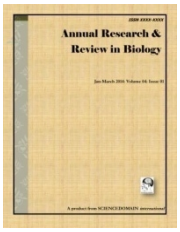


Annual Research & Review in Biology
4(17): 2779-2794, 2014

SCIENCEDOMAIN *international*
www.sciencedomain.org

Structures of Phospho*enol*pyruvate carboxylase (PEPC) Gene Promoter from C_4 and C_3 *Flaveria* species Using Sequence Analysis by Bioinformatics Tools

Eghtedary Naeini Abdolreza[1*], Etedali Fatemeh[1*], Mortazavi Seyed Elyas[2], Sobhani Najafabadi Ahmad[3], Golabadi Maryam[1], Gharyazi Behzad[2] and Malboobi Mohammad Ali[4]

[1]Department of Agronomy, Plant Breeding and Biotechnology, Khorasgan Isfahan Branch, Islamic Azad University, Isfahan, Iran.
[2]Agricultural Biotechnology Research Institute of Iran, Karaj, Iran.
[3]Department of Biotechnology, University of Tehran, Tehran, Iran.
[4]National Institute of Genetic Engineering and Biotechnology, Tehran, Iran.

Authors' contributions

This work was carried out in collaboration between all authors. Authors EF and SNA designed the study. Authors EF, MSE, ENA, GM and SNA managed the Literature searches. Author EF with guideline authors GB and MMA performed the bioinformatics analysis. All authors read and approved the final manuscript.

Original Research Article

Received 5th March 2014
Accepted 7th April 2014
Published 6th May 2014

ABSTRACT

Aim: This research can be used to design and subsequently to study stable- tissues-specific gene expression in other plants, and potentially may be used in altering plants with C_3 photosystem into C_4 through genetics.
Study design: *In silico* analysis by bioinformatics tools was used to analyze the promoter regions of Phospho*enol*pyruvate carboxylase (PEPC) genes in genus *Flaveria* (Asteraceae).
Methodology: Detailed promoter gene studies of the C_3 and C_4 *ppcA1* genes in *Flaveria sp*. have been deposited by EMBL/GenBank data. To find regulatory elements in promoter sequences, PLANTCARE and PLACE database were applied. Moreover, ClustalW2 &

**Corresponding authors: Emails: anita_19752002@yahoo.com, a.eghtedar@gmail.com;*

MAFFT were used for multiple sequence alignment.
Results: The results revealed that light-inducible promoters of PEPC genes in *Flaveria* species have *cis*-elements for organ- and cell-specific expression, which may be subdivided into six classes: 1) Elicitor specific, 2) Component specific (hormone-responsive elements), 3) Light responsiveness, 4) Binding site specific, 5) Condition specific (stress-responsive elements) and 6) Plant tissues specific.
Conclusion: The distal region (DR) of promoter behaves as an enhancer-like expression module and is able to confer a mesophyll expression while the proximal region (PR) is responsible for a quantitative expression. Coordinated motifs indicate that the response of *pepc* gene is controlled by light in coordination with hormones and stresses and potentially may be used in altering plants with C_3 photosystem into C_4 through genetics.

Keywords: Cis-regulatory element; Flaveria; Phosphoenolpyruvate carboxylase (PEPC) promoter.

1. INTRODUCTION

The C_4 plants are adapted to high light, arid and warm environments and achieve a higher photosynthetic capacity and higher water- and nitrogen-use efficiencies compared to C_3 plants [1]. C_4 photosynthetic carbon cycle is a sophisticated addition to the C_3 photosynthetic pathway and enables C_4 plants to cope well with high light intensities, high temperatures and drought. The high photosynthetic capacity of C_4 plants is due to their unique mode of carbon assimilation which involves two different photosynthetic cell types, mesophyll and bundle-sheath cells but in leaves of C_3 plants, photosynthesis and subsequent carbon and nitrogen metabolism are proceed mainly in the mesophyll cells. However, some C_4 enzymes such as phospho*enol*pyruvate carboxylase (PEPC), pyruvate orthophosphate dikinase, or NADP-dependent malic enzyme is not limited to C_4 plants. In C_4 plants, CO_2 is initially fixed into a C_4 acids malate and/or aspartate by phospho*enol*pyruvate carboxylase (PEPC) in the leaf mesophyll cells into and then transported to the bundle-sheath. In the bundle-sheath cells the C_4 acids are decarboxylated, and the CO_2 is prefixed by ribulose 1,5-bisphosphate carboxylase/oxygenase [2,3].

C_3 species and nonphotosynthetic tissues of C_4 species normally contain nonphotosynthetic isoforms of C_4 enzymes. This fact may indicate that these C_3 isoforms are enrolled as a starting point for the evolution of the C_4 genes [4,5]. The C_4 cycle enzymes differ from their C_3 counterparts in kinetic and regulatory characteristics [6].

A gene transformation was used to overproduce C_4 enzymes of bundle sheath cells in rice, in which cDNA of C_4 enzymes under a strong promoter to an increased enzyme activity up to 110-fold compared to non-transgenic lines [7]. It was shown that the different expression of C_4 genes is largely under transcriptional factors control [8,9]. Light-inducible promoters of C_4-specific PEPC genes have *cis*-elements for organ- and cell-specific expression.

In this research, we focused on 5'flanking sequences of *pepc* gene in genus *Flaveria* (Asteraceae). Genus *Flaveria* involves both C_3 and C_4 and also numerous intermediate species, so it is a very appropriate genus for this evaluation.

The PEPCs of *Flaveria* species are encoded by a gene family comprising three classes: *ppcA*, *ppcB* and *ppcC*. *PpcA* orthologous of PEPC genes are found in all C_3 and C_3-C_4

intermediate *Flaveria* species [3,5,10]. The comparative enzymatic analysis of PEPC proteins from C_3 and C_4 *Flaveria* species revealed that the *ppcA* PEPCs of *F. pringlei* (C_3), *F. cronquistii* (C_3), *F. bidentis* (C_4) and *F. trinervia* (C_4) are typical C_3 and C_4 PEPCs respectively and there are a few differences in amino acid sequences [5]. This study is a short review on characterization of the promoter sequences of *ppcA* gene in four *Flaveria* species and presents a comparative analysis by bioinformatics tools.

2. MATERIALS AND METHODS

Detailed promoter gene studies of the C_3 and C_4 *ppcA1* genes in *Flaveria sp*. have been deposited by EMBL/GenBank data libraries under the accession numbers X64143 (*F. trinervia*)(C_4), AY297087 (*F. bidentis*)(C_4), AY297089 (*F. cronquistii*)(C_3), and X64144 (*F. pringlei*)(C_3).

The bioinformatics tools, *in silico* analysis, were used to analyze the promoter regions of Phospho*enol*pyruvate carboxylase (PEPC) genes. To find regulatory elements in promoter sequences, PLANTCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html) and PLACE database [11] were applied.

