A segment containing a G-box and an ACGT motif confers differential expression characteristics and responses to the Arabidopsis Cytc-2 gene, encoding an isoform of cytochrome c

Elina Welchen, Ivana L. Viola, Hye Jin Kim, Luciana P. Prendes, Raúl N. Comelli, Jong Chan Hong and Daniel H. Gonzalez

J. Exp. Bot. 60:829-845, 2009. First published 19 Dec 2008;
doi:10.1093/jxb/ern331

Journal of Experimental Botany, Vol. 60, No. 3, pp. 829–845, 2009
doi:10.1093/jxb/ern331 Advance Access publication 19 December, 2008

Journal of Experimental Botany
www.jxb.oxfordjournals.org

RESEARCH PAPER

A segment containing a G-box and an ACGT motif confers differential expression characteristics and responses to the *Arabidopsis Cytc-2* gene, encoding an isoform of cytochrome *c*

Elina Welchen[1], Ivana L. Viola[1], Hye Jin Kim[2], Luciana P. Prendes[1], Raúl N. Comelli[1], Jong Chan Hong[2] and Daniel H. Gonzalez[1,*]

[1] Cátedra de Biología Celular y Molecular, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, CC 242 Paraje El Pozo, 3000 Santa Fe, Argentina
[2] Division of Applied Life Science (BK21 program), Environmental Biotechnology Research Center, Gyeongsang National University, 900 Gajwa-dong, Jinju, Gyeongnam 660-701, Republic of Korea

Received 29 September 2008; Revised 24 November 2008; Accepted 24 November 2008

Abstract

Sequences required for the expression of *Cytc-2* (At4g10040), one of two cytochrome *c* genes from *Arabidopsis thaliana*, were characterized using plants transformed with deleted and mutagenized promoter fragments fused to *gus*. These studies indicated that a region containing a G-box and an ACGT motif is essential for expression. Mutation of the ACGT motif causes a complete loss of expression, while mutation of the G-box causes decreased expression in aerial parts and abolishes expression in roots and induction by environmental factors. Upstream located site II elements are required for maximal expression, mainly in reproductive tissues, and maximal induction by different factors. One-hybrid screenings allowed the identification of transcription factors from the bZIP and bHLH families that interact mainly with the G-box. Four of these factors were able to bind to the *Cytc-2* promoter *in vitro* and in transactivation assays in *Arabidopsis*. Analysis of available microarray data indicated that the bZIP transcription factors share expression characteristics with the *Cytc-2* gene, suggesting that they act as mediators of its response to tissue-specific, environmental, and metabolic conditions. Site II elements interact with a TCP family protein and may co-ordinate the expression of the *Cytc-2* gene with that of other respiratory chain components. A model is proposed for the evolution of the *Cytc-2* gene through the incorporation of a segment containing a G-box and an ACGT motif into an ancestral gene that contained site II elements. This may have reduced the importance of site II elements for basal expression and conferred new responses to environmental factors.

Key words: *Arabidopsis thaliana*, cytochrome *c*, mitochondrion, promoter analysis.

Introduction

Mitochondrial biogenesis involves the expression of genes present in two different compartments. In plants, about 2% of mitochondrial components are encoded and synthesized within the organelle, while the rest are encoded in the nuclear genome (Mackenzie and McIntosh, 1999; Adams and Palmer, 2003). Based on this dual location, and on the fact that many mitochondrial proteins participate in related processes, it is generally assumed that the expression of the respective genes must be co-ordinated. This assumption is particularly valid for components of the oxidative phosphorylation machinery. Examples of co-ordination include preferential expression in anthers and roots (Huang *et al.*, 1994; Zabaleta *et al.*, 1998; Elorza *et al.*, 2004; Welchen *et al.*, 2004; Welchen and Gonzalez, 2005; Gonzalez *et al.*,

* To whom correspondence should be addressed: E-mail: dhgonza@fbcb.unl.edu.ar

2007) and induction by carbohydrates (Felitti and Gonzalez, 1998; Welchen *et al.*, 2002; Curi *et al.*, 2003), that operate for a majority of nuclear genes encoding components of the five respiratory complexes (Gonzalez *et al.*, 2007). In turn, expression of the mitochondrial genome does not seem to be co-ordinated with that of nuclear genes (Giegé *et al.*, 2005). Several authors have also described changes in the expression of nuclear genes encoding mitochondrial proteins brought about by signals originated in the organelle (Yu *et al.*, 2001; Sweetlove *et al.*, 2002; Lister *et al.*, 2004). This retrograde regulation indicates the existence of cross-talk between the mitochondrion and the nucleus to adjust the expression of specific genes according to the conditions prevailing in the organelle.

An additional source of complexity in the context of the expression of mitochondrial components is the presence of duplicated genes (Vision *et al.*, 2000; Prince and Pickett, 2002). Gene duplication is common in flowering plants and is usually followed by subfunctionalization or neofunctionalization, most frequently related to the acquisition of new expression characteristics (Prince and Pickett, 2002; Haberer *et al.*, 2004; Duarte *et al.*, 2006). In the case of respiratory chain components, a clear example of specialization is given by the *Arabidopsis SDH2-3* gene, encoding an isoform of the iron–sulphur subunit of succinate dehydrogenase (Complex II). This gene is expressed specifically during seed maturation, at late stages of embryogenesis, while the other two genes encoding the same subunit are expressed in vegetative tissues (Elorza *et al.*, 2004, 2006). In another example, both *Arabidopsis* genes encoding cytochrome *c* (*Cytc-1* and *Cytc-2*) show only partial divergence in expression patterns, with overlapping mainly in anthers, and similar responses to carbohydrates (Welchen and Gonzalez, 2005). Studies on the promoter region of the *Cytc-1* gene have shown that its expression is governed by site II elements (TGGGCC/T) and a *telo* box (AAACCCTA). A survey of similar motifs in other nuclear genes encoding respiratory chain components revealed that many of them, including the *Cytc-2* gene, contain site II elements in their upstream proximal regions, suggesting that proteins that interact with them may participate in the co-ordination of the expression of these genes (Welchen and Gonzalez, 2006). An intriguing question is, however, what is the basis for the different expression patterns displayed by the *Cytc-1* and *Cytc-2* promoters.

In the present work, an extensive analysis of the *Cytc-2* promoter region is described. It is demonstrated that the main expression characteristics of the *Cytc-2* gene are dependent on a segment that contains a G-box and an ACGT motif recognized by a group of basic leucine zipper (bZIP) and bHLH proteins, while the role of site II elements is less important, acting to enhance the expression levels mainly in reproductive organs and to regulate the magnitude of the response to environmental and metabolic cues. The presence of a similar arrangement in promoters of other respiratory chain component genes indicates that the incorporation of the relevant elements may have occurred in one step and suggests a model for the evolution of functional diversity in the promoter regions of genes encoding mitochondrial proteins. In addition, incorporation of the G-box seems to have conferred to the *Cytc-2* gene the capacity to respond to several environmental and metabolic factors through the interaction with different sets of transcription factors.

Materials and methods

Plant material and growth conditions

Arabidopsis thaliana Heyhn. ecotype Columbia (Col-0) was purchased from Lehle Seeds (Tucson, AZ). Plants were grown in soil in a growth chamber at 22–24 °C under long-day photoperiods (16 h of illumination by a mixture of cool-white and GroLux fluorescent lamps) at an intensity of approximately 200 μE m^{-2} s^{-1}. Plants used for the different treatments were grown in Petri dishes containing Murashige and Skoog (MS) medium, 0.8% agar, and different additions, as indicated in the respective figure legends. For UV treatment, 2-week-old plants were exposed for 1 h to a lamp (Phillips TUV30W/G30T8) that emits light in the UV-B spectrum (280–320 nm). Plants were processed 2 h after treatment. To evaluate the effects of extreme temperatures, 2-week-old plants were either incubated at 4 °C for 24 h in a growth chamber or incubated at 38.5 °C for 3 h. GUS activities of plants incubated at high temperature were measured 1 h after treatment.

Reporter gene construct and plant transformation

A fragment spanning nucleotides –967 to +54 with respect to the ATG initiation codon of the *Cytc-2* gene (At4g10040) was obtained by PCR amplification of *Arabidopsis* genomic DNA using primers CATP22 and CATP26 (see Supplementary Table S1 at *JXB* online). Upstream deletions were constructed in a similar way with primers CATP22 and either CATP23, CATP24, CATP29, CATP210, CATP25, CATP211 or CATP212 (see Supplemenary Table S1 at *JXB* online). The resulting fragments, containing *Hin*dIII and *Bgl*II sites, were cloned in-frame in the *Hin*dIII and *Bam*HI sites of pBI101.3. Complementary primers were used to generate scanning mutations along the –188 to –139 promoter region and to mutate specific putative regulatory elements. These primers were used with CATP22 and CATP29 or CATP210 to amplify hybridizing downstream and upstream sequences. The resulting products were mixed in buffer containing 50 mM TRIS-HCl (pH 7.2), 10 mM $MgSO_4$, and 0.1 mM DTT, incubated at 95 °C for 5 min, and annealed by allowing the solution to cool to 24 °C in approximately 1 h. After this, 0.5 mM of each dNTP and 5 units of the Klenow fragment of *E. coli* DNA polymerase I were added, and incubation was followed for 1 h at 37 °C. A portion of this reaction was directly used to amplify the chimeric fragments using primers CATP22 and either CATP29 or CATP210. The sequences of the oligonucleotides used are shown in Supplementary Table S1 at *JXB* online. All constructs were checked by DNA sequencing.

The respective constructs were introduced into *Agrobacterium tumefaciens* strain LB4404, and transformed bacteria were used to obtain transgenic *Arabidopsis* plants by the floral dip procedure (Clough and Bent, 1998). Transformed plants were selected on the basis of kanamycin resistance and positive PCR carried out on genomic DNA with gene-specific primers and the *gus* primer 5'-TTGGGGTTTCTACAGGAC-3′. Thirty independent lines for each construct were further reproduced and homozygous T_3 and T_4 plants were used to analyse *gus* expression. Plants transformed with pBI101.3 were obtained in a similar way.

Agroinfiltration of Arabidopsis thaliana *leaves*

The protocol used for transient *Arabidopsis* plant transformation was modified from http://www.plantsci.cam.ac.uk/Baulcombe/protocols.html. The *Agrobacterium* strains used were LBA4404 and the anti-silencing strain 19 K (Voinnet *et al.*, 2003). *Agrobacteria* from fresh plates were grown at 28 °C in LB medium supplemented with 50 mg ml^{-1} kanamycin and 5 mg ml^{-1} tetracycline to stationary phase. Bacteria were sedimented by centrifugation at 5000 *g* for 15 min at room temperature and resuspended in 10 mM $MgCl_2$ and 150 mg ml^{-1} acetosyringone and the cells were left in this medium for 3 h at room temperature. Individual *Agrobacterium* cultures carrying the *Cytc-2*::*gus* or the 35S::transcription factor constructs were mixed together with those carrying the 35S::p19 construct and co-infiltrated into the abaxial air spaces from 4-week-old wild-type (control) or transgenic *Arabidopsis* leaves. After infiltration, *Arabidopsis* plants were grown under normal conditions. Four days post-infiltration, leaf discs were harvested and total soluble protein extracts were prepared and used for GUS activity measurement.

Constructs expressing transcription factors under the control of the 35S CaMV promoter were obtained by cloning the complete coding region of the respective cDNAs (amplified using primers described in Supplementary Table S1 at *JXB* online) in the binary vector pBI121.

