UNIVERSITA' degli STUDI di SASSARI

Dottorato di Ricerca in Biochimica, Biologia e Biotecnologie Molecolari

XXI Ciclo

Coordinatore Prof. Bruno Masala

Studies on the cell cycle-dependent regulation of plant DHFR/TS genes

Tutor:

Prof. Diego Albani

Tesi di Dottorato di

Dr. Stefania Ghisaura

Anno Accademico 2008/2009

CONTENTS

1.INTRODUCTION

1.1 Folates

The synthesis of numerous compounds and the regulation of many methabolic processes require the addition or removal of one-carbon units (C1-methabolism). These one-carbon (C1) reactions play essential roles in major cellular processes including the synthesis of nucleic acids, methionine and pantothenate and the biogenesis of many products such as choline, lignine and chlorofill (Hanson and Roje, 2001; Sahr et al., 2005). Most of the transfers of C1-units are mediated by tetrahydrofolate (H_4F) derivatives and S-adenosylmethionine (that is a derivative of methionine that is the source of methyl units for the synthesis of choline, chlorophyll etc.).

H_4F is composed of three distinct parts: a pterin ring, a p-ABA (p-aminobenzoic acid) and a polyglutamate chain with a number of residues varying between 1 and 8 (Cossins, 1984).

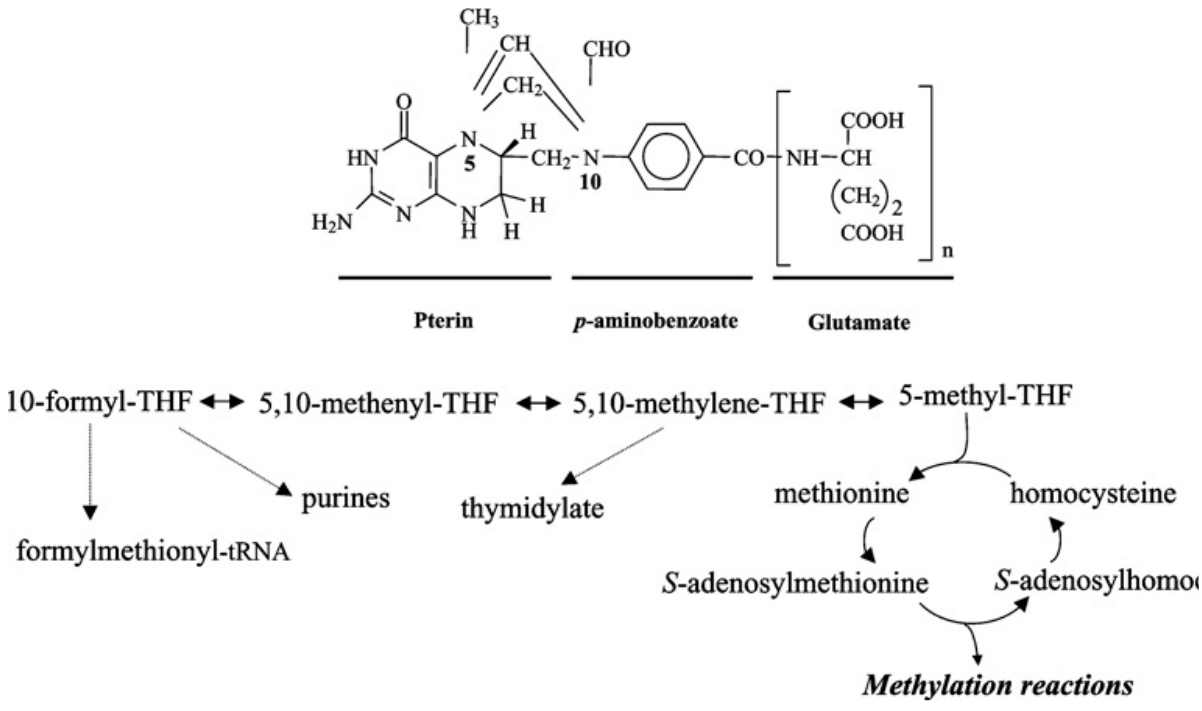


Fig 1.1: Structure of H4Fshowing the different positions of substituted C1 units and major reactions of C1 metabolism

In biological systems, the C1-units exist under various oxidation states (methyl, methylene, methenyl or 10-formyl, from the most reduced to the most oxidized) and the different H_4F derivatives constitute a family of related molecules named indistinctly under the generic term of folates or vitamin B9 (Jabrin et al., 2003).

The biosynthesis of THF in plants and microorganisms requires the sequential operation of five reactions occurring in the mitochondria, as shown in figure 1.2 (Scott et al., 2000; Hanson and Gregory, 2002).

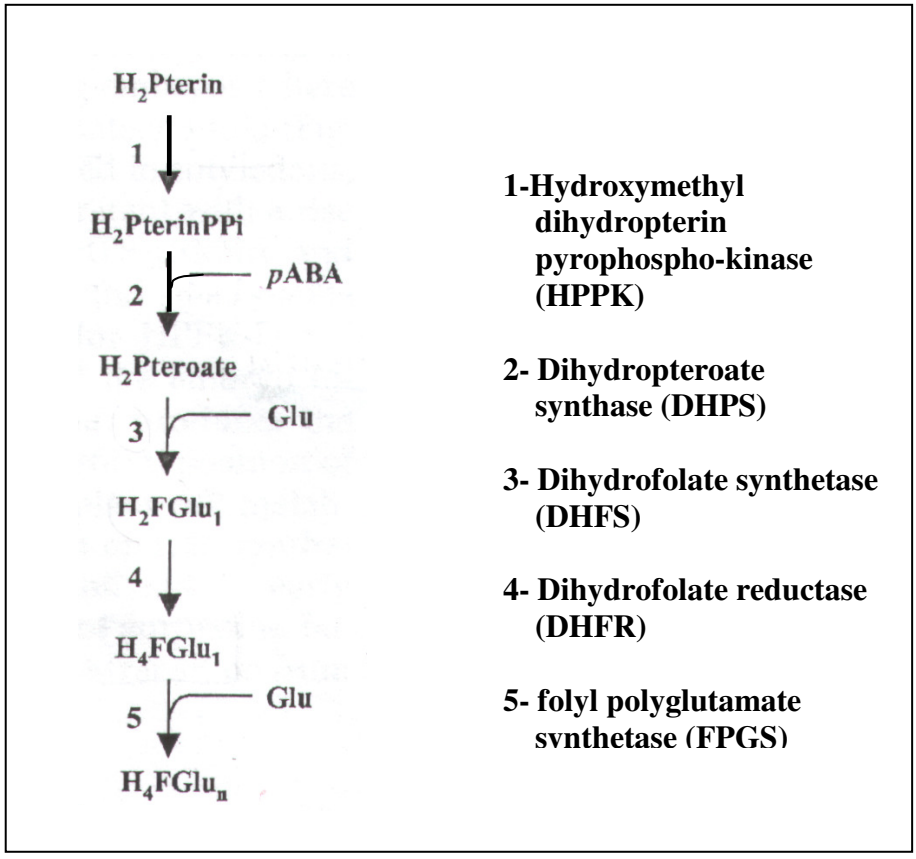


Fig 1.2 Biosynthesis of THF: the sequential operation of five reactions in the mitochondria

The first steps of H_4F synthesis are absent in animals; thus, folate supply in these organisms is ensured by the diet (Bailey and Gregory, 1999; Scott et al., 2000). Folate deficiency may have severe repercussions on human health, for example neural tube defects, heart coronary disease or increased risk of cancer. In contrast, microorganisms and plants are able to synthesize H_4F *de novo*.

1.2 H_4F biosynthesis

As shown in figure 1.3, three subcellular compartments participate in plant folate synthesis, these are: cytosol, plastids and mitochondria (Neuburger et al., 1996; Rebeillé et al., 1997; Ravanel et al., 2001).

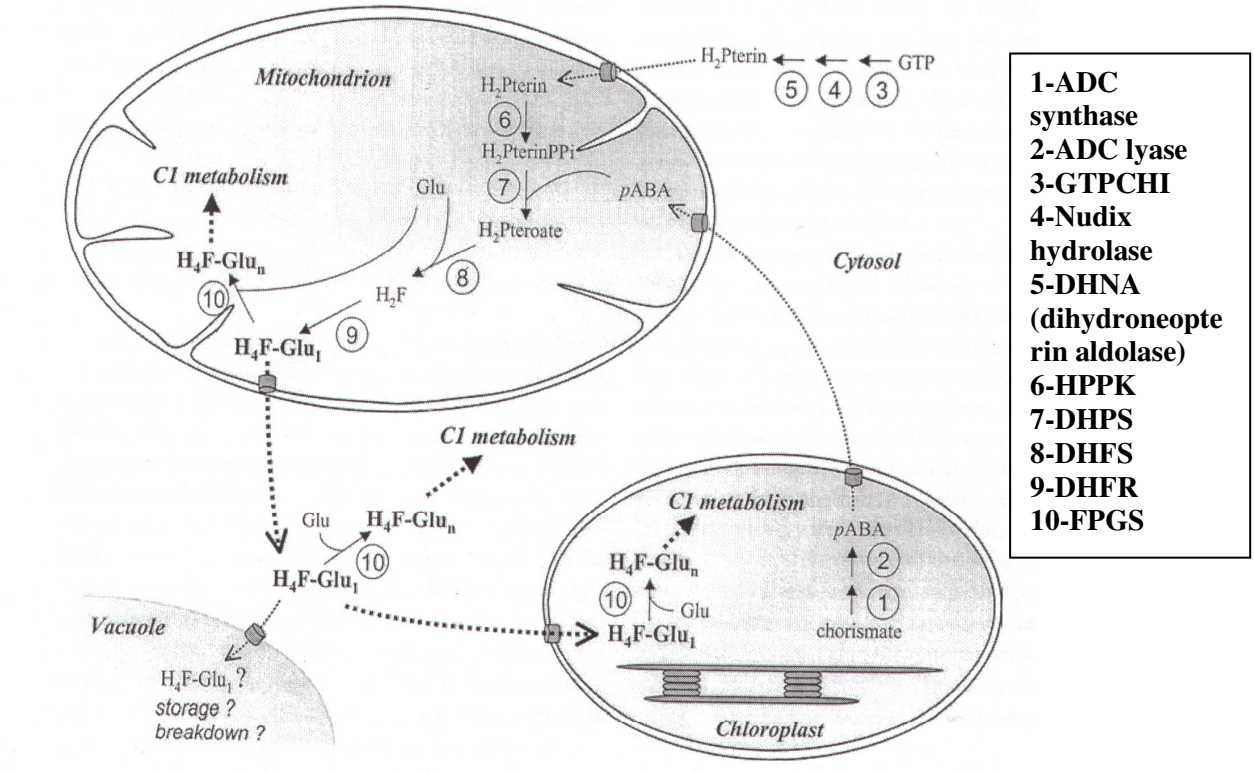


Fig 1.3 The H4F biosynthetic pathway in plants and the enzymes involved in the synthesis of H4F-Glu_1

Cytosolic steps

In the cytosol dihydropterin (or hydroxymethyldihydropterin) is synthesized from GTP in three steps. None of the enzymes required for this synthesis have predicted targeting peptides, suggesting that they are located in the cytosol. The first reaction is catalysed by GTPCHI (GTP-cyclohydrolase I) to form dihydroneopterin triphosphate. GTPCHI is present in organisms synthesizing H_4F de novo and in mammalian cells where it is involved in the synthesis of pteridines such as tetrahydrobiopterin, a coenzyme participating in redox reactions (Werner-Felmayer et al., 2002). The triphosphate side chain of dihydroneopterin triphosphate is further removed to produce dihydroneopterin. Two steps are required for this process, in fact the pyrophosphate group is removed by a specific nudix hydrolase (Klaus et al., 2005), followed by the action of a non-specific phosphatase (Suzuki and Brown, 1974). Lastly, DHNA (dihydroneopterin aldolase) cleaves the lateral side chain of dihydroneopterin to release dihydropterin.

Plastidic steps

p-ABA is synthesized in plastids. There are evidences that p-ABA is a potential limiting step for folate accumulation in plants. The first enzyme responsible for this synthesis in plant, which is a protein fusion of the PabA and PabB domains, possesses an ADC (aminodeoxychorismate) synthase activity and exhibits a transit peptide that is typical of chloroplast targeting. The last step in the synthesis of p-ABA is catalysed by an ADC lyase, a protein also located in plastids (Basset et al., 2004).

Mitochondrial steps

The final steps of H_4F synthesis occur in mitochondria. In fact, leaf mitochondria contain all the required enzymes and the first three steps are presumably exclusively localized in this compartment. Into mitochondria, dihydropterin and p-ABA are combined together with glutamate to produce H_4F. Four reactions are required to produce $H_4F\text{-}Glu_1$ (the monoglutamate form of H_4F). Dihydropterin is activated into its pyrophosphorilated form to be combined with p-ABA in a second step, resulting in dihydropteroate. These two reactions are catalysed in plants by a single bifunctional enzyme, HPPK (Hydroxymethyldihydropterin pyrophospho-kinase) and DHPS (Dihydropteroate synthase) that was only detected in mitochondria (Neuburger et al., 1996; Rebeillé, 1997). A detailed analysis of kinetic properties of the mitochondrial HPPK-DHPS isoform indicates that the DHPS reaction is feedback inhibited by dihydropteroate, H_2F and $H_4F\text{-}Glu_1$. The third step is the ATP-

dependent attachment of glutamate to the carboxyl moiety of p-ABA to form H_2-Glu_1. It is catalysed by DHFS (dihydrofolate synthetase). This enzyme , together with the preceding reaction catalysed by HPPK-DHPS is absent from animals, rendering these organisms autotrophic for folates. H_2F is then reduced to H_4F-Glu_1 by DHFR (dihydrofolate reductase) using NADPH.

Higher plants and protozoa possess a bifunctional enzyme bearing DHFR and TS (thimidilate synthase) activities. TS catalyses the methylation of deoxyuridine-monophosphate into deoxythimidine-monophosphate in the presence of CH_2-H_4F-Glu_n (where H_4F-Glu_n stands for tetrahydro-pteroylpolyglutamate or tetrahydrofolate) (Ivanetich, 1990; Cella and Parisi, 1993). In this reaction CH_2-H_4F-Glu_n acts both as a C1-unit donor and a reducing agent, producing H_2-Glu_n (dihydropteroyl-polyglutamate or dihydrofolate). Thus, the DHFR domain of the bifunctional enzyme is involved in the reduction of H_2F originating from either the *de novo* synthesis pathway (monoglutamate form) or the oxidation of H_4F-Glu_n by TS activity (polyglutamate form).

In all organisms, the polyglutamate tail of H_4F-Glu_n is formed by the sequential addition of γ-linked glutamate residues to H_4F-Glu_1, a reaction catalysed by FPGS (folyl-polyglutamate synthetase). Glutamylation is essential to retain folate in a given compartment of the cell by increasing the anionic nature of folate coenzymes, thus impairing their diffusion through hydrophobic barriers (Appling, 1991). H_4F-Glu_1, once synthesized in the mitochondria, is exported to the other cell compartments before the final glutamylation step. In plant cells, the cytosol, mitochondria and chloroplasts contain folates predominantly in the form of polyglutamylates derivatives (Cossins, 2000).

1.3 C3 plants and folates

Despite its low concentration in plant tissues (Cossins et al, 1984; Jabrin et al, 2005), folate is likely to be of major importance during seedling development due to the housekeeping functions mediated by folate coenzymes, for example the pool of folate in pea (Pisum sativum) cotyledons increased during germination and that the inhibition of *de novo* synthesis of THF using folate analogs blocked seedling development (Roos and Cossins, 1971; Gambonnet et al, 2002).

