs. The DEGS were analyzed through their clustering, chloroplast localizations, functional domains, and their possible involvement in chloroplast development and by other various criteria to provide five sets of candidate genes that have a high probability to be involved in chloroplast biogenesis. Furthermore ~180 genes were cross checked for their expression through HT-qPCR on Fluidigm BIOMARK HD. Moreover the methodology is described in detail.

Keywords: Laser microdissection; Laser capture microdissection; NGS; High throughput quantitative polymerase reaction; Thylakoid development

Abbreviations: SAM: Shoot Apical Meristem; LM: Laser Microdissection; LCM: Laser Capture Microdissection; NGS: Next Gen Sequencing; CEL-Seq: Cell Expression By Linear Amplification and Sequencing; HT-Qpcr-High Throughput Quantitative Polymerase Reaction; DEGS: Differentially Expressed Genes; PSI: Photosystem-I; PSII: Photosystem-II; Cyt B6/F: Cytochrome B6/F Complex; CZ: Central Region/Zone; PZ: Peripheral Zone; LP: Leaf Primordial; NT: No Thylakoids; IT: Intermediate Thylakoids; DT: Developed Thylakoids; L 1/2/3: Layer 1/2/3; FDR: False Discovery Rate; LTP: Lipid Transfer Protein; BA: Bioanalyzer; Arna: Amplified RNA; Cdna-Complementary DNA; DDW: Double Distilled Water; IFC: Integrated Fluidic Circuit; TE: Tris EDTA; ZIP: Leu Zipper; IDT: Integrated DNA Technologies

Introduction

Thylakoids in plant chloroplasts contain photosynthetic complexes such as PSI, PSII, cyt b6/f and ATP synthase. Thylakoid biogenesis is central to the chloroplast formation. In SAM, proplastids present in the CZ of L2/L3 layer also have thylakoids although less developed [1,2]. As these cells divide they develop intermediate level thylakoids in PZ on the sides and form mature thylakoids further away. Therefore samples were isolated with Laser microdissection along this thylakoid gradient i.e. (i) proplastid containing, NT (CZ in published version), (ii) partially developed, IT (PZ in published version) and (iii) highly developed thylakoids, DT (LP in published version), guided by chlorophyll fluorescence based identification of cells. RNA-seq with low amount of RNA was performed by adopting CEL sequencing protocol [3]. In total, ~200 million reads were obtained (ranging from ~5-29 million per sample). Unique mapping for different samples was from 9-38%. The results using stringent criteria {i.e. ≥6 reads in at least one of the tested zones (CZ, PZ or LP), as well as a fold change (increase or decrease) of ≥1.5 in at least one of the three pair-wise comparisons; additional criteria applied was either two-tailed FDR ≤ 0.05, or individual p-value ≤ 0.05, validated by subsequent qPCR analysis} for the determination of DEGs, were published earlier [1]. However this criteria masked most of the genes that expressed lower number of transcripts, nevertheless might be equally important. Therefore a lesser stringent criteria of 1.5 fold expression between samples and minimum 2 mapping reads per gene was applied to get

DEGs. With the criteria, 1266 DEGs were obtained among the three comparisons i.e. NT to IT and NT to DT and IT to DT, including a number of important genes as can be seen in Table 1-5. The DEGs has been submitted to public database as described in and can be retrieved using the cutoff criteria here [1].

DEGs were grouped into 9 clusters (Figure 1) using the same method as given in [1]. These genes and their detailed analysis e.g. trans-membrane domain, PFAM domain, chloroplast localization and interaction of Arabidopsis homologues in STRING database, is provided in Table 2 & 3. The photosynthesis genes were present mostly in clusters 4-7 (Figure 1c) that showed increased expression along the thylakoid gradient (Figure 1b), however some genes were also sorted in the clusters that down-regulated in DT after initial up-regulation in IT. This cluster had genes that are most probably involved in initial thylakoid proliferation that subsequently subsided in DT. It was found that lipid modification of proteins such as Myrsitoylation, Palmitoylation, Prenylation/ Farnesylation are up-regulated during thylakoid development that could help impart hydrophobicity on thylakoid proteins, essential for lipid membrane insertion. This study projects many lipid binding, LTP, lipid binding START domain containing, membrane insertion domain containing proteins such as C2 domains, and some transcription factors such as those with START domain HD-ZIPs (Table 1-4), that help in protein assembly in membranes, as candidate genes for thylakoid formation. Table 1 shows TFs that are putatively involved in chloroplast biogenesis. Many other DEGs that have unknown function or have (un)known domains and putative chloroplast localization are enumerated in Table 2 & 3 and are highly expected be involved in thylakoid biogenesis. Table 4 represents putative "vesicle formation" genes involved in chloroplast biogenesis. Some genes (Table 5) showed very high transcript abundance already in NT zone and gradually increased further. It may be possible that these proteins are the starting point for assembly of functional photosynthetic complexes through addition of subsequent proteins. The authenticity of expression patterns of selected genes was checked with HT-qPCR (Fluidigm BIOMARK HD) (around 95 assays (genes), in triplicate (having 9216 wells per chip) on two chips).