β-*glucuronidase assays*

β-glucuronidase (GUS) activity of transgenic plants was analysed by histochemical staining using the chromogenic substrate 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid (X-gluc) as described by Hull and Devic (1995). Whole plants or separated organs were immersed in a 1 mM X-gluc solution in 100 mM sodium phosphate, pH 7.0, and 0.1% Triton X-100 and, after applying vacuum for 5 min, they were incubated at 37 °C until satisfactory staining was observed (see figure legends for details). For lines transformed with constructs that produce no staining, incubation was prolonged for 24 h, together with the respective negative (plants transformed with pBI101.3 and non-transformed plants) and positive controls (a reference positive line according to the deletion or mutation analysed). Tissues were cleared by immersing them in 70% ethanol.

Specific GUS activity in protein extracts was measured using the fluorogenic substrate 4-methylumbelliferyl β-D-glucuronide (MUG) essentially as described by Jefferson *et al.* (1987). Total protein extracts were prepared by grinding the tissues in extraction buffer (50 mM sodium phosphate, pH 7.0, 10 mM EDTA, 10 mM β-mercaptoethanol) containing 0.1% (w/v) SDS and 1% Triton X-100, followed by centrifugation at 13 000 *g* for 10 min. GUS activity in supernatants was measured in extraction buffer containing 1 mM MUG and 20% methanol. Reactions were stopped at different times with 0.2 M Na_2CO_3 and the amount of 4-methylumbelliferone was calculated by relating relative fluorescence units with those of a standard of known concentration. Controls were made to ensure that activity measurements were within the linear range as a function of time and protein concentration. The protein concentration of extracts was determined as described by Sedmak and Grossberg (1977).

DNA binding assays

The full-length coding sequence of bHLH080 was amplified from the respective clone in pDEST22 and inserted in frame into the *Bam*HI and *Eco*RI sites of the expression vector pGEX-3X (Smith and Johnson, 1988). Amplifications were performed using oligonucleotides HLH80-1 and HLH80-2 (see Supplementary Table S1 at *JXB* online). In a similar way, fragments encoding AREB2/ABF4 (amino acids 331 to 410), GBF3 (amino acids 201 to 300) and the GBF-like protein (amino acids 276 to 389) were amplified with specific primers (see Supplementary Table S1 at *JXB* online) and cloned in the same vector. All constructions were checked by DNA sequence analysis. For expression, *E. coli* cells bearing the corresponding plasmids were grown and induced as described previously (Palena *et al.*, 1998). Purifications of recombinant proteins were performed as indicated by Smith and Johnson (1988), with modifications described by Palena *et al.* (1998). Nuclear extracts were prepared from cauliflower buds (obtained from a local market) as described by Maliga *et al.* (1995).

For EMSAs, aliquots of extracts (10 μg) or purified recombinant proteins (100 ng) were incubated with double stranded DNA (10 000 cpm) obtained by amplification of the corresponding fragments with primers CATP213 and CATP214, followed by restriction enzyme cleavage and labeling with [α-^{32}P]dATP by filling-in the 3′-ends using the Klenow fragment of DNA polymerase. Binding reactions (20 μl) contained, in addition to labelled DNA, 20 mM HEPES (pH 7.5), 50 mM KCl, 2 mM $MgCl_2$, 0.5 mM EDTA, 1.0 mM dithiothreitol, 0.5% Triton X-100, 10% glycerol, and 1.5 μg poly(dI-dC). Reactions were incubated for 20 min on ice, supplemented with 2.5% Ficoll and immediately loaded onto a running gel (5% acrylamide, 0.08% *bis*-acrylamide in 0.5× TBE plus 2.5% glycerol; 1× TBE is 90 mM TRIS-borate, pH 8.3, 2 mM EDTA). The gel was run in 0.5× TBE at 30 mA and 4 °C for 1.5 h and dried prior to autoradiography.

One-hybrid analysis in yeast

To obtain a yeast strain carrying the *Cytc-2* promoter region comprising nucleotides –189 to –139 in front of

the *HIS3* reporter gene, three tandem copies of the fragment were cloned in the *Bam*HI-*Eco*RI sites of vector pHIS3NX (Meijer *et al.*, 1998) that contains the *HIS3* coding region preceded by its own minimal promoter. The entire cassette was then transferred to the *Not*I-*Xba*I sites of the integrative yeast plasmid pINT1, which confers resistance to the antibiotic G418. An *Nco*I-*Sac*I fragment of the clone in pINT1 was introduced into the *PDC6* locus of the yeast strain Y187 as described by Meijer *et al.* (1998). The presence of the fragment of interest in the genome of cells resistant to G418 was analysed by PCR with specific oligonucleotides. Alternatively, the *Cytc-2* promoter fragment was placed in front of the *lacZ* reporter gene contained in plasmid pLacZi (Clontech). Plasmid linearized in its *Nco*I site was introduced into the *URA3* locus of the yeast aW303 strain.

Proteins interacting with the *Cytc-2* promoter were identified using a DNA library carrying a 1050 *Arabidopsis* transcription factor (TF) ORFeome collection in the prey vector pDEST22 (Invitrogen) (HJ Kim *et al.*, unpublished data). The library contains 67 and 101 members of the bZip and bHLH families, respectively. To construct the *Arabidopsis* TF library, the predicted coding region of each TF was PCR amplified using RNA from various *Arabidopsis* tissues with gene-specific primer sets carrying 12 nucleotide *att*B1 (5′-AAAAAGCAGGCTCA-3′; 5′-specific primer) or 12 nucleotide *att*B2 (5′-AGAAAGCTGGGTA-3′; 3′-specific primer) sites in the gene-specific primer sequence. The amplified fragments were cloned into the entry vector pDONOR 201 using BP clonase (Invitrogen). The identity of each TF was confirmed by DNA sequencing. The plasmids carrying TF coding regions in the entry vector were pooled and mixed in equal ratios and the resulting TF ORFeome library was transferred to pDEST22 using LR clonase (Invitrogen) to make the yeast two-hybrid prey TF library. Plasmid DNA from the library (10 μg) was introduced into yeast using the standard lithium acetate transformation method (Gietz *et al.*, 1992). A total of 7.5×10^5 transformants were plated on SC-His-Trp medium containing 0.1 mM 3-AT and the resulting putative positive clones from the screen were streaked on fresh SC-His-Trp +0.1 mM 3-AT medium to purify colonies. The plasmid DNAs containing TF ORFs were rescued and retransformed into yeast carrying the *lacZ* reporter for further confirmation.

Analysis of publicly available microarray data

Expression data for the *Cytc-2* and transcription factor genes in different tissues or after different treatments were obtained from the BAR Expression Browser (http://www.bar.utoronto.ca/; Toufighi *et al.*, 2005). To analyse time-dependent expression changes after treatments, raw microarray data were obtained from the NASC microarray database (http://affymetrix.arabidopsis.info/narrays/experimentbrowse.pl) and mean expression values (±SD) for the different genes at each data point were calculated.

Results

Deletion and mutagenic analysis of the Cytc-2 *promoter*

It has previously been shown that a 1 kbp region of the *Cytc-2* gene located upstream of the translation start site directs expression of a *gus* reporter gene in vascular tissues of roots, hypocotyls, cotyledons, leaves, and petals, in developing anthers and pollen, and at the junction of flowers and siliques with pedicels (Welchen and Gonzalez, 2005). To characterize the minimal promoter regions involved in regulating expression of the *Cytc-2* gene, progressive deletions from the upstream portion of the 1 kbp fragment (Fig. 1A) have been performed and the activity of the respective constructs have been analysed using plant–*Agrobacterium*-mediated transformation followed by analysis of GUS expression in several independent lines (Fig. 1B, C). Fluorometric assays using total protein extracts from 2-week-old seedlings showed that deletion of fragments from –967 to –605 and –315 to –259 produced a partial, but significant, reduction in expression levels, suggesting that positive regulatory elements are present in these regions (Fig. 1B). In addition, deletion of a fragment located between –189 and –139 completely abolished activity, producing values similar to those observed with the promoterless *gus* gene (Fig. 1B). Histochemical assays were in agreement with this observation, since no staining was observed in plants transformed with the –139 construct (Fig. 1C, panel c), while plants containing larger promoter fragments displayed GUS activity mainly in cotyledons (Fig. 1C, panels a, b, h–j). In addition, for the –189 construct, a net reduction of GUS staining could be observed (Fig. 1C, panel j).

The effect of promoter deletions on expression in different organs of adult plants was also analysed. As observed in seedlings, deletion of nucleotides down to –139 completely abolished GUS activity in all organs (siliques are shown as an example in Fig. 1C, panels o, p), suggesting that elements essential for expression are located in this region.

To identify discrete elements responsible for *Cytc-2* gene expression, the region from –189 to –139 was analysed further using a set of five 10 bp scanning mutations arbitrarily named mut1 to mut9 (odd numbers, upstream to downstream; see Fig. 1A for details). Histochemical and fluorometric analysis of GUS expression in plants transformed with promoter fragments comprising nucleotides –259 to +54 mutagenized as described indicated that sequences located in the two downstream 10 bp segments (–158 to –139) are essential for expression of the *Cytc-2* gene, since no activity was detected with constructs mut7 and mut9 (Fig. 1B, C, panels g, l, n, r). Mutation of an additional fragment located immediately upstream of –158 (mut5) produced a noticeable decrease in the level of GUS activity in 2-week-old *Arabidopsis* plants (Fig. 1B, C, panel f).

Informatic analysis of the mutagenized regions indicated the presence of a G-box motif present around –169, partially comprised the fragment that originated a decrease in expression when modified (mut5). Since G-box motifs are