In plants THF is also involved in the photorespiratory cycle, a specific pathway that occurs at very high rates in green leaves from C3 plants. Photorespiration relies on two THF-dependent

enzymes present in the matrix space of leaf mitochondria, the Gly-decarboxylase complex (GDC) and Ser-hydroxymethyltransferase (SHMT). A continuous synthesis of THF is essential to maintain high rates of Ser synthesis through the mitochondrial activities of GDC and SHMT in *Arabidopsis* (Prabhu et al, 1996).

Folate turnover could be a potential limiting factor in proliferating cells where there is a high request of nucleotide synthesis and also of THF synthesis. In addition to actively dividing tissues, folate synthesis and accumulation were found to be elevated in photosynthetic leaves. During leaf development in the light, the expression pattern of the mRNA for HPPK-DHPS follows the accumulation of the transcripts coding proteins involved in photosynthesis and photorespiration, for example the small subunit of Rubisco and the constituents of the Gly cleavage system (Vauclare et al, 1996). Upon light exposure, the HPPK-DHPS mRNA are inducted, suggesting that the observed accumulation is an indirect response to illumination and most probably is a consequence of the stimulation of photosynthesis and photorespiration by light (Oliver et al, 1994). These observation suggest that a huge demand for folate and thus for C1-transfer reactions is associated with leaf development in the light.

1.4 DHFR-TS enzyme: role in the metabolism of nucleotides

Thymidilate synthase (TS) and Dihydrofolate reductase (DHFR) are ubiquitous enzymes which play central roles in DNA precursor biosynthesis both in prokaryotic and eukaryotic cells. Thymidylate synthase is exclusively involved in *de novo* dTMP biosynthesis and is closely linked in function to DHFR. DHFR can have two different roles depending on the source of dihydrofolate: *de novo* synthesis of tetrahydrofolate or recycling of the dihydrofolate released as one of the end products of the TS catalysed reaction. Both functions are essential in maintaining the tetrahydrofolate level, a critical intermediate for thymidylate, purine biosynthesis and in various reactions of one carbon transfer metabolism because inhibition of DHFR or TS results in the depletion of tetrahydrofolate and dTMP pools and causes the subsequent cessation of DNA synthesis.

Both enzymes are obvious targets for chemotherapy in malignant diseases, for example 5-fluoruracyl is an inhibitor of TS and metotrexate, that is an analogue of DHFR, functions as a competitor of this enzyme.(Lazar et al, 1993). Thymidylate synthase (TS) is involved in the *de novo* synthesis of thymidine monophosphate (dTMP), a precursor of thymidine triphosphate (dTTP), in both eukaryotic and prokaryotic cells. Thymidylate synthase catalyzes

the transfer of a methyl group from 5,10-methylene tetrahydrofolate to dUMP resulting in the formation of dTMP and 7,8-dihydrofolate. (Cox et al, 1999). These reactions are summarized in a simple equation:

$$\mathbf{dUMP + 5,10\text{-}CH_2\text{-}THF \rightleftharpoons dTMP + DHF}$$

In this reaction 5,10-methylenetetrahydrofolate acts both as donor of the methyl group and as a reducing agent, giving rise to DHF. (Luo et al, 1993). In a subsequent reaction, dihydrofolate reductase (DHFR) catalyzes the reduction of 7,8-dihydrofolate to tetrahydrofolate (Luo et al, 1993). TS is dependent on DHFR for regeneration of tetrahydrofolate, which is in turn necessary for the formation of 5,10-methylene-tetrahydrofolate. These two enzymes, DHFR and TS, are functionally correlated, in fact in protozoa and in plants these two activities are restricted in a single polypeptide (Cella and Parisi, 1993; Luo et al, 1997).

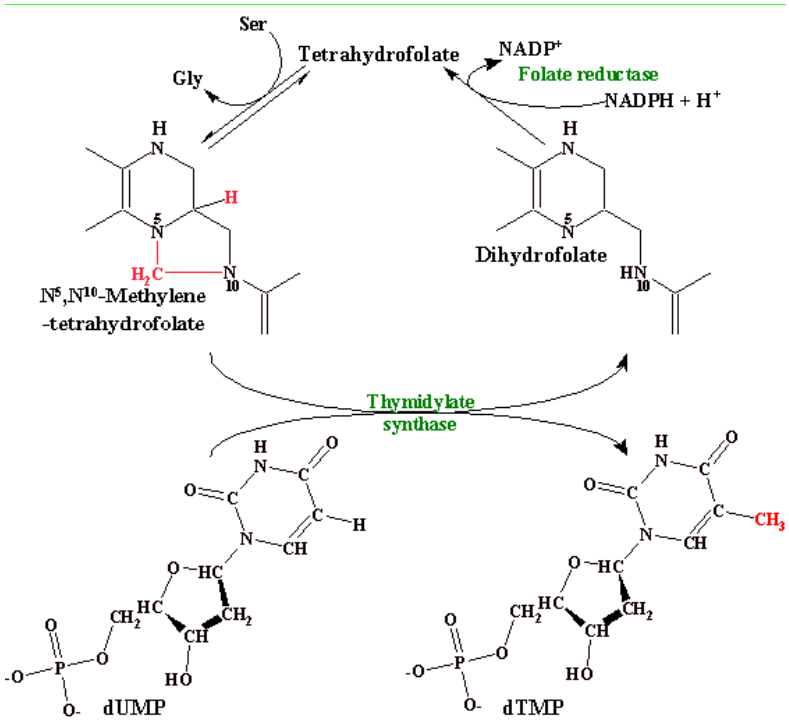


Fig 1.4: Biosynthetic pathway of dTMP

1.5 DHFR-TS: A bifunctional enzyme

In bacteriophages, viruses and most organisms (bacteria, fungi, vertebrates and mammals) DHFR and TS occur as distinct monofunctional polypeptides (Luo et al., 1997). In these organisms DHFR occurs as a protein of 18-22 KDa (Luo et al., 1993, Schweitzer, et al., 1990) whyle TS occurs as a dimer of 60-75 KDa . In contrast studies in protozoa revealed the presence of a bifunctional enzyme DHFR-TS with both activities on a single polypeptide (Ivanetich et al., 1990). For example, in *Chritidia fasciculata* (Ferone and Roland, 1980) and in *Leishmania tropica* (Meek et al., 1985), the enzyme is present as a dimer of 110-140 KDa and possesses both catalytic activities. The size of bifunctional polypeptide is close to the sum of monofunctional DHFR and TS and the gene encoding DHFR-TS may have resulted from the fusion of monofunctional genes.

For plants, both monofunctional (Toth et al., 1987) and bifunctional forms have been described (Cella and Parisi, 1988, 1993). In the case of the green algae *Scenedesmus obliquus* (Bachmann et al., 1987), the two enzymes co-migrate as a single protein band in polyacrylamide gel electrophoresis under non-denaturing conditions. The co-migration of the two enzymes in a single protein suggest a communication between the two domains by channelling DHF from TS to DHFR (Meek et al., 1985). There may be a selective advantage for the fusion of genes encoding enzymes in the same pathway (Kim et al., 1998). Although DHFR and TS both participate in dTMP formation, DHFR has other roles in metabolism that are independent of TS reactions (Cossins and Chen, 1997).

In higher plants bifunctional DHFR-TS genes have been identified in *Arabidopsis* (Lazar et al., 1993), carrot (Luo et al., 1993, 1997), soybean (Wang et al., 1995) and also in maize (Cox et al., 1999). Sequence alignments suggest that the two known groups of bifunctional genes from protozoa and higher plants may have different evolutionary precursors. Nucleotide and amino-acid sequence comparisons indicate that DHFR and TS domains of the bifunctional genes evolved at different rates, following the evolutionary history of their monofunctional counterparts. Intron position comparisons support the hypothesis that the bifunctional plant genes and the monofunctional DHFR of vertebrates evolved from a common intron containing ancestor. Bifunctionality could have arisen independently during plant evolution or been derived from a common ancestor with the protozoa.

In eubacteria, animals and fungi both genes DHFR and TS are separately translated, often in one operon, TS preceding DHFR (Fig.1.5), presumably the original eukaryotic condition (Philippe et al., 2000).

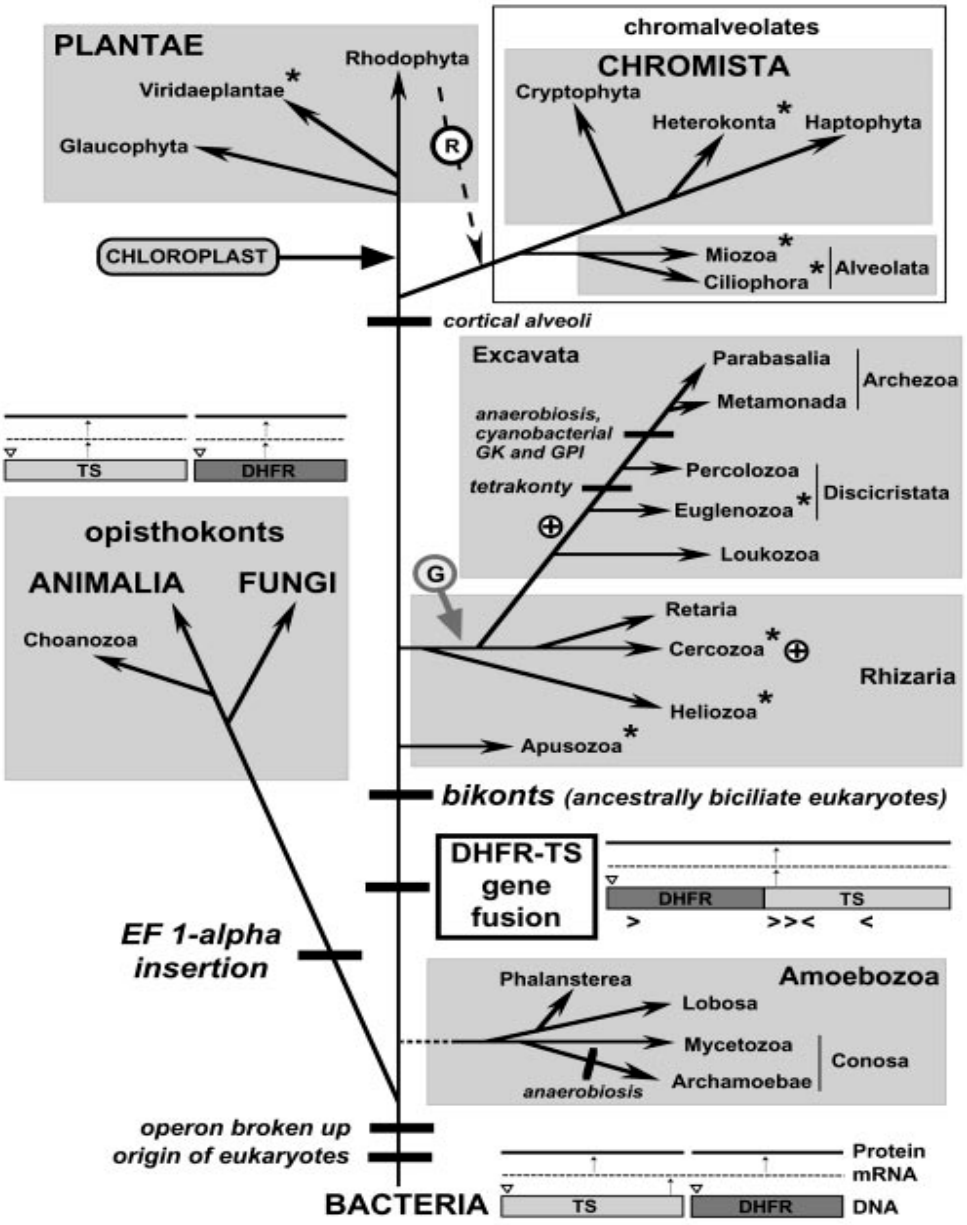


Fig 1.5: Phylogenetic relationships of the major eukaryote groups: asterisks mark all eight groups positive for the bifunctional DHFR-TS fusion gene.

Plants, alveolates, and Euglenozoa instead have a bifunctional fusion gene with both activities in one protein (Philippe *et al* 2000). It suggests that the eukaryote tree's root must be below the common ancestor of plants, alveolates and Euglenozoa (Philippe et al., 2000).
Stechmann et al amplified and sequenced DHFR-TS fusion genes from four groups: from heterokonts and from three protozoan phyla; plus, as positive controls, additional Euglenozoa and Ciliophora. Multiple allignment shows that all are authentic DHFR-TS fusion genes with one open reading frame. Only in one protist phylum (Amoebozoa) they detected no fusion gene. This phylogenetic distribution of DHFR-TS fusion genes is shown in fig1.5; their origin coincides with that of the biciliate condition. All organisms above the apparent point of origin of the fusion protein in fig.1.5 are ancestrally biciliate and collectively called Bikonts (Stechmann and Cavalier-Smith, 2000). The separate Opisthokont DHFR and TS genes are distinctly more similar to the fusion genes than to the separate bacterial genes. A unique insertion in EF-1α of animals and fungi (Baldauf, et al., 1999) is a derived character indicating that the root cannot be within opisthokonts. So the root may be between Opisthokonts and Amoebozoa/Bikonts (Stechmann and Cavalier-Smith, 2000).

1.6 dhfr-ts genes in higher plants

Biochemical analyses have revealed that, like protozoa, higher plants contain bifunctional dihydrofolate reductase-thymidylate synthase proteins (Cella and Parisi, 1993). Genomic and cDNA sequences coding for bifunctional DHFR-TS enzymes have been identified for the first time in *Arabidopsis thaliana* by Lazar et al. in 1993. These investigators amplified two genomic fragments encoding thymidylate synthase and isolated the corresponding genomic clones. Upon characterization of these clones, DHFR-specific sequences were identified upstream of the region coding for TS. Cloning the cDNAs confirmed that the sequences of the DHFR and TS domains reside on single transcripts encoding bifunctional gene products.
Following this report, bifunctional DHFR-TS cDNA sequences have been identified also in other higher plants such as *Daucus carota* (Luo et al., 1993), *Glycine max* (Wang et al., 1995) and *Zea mays* (Cox et al., 1999). This represents the first *DHFR-TS* gene cloned from a monocotyledonous plant.
Comparison of the deduced amino acid sequences revealed over 75% similarity and the conservation of motifs typical of DHFR and TS proteins, for example the metotrexate, folate

and dUMP binding sites. Figure 1.6 represents a sequence alignment of these higher plants DHFR-TS sequences.