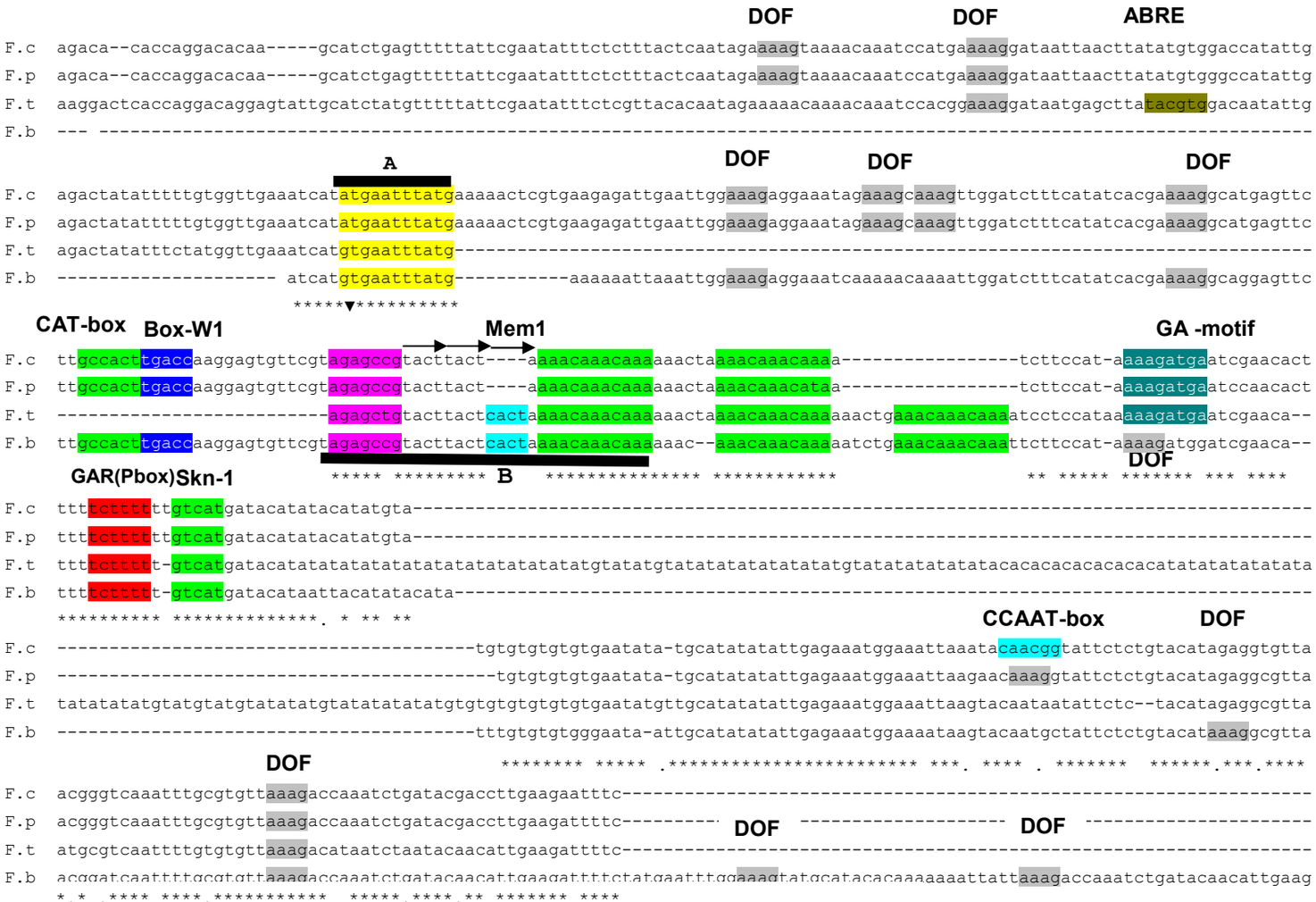
Moreover, ClustalW2 & MAFFT version 7 (http://mafft.cbrc.jp /alignment / server) were used for multiple sequence alignment.