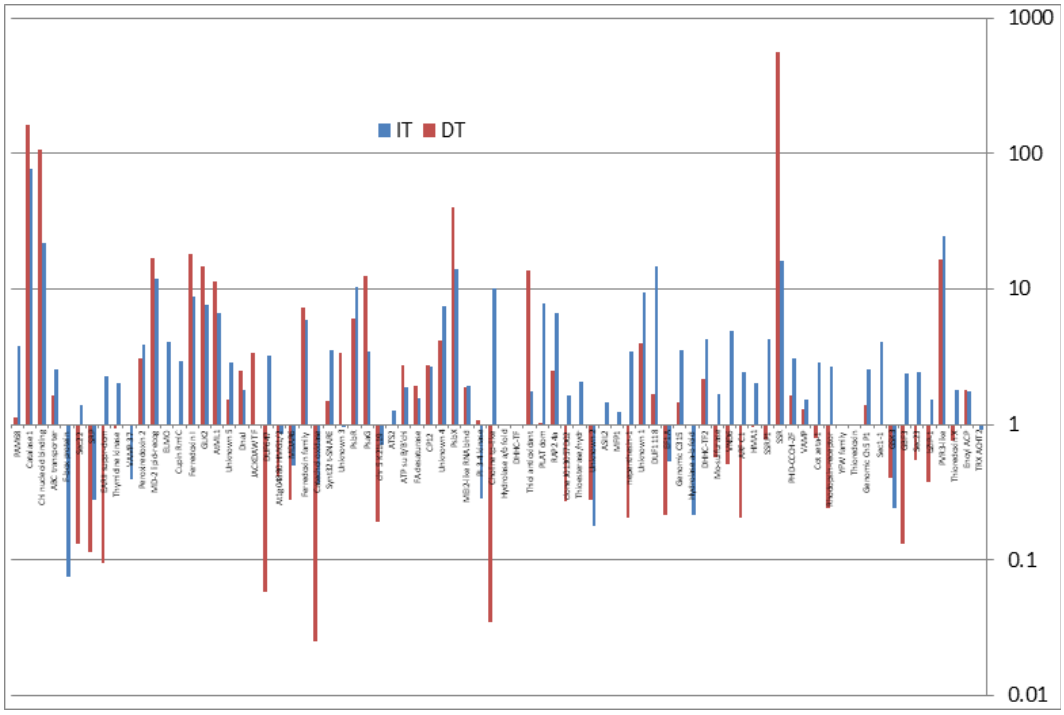

Figure 1: HT-qRT-PCR results; Fold change of selected genes, involved in thylakoid biogenesis, in IT and NT samples as compared to NT (least developed) sample.

High Throughput (HT) Real Time PCR

The RNA was isolated as mentioned earlier [1]. cDNA synthesis and RNA amplification was performed with Ambion/Life Technology (Invitrogen) Message Amp™ II aRNA Amplification Kit (AM1751) as described by the manufacturer using 1/10 of the reaction volume as per [4], however for two rounds, instead of one. The aRNA obtained from the first amplification was concentrated to 2µl for each sample and used for second amplification. The eluted volume of aRNA in each sample was 18µl. The aRNA was checked on BA (Agilent). 20ng of aRNA in 2.5µl was used for cDNA synthesis for each sample as per kit instructions.

qPCR Assays Design

Primer pairs were designed for 95 transcripts. The primers were designed with IDT (Syntezza, Jerusalem) in-house software Primer Quest, towards the 3' end, so that they could also work with degraded RNA. The Tm for all the primers was chosen around 60°C, taking into account the default PCR mix concentration of 0.8 mM dNTP and 3mM Mg2+. Primers were not specifically designed across introns, since the two rounds of amplifications were earlier performed with polyA mRNA tail priming, which would exclude the genomic DNA. The primers were delivered by IDT in 96-wells plate. Primers were dissolved at a concentration of 100µM in nuclease free water (Sigma). For each assay, a Primer pair mix was prepared containing 50µM Forward+50µM Reverse primers mixing 5µL of each.

Two assays were run initially with the regular PCR to check if the cDNA was formed properly. HT-qPCR was performed at The Hebrew University of Jerusalem, Genome Centre.

Dilution Curve for Assay's Efficiency

It is important that the qPCR assays work good under different concentrations of transcript in the sample: efficiency close to 100% ensure this. Each primer pair may perform differently in this regard, therefore a dilution curve is made for each primer pair to check their efficiency. For checking the efficiency of primers, standard curves were made with dilution series of 1:5, 1:25, 1:125, 1:625, 1:3125 dilutions

with NT (one replicate) sample, for all assays. These diluted samples were also subjected to preamplification along with other samples. Standard curves were made with Fluidigm Real-Time PCR Analysis software. Efficiency of each assay was calculated based on slope of standard curve as Efficiency= (10^-1/slope -1)*100 (Figure 2) [5]. Assessment of standard curves (constructed from Ct values plotted against starting quantity of DNA) indicated linearity over at least 5 logs in all the tests. Using a constant Ct threshold of 0.1, all the assays showed standard curve slopes within the acceptable range of -3.6 to -3.1, corresponding to amplification efficiencies of 90% to 110%.