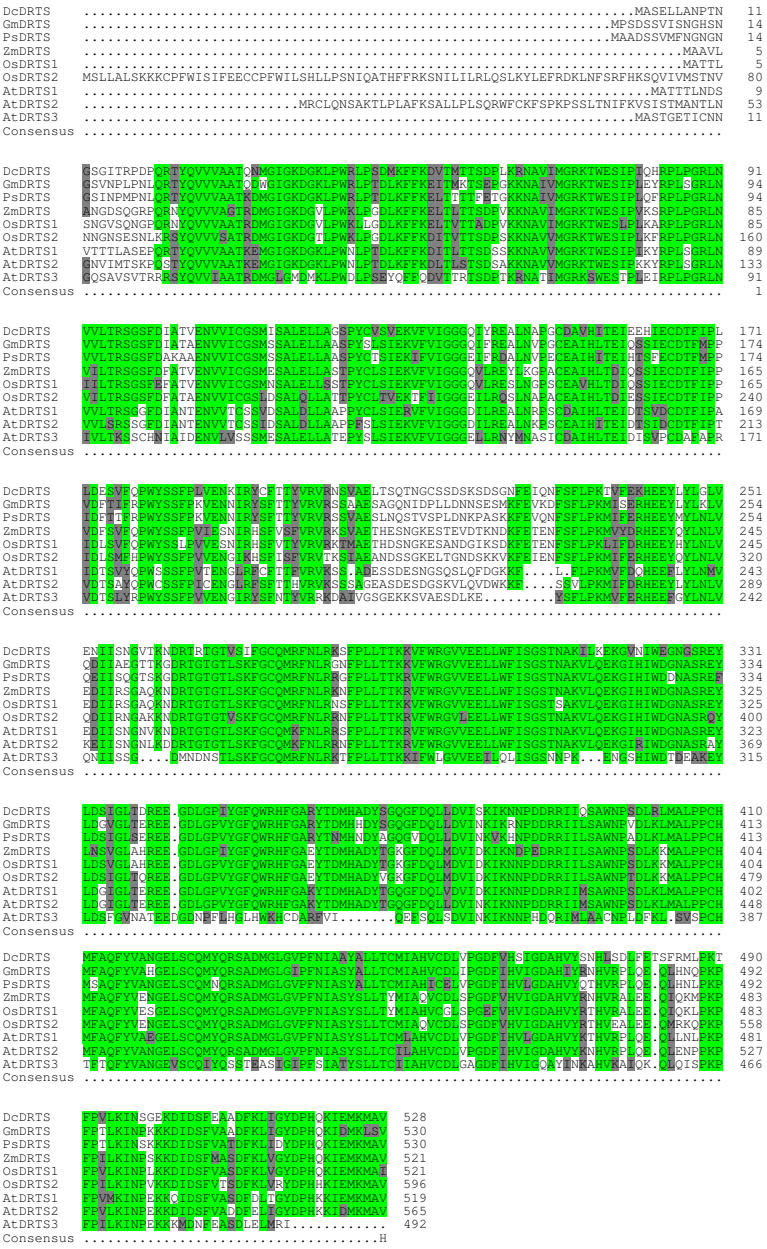


Fig 1.6: Comparison of DHFR-TS amino acid sequenze from plants; amino acids underlined in green are identical while amino-acids in grey are similar; Dc Daucus carota, Gm Glycine max, Ps Pisum sativum, Zm Zea mays, Os Oriza sativa

In 1997 Luo et al. mapped the 5' end of the carrot *dhfr-ts* gene by primer extension and by rapid amplification of 5' cDNA ends (RACE). This work revealed the production of two major classes of transcripts derived from alternative promoters (Luo et al., 1997). Moreover,

sequencing of the 5' flanking genomic region confirmed the presence of two well defined TATA box sequences located to 25 to 27 bp upstream of the most proximal transcription start points (Luo and Cella, 1998).

1.7 DHFR-TS expression models in higher plants

Biochemical studies on TS and DHFR in plants demonstrated enzyme activity to be high in proliferating cells and virtually undectable in quiescent cells (Vandiver and Fites, 1979). These findings imply that the enzymes are present in tissues where DNA synthesis is active.
To determine if maize has more than one form of the genes, Cox et al. analyzed transcripts by RNA gel blot hybridization. Because TS gene expression in mammals and yeast is confined to S-phase of the cell cycle (Storms et al., 1984; Jenh et al., 1985; Ayusawa et al., 1986), high levels of *ZmDHFR-TS* expression were expected in meristematic tissues and in tissues undergoing endoreduplication, both of which require abundant pools of nucleotides. In this study, *ZmDHFR-TS* expression was also analysed in the endosperm tissue of maize kernels that exhibits developmentally controlled endoreduplication (Kowles et al., 1988, 1997). DNA synthesis in developing kernels is reported to be highest between 10 and 16 DAP (days after pollination), but is still detectable at 28 DAP (Kowles et al., 1988, 1997; Grafi and Larkins, 1995). Thus, *ZmDHFR-TS* gene expression was investigated in kernels at different stages of development from 8 to 35 DAP. RNA gel blot hybridizations of total RNA from each set of kernels in the time course showed high levels of *ZmDHFR-TS* expression from 8 to 20 DAP, followed by a decrease at 24 DAP. *ZmDHFR-TS* expression was maintained in both embryo and endosperm tissue at 35 DAP. Although the relative amount of *ZmDHFR-TS* mRNAs appeared to be similar in both 35 DAP endosperm and embryo tissues, in these samples *ZmDHFR-TS* mRNA levels were lower than in the 12 and 18 DAP kernels. At 35 DAP the embryo is still developing (Abbe and Stein, 1954) and requires DNA synthesis whereas cell division and endoreduplication are generally completed in the endosperm tissue by this stage (Kowles and Phillips, 1988). The persistence of *ZmDHFR-TS* mRNA in endosperm at 35 DAP may be due to a requirement of nucleotides by mitochondria or to the onset of desiccation in the endosperm.
To compare the expression of *ZmDHFR-TS* in maize meristematic tissues and differentiated tissues, RNA was extracted from root tips, root elongation zones, and from mature leaf tissues. The root tips provided a good source of proliferating cells while the root elongation

zones and mature leaf tissue provided two different sources where cell division was expected to be low or absent. *ZmDHFR-TS* RNA levels were low in the root elongation zone and in mature leaf tissues.

Also studies from *Daucus carota* confirmed that the major quantity of *DHFR-TS* transcripts is located in highly dividing meristematic tissues. Albani et al. (2005), with in situ hybridization analyses, saw a strong expression in pro-embryogenic masses as well as in shoot and root meristems of somatic embryos, and a well defined expression in meristems and in expanding cotyledons at the torpedo/plantlet stage. Moreover, northern hybridization experiments have confirmed a higher accumulation of *DHFR-TS* transcripts in proliferating suspension cells compared to cells in stationary phase or cells blocked with propyzamide and also revealed a low expression of the *DHFR-TS* gene in carrot leaves. The expression analyses performed in both carrot and maize contrast with the data reported by Neuburger et al (1996) which described strong accumulation of DHFR-TS in the mitochondria of mature pea leaves. This discrepancy could derive from the presence of paralog genes that might be differentially expressed during development in different tissues.

1.8 Organization of the DHFR-TS genes in Arabidopsis thaliana

The genome of *Arabidopsis thaliana* has been fully sequenced and contains three *DHFR-TS* genes called *AtDRTS1, AtDRTS2* and *ATDRTS3*. These genes have been also named *THY1, THY2* and *THY3*.

AtDRTS1, annotated with the sequential number At2g16370 and contained in clone BACF16F14, is located in chromosome 2 and extends 2654 bp in length, from position 7088985 to 7091639. The gene contains 9 introns and the resulting cDNA is 1660 bp long. The ATG start codon is located in the second exon at position 225. The coding region translates into a protein of 519 aa with a MW of 58KDa.

AtDRTS2, annotated with the sequential number of At4g34570 and identified in clone BACT4L20, is located in chromosome 4 extending 3303 bp in length from position 16511012 to 16514316. This gene contains 11 introns and the resulting cDNA is 1683 bps long. The ATG start codon is located in the second exon at position 201 and the coding region translates into a 565 aa protein with MW of 63 KDa.

AtDRTS3, with the sequential number At2g21550 and identified in clone BACF2G1, is located in chromosome 2. It extends 2980 bp from position 9234289 to 9237269 and contains

9 introns. The resulting cDNA is 1469 bp long starting from the ATG codon. The coding region translates into a 492 aa protein with a MW of 54 KDa.

These data are reported in the TAIR (The Arabidopsis Information Resource) database at the site *www.arabidopsis.org*, while the gene structures shown in Fig 1.7 were extracted from the Arabidopsis Massively Parallel Signature Sequencing database at the site *http://mpss.udel.edu/at/GeneQuery.php*.

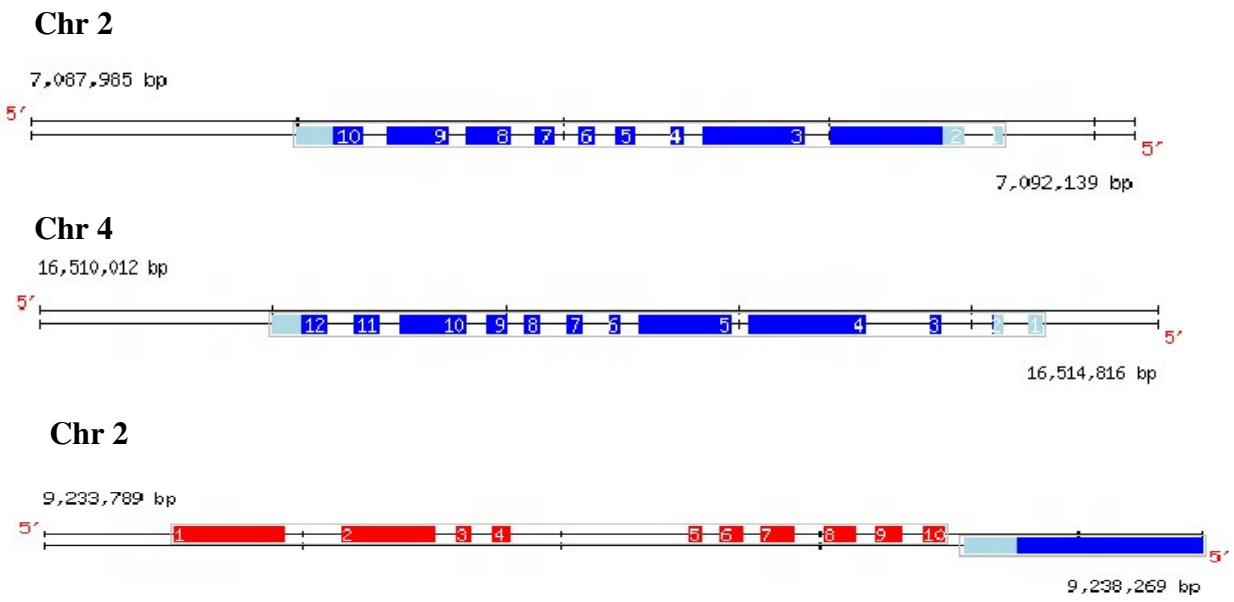


Fig 1.7: Structure of dhfr-ts genes of Arabidopsis thaliana; from up to down AtDRTS1, ATDRTS2 andATDRTS3

The reported structures of the *AtDRTS1* and *AtDRTS2* genes derive from the analysis of cDNA sequences, whereas the genomic organization of *AtDRTS3* is a putative gene model based on sequence homology.

Concerning the *AtDRTS1* gene, unpublished 5'RACE analyses performed in our laboratory have revealed the occurrence of events of alternative splicing that gives rise to transcripts lacking the second exon. Because this exon contains the first ATG codon of the gene, assuming this start codon is functional, its absence in the alternative transcripts would result in the translation of a smaller isoform of the AtDRTS1 protein, starting from the ATG sequence located in the fourth exon. Thus, this alternative start codon, which is the same ATG described by Lazar et al. in 1993, could give rise to a isoform of the AtDRTS1 protein potentially lacking or possessing different signal peptides allowing organellar targeting of the enzyme.

Concerning *AtDRTS3*, as mentioned before, this gene has not yet been described in scientific publications and its genomic organization corresponds to the genic model reported in the TAIR databases. However, it is unlikely that this putative gene model represents the real *AtDRTS3* transcripts because the unique cDNA sequence reported in the databases appears to retain the fourth intron of the gene model and terminates in its center. Moreover, the fact that the fourth putative intron cannot be spliced out could be linked to the insertion of a transposon-like element, which also explains its large size. The retained intron introduces a TAG stop codon in the reading frame of the gene and causes a premature termination of the transcripts. Thus, unless the gene model can be confirmed, it is likely that the *AtDRTS3* gene can only code for a partial DHFR-TS protein truncated in the TS domain at the C-terminus. However, it is not known whether the transcripts corresponding to the *AtDRTS3* cDNA are actually translated into a product and whether the resulting protein possesses only DHFR activity or is having also partial TS activity.

Interestingly, the analysis of the genomic organization of the *AtDRTS* loci has revealed the presence of three members of the sec14-like *SFH* gene family upstream of all three *DHFR-TS* genes, oriented in opposite direction with respect to the *AtDRTS* coding sequences, as shown in figure 1.8. In chromosome 2, *AtSFH7* (At2g16380) is upstream of *AtDRTS1* (At2g16370) and *AtSFH3* (At2g21540) is upstream of *AtDRTS3* (At2g21550) whereas *AtSFH1* (At4g34580) is located upstream of *AtDRTS2* (At4g34570) in chromosome 4.

SFHs (Sec Fourteen Homologues) are a diverse group of proteins with distinct subcellular localizations and varied physiological functions related to lipid metabolism, phosphoinositide mediated signalling and membrane trafficking.

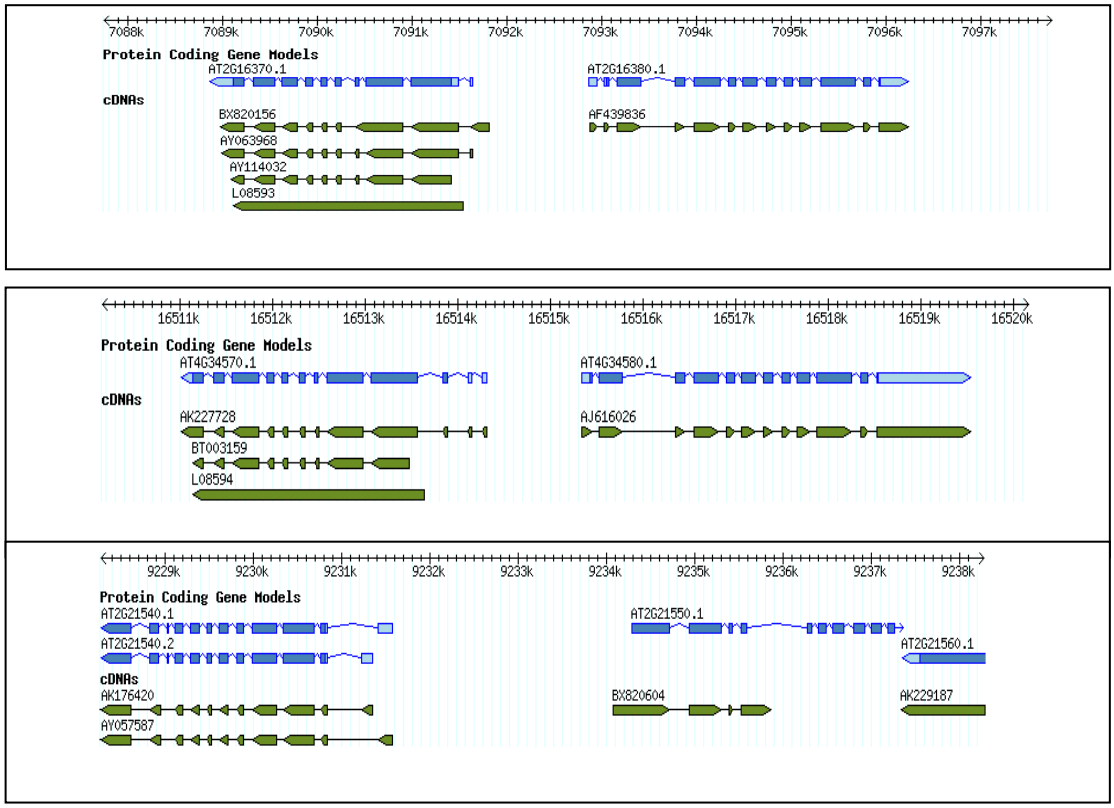


Fig 1.8: Organization of the AtDRTS and AtSFH genes and cDNAs. From up to down AtDRTS1-AtSFH7; AtDRTS2-AtSFH1 and AtDRTS3-AtSFH3

1.9 The E2F transcription factors

In plants, like in other eukaryotes, cell division and differentiation are coordinated by the activity of numerous cell cycle genes (den Boer and Murray,2000; Stals and Inzé, 2001). A pivotal role in the control of cell cycle progression is played by members of the E2F family of transcription factors. The E2F transcription factors are key components of the cyclin D/retinoblastoma/E2F pathway. Retinoblastoma (pRB) acts as repressor of E2F activity and its deregulation causes tumor formation in mammals (Harbour and Dean, 2000; Trimarchi. and Lees, 2002). The first E2F protein has been discovered as a cellular factor required for transcriptional activation of the adenovirus E2 promoter (Kovesdi et al, 1987).