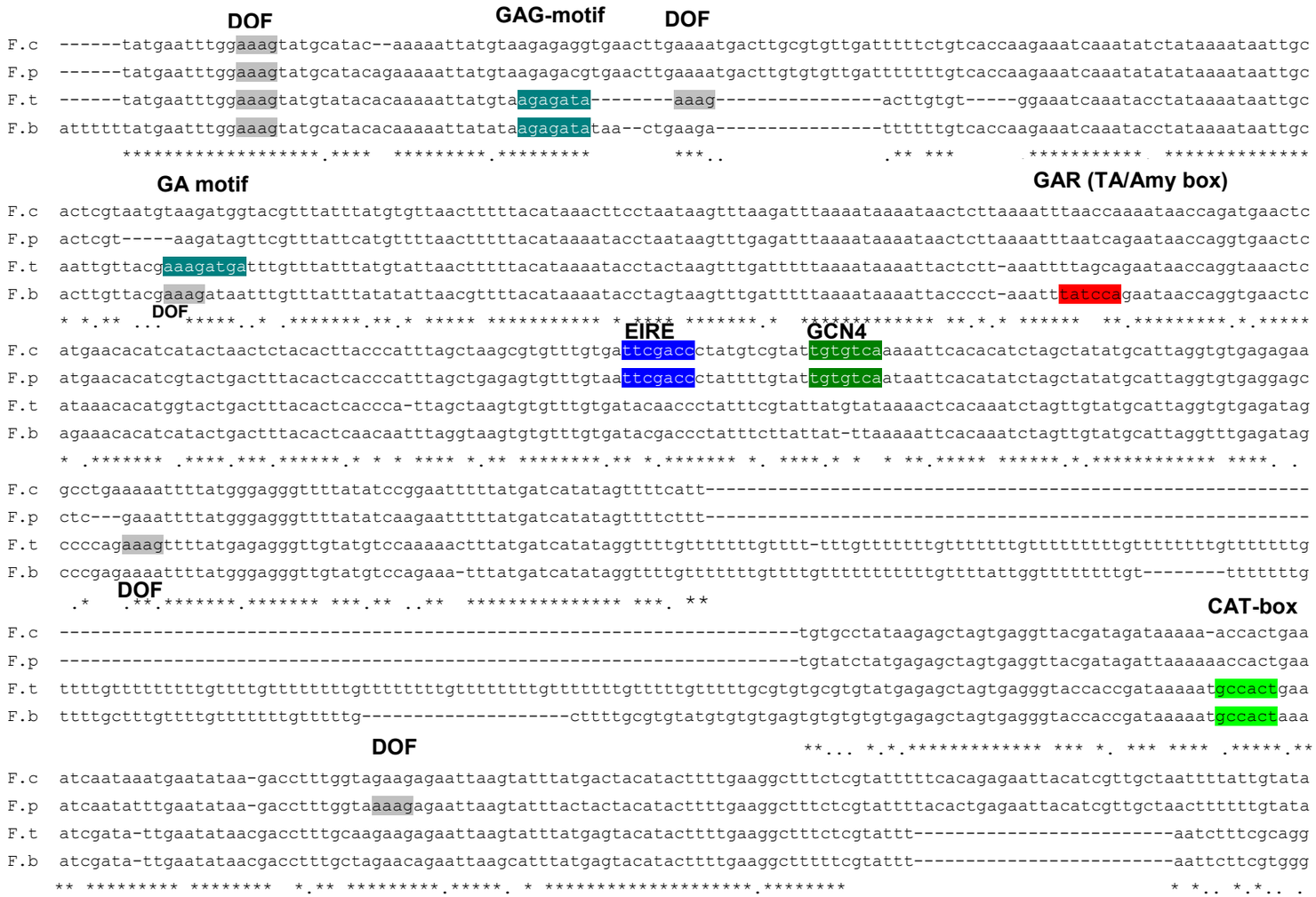
3. RESULTS AND DISCUSSION

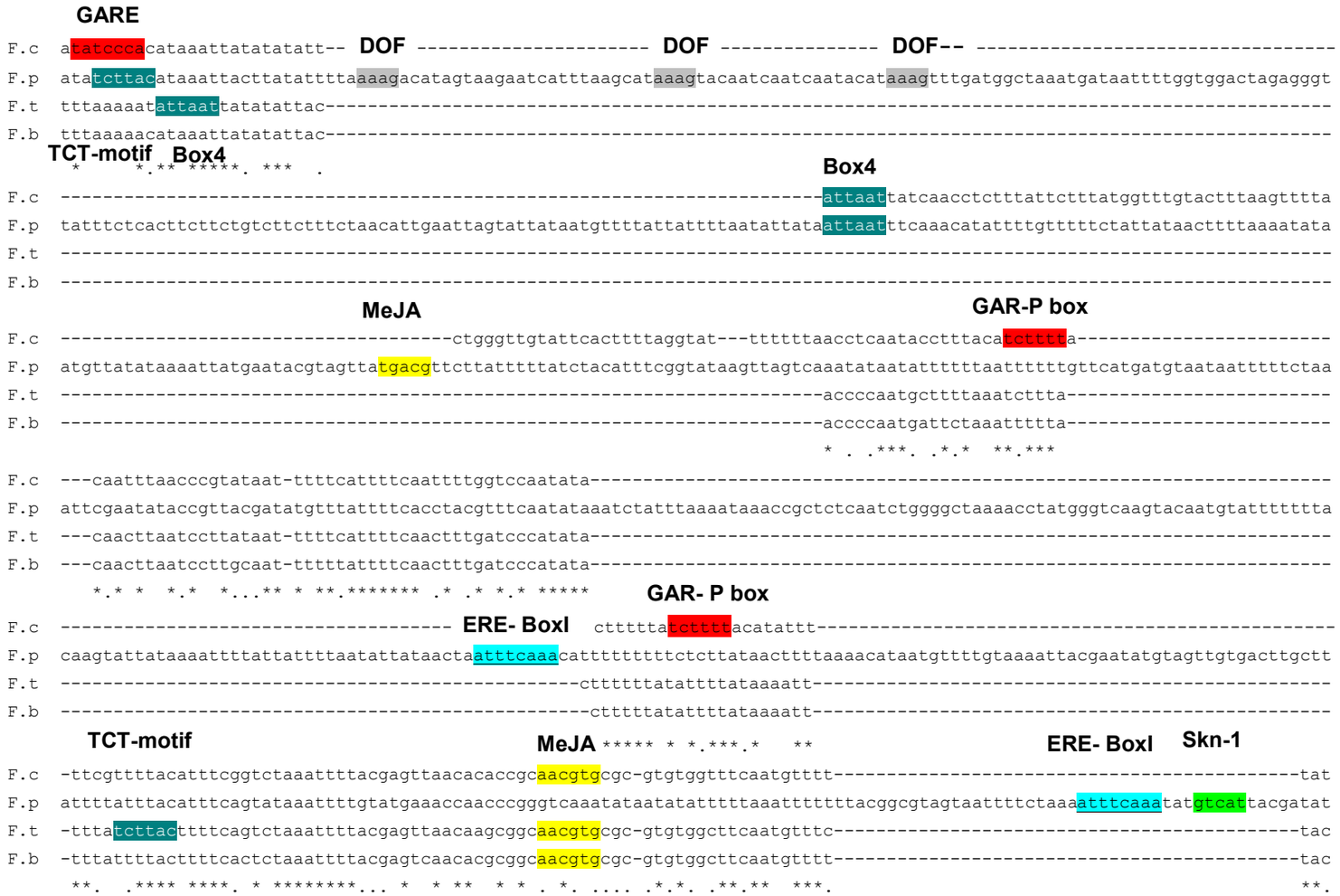
Photosynthesis cycle genes in C_4 and C_3 plants are largely regulated by transcription factors. This type of gene regulation involves sets of *cis*-regulatory modules and their corresponding trans-regulatory factors that interact and thereby control the specific expression of the genes in either mesophyll or bundle-sheath cells [9]. *pepcA* Promoter of *Flaveria* species contains two main regions, a proximal (PR) and a distal region (DR), Both regions are necessary and sufficient for a high mesophyll-specific expression. These regions do not act as separate modules and have additive effect as a synergistic transcriptional controlling system that evolves each other. The DR behaves as an enhancer-like expression module and would be able to confer mesophyll expression while the PR is responsible for quantitative expression and harbors *cis*-regulatory determinants which account for high levels of PEPC expression in the leaf [5,12].

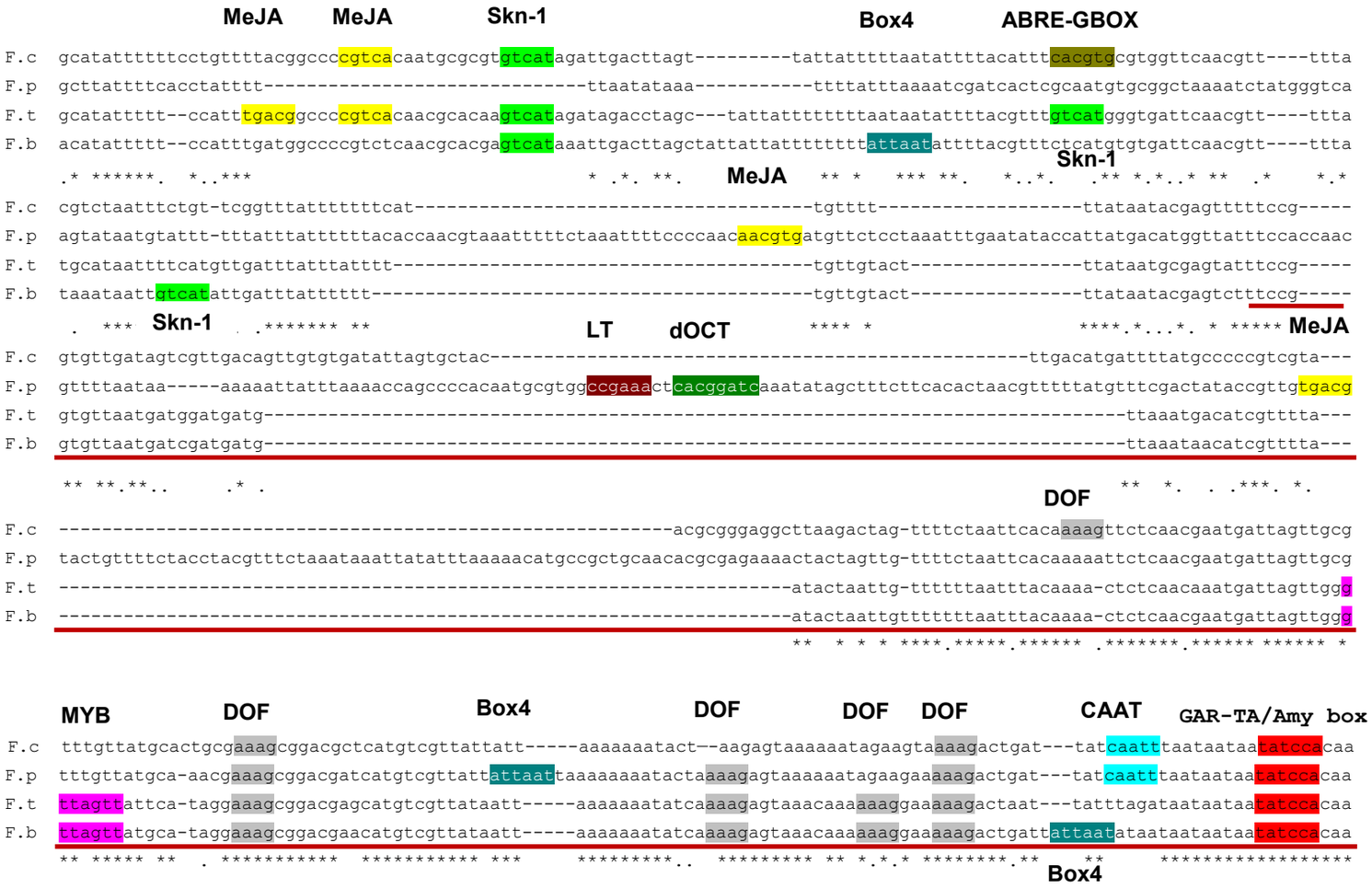
Translation initiation signals or the Kozak consensus sequences are species-specific and recognized by the ribosome as the translational start site. The consensus sequence GTA in four *Flaveria* species are recognized after 5'UTR [13,14] appears as 3 bp upstream of the translational start codon ATG (Fig. 1).