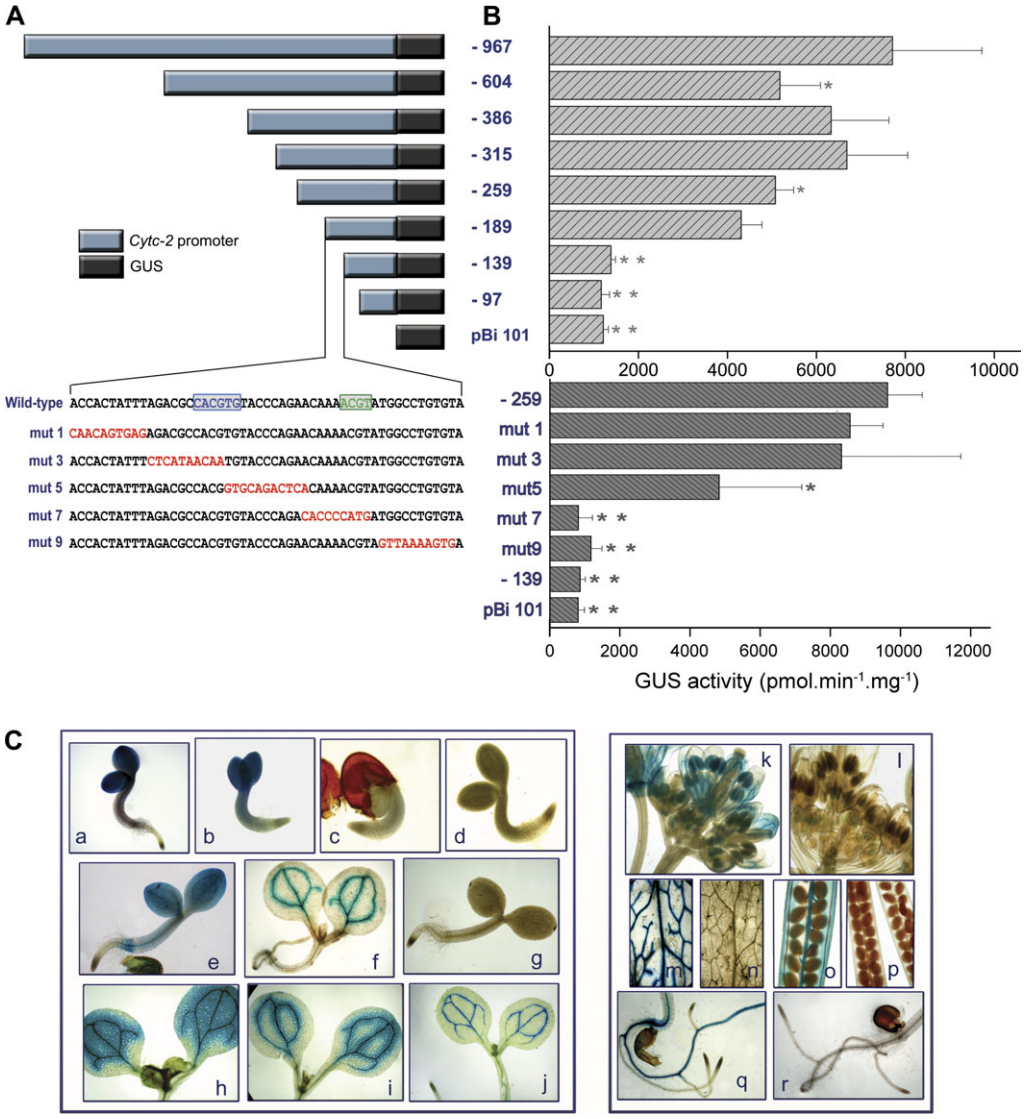

Fig. 1. Analysis of GUS activity driven by truncated and mutagenized forms of the *Cytc-2* promoter. (A) Schematic representation of different constructs used to transform *Arabidopsis* plants. The numbers indicate the upstream end of the promoter fragment present in each construct with respect to the translation start site; the downstream end was at +54 for all constructs. The sequences of progressive scanning mutations between positions –189 and –139 are shown in red. Blue and green letters indicate the positions of the G-box and ACGT motif, respectively. (B) GUS expression levels measured in total protein extracts from 2-week-old *Arabidopsis* plants transformed with constructs carrying progressive deletions or mutagenized forms of the *Cytc-2* promoter. The significance of changes produced after each deletion was assessed using Student's *t* tests (*P <0.05, **P <0.01). Plants transformed with the promoterless *gus* gene (pBI101) were also used. The results indicate the mean (±SD) of five independent lines. (C) Histochemical localization of GUS activity in *Arabidopsis* plants transformed with wild-type or mutagenized promoter forms fused to the *gus* reporter gene: 1-d-old seedlings (a: –967; b: –259; c: –139; d: mut7); 3–7-d-old seedlings (e: mut1; f: mut5; g: mut9; h: –386; i: –315; j: –189); flowers (k: –967; l: mut7); mature leaf sections (m: –259; n: mut7); siliques (o: –259; p: –139); primary and secondary roots (q: –315; r: mut7). The images are representative of 30 lines analysed for each deletion or mutagenized promoter construct. Incubation time in the staining solution was 4 h for plants of all stages and organs (except roots, that were incubated for 12 h).

known to modulate the expression of several genes (Block *et al.*, 1990; Donald and Cashmore, 1990; Salinas *et al.*, 1992; Kim *et al.*, 1992; Menkens *et al.*, 1995; Ishige *et al.*, 1999; Hudson and Quail, 2003), a 3 bp mutation (CACGTG to CAATGG) was introduced in the central part of this motif (Fig. 2A). As shown in Fig. 2, mutation of the *Cytc-2* G-box produced a significant decrease in expression in seedlings (Fig. 2B, D, panels c, d, g) and organs from adult plants (Fig. 2C, D, panels k, o). Particularly in roots, mutation of the G-box completely abolished activity, indicating that it is essential for expression in this organ.

Regarding the effect of mutations in the region located downstream of the G-box (mut7), a search for known motifs present in this region using the PLACE database (Higo *et al.*, 1999) yielded the sequence ACGT (–153 to –150), containing

the central core of a G-box. Mutation of the ACGT motif completely abolished expression in seedlings and originated very low expression levels in adult plants (Fig. 2B, C, D, panels d, h, l, p). This residual activity was completely absent in a double-mutant of the G-box and ACGT motifs (Fig. 2B, C). The ACGT motif present around –152 has then a very important role in the expression of the *Cytc-2* gene in all organs and developmental stages.

A mutagenic analysis was also performed to investigate the role of the site II motifs present in the *Cytc-2* promoter, since similar elements were described as essential for expression of the *Cytc-1* gene and other genes involved in respiration (Welchen and Gonzalez, 2005, 2006). Two copies of the site II element are located within the *Cytc-2* promoter region from –315 to –259 (Fig. 2A) whose deletion produces a moderate decrease in expression. Mutation of site II elements within the context of the –315 promoter region produced a significant decrease in expression in seedlings and organs from adult plants, except for roots, where no changes were observed (Fig. 2B, C, D, panels b, f, j, n). The decrease in expression was similar to or even greater than the one observed upon deletion of the fragment from –315 to –259 (Fig. 2B, C), indicating that site II elements are the main regulatory motifs present in this region. Careful examination of expression values in different organs suggests that site II elements act as regulators of the magnitude of expression mainly in reproductive organs (Fig. 2C). According to these results, site II elements seem to have a less important role in expression of the *Cytc-2* gene as compared with *Cytc-1*, where mutation of site II elements completely abolishes expression.

The Cytc-2 *promoter responds to several factors*

It was previously reported that incubation of plants with metabolizable carbohydrates produces an increase in *Cytc-2* transcript levels (Welchen *et al.*, 2002). Accordingly, reporter gene expression levels in plants transformed with *Cytc-2* promoter/*gus* fusions were analysed after incubation in the presence of sucrose. As shown in Fig. 3A, seedlings of plants transformed with the –315 construct grown in MS medium with 200 mM sucrose for 1 week have GUS activity levels that are three times higher than those grown in its absence or in the presence of a similar concentration of mannitol as a control of the effect of osmotic potential. Similar results were obtained with the largest (–967) construct in seedlings and in isolated leaves incubated in solutions containing sucrose (not shown). Analysis of the response of plants carrying further deletions or mutagenized fragments indicated that mutation of site II elements originates a significantly lower response to sucrose that is completely lost upon mutation of the G-box (Figs 2E, 3A). Growth of plants in the presence of the cytokinin 6-benzylaminopurine (BAP) also produced *c.* 4-fold induction of *Cytc-2* promoter-dependent expression (Figs 2E, 3B). As observed in the case of sucrose, the effect of cytokinin was considerably reduced upon mutation of site II elements or deletion of the fragment that contains these elements and abolished after mutation of the G-box (Fig. 3B). Site II elements within the context of the *Cytc-2* promoter participate in the induction by sucrose and cytokinin and are required for a maximal response. The G-box seems to play an additional role in establishing basal induction levels. It is interesting to note that the *Cytc-1* promoter, that contains site II elements, is also induced by sucrose and cytokinin (Welchen and Gonzalez, 2005).

It has also been observed that GUS activity driven by the *Cytc-2* promoter is lower in seedlings grown in darkness for 1 week with respect to controls grown in the light (Figs 2E, 3C). This effect is reversed after several hours of illumination of dark-grown seedlings (Fig. 3C). Incubation of light-grown seedlings for more than 24 h in the dark also produced a decrease in expression levels (Fig. 2E). The effect of light conditions on expression was conserved after elimination or mutation of site II elements, but was completely abolished by mutation of the G-box (Fig. 3C), indicating that this motif is responsible for the observed behaviour.

Analysis of data from public microarray experiments indicates that the *Cytc-2* gene is induced under several stress conditions, among them cold, heat, UV-B and pathogens. The response of constructs containing different fragments of the *Cytc-2* promoter to some of these conditions was then examined. Figure 3D shows that GUS expression under the control of the *Cytc-2* promoter is induced after incubation of plants at either 4 °C or 38.5 °C and by treatment with UV light. Similar responses were observed for most of the constructs analysed, except for those carrying mutations in the G-box or the ACGT motif. This suggests that the G-box located at –169 participates in the induction by these treatments. Regarding the ACGT motif, the lack of induction may be caused by the fact that mutation of this element almost completely abolishes expression of the gene. GUS expression data indicate, then, that the G-box present in the *Cytc-2* promoter is responsible for the response of this gene to several environmental and metabolic factors, while site II elements participate in maximal induction by most of these factors.

Nuclear proteins bind to the Cytc-2 *promoter region required for expression*

The presence of proteins able to recognize the *Cytc-2* promoter regions required for expression was analysed using nuclear extracts prepared from cauliflower inflorescences, as a source of proteins from a species closely related to *Arabidopsis*. Electrophoretic mobility shift assays (EMSAs) using a labelled fragment comprising nucleotides –189 to –139 showed the presence of two retarded bands, indicating the existence of complexes with different mobility (Fig. 4). The band with lowest mobility represents a complex formed by protein(s) bound to the G-box, as indicated by the fact that this band disappeared when fragments with a mutagenized G-box were used (Fig. 4A, lanes 4 and 8). Mutations in either the ACGT motif or the 10 bp segment from –158 to –149 (mut7, that comprises the ACGT motif) did not produce any