E2F factors control the transcription of a wide range of genes involved in cell cycle progression but also in DNA synthesis, replication and repair (Menges *et al,* 2002). The accumulation and the activity of the E2Fs are regulated through different mechanisms such as

transcriptional and post-transcriptional control, proteolysis, phosphorylation, subcellular localization and the binding to different RB-related pocket proteins.

E2Fs are known to regulate G1 to S-phase transition of the cell cycle. In growth arrested cells (G0) and during the early G1 phase E2F activity is repressed by pRB. Upon growth stimulation and/or cell cycle progression, pRB is phosphorylated in the late G1 phase by cyclin-dependent kinases (CDKs) and, consequently, loses its affinity for E2F. The release of pRB triggers the activation of target genes by E2F, which irreversibly commits cells to undergoing DNA replication (S phase).

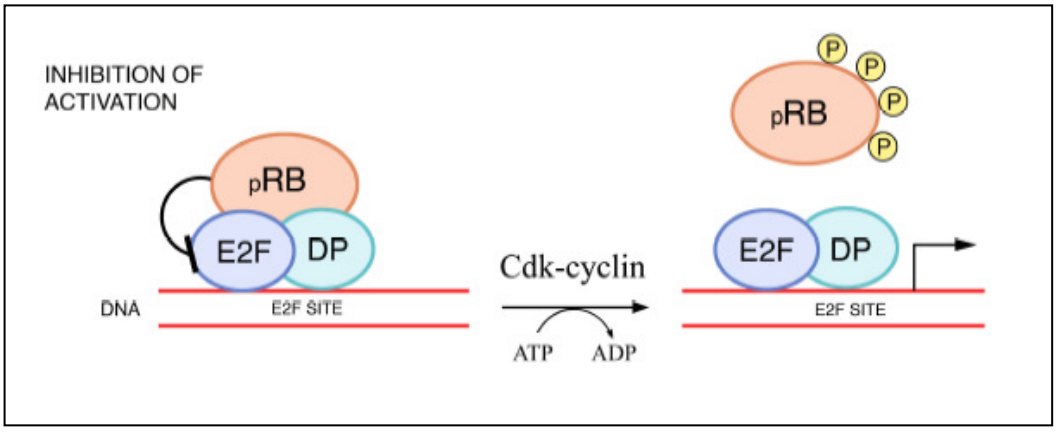


Fig. 1.9: General mechanisms that control the repressive activity of E2F.

1.9.1 Animal E2F factors

In mammals the E2F family comprises eight different factors (E2F1-8), most of which (E2F1-6) heterodimerize with a DP partner in order to bind DNA (DP1 or DP2) (de Bruin et al, 2003; Attwooll et al, 2004; Dimova and Dyson, 2005; Maiti et al, 2005; Christensen et al, 2005). In fact, E2F1-6 are considered typical E2Fs because possess highly conserved domains for sequence specific DNA binding and dimerization with DP proteins. Moreover, distinctive features of the various E2F factors allow a subdivision of this family into subgroups on the basis of their primary structure, their need to form complex with members of the DP family, their affinity for distinct members of the pRB family, their expression pattern and their capacity to either activate or repress E2F-responsive genes (Christensen et al, 2005). The first three factors, E2F1-3, are considered transcriptional activators which induce the expression of E2F responsive genes at G1/S transition and, when overexpressed, can drive quiescent cells to

re-enter into cell cycle (Attwooll et al, 2004). E2F4-6 act differently and are able to repress transcription of specific target genes by interacting with the p107 and p130 pocket proteins, related to pRB, which can tether chromatin remodelling co-repressors to the E2F-regulated promoters. E2F7 and 8, instead, are considered atypical E2Fs which have been identified more recently (deBruin et al, 2003; Di Stefano et al, 2003; Christensen et al, 2005), following the initial discovery of atypical E2Fs in plant cells (Mariconti et al., 2002). These E2Fs possess duplicated DNA binding domains that allow recognition of E2F cis elements independently of DP proteins. Moreover, they lack the conserved regions necessary for dimerization and transactivation as well as the marked box domains. These factors can compete with the typical E2Fs for the binding to DNA and are believed to inhibit the transcription of E2F responsive genes.

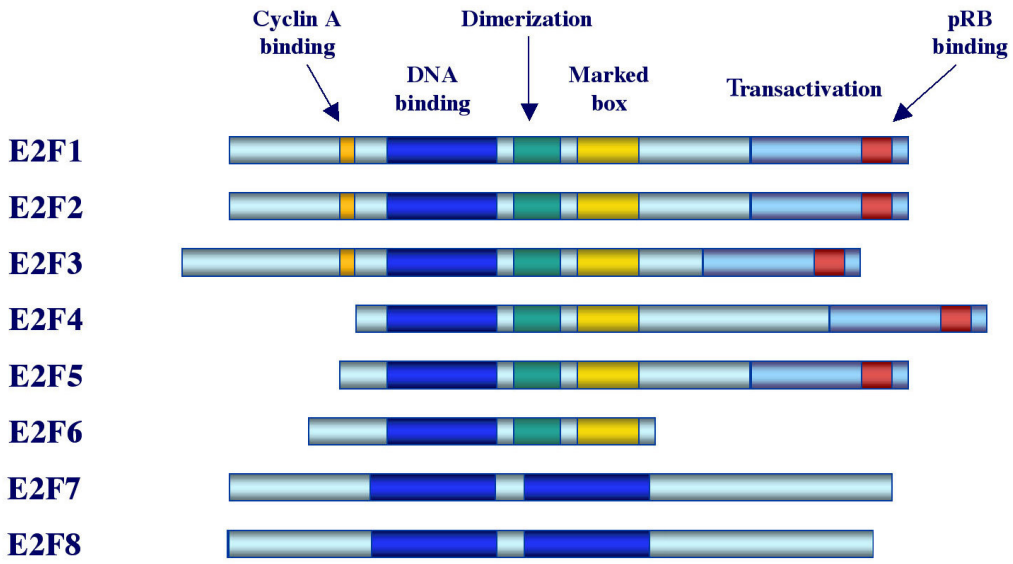


Fig 1.10: Schematic comparison of the conserved domains of the mammalian E2F family members.

1.9.2 Plant E2F factors

Following the initial discovery of pRB-related proteins, typical E2F factors were identified also in plants. Plant E2Fs were initially isolated from wheat (*TmE2F*), tobacco (*NtE2F*) and carrot (*DcE2F*) (Ramirez-Parra, 1999; Sekine et al, 1999; Albani et al, 2000). DNA binding

assays have revealed that, like the animal counterparts, also the typical plant E2Fs need to interact with DP partners in order to bind E2F consensus sites (Albani et al, 2000).

Shortly after these reports, studies in *Arabidopsis thaliana* allowed the identification of three typical E2Fs (AtE2Fa-c) and, for the first time, the discovery of three atypical E2Fs (AtE2Fd-f) in this plant (Mariconti et al, 2002). Further studies have revealed that the AtE2F factors can be divided in three groups on the ground of their structural and functional characteristics. AtE2Fa and AtE2Fb are considered activators of S-phase genes and stimulate cell division (Rossignol et al, 2002; De Veylder et al, 2002; Magyar et al, 2005; Sozzani et al, 2006). Different is the role of AtE2Fc, which features a truncated transactivation domain but possesses a conserved pRBR binding region in the C-terminus. By interacting with pRBR, AtE2Fc appears to be able to act as a negative regulator of cell division (del Pozo et al, 2002). The atypical AtE2Fd-f, due to the lack of transactivation domains, are believed to inhibit and/or repress the expression of specific target genes.

AtE2Fa plays an important role in cell cycle progression and is maximally expressed in G1 and early S-phase (Mariconti et al, 2002). It is a physiological activator of E2F responsive genes and transient expression of *AtE2Fa* induces quiescent leaf cells to enter S-phase (Rossignol et al, 2002). Using transgenic *Arabidopsis* plants it was shown that ectopic overexpression of a cDNA encoding AtE2Fa induces cell proliferation, as shown by the increased number of smaller epidermal cells, and increases also the level of endoreduplication (De Veylder e al, 2002). Moreover plants overexpressing ectopically *AtE2Fa* together with *AtDPa* show up-regulation of S-phase specific genes such as *AtDNApolα*, *AtCDC6* and *AtMCM*. Partly similar results were obtained when *AtE2Fa* and *AtDPa* cDNAs were overexpressed in tobacco transgenic plants (Kosugi and Ohashi, 2003).

Also *AtE2Fb* transcripts accumulate maximally at G1/S transition. In tobacco BY-2 cells, *AtE2Fb* overexpression shortens cell cycle, reduces cell size and promotes cell division in the absence of auxin (Magyar et al, 2005). AtE2Fb seems to be a strong activator of G2/M genes (Vandepoele et al, 2005). Additional studies in transgenic plants have also shown that *AtE2Fb* overexpression promotes cell division but, contrary to *AtE2Fa*, does not affect ploidy levels (Sozzani et al, 2006). These studies have revealed that AtE2Fa upregulates *AtE2Fb* which is likely to be main activator of cell proliferation (Magyar et al, 2005; Sozzani et al, 2006). Moreover, accumulation of AtE2Fb causes a diminuition of AtE2Fa protein suggesting an autoregulation mechanism.

Contrary to AtE2Fa and AtE2Fb, AtE2Fc acts as a negative regulator of cell division (del Pozo et al, 2002). Pulldown experiments revealed a strong interaction between AtE2Fc and

ZmRBR. Plants overexpressing *AtE2Fc* produced larger ad irregular cells confirming a cell division inhibitory role for this factor. Moreover AtE2Fc is able to downregulate the early S-phase gene *AtCDC6* suggesting that this factor might act as a repressor of genes required for cell cycle progression (del Pozo et al, 2002). Accumulation of AtE2Fc depends on the degradation by the ubiquitin-proteasome proteolytic pathway in response to light. Interestingly, AtE2Fc appears to be part of an autoregulation mechanism in which AtE2Fa activate the expression of *AtE2Fc* which together with *AtRBR* act as repressors of cellular proliferation. Moreover, it has been shown that the inactivation of pRBR or the downregulation of *AtE2Fc* (Desvoyes et al, 2006; Del Pozo et al, 2006) can increase *AtE2Fa* expression suggesting that the AtE2Fc/pRBR complex is also a negative regulator of the expression of *AtE2Fa*.

Like mammals E2F7 and 8, the atypical AtE2Fd- f factors contain duplicated DNA binding domains which allow these proteins to recognize and bind a consensus E2F site without the need to heterodimerize with a DP partner. The atypical AtE2Fs, which are also known as DEL1, DEL2, DEL3 (de Jager et al, 2001; Vandepoele et al, 2002) or E2L1, E2L2, E2L3 (Kosugi and Ohashi, 2002), are able to inhibiting the transactivation of a E2F-responsive gene by AtE2Fa and AtE2Fb (Mariconti et al 2002; Kosugi and Ohashi, 2002). These atypical E2Fs cannot activate gene expression; they are targeted towards the nucleus and interfere with the activity of the activating AtE2Fs competing for the same DNA target sites. *AtE2Fe/DEL1* controls the endocycle. Its overexpression decreases endoreduplication while, conversely, a *del1* insertional mutant is characterized by increased ploidy levels (Vlieghe et al, 2005). Conversely, *AtE2Ff* is not involved in cell cycle progression but rather controls cell elongation of root and hypocotyl cells. This factor binds E2F consensus sites present in the promoters of genes involved in cell wall biosynthesis thus inhibiting their transcription. AtE2Fe and f recognize only a subset of E2F-responsive genes confirming the selectivity of recognition of E2F *cis* elements by plant E2Fs (Egelkrout et al, 2002).

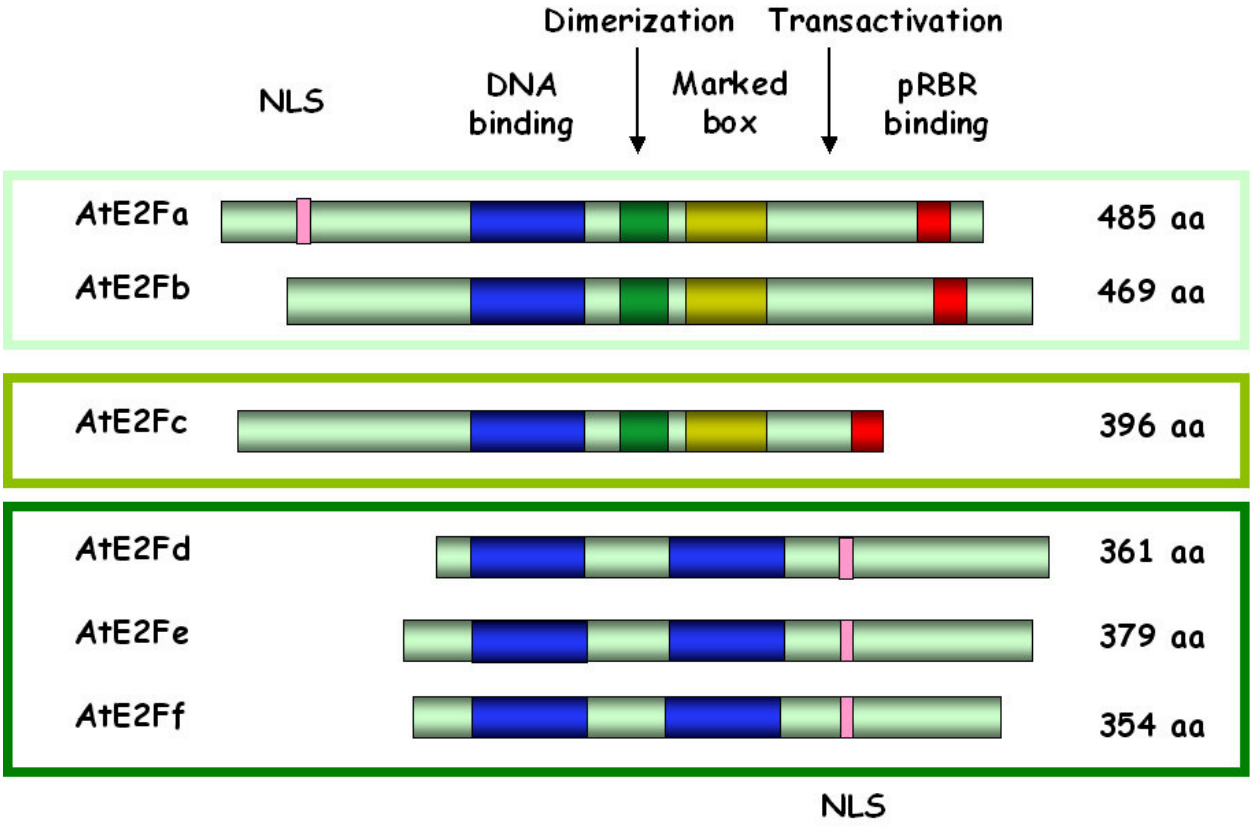


Fig 1.11: Structural features of the AtE2F proteins.