After analyzing of the *pepc* promoter regions, we revealed that the light-inducible promoters of PEPC genes in *Flaveria* species have *cis*-acting elements for organ- and cell-specific expression, subdivided to six classes: 1) Elicitor specific, 2) Component specific (hormone-responsive elements), 3) Light responsiveness, 4) Binding site specific, 5) Condition specific (stress-responsive elements) and 6) Plant tissues specific.


DOF
DOF
ABRE
A
DOF
DOF
DOF
CAT-box
Box-W1
Mem1
GA -motif
DOF
GAR(Pbox)Skn-1
B
CCAAT-box
DOF
DOF
DOF
DOF


DOF
GAG-motif
DOF
GA motif
GAR (TA/Amy box)
DOF
EIRE
GCN4
DOF
CAT-box
DOF


GARE
DOF
DOF
DOF--
TCT-motif
Box4
Box4
MeJA
GAR-P box
GAR- P box
ERE- BoxI
TCT-motif
MeJA
ERE- BoxI
Skn-1


MeJA
MeJA
Skn-1
Box4
ABRE-GBOX
F.c gcatatttttttcctgttttacggcccgtcacaatgcgcgtgtcatagattgacttagt---------tattatttttaatattttacatttcacgtgcgtggttcaacgtt----ttta
F.p gcttattttcacctatttt-----------------------------ttaatataaa-----------ttttatttaaaatcgatcactcgcaatgtgcggctaaaatctatgggtca
F.t gcatattttt--ccatttgacggcccgtcaacgcacaagtcatagatagacctagc---tattatttttttttaataatattttacgtttgtcatgggtgattcaacgtt----ttta
F.b acatattttt--ccatttgatggccccgtctcaacgcacgagtcataaattgacttagctattattatttttttttattaatattttacgtttctcatgtgtgattcaacgtt----ttta
Skn-1
MeJA
F.c cgtctaatttctgt-tcggtttatttttttcat---tgtttt-------------------ttataatacgagtttttccg-----
F.p agtataatgtattt-tttatttatttttttacaccaacgtaaatttttctaaatttccccaacaacgtgatgttctcctaaatttgaatataccattatgacatggttatttccaccaac
F.t tgcataattttcatgttgatttatttatttt---tgttgtact----------------ttataatgcgagtatttccg-----
F.b taaataattgtcatattgatttatttttt---tgttgtact----------------ttataatacgagtctttccg-----
Skn-1
LT
dOCT
MeJA
F.c gtgttgatagtcgttgacagttgtgtgatattagtgctac---ttgacatgattttatgcccccgtcgta---
F.p gttttaataa-----aaaaattatttaaaaccagccccacaatgcgtggccgaaactcacggatcaaatatagctttcttcacactaacgtttttatgtttcgactataccgttgtgacg
F.t gtgttaatgatggatgatg---ttaaatgacatcgtttta---
F.b gtgttaatgatcgatgatg---ttaaataacatcgtttta---
DOF
F.c --acgcgggaggcttaagactag-ttttctaattcacaaaagttctcaacgaatgattagttgcg
F.p tactgttttctacctacgtttctaaataaattatatttaaaaacatgccgctgcaacacgcgagaaaactactagttg-ttttctaattcacaaaaattctcaacgaatgattagttgcg
F.t ---atactaattg-ttttttaatttacaaaa-ctctcaacaaatgattagttggg
F.b ---atactaattgttttttttaatttacaaaa-ctctcaacgaatgattagttggg
MYB
DOF
Box4
DOF
DOF
DOF
CAAT
GAR-TA/Amy box
F.c tttgttatgcactgcgaaagcggacgctcatgtcgttattatt-----aaaaaaatact--aagagtaaaaaatagaagtaaaagactgat---tatcaatttaataataatatccacaa
F.p tttgttatgca-aacgaaagcggacgatcatgtcgttattattaattaaaaaaaatactaaaagagtaaaaaatagaagaaaaagactgat---tatcaatttaataataatatccacaa
F.t ttagttattca-taggaaagcggacgagcatgtcgttataatt-----aaaaaaatatcaaaagagtaaacaaaaaaggaaaaagactaat---tatttagataataataatatccacaa
F.b ttagttatgca-taggaaagcggacgaacatgtcgttataatt-----aaaaaaatatcaaaagagtaaacaaaaaaggaaaaagactgattattaatataataataatatccacaa
Box4