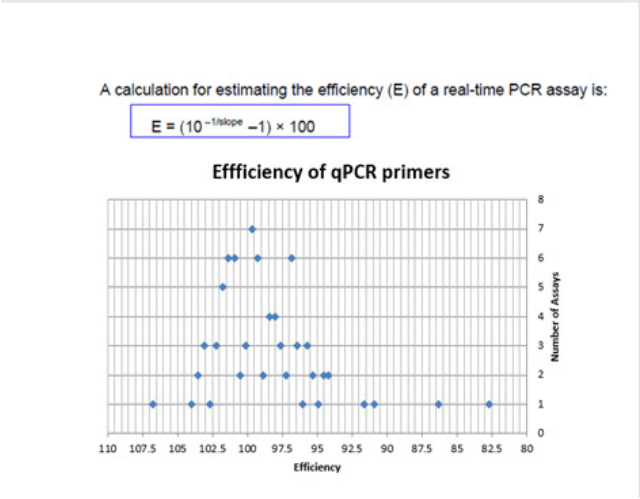

Figure 2: Efficiency of all assays. Nine assays (efficiency not between 104% and 94%) were removed from efficiency analysis due to poor performance. Out of rest 86, 28 assays performed very well with efficiencies ranging from 98-100.5%. 92% of the assays showed efficiency distribution between 104% and 94%.

Preamplification

For first preamplification, Primer Mix (500 nM each primer) was prepared by taking 2µL from each pair 95 genes (total 190µL) and adding 10µL of DDW to make it 200µL. For second preamplification with 83 genes, Primer Mix (500nM each primer) was prepared by taking 2µL from each pair of 83 genes (total 166µL) and adding 34µL of DDW to make 200µL. Preamp mix was prepared with 0.05µL Primer mix (10x), 2.5µL of 2X TaqMan PreAmp Master Mix (AB) and 0.7µL of DDW per assay. 1.25µL of cDNA was added to the above mentioned 3.75µL preamp mix. Preamplification was carried out by subjecting the reaction at 950C for 10 min and for 14 cycles at 950C for 15 sec and 600C for 4 min. Unused primers were cleaned by treating with 0.4ul Exonuclease I (20U/µL) per reaction at 37 0C for 30 min and inactivation at 80°C for 15 min. Cleaned up product was diluted 1:5 with TE buffer (0.5 ml of 1M Tris pH 8.0, 0.1 ml 0.5M EDTA, 49.6ml DDW)

qPCR

Preamplifier cDNA samples were analyzed by qPCR using 96.96 Dynamic Array™ IFCs and the BioMark™ HD System from Fluidigm. Processing of the IFCs and operation of the instruments were performed according to the manufacturer's instructions. In order to prepare samples for loading into the IFC, 2.5µL 2X Sso Fast EvaGreen Supermix with Low ROX (BioRad 172-5212), 0.25µL 20X DNA Binding Dye Sample Loading Reagent (Fluidigm 100-3738) were added. 2.75µL of this mix was dispensed to each well of a 96-well assay plate (Axygen Scientific, P-96-450-V-C). 2.25µL of Exo I treated diluted preamplified cDNA sample was added to each well and the plate was briefly vortexed and centrifuged. Following priming of the IFC in the IFC Controller HX, 5µL of the cDNA sample + reagent mix were dispensed to each Sample Inlet of the 96.96 IFC. Assay mix was prepared by adding 2.5µL Assay loading reagent, 2µL of 1X low TE buffer and 0.5µL of 50 µM (F+R) primer mix. For gene assays, 5µL of each 95 assay-mix were dispensed to each Detector Inlet of the 96.96 IFC. After loading the assays and samples into the IFC in the IFC Controller HX, the IFC was transferred to the BioMark HD and PCR was performed using the thermal protocol GE 96 x 96 Fast PCR + Melt 40 cycles. This protocol consists of a Thermal Mix of 70˚C, 40 min; 60˚C, 30 s, Hot Start at 95˚C, 1 min, PCR Cycle of 40 cycles of (96˚C, 5 s; 600C, 20 s), and Melting using a ramp from 60˚C to 95˚C at 1˚C/3 s.

Data Processing and Analysis

Data was analyzed using Fluidigm Real-Time PCR Analysis software using the Linear (Derivative) Baseline Correction Method and the Auto (Global) Ct Threshold Method. The Ct values determined were exported to Excel for further processing. Average Ct values (Geometric mean in linear mode) of four reference genes was used to normalize the expression values (Ct) of each assay. IT and DT samples were compared with NT for fold change calculation with delta delta Ct (ddCt) method.

Most of the assays showed efficiencies near to 99% (Figure 2). A single peak on the dissociation curve was obtained [6]. Many known chloroplastic genes and candidate genes showed expected expression i.e. increased expression from NT to IT and further in DT (Figure 3). When these results were correlated with CEL-Seq: highly expressed genes showed higher correlation with CEL-seq expression pattern, however a fair level of correlation was also observed among medium or lowly expressing genes.