1.10 E2F-dependent regulation of gene expression

The E2F transcription factors are known to regulate genes that are implicated in DNA replication, cell-cycle regulation, defense response and signalling (Ramirez and Parra, 2003). Example of E2F responsive genes are PCNA, CDC6, RNR, and a few other such as CDKB1 (Chaboutè et al, 2000-2002; de Jager et al, 2001; Egelkrout et al, 2001-2002; Kosugi and Ohashi, 2002; Stevens et al, 2002; Boudolf et al, 2004).

The E2Fs are believed to recognize a canonical E2F cis element of sequence TTTSSCGS (S correspond to C or G) found in one or more copies (usually two) in the promoter region or in the 5'-UTR of E2F responsive genes. Conserved E2F binding sites are found in the promoters of a large number of potential E2F target genes. These E2F sites have been shown to mediate transcription activation or repression depending on the stage of development (Egelkrout et al, 2002).

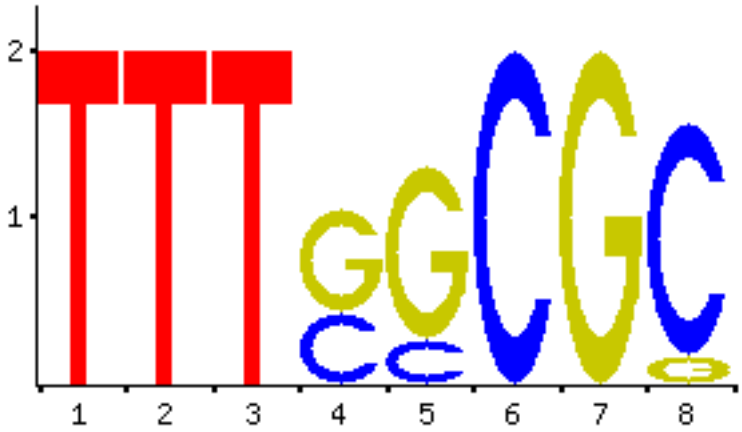


Fig. 1.12: Matrix of the E2F consensus site.

Most of the information concerning the function and binding site preference of E2F factors derives from studies on animal E2Fs whereas considerably less is known about how E2F factors regulate gene expression in plant. Plant promoters containing E2F consensus sites which have been investigated are the *RNR1*, *RNR2* and *PCNA* promoters of tobacco and the *CDC6* and *MCM3* promoters of *Arabidopsis thaliana*.

The promoter of the *RNR2* gene of tobacco was the first plant promoter shown to be regulated by E2F *cis* elements (Chaboutè et al, 2000). This promoter contains two E2F consensus sites involved in activation of expression at the G1/S transition. Both sites are necessary for up-regulation of the promoter but one of them behaved also like a repressor outside S phase (Chaboutè, et al 2000). Also the promoter of *Nicotiana benthamiana* PCNA gene has two E2F binding sites. Egelkrout et al. (2002) demonstrated that one these E2F elements activates the transcription in proliferating cells but the other is a repressor in differentiated tissues. E2F consensus sequences in the rice and tobacco PCNA promoters are involved in the activation of a reporter gene in both cultured cells and in whole plants (Kosugi and Ohashi, 2002). Also the *CDC6* promoter of *Arabidopsis thaliana* contains an E2F consensus site to which E2F factors have been shown to bind (de Jager et al, 2001).

By sequence analysis, Lincker et al. (2004) identified two E2F sites on the promoter of the *RNR1* gene of tobacco. The two elements interact with purified tobacco E2F factors as well as with specific nuclear complexes and functional analyses have revealed that both are important

for *RNR1* promoter induction at the G1/S transition in synchronized tobacco BY-2 cells as well as in response to UV irradiation in dividing non-synchronized cells and S-phase cells. Mutation of both elements decreased *RNR1* promoter activity in growing cells but other *cis*-elements present on the *RNR1* promoter might act as positive regulators. Indeed, a telo-box is found next to E2F sites in promoters of tobacco and Arabidopsis genes, such as PCNA, DHFR and RNR (Tremousaygue et al, 2003), that are similarly induced at the G1/S transition. This raises the possibility that the telo-box might be involved in a common regulation pathway in synergy with E2F for genes involved in DNA synthesis. The diversity of the E2F family in plants could reflect the complexity of the regulation of expression of various E2F-dependent genes. The six E2F proteins found in Arabidopsis can be classified as activators or repressors of gene expression (Mariconti et al, 2002, Kosugi and Ohashi 2002). Analysis of the specific E2F complexes associated with the E2F elements of *RNR1* and of other promoters could explain the mode of action of the various plant E2F factors in concert with specific co-regulators.

Other genes known to be under the control of E2F factors are those coding for the mini-chromosome maintenance (MCM) proteins which play important roles in the initiation of replication (Ohtani 1999). The promoter of the Arabidopsis *MCM3* gene is transcriptionally regulated at S-phase. The 5' region of this gene contains several E2F consensus binding sites two of which (D1 and D2) match the human consensus binding site. The D1 site is responsible for the G2-specific repression of promoter activity in synchronized cell suspensions, while the second site (D2) is crucial for the meristematic expression of the *MCM3* gene because deletion of this site eliminates the activity of the promoter in root and shoot meristems (Stevens et al., 2002). Also in mammalian cells, two highly similar E2F binding sites in the promoter of the *MCM3* gene are responsible for different cell cycle-regulated or developmental expression patterns depending on the cellular environment (Stevens and Farnham, 1999).

2.AIMS OF RESEARCH

The aim of this research is to study the patterns of expression of the three *DHFR-TS* genes of *Arabidopsis thaliana*, called *AtDRTS1*, *AtDRTS2* and *AtDRTS3*, and to analyse regulatory regions of these promoters. These genes are expected to play central roles in DNA precursor biosynthesis and are expected to be expressed in the G1/S phase of the cell cycle. Studies from *Daucus carota* indicate that the *DHFR-TS* transcripts accumulate in highly dividing meristematic tissues (Albani et al, 2005) and also northern hybridization experiments have confirmed a higher accumulation of *DHFR-TS* transcripts in proliferating suspension cells compared to cells in stationary phase. Thus, this thesis will focus mainly on the study of the meristematic expression of the *DHFR-TS* genes but will investigate also their expression during plant development. In particular, studies on the DNA cis-elements that could be involved in the cell cycle dependent regulation of expression of these genes will be carried out in order to identify regulatory regions responsible for the meristematic expression of plant genes.

3. MATERIAL AND METHODS

3.1 Plant material

Wild type or transgenic *Arabidopsis thaliana* ecotype Columbia seeds were surface sterilized for 8/12 hours in 2% v/v PPM® (Plant Preservative Mixture, Plant Cell Thnology) supplemented with 50 mg/L magnesium salts (magnesium sulphate). Seeds were imbibed over-night in 0,1 % agarose at 4°C in the dark and then germinated in a growth cabinet at 22°C under long day conditions of 16 h of light and 8 h of dark on petri plates with MS salts (Duchefa Biochemie) and Phyto Agar (8g/l), supplemented with Sucrose (10g/l). The transgenic *Arabidopsis* lines used in this study were generated by the floral dip method (Clough and Bent, 1998) using *Agrobacterium tumefaciens* GV3101/pMP90 and EHA105 strains (Koncz, C. and Schell, J., 1986). Progeny plants were selected on MS plates containing the resistence antibiotic (Hygromicin, 30mg/l) and PPM®. After two weeks on selection, the transformed plants were transferred to recovery plates without the selection agent and at four weeks of age transferred to soil and grown to maturity in growth cabinets set at long day conditions of 16 h of light (22±3°C) and 8 h of dark (22±3°C) with 70% relative humidity.

3.2 Isolation of the promoter regions and creation of the bidirectional AtDRTSs::GFPGUS/AtFSH::eqFP611 constructs

The promoter regions upstream of *AtDRTS1* (reported in BAC clone F16F14), *AtDRTS2* (found in BAC clone T4L20) and *AtDRTS3* (found in BAC clone F2G1) were amplified from *Arabidopsis* genomic DNA using primers designed to amplify the entire intergenic region comprised between the putative ATG codons of the *AtDRTS* and *AtSFH* genes. For the *AtDRTS2* promoter construct the second ATG codon, which is located in the fourth exon of the gene, was chosen. The sequence of the primers is reported below. To amplify the *AtDRTS1/AtSFH7* intergenic region the primer F16F1, which pairs next to the *SFH1* gene, contains an *Hin*dIII site (underlined) whereas the primer F16F2, pairing next to the *DRTS1* gene, contains a *Bam*HI site (underlined). For the *AtDRTS2/AtSFH1* promoter fragment, the primer T4L1, pairing next to the *SFH* gene, contains an *Xba*I site (underlined) overlapping a *Bgl*II site (shown in red), which will allow the insertion of a reporter gene, whereas the primer T4L2, which pairs next to the *DRTS* gene, contains a *Nco*I site. Finally, to amplify the

AtDRTS3/AtSFH3 intergenic region, the primer F2G1 pairs next to the *SFH* gene and contains a *Pst*I site (underlined) whereas the primer F2G2, pairing next to the *DRTS* gene, contains a *Nco*I site overlapping a *Msc*I site (shown in red).

AtDRTS1/AtSFH7 promoter primers
F16F1: 5' –ACAAAGCTTGATTGGTTTACATTAACATTTTAG-3'
F16F2: 5' –GTTGGATCCATTGTGGAAATCAAAACCTTG- 3'

AtDRTS2/AtSFH1 promoter primers
T4L1: 5' –TTTTCTAGATCTGGTTAGATGAGTTTTAAGCAC -3'
T4L2: 5' -TGTTTGCCATGGTTGAAATTGAAACCTTGA - 3'

AtDRTS3/AtSFH3 promoter primers
F2G1: 5'- ATCCTGCAGGGTTAAAGTCTGGATTTAAGATTT-3'
F2G2: 5' –CAGCCATGGCCATATTCTGAAACTTAA AAA TC-3'

Following PCR amplification, performed using annealing temperatures between 62°C and 64°C and using a high fidelity Taq (Pfx Polymerase, Invitrogen), the resulting PCR fragments were cloned into the pBS-KS and pGEM-T Easy vectors. The *AtDRTS1/AtSFH7* promoters fragment was cloned in pBS-KS as a *Hin*dIII-*Bam*HI fragment, to yield the pBS-F16F14 plasmid (based on the name of the corresponding BAC clone). The *AtDRTS2/AtSFH1* promoters were cloned in pGEM-T Easy as an *Xba*I-*Nco*I fragment (into the *Spe*I and *Nco*I sites of the polylinker), to yield the pGEM-T4L20 plasmid, whereas the *AtDRTS3/AtSFH3* promoters were cloned in pGEM-T Easy as a *Pst*I-*Nco*I fragment, to yield the pGEM-F2G1 plasmid. All the clones were then sequenced to verify the fidelity of the PCR reactions. These intergenic DNA fragments were used to produce bidirectional constructs in which a chimeric *GFP/GUS* reporter gene was placed downstream of the *AtDRTS* promoters whereas the reporter gene coding for the red fluorescent protein eqFP611 was placed downstream of the *AtSFH* promoters.

As a first step, the *AtDRTS* promoters were placed upstream of the *GFP/GUS* reporter gene contained in the binary vector pCAMBIA 1304. The pBS-F16F14 plasmid was digested with

*Hin*dIII-*Bam*HI and the resulting *AtDRTS1/AtSFH7* fragment was inserted into the *Hin*dIII-*Bgl*II sites of pCAMBIA 1304 giving rise to the pCAMBIA-F16F14 plasmid. The *AtDRTS2/AtSFH1* and *AtDRTS3/AtSFH3* fragments were isolated from the pGEM-T4L20 and pGEM-F2G1 plasmids digested with *Pst*I-*Nco*I and the resulting fragments were inserted into the corresponding sites of pCAMBIA 1304 to give rise to pCAMBIA-T4L20 and pCAMBIA-F2G1, respectively.

As a second step, the reporter gene coding for the red fluorescent protein eqFP611 was placed downstream of the *AtSFH1* and *AtSFH7* promoters. To obtain suitable fragments containing the *eqFP611* sequence, the corresponding cDNA inserted in the original pQ32 vector (Wiedenmann et al, 2000) was amplified using different sets of primers which create new restriction sites at the ends of the fragment (shown underlined in the table below).

eqFP611 primers
eqFP1: 5'-CCAGTCGACATGAACTCACTGATCAAGGAAAA-3'
eqFP2: 5'-CCAGGTACCTCAAAGACGTCCCAGTTTGG-3'
eqFP3: 5'-GAAGATCTAGAATGAACTCACTGATCAAGGAAA-3'
eqFP4: 5'-CCAGTCGACTCAAAGACGTCCCAGTTTGG-3'

The primer eqFP1 creates a *Sal*I site at the 5'end of *eqFP611* whereas the primer eqFP2 introduces a *Kpn*I site at the 3'end of the sequence. Conversely, the primer eqFP3 introduces a *Bgl*II site overlapping an *Xba*I site (shown in red) at the 5' end whereas eqFP4 creates a *Sal*I site at the 3'end. The fragment amplified using the eqFP1 and eqFP2 primers was digested with *Sal*I and *Kpn*I and inserted into the corresponding sites of pBS-F16F14 in order to obtain pBS-F16F14eq whereas the PCR fragment obtained with the eqFP3 and eqFP4 primers was inserted into the *Bgl*II and *Sal*I sites of the pGEM-T4L20 plasmid to yield pGEM-T4L20eq.

To place a terminator cassette dowstream of the *eqFP611* sequences, the promoter fragments with the linked *eqFP611* gene were transferred into the pFF19 vector (Timmermans et al.,1990), which contains a 35S terminator sequence. The *AtDRTS1/AtSFH7::eqFP611* promoter fusion was cloned as an *Xba*I/blunt-*Kpn*I fragment into the *Hin*dIII/blunt-*Kpn*I sites of pFF19, obtaining the pFF-F16F14eq-Ter plasmid, whereas the

AtDRTS2/AtSFH1::eqFP611 promoter fusion was cloned as an *Apa*I/blunt-*Sal*I fragment into the *Hin*dIII/blunt-*Sal*I sites of pFF19 to give rise to the pFF-T4L20eq-Ter plasmid.

Finally, the bidirectional constructs, containing the *GFP/GUS* reporter gene downstream of the *AtDRTS* promoters and the *eqFP611* reporter gene downstream of the *AtSFH* promoters, were assembled. The *Bgl*II/*Eco*RI fragment isolated from the pFF-F16F14eq-Ter plasmid was inserted into the corresponding sites of pCAMBIA-F16F14 to give rise to the pSFH7/DRTS1-GGEQ vector which contains the *AtDRTS1::GFPGUS/AtSFH7::eqFP611* construct. Similarly, the *Bam*HI/*Eco*RI fragment isolated from the pFF-T4L20eq-Ter plasmid was inserted into the corresponding sites of pCAMBIA-T4L20 to give rise to the pSFH1/DRTS2-GGEQ vector containing the *AtDRTS2::GFPGUS/AtSFH1::eqFP611* construct. Finally, the pSFH3/DRTS3-GGEQ vector carrying the *AtDRTS3::GFPGUS/AtSFH3::eqFP611* construct, was obtained inserting the *Bgl*II/*Eco*RI fragment from the pFF-T4L20eq-Ter plasmid, carrying the *eqFP611* sequence linked to the 35S terminator, into the *Bam*HI/*Eco*RI sites of the polylinker found downstream of the *AtSFH* promoter in the pCAMBIA-F2G1 vector. The final binary vectors were then transferred to *Agrobacterium tumefaciens* and used for the stable transformation of *Arabidopsis* plants.