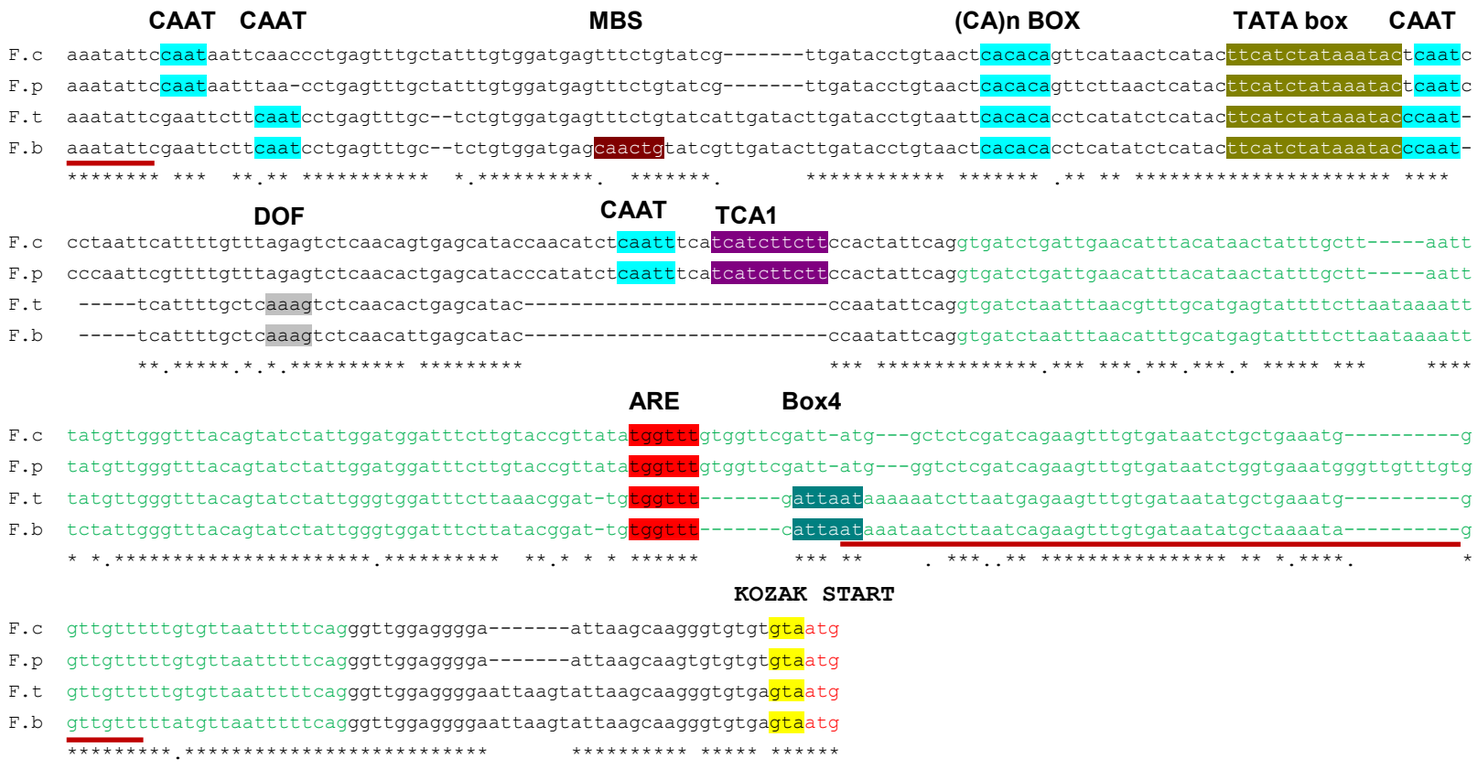


Fig. 1. Sequence alignment of conserved regulatory elements of Phospho*enol*pyruvate carboxylase A (*ppcA1)* gene promoter from C_4 and C_3 *Flaveria* species such as *F. trinervia* (F.t, C_4), *F. bidentis* (F.b, C_4), *F. cronquistii* (F.c, C_3) and *F. pringlei* (F.p, C_3) using sequencing analysis by bioinformatics tools.

3.1 Elicitor Specific

3.1.1 Box-W1

The W-box [(T)TGAC(C/T) motif] is the binding site for members of the WRKY family of transcription factors. There is increasing evidence that W boxes are a major class of *cis*-acting elements responsible for the pathogen inducibility of many plant genes as an elicitor [15]. The importance of W boxes was illustrated recently by studies on the Arabidopsis transcriptome during systemic acquired resistance [16,17]. The W1-box (TTGACC motif, by blue box and White nucleotides, Fig. 1) appears as 6 bp upstream sequences of the Mem1 between part A & B (black bars) in DR of both C_3 and C_4 plants except for *F. trinervia*.

3.1.2 EIRE

The *Cis*-acting regulatory element, EIRE box (TTCGACC motif [18], marked with blue box and white nucleotides) is involved in elicitor-responsive element of C_3 plants such as *F. pringlei* and *F. cronquistii* (Fig. 1).

3.2 Component Specific (Hormone-responsive Elements)

3.2.1 Ethylene-responsive element binding protein (EREBP or ERE box1)

ERE is a homeobox gene from *Arabidopsis thaliana* and other plants which encodes a transcription factor for response to plant hormone ethylene [19]. Two ERE *cis*-acting regulatory elements regions (ATTTCAAA motif, [18]) are marked by turquoise boxes only in *F. pringlei* (C_3) between PR and DR of *pepc* promoter (Fig. 1).

3.2.2 Jasmonic acid-responsive element (MejA)

The CGTCA- and TGACG- motifs [18,20] in *Hordeum vulgare* and the T/G-box (AACGTG) in *Nicotiana benthamiana and Brassica juncea* [21] are *cis*-acting regulatory elements involved in the methyl Jasmonate (MeJA)-responsiveness. These sequences are indicated by yellow boxes (Fig. 1). T/G-box is present in all C_3 and C_4 *Flaveria* species whereas the 5' flanking of *F. bidentis ppcA1* promoter does not contain detectable binding sites for the CGTCA- or TGACG-motifs.

Deletion analysis in *Nicotiana benthamiana* showed that the removal of a conserved T/G-box (AACGTG) in *ppcA1* promoter greatly reduced the induction by MeJA [21].

3.2.3 Abscisic acid-responsive element (ABRE)

The abscisic acid-responsive *cis*-acting element ABRE represents the motifs TACGTG [18] and CACGTG [22,23] which belong to the G-box family. ABRE contain an ACGT core, a sequence known to be recognized by plant bZIP proteins [24]. The ABRE are indicated by dark yellow box in *F. trinervia* and *F. cronquistii* (Fig. 1).

3.2.4 Salicylic acid responsive element (TCA1)

The TCA-1 type *cis*-acting elements (TCATCTTCTT motif, [25] can respond to salicylic acid, a well-characterized signaling molecule in plants. The motif is marked with violet box

and white nucleotides in proximal regions of C_3 plants such as *F. pringle* and *F. cronquistii* (Fig. 1).

3.2.5 Gibberellins-responsive (GAR)

Gibberellins response complex (GARC) includes O2S/W box, pyrimidine (P) box (C/TCTTTT), TA/Amy box (TATCCA) and GA response element (GARE; C/TAACC/GG/AA/CC/A, TATCCCA) [26,27,28]. P box and TA/Amy box are present in all C_3 and C_4 *Flaveria* species whereas GA response element (GARE) is marked in 5' flanking of *F. cronquistii ppcA1* promoter (by red boxes, Fig. 1).

3.3 Light Responsiveness

BOX4 (ATTAAT motif, by teal box) in *Petroselinum crispum* [18], BOXI (TTTCAAA motif, by turquoise box) in *Pisum sativum* [23,29], G BOX (CACGTG motif, by dark yellow box) in *Arabidopsis thaliana* and *Pisum sativum* [23,29], GA-motif (AAAGATGA motif, by teal box) in *Helianthus annuus* [23,30], GAG-motif (AGAGATG motif, by teal box) in *Spinacia oleracea* [23,31] and TCT-motif (TCTTAC motif [18], by teal box) were involved in light responsiveness. This study show that BOX4 is present in all C_3 and C_4 *Flaveria* species whereas *F. trinervia* has BOX4, G BOX, GA-motif, GAG-motif and TCT-motif as light responsive elements.