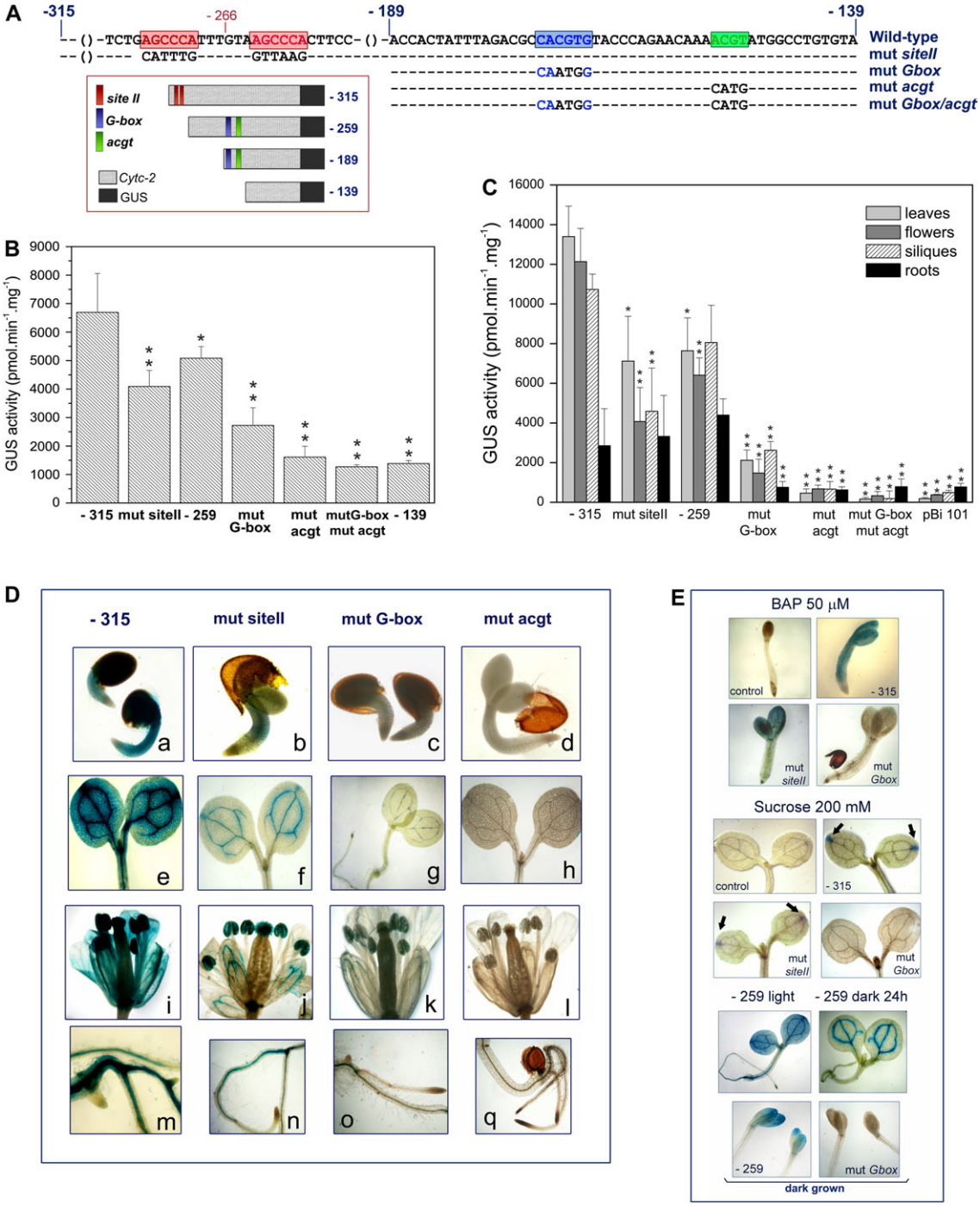

Fig. 2. Analysis of reporter gene expression in *Arabidopsis* plants transformed with mutagenized promoters fused to the *gus* gene. (A) Schematic representation and sequence of elements that were mutagenized within the context of the –315 and –259 *Cytc-2* promoter (see Fig. 1). On the right, the name used for each construct as cited in the Figures and the text is shown (red: site II element, blue: G-box, green: ACGT motif). (B) GUS expression levels measured in total protein extracts from 2-week-old plants. Numbers below the bars indicate the respective wild-type or mutagenized fragment analysed by the fluorometric assay. (C) Specific GUS activity in protein extracts from different organs of plants transformed with either wild-type or mutagenized promoter/*gus* fusions. Activity was also measured in extracts from plants transformed with a promoterless *gus* gene (pBI101). In (B) and (C), bars indicate the mean activity values (±SD) obtained with organs from five independent transformants for each construct. (D) Histochemical analysis of GUS activity. One- and 2-d-old seedlings with either wild-type (a: –315) or mutagenized forms of the *Cytc-2* promoter (b: mut sitell, c: mut G-box, d: mut acgt). Five-day-old plants (e: –315, f: mut sitell, g: mut G-box, h: mut acgt) and flowers (i: –315, j: mut sitell, k: mut G-box, l: mut acgt). Primary and secondary roots from mature plants (m: –315, n: mut sitell, o: mut G-box, p: mut acgt). Incubation time in the staining solution was 4 h for plants of all stages and organs (except roots, that were incubated for 12 h). The images are representative of 30 different lines from each construct. (E) Histochemical analysis of transgenic 7-d-old plants that were grown in the presence of 50 µM

noticeable effect (Fig. 4A, lanes 6 and 10). Interestingly, mutation of the 10 bp segment located immediately downstream (mut9, –148 to –139), that originated a strong decrease in expression of the *gus* reporter gene, also produced a significant decrease in the amount of the complex with lowest mobility (Fig. 4A, lane 12). This result indicates that either the integrity of this region is required for efficient binding of proteins to the G-box motif, or that the binding of a factor to this region is dependent on the integrity of the G-box, suggesting the existence of co-operative interactions between these two regions. The low mobility complex reflects, then, the presence of nuclear proteins that bind to regions of the *Cytc-2* promoter required for expression. However, there is not a complete correlation between the complexes observed in EMSAs and the expression characteristics of the different constructs. For example, mutations in the G-box completely abolish binding and only partially affect gene expression, while the inverse effect is observed upon mutation of the 3′ element (mut9). Most notably, mutation of the essential ACGT motif does not produce significant changes in the formation of protein–DNA complexes.

The interaction of nuclear proteins with the *Cytc-2* promoter was also studied using competition for the binding to the non-mutagenized labelled fragment by unlabelled DNA with different mutations (Fig. 4B). A 10-fold molar excess of the non-mutagenized fragment competed efficiently with the formation of the low mobility complex. This competition was not observed with fragments that contain a mutagenized G-box or mutations in the 3′ region (mut9), in agreement with the results described above. In turn, the non-mutagenized fragment competed only slightly with the formation of the less retarded band (Fig. 4B). This suggests that this band is largely originated

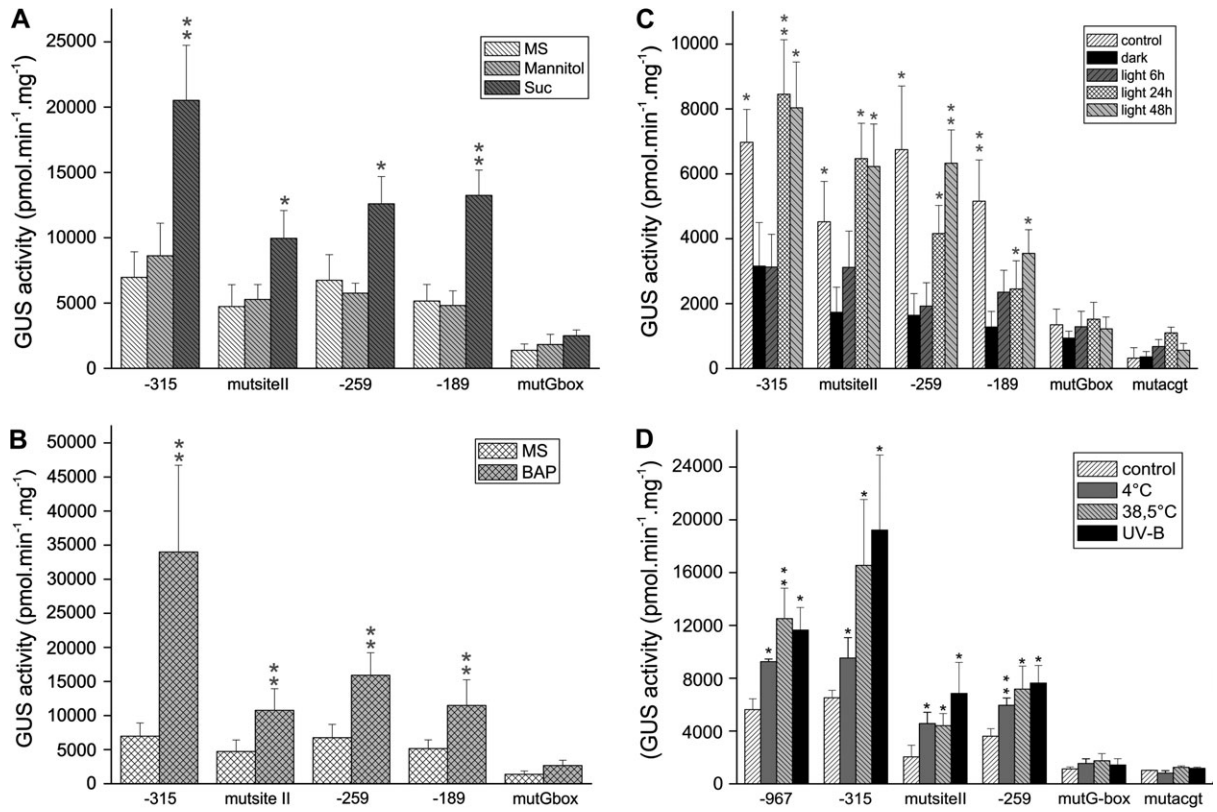

Fig. 3. Regulation of *gus* reporter gene expression by the *Cytc-2* promoter in the presence of sucrose, cytokinin (BAP), and by changes in environmental conditions. GUS activity was measured using the fluorogenic substrate MUG and protein extracts prepared from plants carrying different wild-type or mutagenized *Cytc-2* promoter fragments (see Fig. 2A for details). The results indicate the mean (±SD) values of five independent lines measured in three different experiments. (A, B) Plants were grown for 7 d in MS medium alone or supplemented with either 200 mM mannitol or 200 mM sucrose (A) or 50 μM BAP (B). (C) Plants were grown for 7 d under a normal photoperiod (16 h light/8 h dark; control) or in the dark. In addition, dark-grown seedlings were then exposed to continuous light for different times (6 h, 24 h, 48 h). (D) After growth in MS medium under normal photoperiod (16 h light/8 h dark) and temperature (21–24 °C; control) for 2 weeks, plants were incubated either at 4 °C for 24 h, at 38.5 °C for 3 h, or under UV-B light for 1 h.

BAP, 200 mM sucrose or under different illumination conditions. Incubation time in the staining solution was 1 h for plants incubated with sucrose or BAP (and the respective controls), and 4 h and 8 h for light-grown and dark-grown seedlings, respectively, when the effect of illumination was analysed. See Fig. 3 and discussion in the text for details.

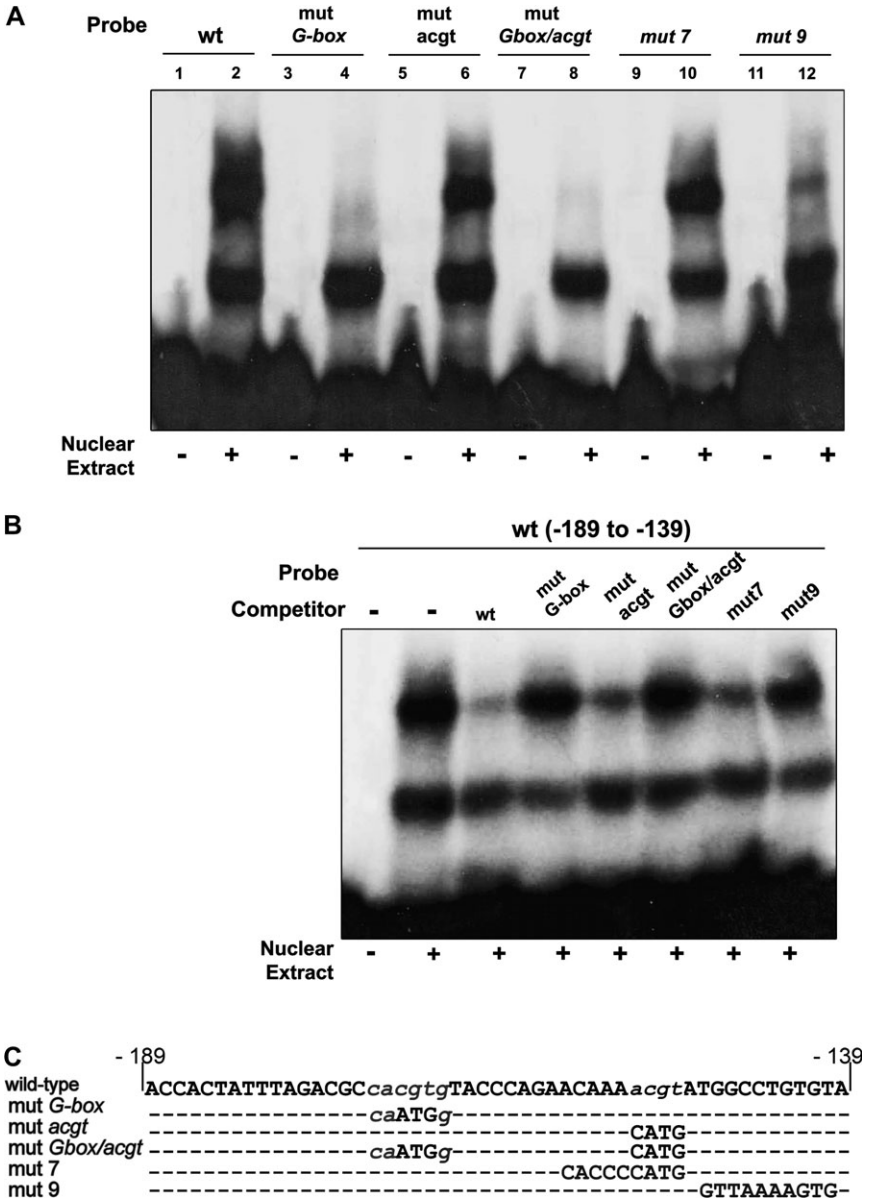

Fig. 4. Nuclear proteins from cauliflower buds specifically bind to the *Cytc-2* promoter. (A) Nuclear extracts (10 μg) from cauliflower inflorescences were analysed by an EMSA for the presence of proteins that bind to labelled DNA spanning nucleotides –189 to –139. The binding was analysed using either a labelled wild-type fragment or different mutagenized forms. (B) The binding of nuclear proteins to a labelled –189 to –139 wild-type fragment was analysed with or without the addition of a 50-fold molar excess of unlabelled wild-type or mutated forms of the same fragment as competitor. (C) Details of the wild-type and mutagenized sequences used in the EMSA experiments.