3.3 Chromatin Immunoprecipitation

Chromatin Immunoprecipitation (ChIP) assays were performed using nuclei extracted from suspension-cultured *Arabidopsis* T87 cells. Nuclei, extracted as previously described (Albani et al., 2000), were treated with 1% formaldehyde at 22°C for 10 min and the cross-linking was stopped by the addition of 0.125 M glycine. Fixed nuclei were resuspended in SDS buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, and 1% SDS) and sonicated to shear DNA in order to obtain DNA fragments of 600-1000 bp. To reduce false positives, sonicated chromatin samples were pre-incubated with 20 µL of preimmune serum for 1 h at 4°C with gentle mixing, transferred to a new tube with 20 µL of Protein A-Sepharose (50% slurry in 15 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, triton X-100 1%,) and incubated with gentle mixing for 1 h at 4°C. Samples were centrifuged at 13000 rpm for 2 min at 4°C. The resulting supernatant was specifically immunoprecipitated with 20 µL of anti-DcE2F serum and further incubated for 2 h at 4°C with gentle mixing. Immunocomplexes were recovered using 20 µL protein A Sepharose (50% slurry) for 2 h at room temperature with gentle mixing, extensively washed and eluted from beads. The immunoprecipitated chromatin was incubated for 5 h at 65°C, added with two volumes of ethanol and centrifuged: the resulting pellet was incubated

with proteinase K (18.5 mg/mL) for 2 h at 42°C and extracted with phenol/chloroform. Upon ethanol precipitation, DNA was resuspended in 10 µL of water and 1 µL was used for PCR analysis using the following primers adjacent to the E2F *cis*-elements.

AtDRTS2 ChIP primers
FT5: ACATCACCACATGGTGAT
FT3: TCAGAGTGAATCTACGCA

AtDRTS3 ChIP primers
F25: TTATGTGGTTGCTCCACT
F23: TATTGCTGCCACTGGATT

AtRNR1b ChIP primers
RN5: AATGGGCTTTAACTCTCTAA
RN3: AAGGGATTTGAAGATTTG

3.4 Mutation of the E2F sites in the AtDRTS2 and AtDRTS3 promoters

Mutations of the E2F binding sites in the *AtDRTS2* and *AtDRTS3* promoters were created by PCR. Portions of the cloned genomic fragments flanking both sides of the E2F site were amplified using primers pairing over the E2F site that introduce the *Eco*RI restriction site in place of the SSCGSS sequence, in combination with universal primers pairing next to the polylinker in the vector. Listed below are the universal and the promoter-specific primers with the underlined nucleotides indicating the *EcoRI* restriction site introduced at the E2F site.

Universal primers
M13FW: 5'-ACGTTGTAAAACGACGGC-3'
M13RV: 5'- GGAAACAGCTATGACCATG-3'

AtDRTS2-ΔE2F *promoter primers*
T4L6: 5'- TCTGAATTCGTTTATACCCTCTCCGAAGC-3'
T4L7: 5'- AACGAATTCAGACGACGGCGACTGAGTCA-3'

AtDRTS3-ΔE2F *promoter primers*
E2F-EN: 5'- CGTGAATTCAACCCGTCCGTAAAAAACTAT-3'
E2F-PE: 5'-GTTGAATTCACGAGTTTGACAGGAAGTTAC-3'

Following amplification of the pGEM-T4L20 and pGEM-F2G1 templates, the PCR fragments were subcloned into the pBS-KS vector. For the *AtDRTS2/AtSFH1* promoters, the portion toward the *DRTS* gene was cloned as an *Eco*RI-*Apa*I fragment whereas the portion toward the *SFH* gene was cloned as an *Eco*RI-*Bam*HI fragment. For the *AtDRTS3/AtSFH1* promoters, the portion toward the *DRTS* gene was cloned as an *Eco*RI-*Apa*I fragment whereas the portion toward the *SFH* gene was cloned as an *Eco*RI-*Sac*I fragment.

To reconstruct the mutated *AtDRTS2/AtSFH1* and *AtDRTS3/AtSFH3* intergenic regions, the EcoRI-*Apa*I fragments containing the portion flanking the *DRTS* genes were inserted into the same sites of the plasmids carrying the portion flanking the *SFH* genes, thus giving rise to the pBS-SFH1/DRTS2-ΔE2F and pBS-SFH3/DRTS3-ΔE2F plasmids. The mutated promoters were then used to replace the wild type sequences in the pSFH1/DRTS2-GGEQ and pSFH3/DRTS3-GGEQ vectors as a *Bam*HI-*Nco*I fragment for the *DRTS2* promoter and as a *Kpn*I-*Nco*I fragment for the *DRTS3* promoter generating the pSFH1/DRTS2-ΔE2F and pSFH3/DRTS3-ΔE2F vectors.

3.5 Mutation of the cis-acting elements Up1 and Up2 in the AtDRTS2 promoter

Mutations of the Up1 and Up2 *cis*-acting elements located in the 5' flanking region and at the beginning of the first intron of *AtDRTS2* were carried out by PCR following the same strategy used for the E2F site. In this case, the PCR reactions were performed using primers that introduce the *Cla*I restriction site in place of the Up1 site and the *Eco*RV restriction site in place of the Up2 site. The sequence of the primers, with the underlined nucleotides indicating the mutation introduced, is reported below.

Up1 site
T4L10: 5'- TGAATCGATATATAATAATGGGTCGATGC-3'
T4L11: 5'- TATATCGATTCAACTGGACCGAAAATAGG-3'

Up2 site
T4L14: 5'-GTTGATATCTTGGAATCTCAGTTTTTTTTTTG-3'
T4L15: 5'-CAAGATATCAACCTTGAGACGAAATCAAAC-3'

Following amplification of the pGEM-T4L20 template using these primers along with suitable universal primers pairing next to the polylinker in the vector, the PCR fragments were subcloned into the pBS-KS vector. For the mutation of Up1, the portion toward the *DRTS* gene was cloned as a *Cla*I-*Apa*I fragment whereas the portion toward the *SFH* gene was cloned as a *Cla*I-*Bam*HI fragment. For the *AtDRTS3/AtSFH1* promoters, the portion toward the *DRTS* gene was cloned as an *Eco*RV-*Apa*I fragment whereas the portion toward the *SFH* gene was cloned as an *Eco*RV-*Sac*I fragment. To reconstruct the mutated *AtDRTS2* promoters, the *Cla*I-*Apa*I and *Eco*RV-*Apa*I fragments containing the portion at the 5'end of *AtDRTS2* were inserted into the corresponding restriction sites of the plasmids carrying the promoter portion flanking the *AtSFH1* gene, thus giving rise to the pBS-SFH1/DRTS2-ΔUP1 and pBS-SFH1/DRTS2-ΔUP2 plasmids. The mutated promoters were then used to replace the wild type sequences in the pSFH1/DRTS2-GGEQ vector as a *Bam*HI-*Nco*I fragment generating the pSFH1/DRTS2-ΔUP1 and pSFH1/DRTS2-ΔUP2 vectors.

3.6 Construction of the AtDRTS2::GFPGUS vectors

To create *AtDRTS2* promoter cassettes lacking the ATG codon of the gene and suitable for a meristematic expression of exogenous genes in transgenic plants, two promoter constructs containing only the *AtDRTS* promoter were prepared by PCR. The first construct lacks the intron with the Up2 site and was obtained amplifying a portion of the SFH1/DRTS2 intergenic region with the distal primer T4L3, which anneals at position -830 relative to the transcription start of AtDRTS2, together with the downstream primer T4L8, which pairs at position +51, near the end of the first exon. Conversely, the second construct contains the first

intron of *AtDRTS2* and was obtained amplifying the SFH1/DRTS2 intergenic region with the primer T4L3 together with the downstream primer T4L9, which pairs at position +175 in the second exon, just upstream of a spurious ATG triplet found in the 5' UTR of *AtDRTS2* which, if mantained, could decrease the translational efficiency of the construct. To facilitate the cloning of the fragments and the assembling of the constructs, the T4L3 primer contains a *Hin*dIII site at the 5'end whereas the T4L8 and T4L9 primers contain an *Eco*RV site. The sequence of the primers with the restriction sites underlined is reported below.

AtDRTS2 promoter
T4L3 5'-AACAAGCTTGAAGCAAGGTGAAAAACATGT-3'
T4L8 5'- GACGATATCAAACGTAGCGTCTTGGC-3'
T4L9 5'-TCCGATATCAGAGTGAATCTACGCATT-3'

The two promoter fragments were then cloned into the *Hin*dIII-*Eco*RV sites of the pBS-KS vector and sequenced to verify the fildelity of the PCR. The final constructs were assembled by inserting the *AtDRTS2* promoter fragments into the *Hin*dIII-*Pst*I sites of the polylinker upstream of the *GUS* reporter gene in the binary vector pCAMBIA-1391 to give rise to the pDRTS2(-830/+51)GUS and pDRTS2(-830/+175)GUS vectors.

3.7 Linker scanning mutagenesis of the first AtDRTS2 intron

To identify the regulatory region in the first intron of *AtDRTS2* responsible for the meristematic expression of the gene, a linker scanning mutational analysis was carried out replacing the 8 internal nucleotides of ten serial segments of 10 bp which span from the Up2 site, next to the splice donor site, to the 3' end of the intron. These changes were created by PCR , as in the case of the E2F and Up sites, introducing specific restriction sites in each mutated region. The specific primers are listed below with the underlined nucleotides indicating the *Nsi*I (ATGCAT), *Nru*I (TCGCGA), *Nde*I (CATATG)and *Sph*I (GCATGC) restriction sites introduced in the mutations.

Site I
T4LinkI 5'-TTTAATGCATACAGTTTTTTTTTTGTAAAA-3'
T4LinkIb 5'-CTGTATGCATTAAACCCTAACCTTGAGACG-3'

Site II
T4LinkII 5'- CTCTTCGCGAATTTTGTAAAAAAACAATAT-3'
T4LinkIIb 5'-AAATTCGCGAAGAGATTCCAAAACCCTAAC-3'

Site III
T4LinkIII 5' TTTAGATATCTAAAACAATATGAATGCTGA-3'
T4LinkIIIb 5'-TTTAGATATCTAAAAAAACTGAGATTCCAA-3'

Site IV
T4LinkIV 5'- AAGCATGCGCTTGAATGCTGAATTTGTTTC-3'
T4LinkIVb 5'-TCAAGCGCATGCTTTACAAAAAAAAAACTG-3'

Site V
T4LinkV 5' AAAACAATATCCTGCAGCAATTTGTTTCTG-3'
T4LinkVb 5'-ATTGCTGCAGGATATTGTTTTTTTACAAAA-3'

Site VI
T4LinkVI 5'- TGAATGCTGATGCATGCACTGTCGATTTTG-3'
T4LinkVIb 5'-CAGTGCATGCATCAGCATTCATATTGTTTT-3'

Site VII
T4LinkVII 5'-AATTTGTTTCATCGCGAATTGCCTCTGTTT-3'
T4LinkVIIb 5'-CAATTCGCGATGAAACAAATTCAGCATTCA-3'

Site VIII
T4LinkVIII 5'- CTGTCGATTTAGATATCCTTTGCAATCATG-3'

T4LinkVIIIb 5'-AAAGGATATCTAAATCGACAGAAACAAATT-3'

Site IX
T4LinkIX 5'- TTGCCTCTGTAGCATGCGATGAATGCGTAG-3'
T4LinkIXb 5'-CATCGCATGCTACAGAGGCAAAAATCGACA-3'

Site X
T4LinkX 5'- TCAAATGCATCTAGATTCACTCTGAATGGG-3'
T4LinkXb 5' CTAGATGCATTTGATTGCTTTCAGAGGCAA-3'

3.8 Phenotypical analyses of transgenic plants

Histochemical detection of GUS activity was performed on *Arabidopsis* transgenic plants at different developmental stages using 5-bromo-4-chloro- 3-indolyl-β-D-glucuronide (X-Gluc) (Jefferson et al., 1987). Plants were incubated in the GUS solution (50 mM pH 7 phosphate buffer, 1 mg/mL X-Gluc, 1 mM potassium ferricyanide) for at least 1 h at 37°C. After staining, chlorophyll interference was removed treating the samples in 70% ethanol.

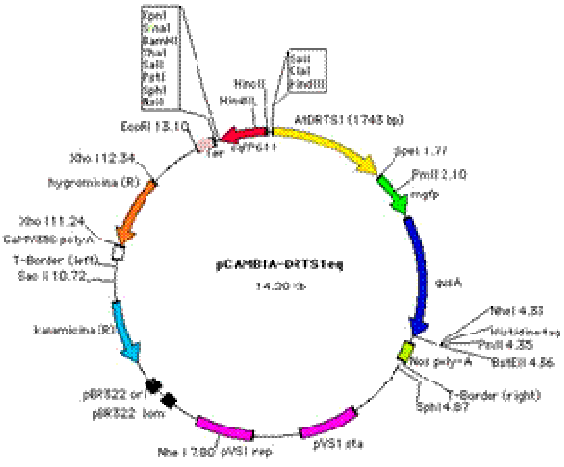
For quantitative analyses, the of GUS activity was detected fluorimetrically using the fluorogenic substrate MUG (4-metil umbelliferil–glucuronide). For this analysis, plantlets of *Arabidopsis thaliana* at the same developmental stage (emergence of the third pair of leaves) were ground in 200 µl of extraction buffer (50 mM NaPO4 pH 7, 10 mM EDTA, 0.1% Triton, 0.1% Sodium Lauryl Sarcosine, 10 mM β-Mercaptoethanol). The homogenate was then centrifuged in a microfuge at 15000 rpm for 10 min at 4°C and 100 µl of the surnatant were recovered and stored at -80°C. To perform the fluorimetric assays, 30 µl of extracts were added to 270 µl of assay buffer (50 mM NaPO4 pH 7, 10 mM EDTA, 0.1% Triton, 0.1% Sodium Lauryl Sarcosine, 10 mM β-Mercaptoethanol, 1mM MUG) and the reactions were incubated at 37 °C. At three different time points, 100 µl of the reaction was transferred to 900 µl of stop buffer (0.2 M Na2CO3) and the amounts of 4MU produced was measured using a fluorimeter.

4. RESULTS

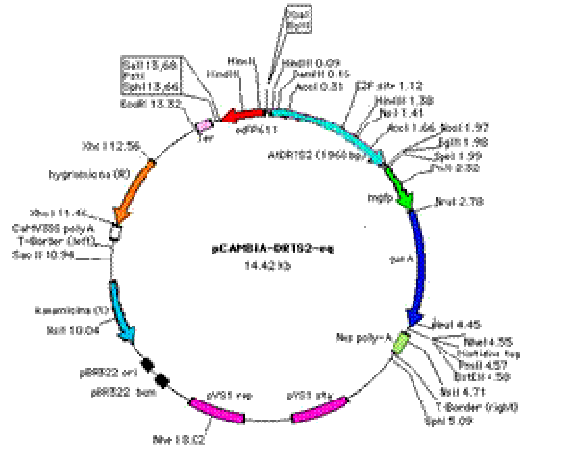
4.1 Functional analyses of AtDRTS promoters reveal distinct patterns of activity

To study the pattern of expression of the *AtDRTS* genes and to verify eventual correlations with the expression of the neighboring *AtSFH* genes, divergent reporter constructs comprising the entire intergenic region located between the putative ATG codons of the *AtDRTS* and *AtSFH* genes were assembled into a pCAMBIA binary vector suitable for Agrobacterium-mediated plant transformation. In these dual promoter constructs a chimeric *gfp/gus* reporter gene was placed under the control of the *DHFR-TS* promoters whereas the *eqFP611* red fluorescent protein gene was placed downstream of the *SFH* promoters. Figure 4.1 shows the maps of the resulting binary plasmids containing the *DRTS1/SFH7* intergenic region of 1743 bp, the *DRTS2/SFH1* intergenic region of 1968 bp or the *DRTS3/SFH3* intergenic region of 3480 bp, flanked by the *gfp/gus* reporter gene and by the *eqFP611* coding region.