3.4 Binding Site Specific

3.4.1 MYB transcription factor binding site

A putative MYB transcription factor binding site (GTTAGTT motif [32], by pink box) is present in all C_4 *Flaveria* species, but is missing in two C_3 species. These sequences are prime candidates for transcription-enhancing *cis*-regulatory elements [33]. MYB proteins comprise one of the largest families of transcription factors in plants with almost 200 different MYB genes present in the Arabidopsis genome. MYB transcription factors regulate a diverse range of pathways including secondary metabolism, signal transduction and defence responses [34,35].

3.4.2 CAAT-box (CCAAT-box)

CAAT-box (CAAT, CAAAT, CAATT, CCAAT motifs) is a common *cis*-acting element in promoter and enhancer region which is generally found in upstream of the TATA box, but at much farther distances from the starting point and can also operate in both directions [23]. Their exact position relative is variable and they may be found in either orientation relative to the direction of transcription from the downstream initiation site. Fang et al. [36] have established that CCAAT-box is recognized by the Nuclear Factor 1/CCAAT transcription factor (NF1/CTF) family and also several other transcription factors/complexes which include; NF-E3, GATA1, NF-Y and C/EBP. They are able to regulate the initiation of transcription by an interaction of CCAAT-binding transcription factors with the basal transcription initiation complex [37]. CCAAT box is present at downstream of TATA box in C_4 *Flaveria* species (by turquoise box, Fig. 1), but PR of C_3 *ppcA1* promoter does not contain detectable binding sites for these transcription-enhancing *cis*-regulatory elements [33]. Also, CAAT and CAATT boxes have been designated at upstream of TATA box in all C_3 and C_4 *Flaveria* species (by turquoise box, Fig. 1).

3.4.3 TATA box

A TATA-box sequence (TTCATCTATAAATAC motif [33], by dark yellow box) has been found in almost all plant *pepc* genes which in turn binds a TATA binding protein and assists the formation of the RNA polymerase transcriptional complex. The TATA box typically lies very close to the transcriptional start site (often within 50 bases) [23,38].

3.4.4 (CA)n box

The (CA)n box sequence (by turquoise box, Fig. 1) has been designated at upstream of the TATA sites in all C_3 and C_4 *Flaveria* species [23,39].

3.4.5 FtHB1 to FtHB4- zinc finger homeodomain proteins (ZF-HD)

Studies were shown that an intron in the 5' untranslated leader region of *ppcA1* gene in transgenic *F. bidentis* (by green nucleotides, Fig. 1) is not essential for the control of expression [5]. The PR region of the C_4 *ppcA1* promoter contains *cis*-regulatory elements so, interacts with homeobox transcription factors of the zinc finger homeodomain proteins (ZF-HD) subclass (named FtHB1 to FtHB4); whereas the PR of the C_3 *ppcA1* promoter does not contain detectable binding sites for these zinc finger homeobox proteins [5,33,40]. Fragments of the *F. trinervia ppcA1* promoter that interact with the FtHB proteins in the yeast one-hybrid system [40] are marked by red bars (Fig. 1). Deletion of intron in 5' untranslated region of C_4 -specific *ppcA1* does not affect the expression pattern and strength, indicating that the previously isolated zinc finger-homeobox transcription factors that specifically interact with this intron are not involved in regulating *ppcA1* expression levels [33].

3.4.6 DOF protein (AAAG)

Dof transcription factors are ubiquitous in plants, and associated with many plant-specific biological processes. In maize, both Dof1 and Dof2 have been shown to bind to an AAAAGG motif in the 5' upstream enhancer region of the C_4 *pepc* gene. Dof1 is a positive regulator of C_4 *pepc* expression, with roles in tissue-specific and light-induced expression. In contrast, the related Dof2 protein was shown to block transcriptional activation of the C_4 *pepc* promoter, possibly through competitive binding interactions with the Dof1 protein [41]. AAAG is Dof core recognition sequence and is so frequent that is regulated by light and tissue specific gene expression [23,42]. In this study, the AAAG Dof Core protein is indicated by bright gray boxes and very frequent in DR and PR of all C_3 and C_4 *Flaveria* species (Fig. 1).

3.5 Condition Specific (Stress-responsive Elements)

These *cis*-acting elements are involved in regulation of gene expression under stress condition.

The LTR box (CCGAAA motif, by dark red box, Fig. 1) as a stress-responsive element involved in the low-temperature responsiveness [18] is present in *F. pringle pepc* promoter. The MBS box ((C/T) AACTG motif, by dark red box) as a MYB binding site involved in drought-inducibility [18] is indicated in 5' upstream region of the *F. bidentis* C_4 *pepc* gene. The ARE motif (TGGTTT) is found in *Z. mays*, necessary for both anaerobic induction and the binding of the transcription factors (TFs) of the MYB family [43,44] has been designated in the 5' flanking of all C_3 and C_4 *Flaveria ppcA1* promoter and the CCAAT-box (CAACGG

motif, by turquoise box) is recognized as a MYBHv1 binding site in *Hordeum vulgare* [18,45] and marked in *F. cronquistii ppcA1* promoter.

3.6 Plant Tissues Specific

The CAT-box (GCCACT motif [18], by green box) in all C_3 and C_4 *Flaveria* and the dOCT motif (CaCGGATC motif [18], by green box) in *F. pringle ppcA1* promoter are *cis*-acting regulatory elements related to meristem expression. The GCN4 motif (TGTGTCA motif [18], by bright green box) only in C_3 *Flaveria* species and the Skn-1 motif (GTCAT [18,23], by bright green box) in DR and PR of all C_3 and C_4 *Flaveria ppcA1* promoter required for endosperm expression.

3.6.1 Mem1 (mesophyll expression module 1)

The distal region of the *ppcA1* promoter in *Flaveria* species has a 41-bp module, named MEM1 (for mesophyll expression module1). Evolutionary and functional analyses identified the tetra nucleotide CACT as a key element of Mem1 in conjunction with the PR, is sufficient for mesophyll-specific expression and critical for Mem1 function. The CACT-containing *cis*-regulatory element is necessary for mesophyll expression but may not be sufficient; this region may be targeted by a basic leucine zipper transcription factor [5,12,33]. MEM1 could be subdivided into two submodules, A and B regions 11 and 30 bp, respectively. The position of Mem1 is indicated by black bars in distal region (DR) of *Flaveria* species (Fig. 1). The tandem duplicated T/CACT repeats are labeled by arrows in all C_4 *ppcA1* promoters but also in the C_3 promoters. The part A (yellow boxes, the 11 bp of the 5' terminal sequences) and part B (pink and bright green boxes, the 30 bp of the 3'part of Mem1) segments in DR are showed in Fig. 1. The comparison of Mem1 between C_3 and C_4 species shows two remarkable features. The A part differs only in one single nucleotide at the very 5' end (labeled with black triangle in Fig. 1). Mem1 of *F. trinervia* (C_4) and *F. bidentis* (C_4) holds a guanine at the outermost 5' position of Mem1; it is an adenine in *F. pringlei* (C_3) and *F. cronquistii* (C_3) [12,33].