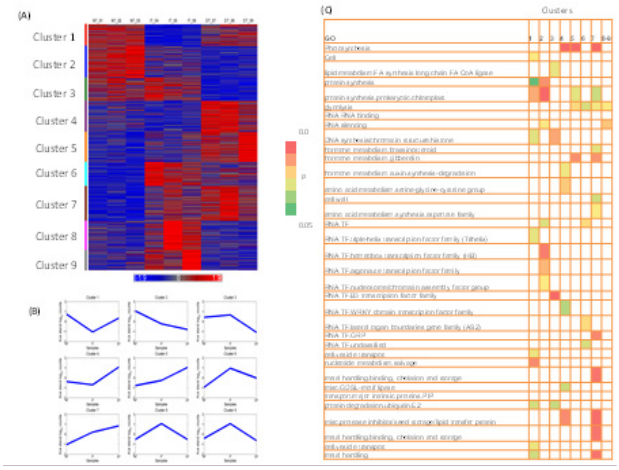

Figure 3: A) Clustering of 1266 DEGs Clusters 4-7 had most of genes for chloroplast localized proteins, B) Representative patterns of each cluster, and C) Heatmap showing some of the enriched GO terms in all nine clusters. Gene Ontology Enrichment from each individual cluster was performed with Ontologizer.

Description of DEGs

Table 6 provides the name of genes that were checked through HT-qPCR and are graphically represented in Figure 2. Functions of some of the important genes among them are discussed here.

a. Glycerolipid and Fatty Acid synthesis during chloroplast biogenesis

Fatty Acids (FAs) have an important role in TB. MGDG (Mono-Galactosyldiacylglycerol) and DGDG (Di-Galactosyldiacylglycerol) galactolipids are the main components of thylakoid membranes [7]. Galactolipids are derived from DAG (Diacylglycerol) which in turn is formed from PA (phosphatidic acid) obtained through "prokaryotic" or "eukaryotic" pathways and PC (phospatidyl chloine). FAs are synthesized in the chloroplast and further galactolipids synthesis can

take place in chloroplasts or FAs can be exported to ER, where PC is synthesized and transported back to plastids. Many important proteins for FA and galactolipid synthesis and transport have been reported e.g. MOD1 (MOSAIC DEATH 1) which encodes for an enoyl-ACP reductase, a component of the FA synthase complex [8]; ATS2, a plastidic Lysophosphatidic Acid Acyltransferase (LPAAT), critical for chloroplasts phosphatidic acid biosynthesis, mutants of which are embryo lethal [9]; CAC1, a component of GPDH; plant Lipid Transfer Domain Protein (LTP); Chlroplastic FAB2/SSI2, an acyl-acyl-carrier-protein and C2 domain membrane targeting proteins. C2 domains show a wide range of lipid selectivity including phosphatidyl serine and phosphatidyl choline. This domain is hypothesized to be involved in Ca2+ dependent phospholipid binding and membrane targeting to produce subcellular protein-gradients (INTERPRO database). Their role in plants are largely undetermined but are believed to play role in defense and stress (INTERPRO).

Two LTPs (Solyc03g121900, PVR3-like/DIR3 and Solyc06g065970, Cortical cell delineating protein) were upregulated. Putative defective in induced resistance 1 (DIR1), a PVR3 domain containing LTP, which is transported through phloem and is involved in Systemic Acquired Resistance (SAR), [10]. LTPs are cationic proteins of approximately 70 and 90 aa, ranging from 6.5-10.5 kD with eight Cys residues. They bind a wide range of lipids including fatty acids (C10–C14), phospholipids, prostaglandin B2, lyso-derivatives, and acyl-coenzyme A, so they are also called non-specific LTPs (ns LTPs). LTPs have a cavity to bind fatty acid and they were earlier believed to transport proteins between membranes [11]. However, now it is doubted due to their presence in mostly extracellular environment. Increasing lines of evidence suggest that nsLTPs do not mediate a simple vectorial lipid transport from one membrane to another, instead, they facilitate lipid transport between membranes in response to their membrane environment. nsLTPs can therefore locally modulate lipid composition and/or fluidity of membranes, and consequently regulate various cellular processes, including vesicular trafficking, signal transduction, and lipid transfer and metabolism [12].

b. Chloroplast antioxidant machinery is induced during thylakoid development

Arabidopsis RAP 2.4a is known to induce chloroplastic antioxidant enzymes like 2CP (2 Cys peroxiredoxin), stromal ascorbate peroxidase and SOD (Csd2) [13] in chloroplastic redox state dependent manner. Arabidopsis RCD1 regulates RAP 2.4a activity, however is independent from chloroplastic redox state [14]. The transcripts for chloroplastic localized antioxidant enzymes such as CAT, CAT2, Arabidopsis MDAR1 orthologue Monodehydroascorbate reductase (NADH)-like protein and Thioredoxins are also up-regulated during thylakoid development. Interestingly the cat2 mutants in Arabidopsis show severe phenotype as compared to cat1 and cat3 which are the mutants for other two Arabidopsis catalases [15]. During chloroplast biogenesis, Chl needs to be formed and assembled in PS complexes. The Chl biosynthesis intermediates and unassembled Chl may produce ROS, which needs to be de-toxified, therefore an up-regulated chloroplastic antioxidant machinery may be responsible for ROS neutralization as observed by [16].

c. Chloroplast metal transporters play important role in thylakoid biogenesis

Metal transport proteins e.g. PAA6/HMA1/HMA6 (heavy-metal-associated-domain containing) and MATE-efflux family proteins are important transporters involved in Mg, Zn2+, Cu2+, Cd2+ and Fe2+ ion transport. Many photosynthetic proteins contain 4Fe-4S (Iron-Sulphur) clusters. Cu2+ and Zn2+ are present in SOD (H2O2 detoxifying antioxidant enzymes). Cu2+ is also present in plastocyanin. Besides, the most prominent pigment of thylakoids, Chlorophyll also contain Mg2+ in its tetrapyrrole ring. Up-regulation of these transcripts corroborates with huge requirement of metal import and homeostasis in the chloroplast during thylakoid biosynthesis [1]. HMA1 and HMA6 transport Cu2+ through distinct pathways in addition to known pathways [17]. Their mutants show chloroplastic defects under Cu2+ deficiency [18]. Different transporters may be substrate specific or may function redundantly [19].