by unspecific binding of abundant proteins. However, this slight competition was abolished when the sequence ACGT was mutated, probably indicating that a low abundance complex formed by proteins that bind to the ACGT motif and a complex established by unspecific interactions with DNA co-migrate in this region. It was then speculated that at least two different complexes of nuclear proteins interact with the region of the *Cytc-2* promoter required for expression. One of the complexes requires the integrity of the G-box motif and sequences located in the downstream portion of the fragment, while the other one is formed with the region containing the ACGT motif.

Identification of transcription factors that interact with the Cytc-2 *promoter*

A yeast one-hybrid screen was performed to identify *Arabidopsis* transcription factors that interact with the *Cytc-2* promoter region relevant for expression. For this purpose, three tandem copies of the segment comprising nucleotides –189 to –139 were cloned in front of the *HIS3* reporter gene containing a minimal promoter and introduced into the yeast genome. The resulting strain was transformed with a full-length cDNA library of *Arabidopsis* transcription factors fused to the GAL4 activation domain

(HJ Kim *et al.*, unpublished). Twenty-four colonies able to grow in the absence of histidine were identified and analysed further by retransformation of yeast and DNA sequencing. Twenty-three of the recovered clones, representing six different transcription factors from the bZIP and bHLH families (Fig. 5A), were able to transactivate the *HIS3* gene when reintroduced into the yeast strain used for the screening.

The efficiency of individual clones to recover growth in the absence of histidine was analysed using serial dilutions of the corresponding yeast cultures. Figure 5B shows that yeasts containing clones encoding each of the transcription factors were able to grow in the absence of histidine at higher dilutions than the control expressing only the GAL4 activation domain. Clones encoding AREB2/ABF4, GBF1, and the GBF-like protein were the most effective since the respective yeast strains were able to grow at higher dilutions, suggesting that they produce a more efficient activation of the reporter gene. For GBF3 and bHLH080, slower growth in the absence of histidine was observed.

A quantitative measurement of activation was obtained by introducing the different clones expressing the GAL4 activation domain fusions into a yeast strain that contained the *Cytc-2* promoter fragment in front of the *lacZ* reporter gene (Fig. 5C). β-galactosidase activity measurements indicated that AREB2/ABF4 produces *c.* 10-fold activation respective to the control strain, while GBF1 and the GBF-like protein produce 3–4-fold activation. For GBF3 and bHLH080, the increase in activity was not enough to be significant with respect to the control strain with this system. The results obtained with both reporter genes are in agreement, showing that AREB2/ABF4, GBF1, and the GBF-like protein are more efficient transactivators in the yeast one-hybrid assay, while GBF3 and bHLH080 display rather poor activation. Since these two proteins are able to confer better growth capacity in the absence of histidine, it was decided to analyse their role as candidate activators of the *Cytc-2* promoter in other systems further.

Binding of some of the proteins identified in the one-hybrid screening to the *Cytc-2* promoter *in vitro* was also analysed. For this purpose, regions comprising the bZIP domains of AREB2/ABF4, GBF3, and the GBF-like protein, as well as the entire bHLH080, were expressed as recombinant proteins in *E. coli*. EMSAs of binding of the purified proteins to wild-type and mutated variants of the *Cytc-2* promoter fragment spanning nucleotides –189 to –139 indicated that the four proteins recognize the G-box, since binding was abolished or considerably reduced upon mutation of this element (Fig. 6A, lanes 2 and 4). With GBF3 and bHLH080, overexposure of the gel indicated the presence of a small amount of binding to the fragment with a mutated G-box (Fig. 6A, lane 2). The fact that this binding disappeared when the ACGT motif was also mutated (Fig. 6A, lane 4), may indicate that these proteins also recognize the ACGT motif, although much less efficiently.

It has previously been proposed that site II elements present in other promoters, including the *Cytc-1* promoter, are recognized by proteins from the TCP family (Kosugi

A

AGI code	Name	Family	Reference
At1g35460	bHLH80	bHLH	Heim et al. (2003) Toledo-Ortiz et al. (2003)
At4g09180	bHLH81	bHLH	Heim et al. (2003) Toledo-Ortiz et al. (2003)
At4g36730	GBF1	bZip	Schindler et al. (1992)
At2g46270	GBF3	bZip	Schindler et al. (1992)
At1g32150	GBF-like	bZip	---
At3g19290	AREB2/ABF4	bZip	Uno et al. (2000) Kang et al. (2002)

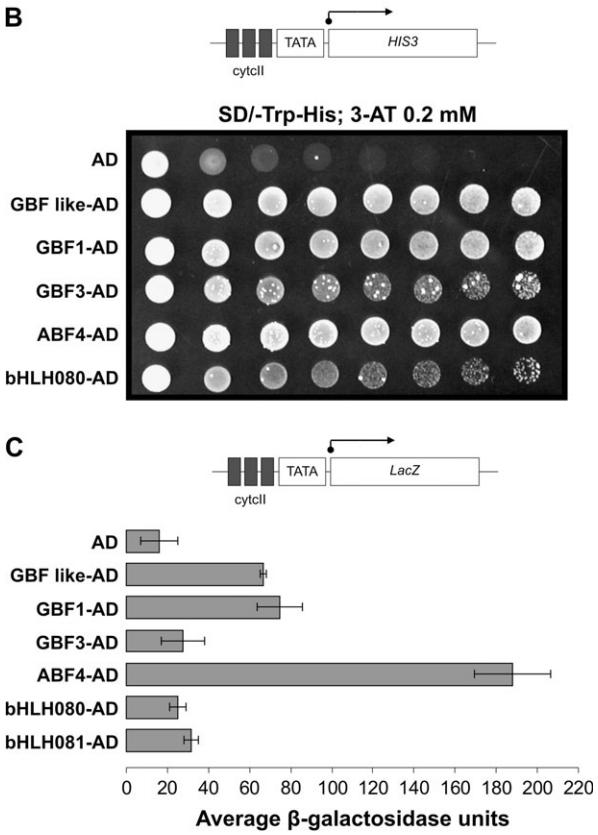

Fig. 5. Isolation and identification of *Arabidopsis* transcription factors that interact with the *Cytc-2* promoter. (A) List of the *Arabidopsis* transcription factors identified during the one-hybrid screening. (B) Activation of the *HIS3* gene by the selected clones. A construct comprising three tandem copies of the –189 to –139 *Cytc-2* promoter fragment fused with a minimal promoter and the *HIS3* reporter gene was introduced into yeast cells. The identified clones, expressing different transcription factors fused to the GAL4 activation domain (AD), were tested for their ability to recover growth in the absence of histidine (minimal medium without histidine and with 0.2 mM 3-amino-1,2,4-triazole) using serial 10-fold dilutions of the corresponding yeast cultures, at an initial OD_{600}=0.8. (C) β-galactosidase activity measurements in yeasts containing the *lacZ* reporter gene fused to the *Cytc-2* promoter fragment used for the screening and transformed with the clones shown in (A). The bars represent mean β-galactosidase specific activity (±SD). In (B) and (C), AD represents a yeast strain transformed with a construct expressing only the GAL4 activation domain.

and Ohashi, 2002; Trémousaygue *et al.*, 2003; Welchen and Gonzalez, 2005). Recombinant protein AtTCP20 were then used to analyse binding *in vitro* to the site II elements present in the *Cytc-2* promoter. In Fig. 6B, the formation of specific retarded bands was observed only when a *Cytc-2* promoter fragment with site II elements was used. In a previous work, it was determined that site II elements present in the *Cytc-1* promoter also interact with AtTCP20 (Welchen and Gonzalez, 2005). This indicates that this protein or other members of the TCP family probably exert a concerted regulation of both genes.

Activation of the *Cytc-2* gene expression by the native forms of the identified transcription factors was studied *in planta*. For this purpose, a transient agroinfiltration procedure of *Arabidopsis* leaves, based on the one commonly used with tobacco leaves (see Materials and methods), was set up. Validity and reproducibility of the method were verified by infiltrating non-transformed plants with the construct expressing the –319 *Cytc-2* promoter/*gus* fusion. As shown in Fig. 7B, a 10-fold increase in specific GUS activity was observed in leaves after agroinfiltration. This value was similar to that obtained when using tobacco leaves (not shown). To analyse the effect of transcription factors on *Cytc-2* gene expression, *Arabidopsis* lines stably transformed with different *Cytc-2* promoter/*gus* constructs (–315, –259, and mutants in the G-box and ACGT motif) were infiltrated with constructs expressing the respective factors under the control of the 35S CaMV promoter. Figure 7A shows that agroinfiltration of constructs expressing AREB2/ABF4, GBF3, bHLH080, and the GBF-like protein produced an increase in GUS expression levels with respect to controls infiltrated with 10 mM $MgCl_2$ when plants containing either the –315 or the –259 fragment were used. By contrast, agroinfiltration with a construct expressing the sunflower HD-Zip transcription factor Hahb-4 (Dezar *et al.*, 2005), that has no putative binding sites within the *Cytc-2* promoter, does not produce changes in expression levels. Activation by the different transcription factors was not observed when agroinfiltration was performed in plants carrying *Cytc-2* promoter/*gus* fusions with mutations in either the G-box or the ACGT motif (Fig. 7). Based on the results, it was assumed that the different transcription factors