A tetra nucleotide (CACT) is present in the Mem1 of *F. trinervia* (C_4) and *F. bidentis* (C_4) but is absent in the *F. pringlei* (C_3) and *F. cronquistii* (C_3) sequence.

The AAACAAACAAA sequence (green boxes in DR) repeats two times in the 3'part of Mem1 in C_3 species but three times in C_4. A comparison of the 5' flanking regions identified in each case Mem1 homologous sequences where the A and B parts were separated by 97 to 108 bp except *F. trinervia* (Fig. 1) this suggests that the contiguous arrangement of the A and B parts is not of functional importance, so the A and B parts should form separate *cis*-regulatory units [5].

The results revealed that light-inducible promoters of PEPC genes in *Flaveria* species have *cis*-elements for organ- and cell-specific expression, which may be subdivided into six classes that controlled by light in coordination with hormones and stresses. The expression pattern of the *ppcA1* gene was modified very early in evolution from C_3 to C_4 [45,46] so that comparatively small changes in the nucleotide sequence should be responsible for these changes that give rise to a novel mode of expression. It will be interesting to see whether the evolution of mesophyll- or bundle-sheath cell-specific gene expression in *Flaveria* always relied upon the same set of *cis*- and trans-regulatory elements.

By fusing the DR of the *ppcA1* promoter of *F. trinervia* (C_4) to the *ppcA1* promoter of *F. pringlei* (C_3), a mesophyll expression component is added to that promoter, but the overall promoter strength does not alter substantially because the DR of the C_4 *ppcA1* promoter provides mesophyll specificity, while the PR is responsible for a quantitative expression. The DR of the C_4 *ppcA1* promoter may contain transcription repressing sequences that reduce *ppcA1* expression in the bundle-sheath cells and the vascular bundle and that thereby relatively increase mesophyll expression potential only when it is combined with a PR from either the C_4 or the C_3 *ppcA1* promoter [5,12,33].

4. CONCLUSION

The key enzymes of photosynthetic carbon assimilation, Phospho*enol*pyruvate carboxylase (PEPC), in C_4 plants have evolved from C_3 isoforms which were present in the C_3 ancestral species. The orthologous *ppc* genes (*ppcA*) of the C_3 *Flaveria* are only weakly expressed and their transcripts do not accumulate in a leaf-specific manner but, rather, are present in all plant organs. Changes in the upstream regions of the *ppcA* genes from C_3 and C_4 *Flaveria* are caused the differences in the expression levels. Sequence alignment of conserved regulatory elements of *ppcA1* promoter in *Flaveria* identified that some 5′ flanking sequences such as TATA box, CAT box, CCAAT, MEM1 (A and B regions), Meja, BOX4, ARE, DOF, GAR(P box), SKN-1, GAR (TA/Amy box) and kozak sequence are essentially homologous in both C_3 and C_4 *Flaveria*. This research also provides evidence for the new *cis*-acting sequences in the C_4 *Flaveria* species such as MYB and CACT motif of MEM1 in distal region of *ppcA1* promoter that enhance mesophyll expression and concomitantly represses expression in bundle sheath cells and vascular bundles. Changes at these positions in the C_3 sequence may be necessary and sufficient to create a mesophyll-specificity element during C_4 evolution.

The putative promoter regions share several identical sequence motifs only in C_3 *Flaveria* (TCA1, EIRE and GCN4 motifs). Additionally, alterations in elements that could contribute to differences in expression rates and light regulation. This research was concluded that promoter studies of the C_3 and C_4 *ppcA1* genes in *Flaveria sp*. can be used to design and subsequently to study stable- tissues- specific gene expression in other plants, and potentially may be used in altering plants with C_3 photosystem into C_4 through genetics.

ACKNOWLEDGMENTS

This research was supported by agriculture jahad organization of the Isfahan province and in part by a Grant-in- Aid from the Khorasgan (Isfahan) Branch, Islamic Azad University of IRAN.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Black Jr CC. Photosynthetic carbon fixation in relation to net CO_2 uptake. Annual Review of Plant Physiology. 1973;24:253-286.
2. Hatch MD. C_4 photosynthesis: a unique blend of modified biochemistry, anatomy and ultrastructure. Biochimica et Biophysica Acta. 1987;895:81-106.

3. Westhoff P, Gowik U. Evolution of C_4 phospho*enol*pyruvate carboxylase. Genes and proteins: a case study with the genus *Flaveria*. Annals of Botany. 2004;93:13-23.
4. Monson RK. The origins of C_4 genes and evolutionary pattern in the C_4 metabolic phenotype. In: Sage RF, Monson RK, eds. C_4 Plant Biology. San Diego: Academic Press. 1999;377-410.
5. Gowik U, Burscheidt J, Akyildiz M, Schlue U, Koczor M, Streubel M, Westhoff P. *cis*-regulatory elements for mesophyll-specific gene expression in the C_4 plant *Flaveria trinervia*, the promoter of the C_4 phospho*enol*pyruvate carboxylase gene. Plant Cell 2004;16:1077-1090.
6. Svensson P, BlaÈsing OE, Westhoff P. Evolution of C_4 phospho*enol*pyruvate carboxylase. Archives of Biochemistry and Biophysics. 2003;414:180-185.
7. Ku MSB, Agarie S, Nomura M, Fukayama H, Tsuchida H, Ono K, Hirose S, Toki S, Miyao M, Matsuoka M. High-level expression of maize phospho*enol*pyruvate carboxylase in transgenic rice plants. Nature Biotechnology. 1999;17:76-80.
8. Dengler NG, Nelson T. Leaf structure and development in C_4 plants. In: Sage RF, Monson RK, eds. C_4 Plant Biology. San Diego: Academic Press. 1999;133-172
9. Sheen J. C_4 gene expression. Annual Review of Plant Physiology and Plant Molecular Biology. 1999;50:187-217.
10. Hermans J, Westhoff P. Homologous genes for the C_4 isoform of phospho*enol*pyruvate carboxylase in a C_3- and a C_4-*Flaveria* species. Mol. Gen. Genet. 1992;234:275-284.
11. Higo K, Ugawa Y, Iwamoto M, Korenaga T. Plant *cis*-acting regulatory DNA elements (PLACE) database. Nucleic Acids Res. 1999;27:297-300.
12. Akyildiz M, Gowik U, Engelmann S, Koczor M, Streubel M, Westhoff P. Evolution and function of a *cis*-regulatory module for mesophyll-specific gene expression in the C_4 dicot *Flaveria trinervia*. The Plant Cell. 2007;19:3391-3402.
13. Cavener DR. Comparison of the consensus sequence flanking translational start sites in Drosophila and vertebrates. Nucleic Acids Res. 1987;15:1353-1361.
14. Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. 1987;15:8125-8148.
15. Rushton PJ, Torres JT, Parniske M, Wernert P, Hahlbrock K, Somssich IE. Interaction of elicitor-induced DNA binding proteins with elicitor response elements in the promoters of parsley PR1 genes. EMBO J. 1996;15:5690-5700.
16. Maleck K, Levine A, Eulgem T, Morgen A, Schmid J, Lawton K, Dangl JL, Dietrich RA. The transcriptome of Arabidopsis thaliana during systemic acquired resistance. Nat. Genet. 2000;26:403-410.
17. Petersen M, Brodersen P, Naested H, Andreasson E, Lindhart U, Johansen B, Nielsen HB, Lacy M, Austin MJ, Parker JE, et al. Arabidopsis MAP kinase 4 negatively regulates systemic acquired resistance. Cell. 2000;103:1111-1120.
18. Mongkolsiriwatana C, Pongtongkam P, Peyachoknagul S. *In silico* promoter analysis of photoperiod-responsive genes identified by DNA microarray in rice *(Oryza sativa L.)*. Kasetsart J. (Nat. Sci.). 2009;43:164-177.
19. Büttner M, Singh KB. Arabidopsis thaliana ethylene-responsive element binding protein (AtEBP), an ethylene-inducible, GCC box DNA-binding protein interacts with an ocs element binding protein. Proc. Natl. Acad. Sci. U.S.A. 1997;94 (11):5961-6.
20. Rouster J, Leah R, Mundy J, Cameron-Mills V. Identification of a methyl jasmonate responsive region in the promoter of a lipoxygenase 1 gene expressed in barley grain. Plant J. 1997;11(3):513-523
21. Wu XF, Wang CL, Xie EB, Gao Y, Fan YL, Liu PQ, Zhao KJ. Molecular cloning and characterization of the promoter for the multiple stress-inducible genes *BjCHI1* from *Brassica juncea.* Planta. 2009;22(6):1231-1242.