Ubiquitin mediated protein degradation marks the importance of Ubiquitin Proteasome System (UPS) in thylakoid formation or the ensuing developmental processes. RING type E3 ligase SP1 (suppressor of ppi), a UPS component present in chloroplast outer membrane, was shown to be involved in thylakoid biogenesis and plastid inter-conversion by regulating the TOC receptor complex [20]. DNAJ domain proteins of chloroplasts usually act as co-chaperones to impart specificity to chaperones [21]. A recent study found involvement of DNAJ domain protein in PSI assembly [22].

d. Potential transcription factors regulating TB

Transcription Factors (TFs) bind to specific cis-elements in the promoters of genes [23] and are important components regulating most processes in living systems. It is expected that thylakoid biogenesis will also be regulated through a network of various TFs. There are only a few TFs that are known to directly regulate chloroplast biogenesis e.g. GLK1 and GLK2. Only a few nuclear encoded TFs have been reported in chloroplasts (e.g. PHD). There are bioinformatics studies that point out at many probable chloroplast localized TFs [24].A few TF are weakly associated with chloroplast biogenesis e.g. ABI3 (a B3 class TF) mutants show some thylakoid developmental even in dark or when plants are transferred to dark from light [25]. abi3 plants also fail in de-greening of seeds during maturation. PIF3 relays the PHYTOCHROME signal to nucleus for regulation of photomorphotgenesis and chloroplast development [26]. PHD, a separate superfamily member, present in the envelope membrane has been found to have role in relaying chloroplast signals to nucleus [27]. S1 domain containing TF, STF1 was found to be localized to chloroplasts and regulate the chloroplast biogenesis by modifying chloroplastic photosynthesis genes expression in Nicotiana [28]. Auxin/Cytokinin regulated TFs CGA1 and GNC control chlroplast biogenesis [29]. Arabidopsis RAP 2.4a is known to regulate the chloroplastic antioxidant gene expression according to the redox status of chloroplasts [30,31].

TFs HB-1, ASIL2, VND6 are putative repressor of photosynthesis genes. TF HB-1, that bind to CAATTATTG sequence, heterodimerizes with its paralogue HB-2 [32]. Ectopic expression of full length HB-1 and the fusion of its DNA binding domains with activation domains of VP1 and GAL4 caused the yellowing of leaf portions [33]. Thylakoid-associated, nucleoid-binding protein; MAR (Matrix attachment Region)-binding Filament-like Protein 1 (MFP1) was originally identified in tomato where it was found to bind matrix attachment region of DNA [34]. In Arabidopsis, it accumulates to high levels in shoots of light-grown seedling and to intermediate levels in dark-grown shoot which correlates with thylakoid levels in plants [35]. Chloroplast nucleoid protein/TF, MFP1 is uniquely long coiled-coil protein and is an unusual target of TAT pathway. MFP1 has DNA-binding activity in Arabidopsis chloroplasts and binds to several regions of the chloroplast DNA with equal affinity. It could be acting as a link between thylakoid biogenesis and maintenance of chloroplast nucleoprotein in chloroplast nucleoids. Recently PEP-Related Development Arrested 1 (PRDA1), a chloroplastic nucleoid protein was found to regulate early chloroplast development [36] by reducing the PEP-dependent chloroplast gene expression.

Some other TFs such as ASIL1 and ASIL2, and HDA6 are downstream targets of miRNA and suppress early embryo maturation including chloroplast development [37,38]. The Arabidopsis mutants of asil2 also showed early chlorophyll formation and chloroplast development in embryos [39] indicating its suppressive role on photosynthesis gene expression. ASIL2 interacts with two histone deactylases (HDA3 and HDA05; STRING-DB) indicating its role in chromatin modification and Arabidopsis Meristem Receptor like kinase (MRLK) [40] showing its important role in SAM. Arabidopsis VASCULAR-RELATED NAC-DOMAIN 6 (VND6), a CCHC type Zn finger TF,