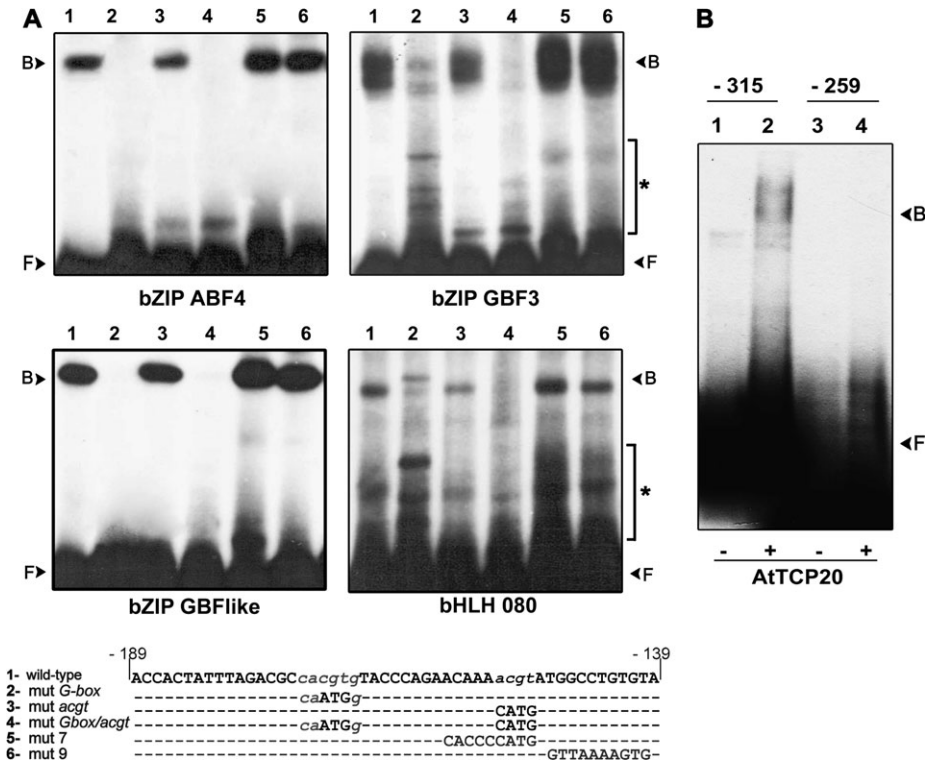

Fig. 6. DNA-binding properties of recombinant transcription factors to the *Cytc-2* promoter. (A) Binding of AREB2/ABF4, GBF3, GBF-like and bHLH080 was analysed by an EMSA. One-hundred nanograms of the different recombinant proteins were incubated with different forms (wild-type or mutagenized) of the *Cytc-2* fragment comprising nucleotides –189 to –139. Lanes 1 to 6 correspond to different labelled *Cytc-2* promoter fragments used as probes. Free and bound DNA are indicated with F and B, respectively. Bands labelled with an asterisk correspond to minor species already present in the free probes as deduced from lanes in which no protein was added. The sequences of the fragments are shown at the bottom. (B) Binding of recombinant protein AtTCP20 to a *Cytc-2* promoter fragment containing two copies of site II elements. Two different fragments of the *Cytc-2* promoter were analysed for binding by an EMSA. Fragment –315 (from –315 to –139; lanes 1 and 2) contains the site II elements, while these elements are not present in fragment –259 (–259 to –139; lanes 3 and 4). In lanes labelled (–), no protein was added; in lanes labeled (+), 200 ng of recombinant protein was included.

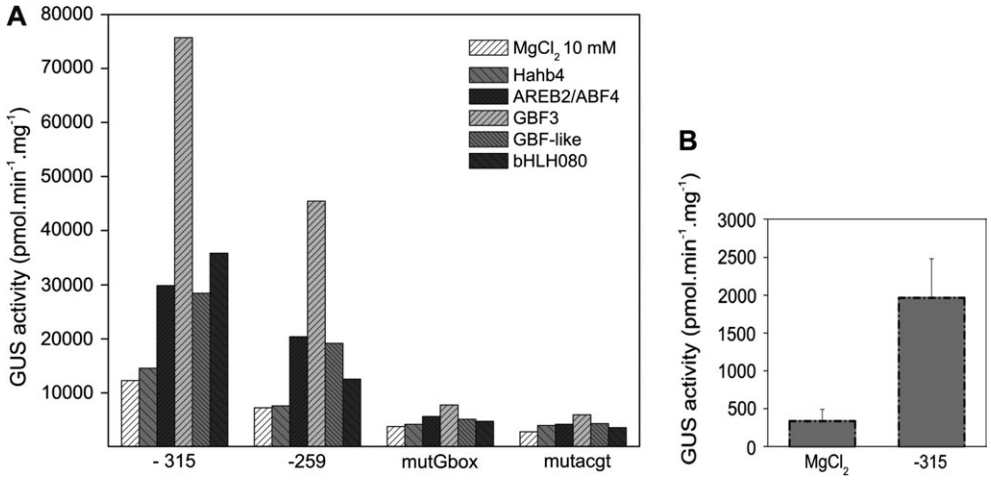

Fig. 7. Activation of the *Cytc-2* promoter by the identified transcription factors *in planta*. (A) Individual *Agrobacterium* cultures carrying full-length cDNAs of the transcription factors identified in the one-hybrid screen cloned in the pBI121 binary vector were mixed with cells carrying the 35S::p19 construct (Voinnet *et al.*, 2003) and infiltrated into 4-week-old *Arabidopsis* transgenic leaves. The lines used for the transient activation analysis were the stable *Arabidopsis* plants previously transformed with promoter constructs spanning nucleotides –315 to +54 (–315), –259 to +54 (–259) and mutagenized forms of the last fragment, either in the G-box (mutGbox) or in the ACGT motif (mutacgt; see Fig. 2A for details). GUS activity was measured in total protein extracts from leaves 4 d post-infiltration. Transgenic plants infiltrated with 10 mM $MgCl_2$ and with a 35S::Hahb4 construct, expressing an HD-Zip transcription factor, were used to evaluate possible unspecific activation effects. The bars indicate the mean activity of 10 individually infiltrated plants. The figure shows a representative experiment of three repetitions performed with different lines of the same constructs that produced similar results. Activation values for the –315 and –259 constructs with respect to plants infiltrated with $MgCl_2$ varied from 1.1 to 1.3 for Hahb-4, 1.8 to 2.4 for AREB2/ABF4, 2.9 to 6.2 for GBF3, 2.2 to 2.8 for GBF-like, and 2.0 to 2.9 for bHLH080 along these experiments. (B) Control of the transformation procedure. Wild-type *Arabidopsis* leaves were infiltrated with either 10 mM $MgCl_2$ or with an *Agrobacterium* culture carrying the –315/*gus* construct.

activate the *Cytc-2* gene mainly through interaction with the G-box. The lack of activation observed upon mutation of the ACGT motif may be due to the fact that this mutation almost completely abolishes gene expression.

Expression analysis of genes encoding the different transcription factors identified by one-hybrid screening using available microarray data showed that some of them share expression characteristics with the *Cytc-2* gene. For example, GBF3 is preferentially expressed in stamens at the F_{12} stage and in mature pollen, in a similar way as the *Cytc-2* gene, as observed in microarray experiments (http://www.bar.utoronto.ca/) and histochemical assays (this work). GBF3 is also induced by carbohydrates, heat, cold, osmotic shock, UV-B, and pathogens, treatments that produce an increase in *Cytc-2* gene expression. Moreover, the time-dependent responses observed for induction of the *GBF3* and *Cytc-2* genes by UV-B, cold and heat are very similar (see Supplementary Fig. S1 at *JXB* online). Similar observations can be made for AREB2/ABF4 after cold and heat treatments (see Supplementary Fig. S1B, C at *JXB* online) and for GBF1 after heat treatment (see Supplementary Fig. S1C at *JXB* online). Induction by cycloheximide is observed for *Cytc-2* and the gene encoding the GBF-like protein. Considering that mutation of the G-box abolishes most of the responses, it can be speculated that induction of the *Cytc-2* gene may be the consequence of changes in the amount of defined bZIP transcription factors under the conditions described above.

Discussion

Arabidopsis thaliana contains two genes encoding cytochrome *c*. Phylogenetic analysis suggests that duplication of these genes is relatively recent, but occurred before the separation of *Arabidopsis* and *Brassica*, since both gene forms can be recognized in *B. oleracea* and corresponding ESTs were reported for *B. rapa* and *B. napus* (see Supplementary Fig. S2 at *JXB* online). Duplication of genes is a common feature in angiosperms and may lead to either neo- or subfunctionalization (i.e. the acquisition of new functions or the partitioning of previous functions between the pair of duplicated genes; Force *et al.*, 1999; Hughes, 2002; Prince and Pickett, 2002). Most often, changes are related to the expression characteristics of the respective genes rather than to the functional properties of the encoded proteins, so that neo- or subfunctionalization are thought to occur through the incorporation of new transcriptionally active elements in promoter regions (Gu *et al.*, 2002; Haberer *et al.*, 2004). Studies using promoter-*gus* fusions showed that the upstream regions of the *Cytc-1* and *Cytc-2* genes produce partially different expression patterns (Welchen and Gonzalez, 2005), suggesting that these promoters have acquired novel regulatory elements after duplication. A detailed analysis of the *Cytc-1* promoter performed previously indicated that expression of this gene depends on site II motifs and a *telo* box (Welchen and Gonzalez, 2005). Notably, two site II motifs are also

present in the *Cytc-2* upstream region, raising the question of what were the changes that originated the differences in expression patterns.

Studies on the *Cytc-2* gene upstream region described in this work revealed that site II motifs do not have an essential role in expression, as in the *Cytc-1* gene, since only a partial decrease in GUS activity levels, mainly in reproductive organs, was observed when these motifs were mutated. In addition, mutation of site II elements produced a general decrease in the magnitude of the response to sucrose and cytokinin, two factors that also regulate *Cytc-1* gene expression (Welchen and Gonzalez, 2005). As in the *Cytc-1* promoter, the site II elements present around –266 of the *Cytc-2* gene were specifically recognized by the *Arabidopsis* TCP20 recombinant protein, a member of the TCP family of transcription factors (Cubas *et al.*, 1999; Kosugi and Ohashi, 2002; Trémousaygue *et al.*, 2003; Navaud *et al.*, 2007; Yao *et al.*, 2007). TCP transcription factors seem to have a central role in regulating growth and cell division control pathways (Li *et al.*, 2005; Koyama *et al.*, 2007; Tatematsu *et al.*, 2008). In this sense, although not essential for gene expression, the site II elements present in the *Cytc-2* promoter may establish a link between the biogenesis of the mitochondrial machinery involved in energy production and cell proliferation and growth (Fig. 8A). Site II elements may also be used to co-ordinate some aspects of *Cytc-2* gene expression with the expression of other nuclear-encoded respiratory genes (Welchen and Gonzalez, 2006; Gonzalez *et al.*, 2007).

The most relevant elements for *Cytc-2* gene expression were mapped to a fragment of about 50 bp located more proximal to the transcription start site and include a G-box, an ACGT motif, and a 10 bp DNA segment located downstream of the ACGT motif. The importance of these elements seems to correlate with their position in the promoter, since mutation of elements located more proximal to the transcription start site produces larger effects on expression. One possibility is that the proteins that bind upstream (to the G-box, for example) exert their effect on transcription through protein–protein interactions with components bound to downstream regions, so that their influence is lost when these components are not able to interact with the promoter. In this sense, protein–DNA interaction studies showed that binding of nuclear proteins to the G-box is dependent on the integrity of a fragment located downstream, the same fragment that abolishes expression when mutated.

The *Cytc-2* promoter does not contain a TATA-box and has an initiator element (PyTCANTPyPy; Nakamura *et al.*, 2002) around the putative transcription start site, located at –85 from the translation start codon according to mRNA sequence data. It is unlikely that the mutations that abolish *Cytc-2* gene expression have disrupted core promoter elements, usually located downstream of –50 (Molina and Grotewold, 2005), since the elements important for *Cytc-2* expression described here are located in the region from –189 to –53 from the transcription start site. Therefore, it can be concluded that the *Cytc-2* gene contains a rather compact promoter, even if deletion of upstream portions produce moderate changes in expression levels. This compactness is clearly not the consequence of the presence of a nearby gene, since the nearest gene (At4g10030) is present at 4 kbp from the *Cytc-2* start site. Similar observations were made when analysing promoters of other nuclear genes encoding respiratory chain components (Zabaleta *et al.*, 1998; Elorza *et al.*, 2004, 2006; Welchen *et al.*, 2004; Dojcinovic *et al.*, 2005; Welchen and Gonzalez, 2005), suggesting that this may be a general property of this type of promoter. The compactness of the *Cytc-2* promoter is also reflected by the fact that mutation of individual elements produces large effects on expression.