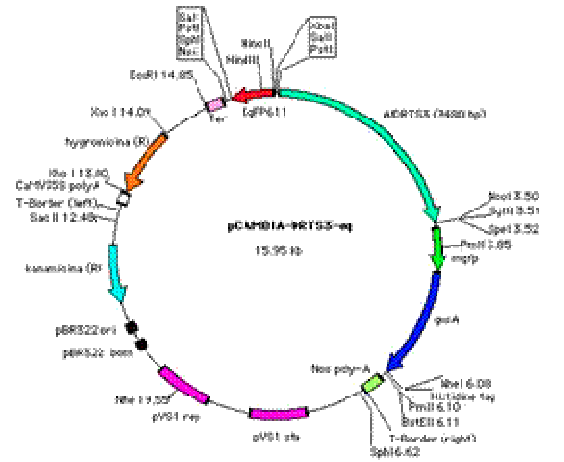
Following *Arabidopsis* transformation, several primary (T1) transformants were obtained. These plants were grown to maturity and allowed to self in order to produce T2 seedlings. To study the patterns of expression of the promoters of the three *DHFR-TS* and *SFH* genes histochemical GUS assays and fluorescence analyses were performed on the transgenic T2 lines.



pCAMBIA-DRTS1 eq



pCAMBIA-DRTS2 eq



pCAMBIA-DRTS3 eq

Figure 4.1. Structure of the divergent promoters constructs. The DRTS1/SFH7 intergenic region is coloured yellow, DRTS2/SFH1 is light blue and DRT3/SFH3 is green.

4.1.1 Patterns of expression conferred by the AtDRTS promoters

To verify the activity of the three *AtDRTS* promoters in transgenic lines of the T2 generation histochemical GUS staining at different stages of development were carried out. As shown in figure 4.2, the histochemical GUS assays of plants transformed with the pCAMBIA-DRTS1/SFH7 construct revealed that the *AtDRTS1* promoter is able to confer a widespread pattern of expression, with particularly strong activity in the vascular tissues of plantlets at one, two and three weeks of age (figure 4.2 panels a, b, c). In the flowers of mature plants the expression is also strong in the styles and in the vascular tissue of filaments and sepals (panels d and e). In mature plants the GUS protein is accumulatig also in siliques, especially in the funiculus, but no clear GUS staining can be found in the seeds (figure 4.2 panel e). Overall, this pattern of activity show good agreement with the microarrays data described in the Genevestigator database (https://www.genevestigator.com/) that suggest a wide expression of *AtDRTS1* in all tissues and organs examined (figure 4.3).

In comparison, as shown in figure 4.4, the *AtDRTS2* promoter appears to drive a narrower pattern of expression in the pCAMBIA-DRTS2/SFH1 lines in which the GUS activity is mainly restricted to organs with proliferating cells, such as shoot and root apical meristems and the basal part of developing leaves, whereas mature leafs show only very low GUS activity. This pattern of expression is constant during development as seen in one, two and three weeks old plants (figure 4.4 panels a, b, c). A clear promoter activity is detected also in the ovaries of flowers and in developing seeds of the siliques (panels d, e, f). Microarrays data partly confirm this pattern of expression showing high accumulation of *AtDRTS2* mRNA in seeds, inflorescences, seedlings and roots, whereas in flowers and rosettes the expression appears to be low (figure 4.5).

Concerning the third *AtDRTS* promoter, the analysis of lines transformed with the the pCAMBIA-DRTS3/SFH3 construct revealed that the *AtDRTS3* promoter can confer strong expression in root caps and central cylinders (figure 4.6). A strong GUS activity is found also in shoot apical meristems but not in the root meristems of one, two and three weeks old plants (panels a, b, c). In mature plants, low expression is detected in ovaries and styles (panels d, e) whereas no expression is detected in siliques (panel f). Interestingly, although the Affimetrix genechips used in microarray analyses include probe sets corresponding to the 3'end of the gene models which, as described previously, in the case of AtDRTS3 is not believed to be

transcribed, microarrays data available in the Genevestigator database confirmed a strong expression of *AtDRTS3* in root caps (figure 4.7).

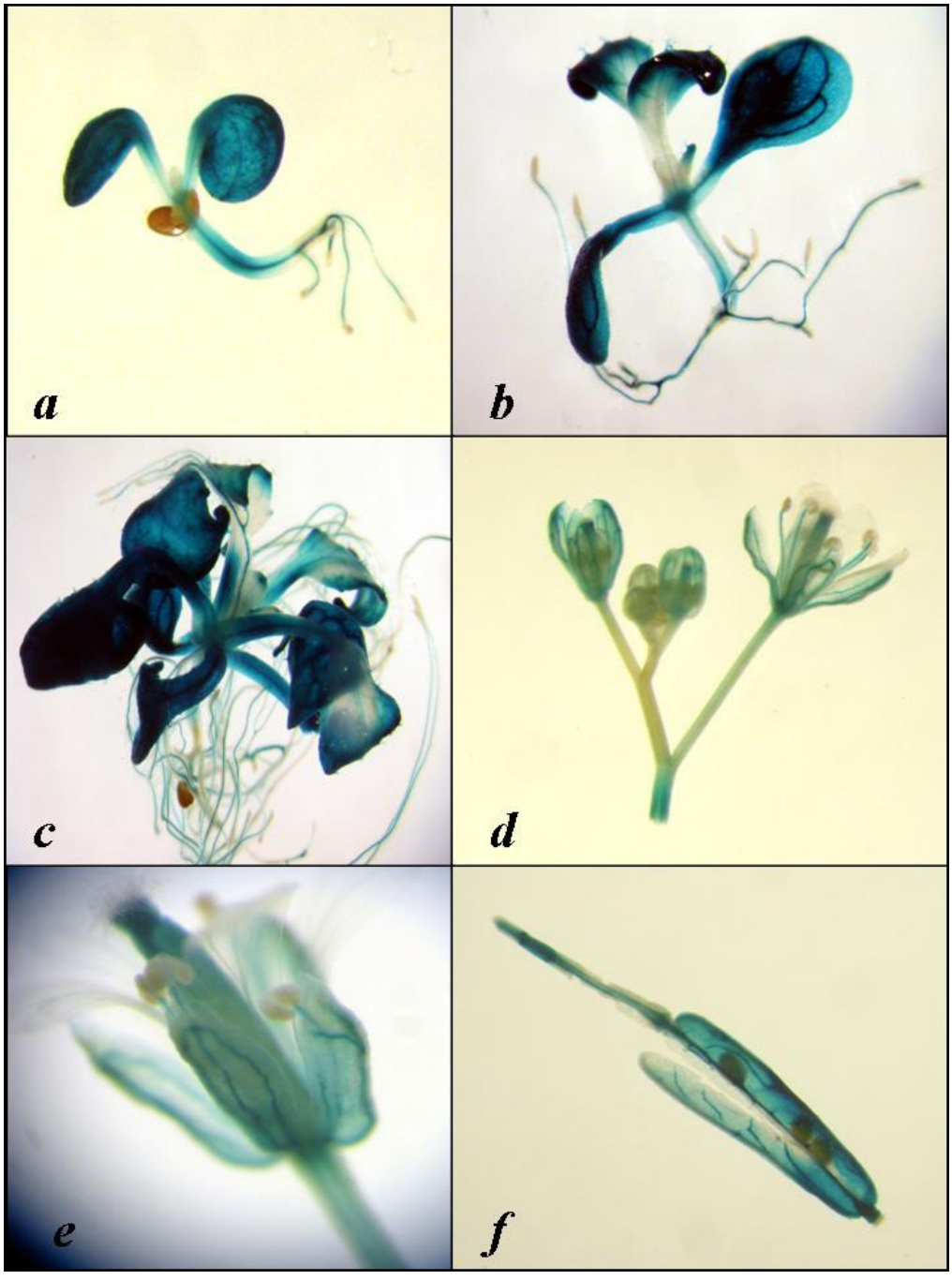


Figure 4.2. Histochemical activity of the AtDRTS1 promoter in pCAMBIA-DRTS1/SFH7 lines. Panels a, b, c: patterns of GUS staining in plantlets of one, two and three weeks, respectively. Panels d and e: GUS staining in inflorescences and flowers. Panel f: GUS staining in siliques.

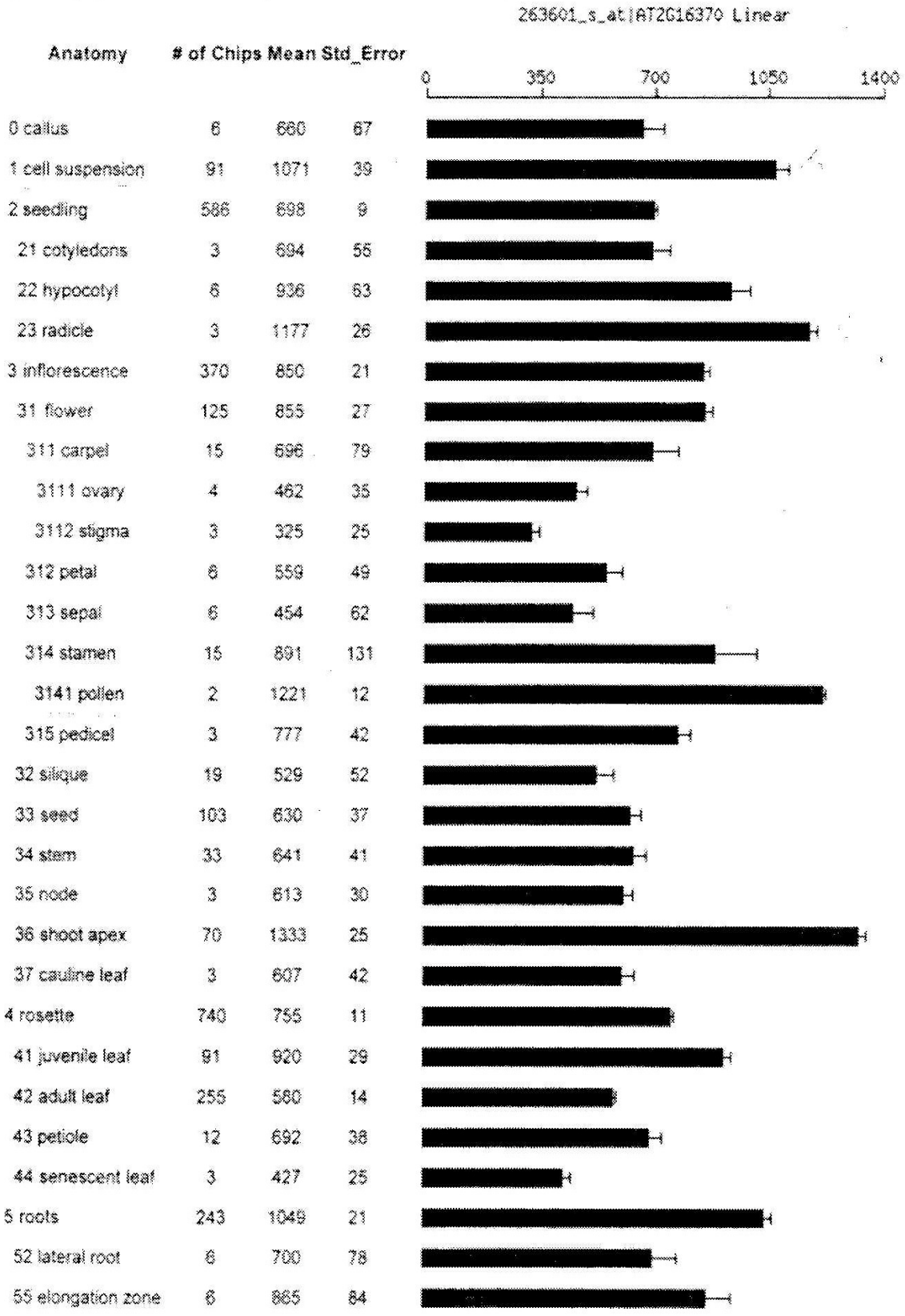


Figure 4.3: Expression of the AtDRTS1 gene according to the microarrays data reported in the Genevestigator database

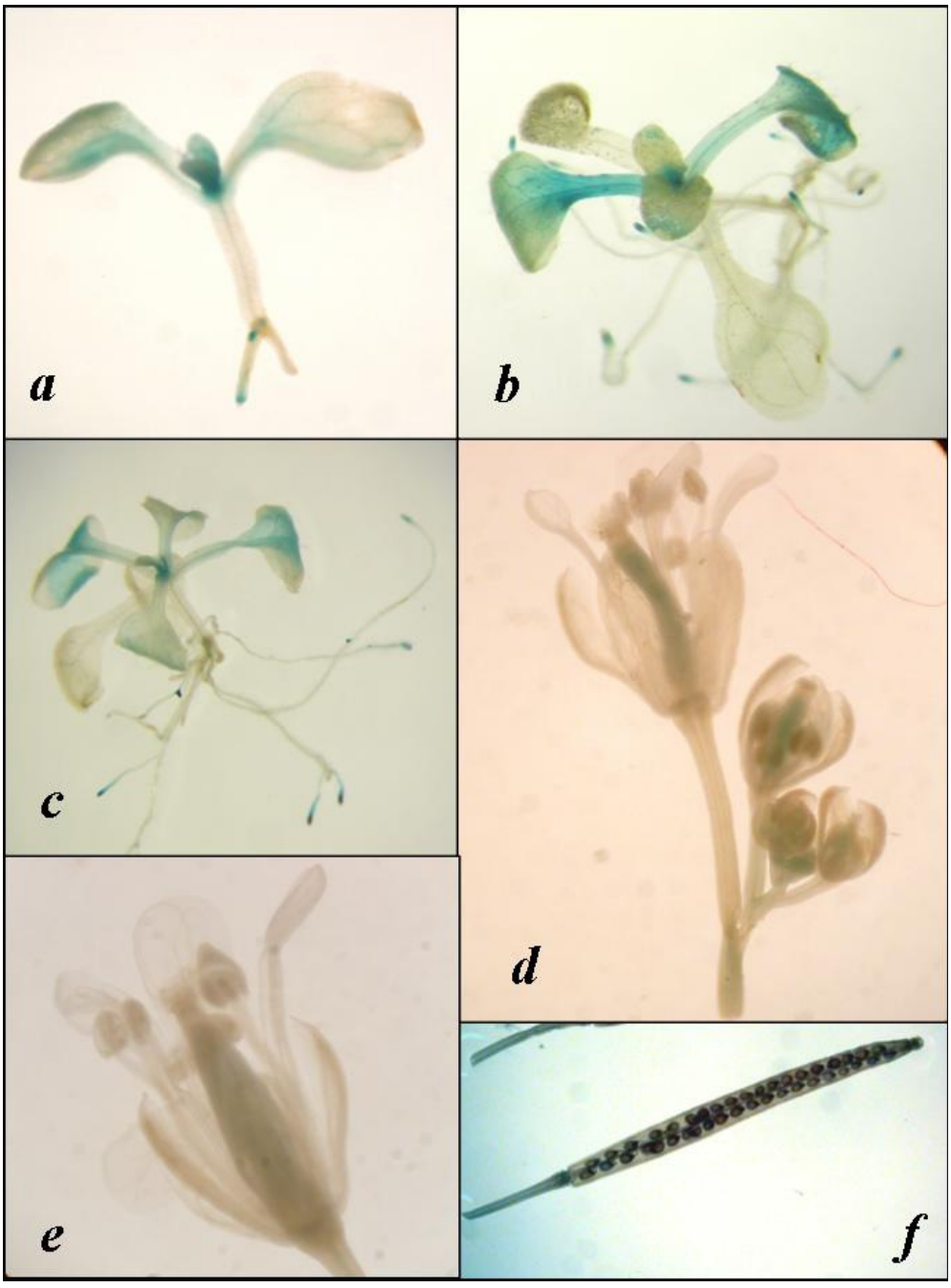


Figure 4.4. Histochemical activity of the AtDRTS2 promoter in pCAMBIA-DRTS2/SFH1 lines.. Panels a, b, c: patterns of GUS staining in plantlets of one, two and three weeks, respectively. Panels d and e: GUS staining in inflorescences and flowers. Panel f: GUS staining in siliques.