also showed chloroplast defects when ectopically expressed with an activation domain [41]. Over-expression of RAP 2.4a, ERF type TF of APETALA2/ERBP superfamily, in Arabidopsis showed high expression of a reporter gene that was fused to the promoter of chloroplastic antioxidant enzyme 2-Cys Peroxiredoxin A (2CPA) when excitation potential builds up in chloroplasts [42]. It also activates expression of many genes of chloroplastic antioxidant system e.g. stromal Ascorbate Peroxidase (sAPx) and Cu/Zn Superoxide Dismutase (Csd2). Mutants of Arabidopsis radical induced cell death 1(RCD1)/ redox imbalanced1 (rimb1) (At1g32230) protein showed up-regulation of chloroplastic antioxidant genes MDAR6 (monodehydro ascorbate reductase6), 2CPA, sAPx, Csd2, peroxiredoxin IIE (PrxIIE) and γ -glutamyl cystein synthase (γ ECS). In Arabidopsis, RCD1 helps RAP 2.4a dependent redox-regulation of chloroplastic genes [43]. Arabidopsis zinc knuckle (CCHC-type) TF (AT3G43590), interacts with RCD1 (Arabidopsis Interactome Mapping Consortium). Complexes of these components might regulate redox mediated chroloplastic gene expression. Most of the developmental processes and plant responses to stimuli are regulated by a recombinatorial network of a number of TFs. In addition the chlorophyll biosynthesis intermediates are phototoxic and partially assembled PS complexes are considered to be phototoxic, therefore it needs to be tightly regulated [44]. It is expected that TFs may also take a comprehensive part in this process and there may be many unidentified TF candidates regulating this.

GBF3 is also an important protein in TB. Homologues of GBF3 in tobacco, CPRF-1 and TAF-1 were proposed to be involved in photosynthesis regulation [45]; GBF3 paralogue, GBF1 interacts with chloroplast biogenesis TFs GLK2 and GLK1 [46]. which supports the role of GBF3 in chloroplast biogenesis in combination with other TFs. bZIP53, a G-box binding TF, interacts with a protein involved in photomorphogenesis. Chloroplast Stem-Loop Binding Protein 41kDa (CSP41b) was shown to control transcription of PEP regulated genes [47]; two chloroplastic nucleoid binding proteins and a RNA-binding protein 45B (RBP45B). RBP45B show interaction with PTF-1 (Plastid Transcription Factor-1) (STRING-DB), which is chloroplast localized and regulates the expression of psaD , PSI-core protein, thus is involved in thylakoid biogenesis.

WRKY TFs are other potent candidates regulating chloroplast proteins. Affymetrix microarray transcriptome analysis of high-light stressed transgenic lines of AtWRKY40- and AtWRKY63- inhibited functions revealed that these factors were involved in regulating many stress-responsive mitochondrial and chloroplast proteins' genes e.g. AOX1a but however had only minor effect on constitutively expressed genes. Aditionally At WRKY13, AtWRKY40 and AtWRKY57 TFs bound to the W-boxes elements in stress induced nuclear encoded photosynthetic genes e.g. LHCB2.4 and HEMA1 [1].

e. Transcription Factors that are also induced in plant embryo

Some TFs e.g. ABI3, bZIP53 and Trihelix TF ASIL2 regulate embryo maturation program and are up-regulated from heart shaped embryo till seed formation [4]. The chloroplast biogenesis also takes place in embryo at heart-shaped stage (although the protoderm/epidermis shows Chl accumulation from the globular stage of embryo). These transcripts are also upregulated in tomato SAM TB gradient, therefore these could be involved in chloroplast biogenesis. bZIP53 interacts with proteins involved in photomorphogenesis (STRING-DB), which hints at the involvement of bZIP53 in chloroplast biogenesis.

DNAJ Domain Containing

DNAJ domain (also known as HSP40) proteins of chloroplasts usually act as co-chaperones to impart specificity to chaperones [30]. Down-regulated (Solyc11g006170.1) gene contains DNAJ domain. DNAJ domain containing EMF1-Interacting Protein 9 (EIP9), Arabidopsis orthologue AT5G64360 of tomato gene (Solyc11g006170) is involved in suppressing flowering. EIP1, EIP6 or EIP9 forms heterodimers with EMF1 to regulate flowering time [16]. eip9 mutants show early flowering [16].

a. Miscellaneous categories of genes involved in TB

Mutants of SPL11/PUB13 which is an E3 ligase [12], PSI (PSK-SIMULATOR) that is co-expressed with PSK receptors [41], Cat1 [29], ClpB-p (HSP 100 family member) [22] show chloroplast development defects. Proteins are modified with lipid attachment (lipidation) which include prenylation, N-myristoylation and acylation that is helpful in imparting hydophobicity to proteins and in membrane insertion of proteins. Similarly, process of lipidation might help in thylakoid protein insertion. Prenylation involves addition of 15-carbon farnesyl or addition of one or two, 20-carbon long geranylgeranyl chain to one or two cys residues of CAAX domain. Proteins with CAAX C-terminal sequence are prenylated and are cleaved between C and AAX by CAAX proteases, followed by methylation of Cys. This processing is essential for proper functioning and localization of such proteins. AtSTE24 is a CAAX protease (Zn-metalloprotease) with broad specificity [6]. G-coupled receptors are integrated into membranes after isopentylation/prenylation or palmitoylation [9]. Arabidopsis contain 11 Rho-related GTPases (ROPs), 10 of which contain CAAX motifs. Plastidial GK, RNR and NUS1 are involved in formation and maintenance of chloroplast genetic system and their mutants show virescent phenotypes [21].

Further description, quality analysis for the method and data is presented in the supplementary file.

Acknowledgment

This work was supported by DEI internal grant to VKD. The work was carried out at The Hebrew University of Jerusalem and collaborating Institutes in Israel.

Author Contribution

The work was conceived, planned and executed by VKD.