G-boxes (CACGTG) have been described in several plant promoters, where they regulate expression in response to different factors, such as light, hormones, and environmental conditions, in addition to conferring tissue-specific expression (Giuliano *et al.*, 1988; Schulze-Lefert *et al.*, 1989; Donald and Cashmore, 1990; Kim *et al.*, 1992; Menkens *et al.*, 1995; Ishige *et al.*, 1999; Hudson and Quail, 2003). G-boxes are usually recognized by a group of bZIP proteins termed GBF (G-box binding factors; Schindler *et al.*, 1992; Menkens *et al.*, 1995; Sibéril *et al.*, 2001) and also by proteins from the bHLH group (Heim *et al.*, 2003; Toledo-Ortiz *et al.*, 2003). One-hybrid screenings using the *Cytc-2* promoter region required for expression identified three GBF and two bHLH proteins able to interact with this fragment, namely GBF1, GBF3, bHLH080, bHLH081, and a GBF-like protein encoded by At1g32150. The bZIP protein AREB2/ABF4, described as a binding factor of the ABRE (abscisic acid responsive element), was also identified (Choi *et al.*, 2000; Uno *et al.*, 2000; Kang *et al.*, 2002). In this sense, it must be noted that the sequence surrounding the *Cytc-2* G-box also conforms to the consensus of the ABRE (C/TACGTGGC). All of the identified proteins were able to transactivate the *HIS3* reporter gene in yeast, while GBF3 and bHLH080 did not produce significant activation with respect to the controls when the β-galactosidase gene was used as a reporter. This apparent discrepancy may arise from differences in the characteristics of the reporter genes or in the assays used to analyse their expression (i.e. while the differential accumulation of β-galactosidase during growth in liquid culture was not enough to reach significance with respect to the control strain, the increased amount of the *HIS3* gene product may have been enough to allow better growth in medium lacking histidine). bHLH080 and GBF3, however, efficiently bound the *Cytc-2* promoter *in vitro* and produced a significant activation of a *Cytc-2* promoter/*gus* construct *in planta*. It is assumed, therefore, that these proteins are true regulators of *Cytc-2* gene expression. bHLH081 was not tested further, but its close similarity with bHLH080 suggests similar behaviour.

It is noteworthy that the four bZIP proteins identified share some expression characteristics with the *Cytc-2* gene. This is an indication that the bZIP proteins may be mediators of the response of the *Cytc-2* gene to several factors. This is supported by the fact that mutation of the G-box abolishes most responses. The responses observed for the *Cytc-2* gene are then the consequence of the

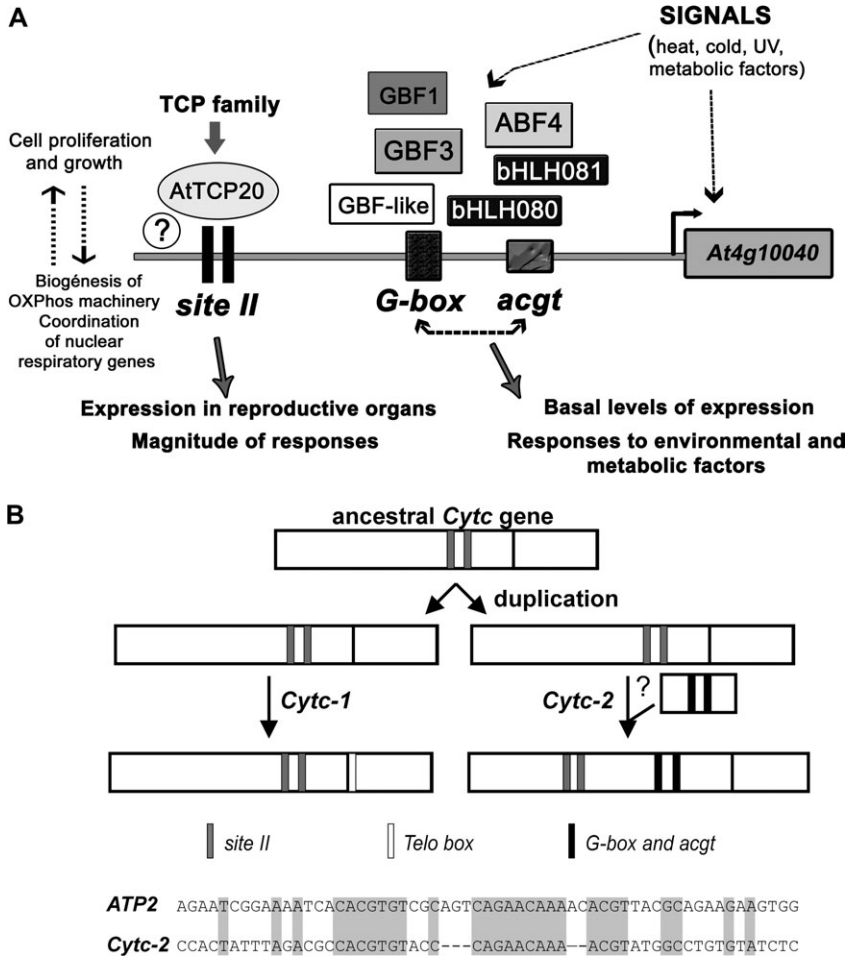

Fig. 8. The expression characteristics of the *Cytc-2* gene are governed by different elements present in its promoter region. (A) Scheme of the *Cytc-2* promoter region showing the presence and location of functional elements described here and the proteins that interact with them. Site II elements interact with TCP domain proteins and are responsible for increased expression in reproductive tissues and responses to sucrose and cytokinin. These elements may co-ordinate the expression of the *Cytc-2* gene with other respiratory chain component genes. The G-box confers induction by several environmental factors and the essential ACGT motif may be involved in basal transcription. (B) Proposed evolutionary scheme for diversification of *Arabidopsis Cytc* genes. After duplication of an ancestral *Cytc* gene that contained site II elements in its promoter, new elements were incorporated, thus producing changes in the expression characteristics of the resulting genes. In *Cytc-2*, incorporation of a fragment containing a G-box and an ACGT motif may have reduced the importance of site II elements for basal expression and conferred responses to several factors. Below, the sequence of the upstream region of the *Cytc-2* gene containing the G-box and the ACGT motif (–188 to –135) is compared with the region from –181 to –123 of the At5g08680 gene, encoding the β subunit of the mitochondrial H^+-ATPase. Identical nucleotides are boxed. Gaps were introduced to optimize the alignment.

interaction of different proteins from the bZIP family with the G-box element located around –169 from the translation start site (Fig. 8A). The G-box may then be viewed as a transducer of tissue-specific, metabolic, and environmental conditions into *Cytc-2* gene expression changes. bHLH proteins, in turn, may be involved in determining the basal transcription levels of the *Cytc-2* gene.

From an evolutionary point of view, it can be speculated that the *Cytc-2* gene incorporated specific elements involved in transcription after duplication, leading to specialized expression patterns (Fig. 8B). The presence of site II motifs in both *Cytc* genes suggests that these motifs were probably present in the ancestral form of the gene. This is also supported by the fact that site II motifs are present in many genes that encode components of the oxidative phosphorylation machinery from both *Arabidopsis* and rice (Welchen and Gonzalez, 2006). Site II motifs are also present in the putative promoters of *Cytc-1* genes from *B. rapa* and *B. oleracea* (no information is available for *Cytc-2* genes). Changes in *Arabidopsis Cytc-2* gene possibly included the incorporation of the G-box and the ACGT motif downstream of the site II elements. Incorporation of the new motifs may have occurred by accumulation of mutations or in a single step, by the introduction of a fragment carrying these elements. Regarding this possibility, it was noticed that other genes encoding isoforms of NDUFA12 (At1g04630),

SDH2 (*SDH2-3*; At5g65165), Cox5b (*Cox5b-1*; At3g15640), and ATP2 (At5g08680), components of Complexes I, II, IV, and V, respectively, have a similar arrangement of motifs. Particularly, the corresponding region of the *ATP2* gene shows homology with the *Cytc-2* gene segment located around the G-box and ACGT motifs (Fig. 8B). A model for the evolution of *Cytc* genes in *Arabidopsis* includes a duplication event and the incorporation in the *Cytc-2* gene upstream region, possibly in one step, of a segment carrying new elements active in transcription. This reduced the importance of site II motifs for expression, either because the new elements act independently or because site II motifs were moved farther away from the transcription start site. Incorporation of new elements has then possibly led to neofunctionalization of the *Cytc-2* gene through the acquisition of novel expression patterns and responses which may confer specific advantages under certain conditions or developmental stages.

Supplementary data

Supplementary data are available at *JXB* online.

Table S1. List of primers used.

Fig. S1. Response of the *Cytc-2* gene and genes encoding bZIP transcription factors that interact with the *Cytc-2* promoter to stress treatments.

Fig. S2. Alignment of *Arabidopsis* and *Brassica* cytochrome *c* protein sequences.

Acknowledgements

We gratefully acknowledge Dr Dominique Trémousaygue (CNRS-INRA, Castanet-Tolosan, France) for sending us a clone expressing protein AtTCP20. This work was supported by grants from CONICET, ANPCyT (Agencia Nacional de Promoción Científica y Tecnológica), and Universidad Nacional del Litoral. EW and DHG are members of CONICET; ILV is a fellow of the same Institution; RNC is a fellow of ANPCyT. JCH was supported by the Basic Research Program (R01-2007-000-11232-0) from KOSEF/MOST, Korea.

References

Adams KL, Palmer JD. 2003. Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Molecular Phylogenetics and Evolution* **29,** 380–395.

Block A, Dangl JL, Hahlbrock K, Schulze-Lefert P. 1990. Functional borders, genetic fine structure, and distance requirements of *cis* elements mediating light responsiveness of the parsley chalcone synthase promoter. *Proceedings of the National Academy of Sciences, USA* **87,** 5387–5391.

Choi HI, Hong JH, Ha JO, Kang JK, Kim SY. 2000. ABFs, a family of ABA-responsive element binding factors. *Journal of Biological Chemistry* **275,** 1723–1730.

Clough SJ, Bent AF. 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* **16,** 735–743.

Cubas P, Lauter N, Doebley J, Coen E. 1999. The TCP domain: a motif found in proteins regulating plant growth and development. *The Plant Journal* **18,** 215–222.

Curi GC, Welchen E, Chan RL, Gonzalez DH. 2003. Nuclear and mitochondrial genes encoding cytochrome *c* oxidase subunits respond differently to the same metabolic factors. *Plant Physiology and Biochemistry* **41,** 689–693.

Dezar CA, Gago GM, Gonzalez DH, Chan RL. 2005. *Hahb-4*, a sunflower homeobox-leucine zipper gene, is a developmental regulator and confers drought tolerance to *Arabidopsis thaliana* plants. *Transgenic Research* **14,** 429–440.

Dojcinovic D, Krosting J, Harris AJ, Wagner DJ, Rhoads DM. 2005. Identification of a region of the Arabidopsis *AtAOX1a* promoter necessary for mitochondrial retrograde regulation of expression. *Plant Molecular Biology* **58,** 159–175.

Donald RGK, Cashmore AR. 1990. Mutation of either G box or I box sequences profoundly affects expression from the Arabidopsis rbcS-lA promoter. *EMBO Journal* **9,** 1717–1726.

Duarte JM, Cui L, Wall PK, Zhang Q, Zhang X, Leebens-Mack J, Ma H, Altman N, dePamphilis CW. 2006. Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of *Arabidopsis*. *Molecular Biology and Evolution* **23,** 469–478.