22. Yamaguchi-Shinozaki K, Shinozaki K. Arabidopsis DNA encoding two desiccation-responsive *rd29* genes. Plant Physiol. 1993;101(3):1119-1120.
23. Nalbandi K, Baghban Kohnehrouz B, Alami Saeedand K Gholizadeh A. Isolating barley (*Hordeum vulgare L.*) B1 hordein gene promoter and using sequencing analysis for the identification of conserved regulatory elements by bioinformatics tools. African Journal of Biotechnology. 2012;11(29):7378-7387.
24. Marcotte WR, Russell SH, Quatrano RS. Abscisic acid-responsive sequences from the *em* gene of wheat. Plant Cell. 1989;1:969-976.
25. Golds brough AP, Albrecht H, Stratford R. Salicylic acid-inducible binding of a tobacco nuclear protein to a 10bp sequence which is highly conserved among stress inducible genes. Plant J. 1993;3:563-571.
26. Lanahan MB, Ho THD, Rogers SW, Rogers JC. A gibberellins response complex in cereal a-amylase gene promoters. Plant Cell. 1992;4:203-211.
27. Rogers JC, Rogers SW. Definition and functional implications of gibberellins and abscisic acid *cís*-acting hormone response complexes. The Plant Cell. 1992;4:1443-1451.
28. Chen PW, Chiang CM, Tseng TH, Yu SM. Interaction between rice MYBGA and the gibberellins response element controls tissue-specific sugar sensitivity of a-amylase genes. The Plant Cell. 2006;18:2326-2340.
29. Persson EL, Brosche M, Strid A. Cultivar differences in the sequences of chalcone synthase promoters in pea. Unpublished. AC (NCBI), AF060237; 1998.
30. Waksman G, Lebrun M, Freyssinet G. Nucleotide sequence of a gene encoding sunflower ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit (rbcs). Nucleic Acids Res. 1987;15(17):7181-7181.
31. Werneke JM, Chatfield JM, Ogren WL. Alternative mRNA splicing generates the two ribulosebisphosphate carboxylase/oxygenase activase polypeptides in spinach and Arabidopsis. Plant Cell. 1989;1(8): 815-825.
32. Chakravarthy S, Tuori RP, D'Ascenzo MD, Fobert PR, Despres C,Martin GB. The tomato transcription factor Pti4 regulates defense-related gene expression via GCC box and non-GCC box *cis*- elements. Plant Cell. 2003;15:3033-3050.
33. Engelmann S, Zogel C, Koczor M, Schlue U, Streubel M, Westhoff P. Evolution of the C_4 phospho*enol*pyruvate carboxylase promoter of the C_4 species *Flaveria trinervia:* the role of the proximal promoter region. BMC Plant Biology. 2008;8(4):1-8.
34. Yanhui C, Xiaoyuan Y, Kun H, Meihua L, Jigang L, Zhaofeng G, Zhi-qiang L, Yunfei Z, Xiaoxiao W, Xiaoming Q, et al. The MYB transcription factor super family of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. Plant Mol. Biol. 2006;60:107-124.
35. Gubler F, Raventos D, Keys M, Watts R, Mundy J, Jacob-sen JV. Target genes and regulatory domains of the GAMYB transcriptional activator in cereal aleurone. Plant J. 1999;17:1-9.
36. Fang X, Han H, Stamatoyannopoulos G, Li Q. Developmentally specific role of the CCAAT box in regulation of human gamma-globin gene expression. J. Biol. Chem. 2004;279(7):5444-5449.
37. Nussinov R. The eukaryotic CCAAT and TATA boxes, DNA spacer flexibility and looping. J. Theor. Biol. 1992;155:243-270
38. Smale T, Kadonaga T. The RNA polymerase II core promoter. Ann Rev. Biochem. 2003;72:449-479.
39. Forde BG, Heyworth A, Pywell J, Kreis M. Nucleotide sequence of a B1 hordein gene and the identification of possible upstream regulatory elements in endosperm storage protein genes from barley, wheat and maize. Nucleic Acids Res. 1985;13:20.

40. Windho¨vel A, Hein I, Dabrowa R, Stockhaus J. Characterization of a novel class of plant homeodomain proteins that bind to the C_4 phospho*enol*pyruvate carboxylase gene of *Flaveria trinervia.* Plant Mol. Biol. 2001;45:201-214.
41. Yanagisawa S, Sheen j. Involvement of maize Dof zing finger proteins in tissues-specific and light- regulated gene expression. Plant Cell. 1998;10:75-89.
42. Vicente-Carbajosa J, Moose SP, Parsons RL, Schmidt RJ. A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. Proc. Natl. Acad. Sci. USA. 1997;94:7685-7690.
43. Walker JC, Howard EA, Dennis ES, Peacock WJ. DNA sequences required for anaerobic expression of the maize alcohol dehydrogenase-1 gene. Proceedings of the National Academy of Sciences of the USA. 1987;84:6624-6628.
44. Manjunath S, Sachs MM. Molecular characterization and promoter analysis of the maize cytosolic glyceraldehyde 3-phosphate dehydrogenase gene family and its expression during anoxia. Unpublished. AC (NCBI). U45858; 2005.
45. Engelmann S, Bla sing OE, Gowik U, Svensson P, Westhoff P. Molecular evolution of C_4 phospho*enol*pyruvate carboxylase in the genus *Flaveria*—A gradual increase from C_3 to C_4 characteristics. Planta. 2003;217:717-725.
46. Stockhaus J, Schlue U, Koczor M, Chitty JA, Taylor WC, Westhoff P. The promoter of the gene encoding the C_4 form of phospho*enol*pyruvate carboxylase directs mesophyll specific expression in transgenic C_4 *Flaveria spp.* Plant Cell. 1997;9:479-489.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=514&id=32&aid=4485