References

1. van Aken O, Zhang B, Law S, Narsai R, Whelan J (2013) AtWRKY40 and AtWRKY63 modulate the expression of stress-responsive nuclear genes encoding mitochondrial and chloroplast proteins. Plant Physiol 162(1): 254-271.

2. Aoyama T, Dong CH, Wu Y, Carabelli M, Sessa G, et al. (1995) Ectopic expression of the Arabidopsis transcriptional activator Athb-1 alters leaf cell fate in tobacco. The Plant cell 7(11): 1773-1785.

3. Baba K, Nakano T, Yamagishi K, Yoshida S (2001) Involvement of a nuclear-encoded basic helix-loop-helix protein in transcription of the light-responsive promoter of psbD. Plant physiology 125(2): 595-603.

4. Barr MS, Willmann MR, Jenik PD (2012) Is there a role for trihelix transcription factors in embryo maturation? Plant signaling & behavior 7(2): 205-209.

5. Boutigny S, Sautron E, Finazzi G, Rivasseau C, Frelet-Barrand A, et al. (2014) HMA1 and PAA1, two chloroplast-envelope PIB-ATPases, play distinct roles in chloroplast copper homeostasis. Journal of experimental botany 65(6): 1529-1540.

6. Bracha K, Lavy M, Yalovsky S (2002) The Arabidopsis AtSTE24 is a CAAX protease with broad substrate specificity. The Journal of biological chemistry 277(33): 29856-29864.

7. Charuvi D, Kiss V, Nevo R, Shimoni E, Adam Z, et al. (2012) Gain and loss of photosynthetic membranes during plastid differentiation in the shoot apex of Arabidopsis. The Plant cell 24(3): 1143-1157.

8. del Pozo T, Cambiazo V, Gonzalez M (2010) Gene expression profiling analysis of copper homeostasis in Arabidopsis thaliana. Biochemical and biophysical research communications 393(2): 248-252.

9. Dalal V, Dagan S, Friedlander G, Aviv E, Bock R, et al. (2018) Transcriptome analysis highlights nuclear control of chloroplast development in the shoot apex. Scientific reports 8: 8881.

10. Franco-Zorrilla JM, Lopez-Vidriero I, Carrasco JL, Godoy M, Vera P, et al. (2014) DNA-binding specificities of plant transcription factors and their potential to define target genes. Proceedings of the National Academy of Sciences of the United States of America 111(6): 2367-2372.

11. Fristedt R, Williams-Carrier R, Merchant SS, Barkan A (2014) A thylakoid membrane protein harboring a DnaJ-type zinc finger domain is required for photosystem I accumulation in plants. The Journal of biological chemistry 289(44): 30657-30667.

12. Fujita H, Takemura M, Tani E, Nemoto K, Yokota A, et al. (2003) An Arabidopsis MADS-box protein, AGL24, is specifically bound to and phosphorylated by meristematic receptor-like kinase (MRLK). Plant cell physiology 44(7): 735-742.

13. Gao MJ, Lydiate DJ, Li X, Lui H, Gjetvaj B, et al. (2009) Repression of seed maturation genes by a trihelix transcriptional repressor in Arabidopsis seedlings. The Plant cell 21(1): 54-71.

14. Hiltscher H, Rudnik R, Shaikhali J, Heiber I, Mellenthin M, et al. (2014) The radical induced cell death protein 1 (RCD1) supports transcriptional activation of genes for chloroplast antioxidant enzymes. Frontiers in plant science 5: 475.

15. Hashimshony T, Wagner F, Sher N, Yanai I (2012) CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell Rep 2(3): 666-673.

16. Jeon Y, Jung HJ, Kang H, Park YI, Lee SH, et al. (2012) S1 domain-containing STF modulates plastid transcription and chloroplast biogenesis in Nicotiana benthamiana. The New phytologist 193(2): 349-363.

17. Jeong SY, Rose A, Meier I (2003) MFP1 is a thylakoid-associated, nucleoid-binding protein with a coiled-coil structure. Nucleic acids research 31(17): 5175-5185.

18. Karlen Y, McNair A, Perseguers S, Mazza C, Mermod N (2007) Statistical significance of quantitative PCR. BMC Bioinformatics 8: 131.

19. Kobayashi K, Baba S, Obayashi T, Sato M, Toyooka K, et al. (2012) Regulation of root greening by light and auxin/cytokinin signaling in Arabidopsis. The Plant cell 24(3): 1081-1095.

20. Kremnev D, Strand A (2014) Plastid encoded RNA polymerase activity and expression of photosynthesis genes required for embryo and seed development in Arabidopsis. Frontiers in plant science 5: 385.

21. Kusumi K, Iba K (2014) Establishment of the chloroplast genetic system in rice during early leaf development and at low temperatures. Frontiers in plant science 5: 386.

22. Lee U, Rioflorido I, Hong SW, Larkindale J, Waters ER, et al. (2007) The Arabidopsis ClpB/Hsp100 family of proteins: chaperones for stress and chloroplast development. The Plant journal: for cell and molecular biology 49(1): 115-127.

23. Lev S (2010) Non-vesicular lipid transport by lipid-transfer proteins and beyond. Nature reviews. Molecular cell biology 11: 739-750.