Elorza A, León G, Gómez I, Mouras A, Holuigue L, Araya A, Jordana X. 2004. Nuclear *SDH2-1* and *SDH2-2* genes, encoding the iron-sulfur subunit of mitochondrial complex II in *Arabidopsis*, have distinct cell-specific expression patterns and promoter activities. *Plant Physiology* **136,** 4072–4087.

Elorza A, Roschzttardtz H, Gomez I, Mouras A, Holuigue L, Araya A, Jordana X. 2006. A nuclear gene for the iron-sulfur subunit of mitochondrial Complex II is specifically expressed during Arabidopsis seed development and germination. *Plant and Cell Physiology* **47,** 14–21.

Felitti SA, Gonzalez DH. 1998. Carbohydrates modulate the expression of the sunflower cytochrome *c* gene at the mRNA level. *Planta* **206,** 410–415.

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151,** 1531–1545.

Giegé P, Sweetlove LJ, Cognat V, Leaver CJ. 2005. Co-ordination of nuclear and mitochondrial genome expression during mitochondrial biogenesis in Arabidopsis. *The Plant Cell* **17,** 1497–1512.

Gietz D, St Jean A, Woods RA, Schiestl RH. 1992. Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Research* **20,** 1425.

Giuliano G, Pichersky E, Malikce S, Timko MP, Scolnik PA, Cashmore AR. 1988. An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. *Proceedings of the National Academy of Sciences, USA* **85,** 7089–7093.

Gonzalez DH, Welchen E, Attallah CV, Comelli RN, Mufarrege EM. 2007. Transcriptional co-ordination of the biogenesis

of the oxidative phosphorylation machinery in plants. *The Plant Journal* **51,** 105–116.

Gu ZL, Nicolae D, Lu HHS, Li WH. 2002. Rapid divergence in expression between duplicate genes inferred from microarray data. *Trends in Genetics* **18,** 609–613.

Haberer G, Hindemitt T, Meyers BC, Mayer KFX. 2004. Transcriptional similarities, dissimilarities, and conservation of *cis*-elements in duplicated genes of Arabidopsis. *Plant Physiology* **136,** 3009–3022.

Heim MA, Jakoby M, Werber M, Martin C, Weisshaar B, Bailey PC. 2003. The basic helix-loop-helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. *Molecular Biology and Evolution* **20,** 735–747.

Higo K, Ugawa Y, Iwamoto M, Korenaga T. 1999. Plant *cis*-acting regulatory DNA elements (PLACE) database:1999. *Nucleic Acids Research* **27,** 297–300.

Huang J, Struck F, Matzinger DF, Levings III CS. 1994. Flower-enhanced expression of a nuclear-encoded mitochondrial respiratory protein is associated with changes in mitochondrion number. *The Plant Cell* **6,** 439–448.

Hudson ME, Quail PH. 2003. Identification of promoter motifs involved in the network of Phytochrome A-regulated gene expression by combined analysis of genomic sequence and microarray data. *Plant Physiology* **133,** 1605–1616.

Hughes AL. 2002. Adaptive evolution after gene duplication. *Trends in Genetics* **18,** 433–434.

Hull GA, Devic M. 1995. The beta-glucuronidase (*gus*) reporter gene system. Gene fusions; spectrophotometric, fluorometric, and histochemical detection. In: Jones H, ed. *Methods in plant molecular biology*, Vol. 49. *Plant gene transfer and expression protocols*. Totowa, NJ: Humana Press Inc, 125–141.

Ishige F, Takaichi M, Foster R, Chua NH, Oeda K. 1999. A G-box motif (GCCACGTGCC) tetramer confers high-level constitutive expression in dicot and monocot plants. *The Plant Journal* **18,** 443–448.

Jefferson RA, Kavanagh TA, Bevan MW. 1987. GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO Journal* **20,** 3901–3907.

Kang JY, Choi HI, Im MI, Kim SY. 2002. Arabidopsis basic leucine zipper proteins that mediate stress-responsive abscisic acid signaling. *The Plant Cell* **14,** 343–357.

Kim SR, Choi JL, Costa M, An G. 1992. Identification of G-box sequence as an essential element for methyl jasmonate response of potato proteinase inhibitor II promoter. *Plant Physiology* **99,** 627–631.

Kosugi S, Ohashi Y. 2002. DNA binding and dimerization specificity and potential targets for the TCP protein family. *The Plant Journal* **30,** 337–348.

Koyama T, Furutani M, Tasaka M, Ohme-Takagi M. 2007. TCP transcription factors control the morphology of shoot lateral organs via negative regulation of the expression of boundary-specific genes in Arabidopsis. *The Plant Cell* **19,** 473–484.

Li C, Potuschak T, Colon-Carmona A, Gutierrez RA, Doerner P. 2005. Arabidopsis TCP20 links regulation of growth and cell division control pathways. *Proceedings of the National Academy of Sciences, USA* **102,** 12978–12983.

Lister R, Chew O, Lee MN, Heazlewood JL, Clifton R, Parker KL, Millar AH, Whelan J. 2004. A transcriptomic and proteomic characterization of the Arabidopsis mitochondrial protein import apparatus and its response to mitochondrial dysfunction. *Plant Physiology* **134,** 777–789.

Mackenzie S, McIntosh L. 1999. Higher plant mitochondria. *The Plant Cell* **11,** 571–586.

Maliga P, Klessig DF, Cashmore AR, Gruissem W, Varner JE. 1995. Identification of promoter sequences that interact with DNA-binding proteins. In: Maliga P, Klessig DF, Cashmore AR, Gruissem W, Varner JE, eds. *Methods in plant molecular biology: a laboratory course manual*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 233–260.

Meijer AH, Ouwekerk PBF, Hoge HC. 1998. Vectors for transcription factor cloning and target site identification by means of genetic selection in yeast. *Yeast* **14,** 1407–1416.

Menkens AE, Schindler U, Cashmore AR. 1995. The G-box: an ubiquitous regulatory DNA element in plants bound by the GBF family of bZip proteins. *Trends in Biochemical Sciences* **20,** 506–510.

Molina C, Grotewold E. 2005. Genome wide analysis of Arabidopsis core promoters. *BMC Genomics* **6,** 25.

Nakamura M, Tsunoda T, Obokata J. 2002. Photosynthesis nuclear genes generally lack TATA-boxes: a tobacco photosystem I gene responds to light through an initiator. *The Plant Journal* **29,** 1–10.

Navaud O, Dabos P, Carnus E, Tremousaygue D, Hervé C. 2007. TCP transcription factors predate the emergence of land plants. *Journal of Molecular Evolution* **65,** 23–33.

Palena CM, Gonzalez DH, Guelman S, Chan RL. 1998. Expression of sunflower homeodomain containing proteins in *Escherichia coli*: purification and functional studies. *Protein Expression and Purification* **13,** 97–103.

Prince VE, Pickett FB. 2002. Splitting pairs: The diverging fates of duplicated genes. *Nature Reviews Genetics* **3,** 827–836.

Salinas J, Oeda K, Chua NH. 1992. Two G-box-related sequences confer different expression patterns in transgenic tobacco. *The Plant Cell* **4,** 1485–1493.

Schindler U, Menkens AE, Beckmann H, Ecker JR, Cashmore AR. 1992. Heterodimerization between light-regulated and ubiquitously expressed Arabidopsis GBF bZip proteins. *EMBO Journal* **11,** 1261–1273.

Schulze-Lefert P, Becker-Andrb M, Schulz W, Hahlbrock K, Dangl JL. 1989. Functional architecture of the light responsive chalcone synthase promoter from parsley. *The Plant Cell* **1,** 707–714.

Sedmak J, Grossberg S. 1977. A rapid, sensitive, and versatile assay for protein using Coomassie Brilliant Blue G-250. *Analytical Biochemistry* **79,** 544–552.

Sibéril Y, Doireau P, Gantet P. 2001. Plant bZIP G-box binding factors. Modular structure and activation mechanisms. *European Journal of Biochemistry* **268,** 5655–5666.

Smith DB, Johnson KS. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase. *Gene* **67,** 31–40.

Sweetlove LJ, Heazlewood JL, Herald V, Holtzapffel R, Day DA, Leaver CJ, Millar AH. 2002. The impact of oxidative stress on Arabidopsis mitochondria. *The Plant Journal* **32,** 891–904.

Tatematsu K, Nakabayashi K, Kamiya Y, Nambara E. 2008. Transcription factor AtTCP14 regulates embryonic growth potential during seed germination in *Arabidopsis thaliana*. *The Plant Journal* **53,** 42–52.

Toledo-Ortiz G, Huq E, Quail PH. 2003. The Arabidopsis basic/helix-loop-helix transcription factor family. *The Plant Cell* **15,** 1749–1770.

Toufighi K, Brady SM, Austin R, Ly E, Provart NJ. 2005. The Botany Array Resource: e-Northerns, Expression Angling, and promoter analyses. *The Plant Journal* **43,** 153–163.

Tremousaygue D, Garnier L, Bardet C, Dabos P, Herve C, Lescure B. 2003. Internal telomeric repeats and 'TCP domain' protein-binding sites co-operate to regulate gene expression in *Arabidopsis thaliana* cycling cells. *The Plant Journal* **33,** 957–966.

Uno Y, Furihata T, Abe H, Yoshida R, Shinozaki K, Yamaguchi-Shinozaki K. 2000. Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. *Proceedings of the National Academy of Sciences, USA* **97,** 11632–11637.

Vision TJ, Brown DG, Tanksley SD. 2000. The origins of genomic duplications in Arabidopsis. *Science* **290,** 2114–2116.

Voinnet O, Rivas S, Mestre P, Baulcombe D. 2003. An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *The Plant Journal* **33,** 949–956.

Welchen E, Chan RL, Gonzalez DH. 2002. Metabolic regulation of genes encoding cytochrome *c* and cytochrome *c* oxidase subunit Vb in Arabidopsis. *Plant, Cell and Environment* **25,** 1605–1615.

Welchen E, Chan RL, Gonzalez DH. 2004. The promoter of the *Arabidopsis* nuclear gene *COX5b-1*, encoding subunit 5b of the mitochondrial cytochrome *c* oxidase, directs tissue-specific expression by a combination of positive and negative regulatory elements. *Journal of Experimental Botany* **55,** 1997–2004.

Welchen E, Gonzalez DH. 2005. Differential expression of the *Arabidopsis* cytochrome *c* genes *Cytc-1* and *Cytc-2*: evidence for the involvement of TCP-domain protein binding elements in anther- and meristem-specific expression of the *Cytc-1* gene. *Plant Physiology* **139,** 88–100.

Welchen E, Gonzalez DH. 2006. Overrepresentation of elements recognized by TCP-domain transcription factors in the upstream regions of nuclear genes encoding components of the mitochondrial oxidative phosphorylation machinery. *Plant Physiology* **141,** 540–545.

Yao X, Ma H, Wang J, Zhang D. 2007. Genome-wide comparative analysis and expression pattern of TCP gene families in *Arabidopsis thaliana* and *Oryza sativa*. *Journal of Integrative Plant Biology* **49,** 885–897.

Yu J, Nickels R, McIntosh L. 2001. A genome approach to mitochondrial-nuclear communication in Arabidopsis. *Plant Physiology and Biochemistry* **39,** 345–353.

Zabaleta E, Heiser V, Grohmann L, Brennicke A. 1998. Promoters of nuclear-encoded respiratory chain complex I genes from *Arabidopsis thaliana* contain a region essential for anther/pollen-specific expression. *The Plant Journal* **15,** 49–59.