24. Li W, Ahn IP, Ning Y, Park CH, Zeng L, et al. (2012) The U-Box/ARM E3 ligase PUB13 regulates cell death, defense, and flowering time in Arabidopsis. Plant physiology 159(1): 239-250.

25. Ling Q, Huang W, Baldwin A, Jarvis P (2012) Chloroplast biogenesis is regulated by direct action of the ubiquitin-proteasome system. Science 338(6107): 655-659.

26. Liu F, Zhang X, Lu C, Zeng X, et al. (2015) Non-specific lipid transfer proteins in plants: presenting new advances and an integrated functional analysis. J Exp Bot 66(19): 5663-5681.

27. Mackay J (2007) Introduction to kinetic (real-time) PCR. Methods Mol Biol 353: 167-176.

28. Meier I, Phelan T, Gruissem W, Spiker S, Schneider D (1996) MFP1, a novel plant filament-like protein with affinity for matrix attachment region DNA. The Plant cell 8(11): 2105-2115.

29. Mhamdi A, Queval G, Chaouch S, Vanderauwera S, Van Breusegem F et al. (2010) Catalase function in plants: a focus on Arabidopsis mutants as stress-mimic models. Journal of experimental botany 61(15): 4197-4220.

30. Miernyk JA (2001) The J-domain proteins of Arabidopsis thaliana: an unexpectedly large and diverse family of chaperones. Cell stress & chaperones 6(3): 209-218.

31. Mitton FM, Pinedo ML, de la Canal L (2009) Phloem sap of tomato plants contains a DIR1 putative ortholog. Journal of plant physiology 166: 543-547.

32. Monte E, Tepperman JM, Al-Sady B, Kaczorowski KA, Alonso JM, et al. (2004) The phytochrome-interacting transcription factor, PIF3, acts early, selectively, and positively in light-induced chloroplast development. PNSA 101: 16091-16098.

33. Mou Z, He Y, Dai Y, Liu X, Li J (2000) Deficiency in fatty acid synthase leads to premature cell death and dramatic alterations in plant morphology. The Plant cell 12(3): 405-418.

34. Park HY, Lee SY, Seok HY, Kim SH, Sung ZR, et al. (2011) EMF1 interacts with EIP1, EIP6 or EIP9 involved in the regulation of flowering time in Arabidopsis. Plant Cell Physiol 52(8): 1376-1388.

35. Qiao J, Li J, Chu W, Luo M (2013) PRDA1, a novel chloroplast nucleoid protein, is required for early chloroplast development and is involved in the regulation of plastid gene expression in Arabidopsis. Plant Cell Physiol 54(12):2071-2084.

36. Rohde A, De Rycke R, Beeckman T, Engler G, Van Montagu M,et al. (2000) ABI3 affects plastid differentiation in dark-grown Arabidopsis seedlings. Plant cell 12(1): 35-52.

37. Ruth Finkelstein (2013) Abscisic Acid Synthesis and Response. Arabidopsis Book

38. Schwacke R, Fischer K, Ketelsen B, Krupinska K, Krause K (2007) Comparative survey of plastid and mitochondrial targeting properties of transcription factors in Arabidopsis and rice. Mol Genet Genomics 277(6):631-646.

39. Sessa G, Morelli G, Ruberti I (1993) The Athb-1 and -2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities. EMBO J 12: 3507-3517.

40. Shaikhali J, Heiber I, Seidel T, Stroher E, Hiltscher H, et al. (2008) The redox-sensitive transcription factor Rap2.4a controls nuclear expression of 2-Cys peroxiredoxin A and other chloroplast antioxidant enzymes. BMC plant biology 8: 48.

41. Stuhrwohldt N, Hartmann J, Dahlke RI, Oecking C, Sauter M (2014) The PSI family of nuclear proteins is required for growth in arabidopsis. Plant molecular biology 86(3): 289-302.

42. Sun X, Feng P, Xu X, Guo H, Ma J, et al. (2011) A chloroplast envelope-bound PHD transcription factor mediates chloroplast signals to the nucleus. Nature communications 2: 477.

43. Tripathy BC, Dalal VK (2013) Modulation of chlorophyll biosynthesis by environmental cues. Advances in Photosynthesis and Photorespiration", Springer Publications, Dordrecht, Netherlands, 36: PP. 601-639.

44. Willmann MR, Mehalick AJ, Packer RL, Jenik PD (2011) MicroRNAs regulate the timing of embryo maturation in Arabidopsis. Plant physiology 155: 1871-1884.

45. Yamaguchi M, Goue N, Igarashi H, Ohtani M, Nakano Y,et al. (2010) VASCULAR-RELATED NAC-DOMAIN6 and VASCULAR-RELATED NAC-DOMAIN7 effectively induce transdifferentiation into xylem vessel elements under control of an induction system. Plant physiol 153(3): 906-914.

46. Yu B, Wakao S, Fan J, Benning C (2004) Loss of plastidic lysophosphatidic acid acyltransferase causes embryo-lethality in Arabidopsis. Plant & cell physiology 45: 503-510.

47. Yu X, Zheng G, Shan L, Meng G, Vingron M, et al. (2014) Reconstruction of gene regulatory network related to photosynthesis in Arabidopsis thaliana. Front Plant Sci 5: 273.