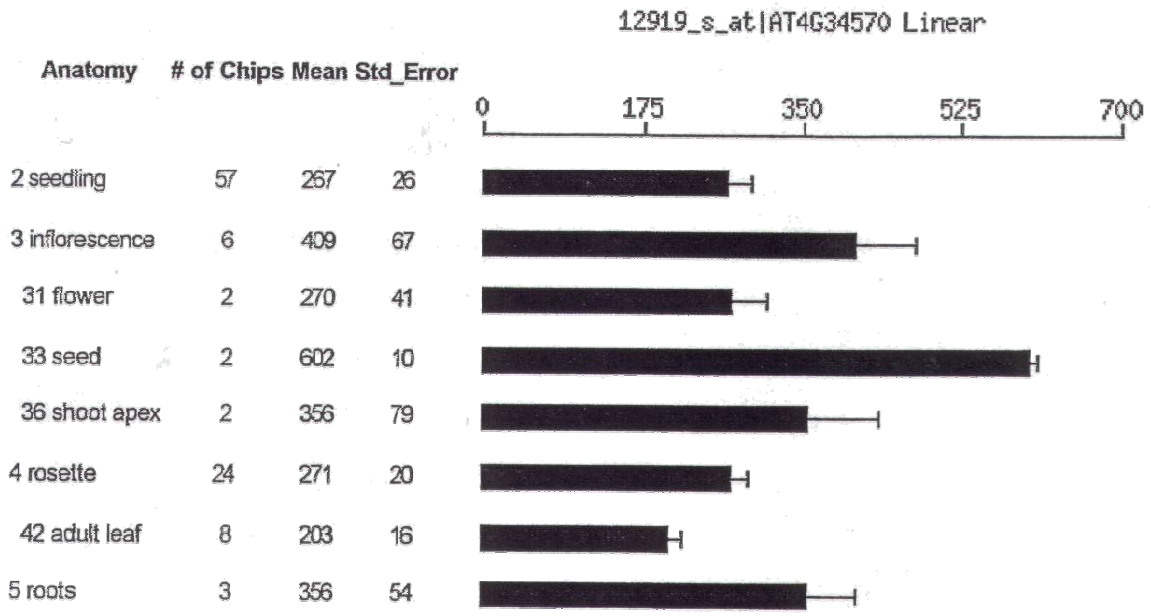


Figure 4.5. Expression of the AtDRTS2 gene according to the microarrays data reported in the Genevestigator database

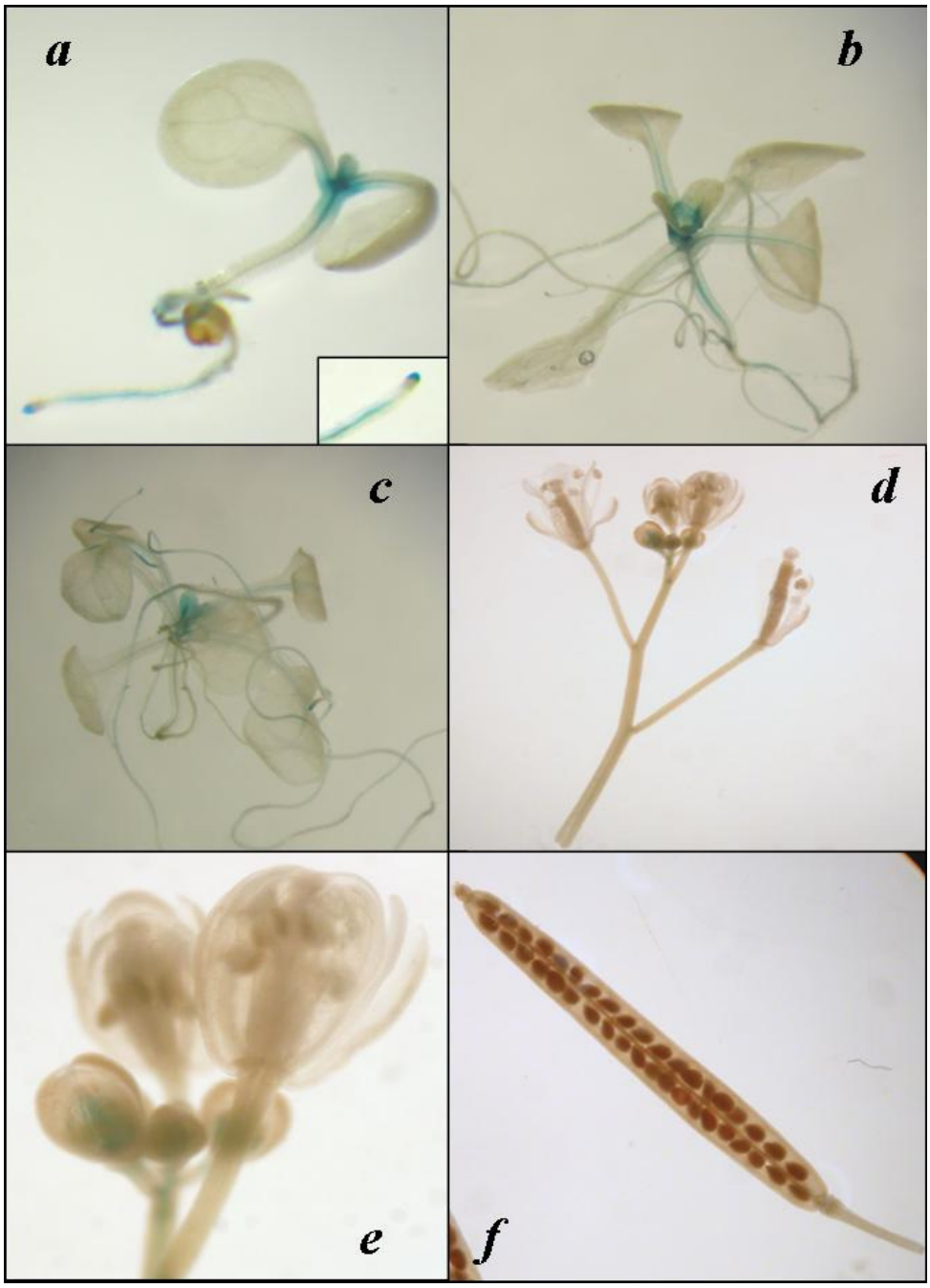


Figure 4.6. Histochemical activity of the AtDRTS3 promoter in plantlets of one, two and three weeks (panels a, b and c), in inflorescences and flowers (panels d and e) and in siliques (panel f).

Anatomy	# of Chips	Mean	Std_Error
0 callus	6	388	28
1 cell suspension	91	315	20
2 seedling	710	309	4
21 cotyledons	9	184	28
22 hypocotyl	6	248	11
23 radicle	9	551	53
3 inflorescence	460	188	5
31 flower	125	153	7
311 carpel	15	127	9
3111 ovary	4	112	24
3112 stigma	3	137	5
312 petal	6	94	10
313 sepal	6	76	14
314 stamen	15	251	32
3141 pollen	2	282	17
315 pedicel	3	131	18
32 silique	19	158	18
33 seed	103	152	9
34 stem	33	282	31
35 node	3	266	9
36 shoot apex	116	253	6
37 cauline leaf	3	129	29
4 rosette	828	202	3
41 juvenile leaf	91	230	10
42 adult leaf	274	199	6
43 petiole	12	383	51
44 senescent leaf	3	237	19
45 hypocotyl	36	225	14
451 xylem	9	170	9
452 cork	9	343	10
5 roots	260	614	15
52 lateral root	6	708	120
53 root tip	4	555	67
54 elongation zone	10	621	105
55 root hair zone	4	490	46
56 endodermis	3	254	80
57 endodermis+cortex	3	238	4
58 epid. atrichoblasts	3	348	60
59 lateral root cap	3	997	184
60 stele	3	338	25

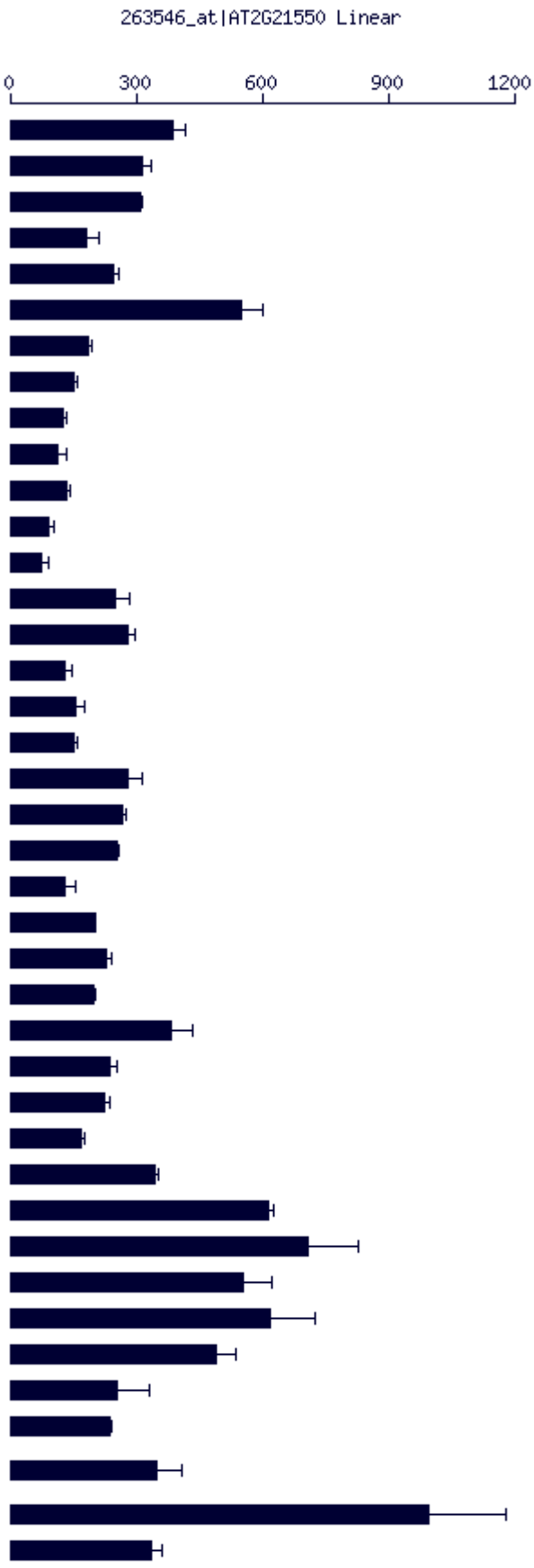


Figure 4.7. Expression of the AtDRTS3 gene according to the microarrays data reported in the Genevestigator database

4.1.2 Patterns of activity of the three AtSFH promoters

Because the three bidirectional reporter constructs included the *AtSFH* promoters driving expression of the *eqFP611* red fluorescent protein gene, fluorescence analyses of primary transformants and T2 progeny were performed to determine the activity of the three *AtSFH* promoters and to verify eventual correlations in the expression of the *SFH* and *DHFR-TS* genes.

Concerning the activity of the *AtSFH7* promoter, located upstream of *AtDRTS1*, analyses of the transgenic plants by fluorescence microscopy did not allow us to detect any accumulation of the eqFP611 red fluorescent protein. However, microarrays data included in the Genevestigator database suggest a weak but widespread expression of the *AtSFH7* gene. Thus, it is possible that the level of activity of the *AtSFH7* promoter is too general and weak to allow its detection using a fluorescent protein as reporter gene. Conversely, confirming results published in other studies, specific patterns of expression were detected for the other *AtSFH* promoters: *AtSFH1* and *AtSFH3,* located upstream of *AtDRT2* and *AtDRTS3* respectively. The *AtSFH1/cow1* gene has been shown to be strongly and specifically expressed in roots where its activity is required for polarized membrane growth during the development of root hairs (Bohme et al, 2004; Vincent et aol, 2005) whereas the *AtSFH3* gene has been shown to be highly expressed in mature and germinating pollen (Mo et al, 2007). Accordingly, strong red fluorescence was detected only in the roots or in the pollen of our transgenic plants (figure 4.8). These results are also consistent with the microarray data deposited in the Genevestigator database (https://www.genevestigator.com/).

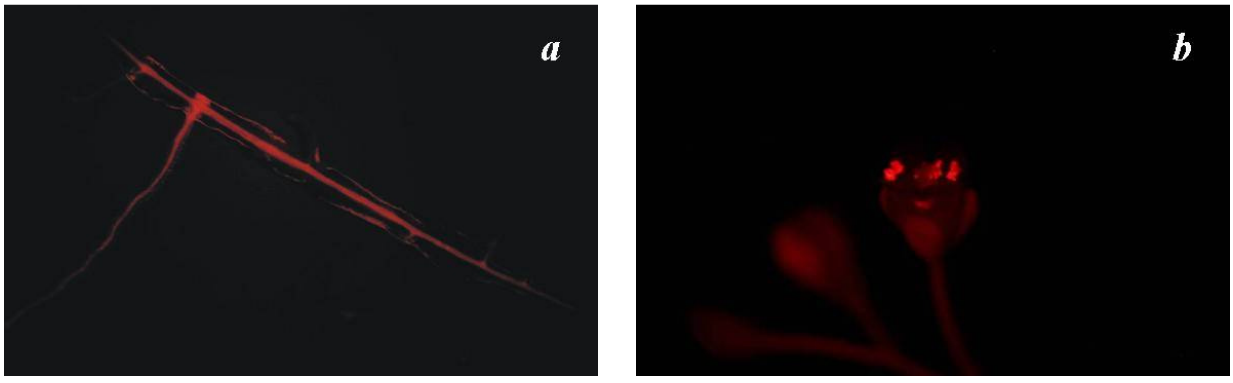


Figure 4.8. Fluorescence analysis of plants transformed with the AtSFH promoters constructs. Panel a: Root-specific expression of the eqFP611 gene driven by the AtSFH1 promoter. Panel b: Pollen specific expression of the AtSFH3 promoter.

4.2 In silico analyses of the three AtDRTS promoters reveal distinct cis-element profiles

I*n-silico* analyses of the promoters of the three *AtDRTS* genes were performed scoring against the PLACE database using the SIGNAL-SCAN software (Prestige, D.S., 1991), available as online resource at the website www.dna.affrc.go.jp/sigscan/signal1.pl. The PLACE (PLAnt Cis-acting regulatory Elements) database includes *cis*-elements that have been specifically associated with plant transcription factors.

Because the *AtDRTS* promoters are flanked by the promoters of *AtSFH* genes and shared sites could be involved in the regulation of both genes, the entire intergenic region located between the ATG codons of the *DRTS* and *SFH* transcripts (see Fig. 4.1; 4.2; 4.3) was investigated *in-silico* in order to identify specific and/or common putative regulatory regions.

Upon evaluation of the results of these analyses, only the most complex and heterogeneous *cis* elements were annotated, as shown in the following figures and tables. In fact, motifs with simpler sequence are generally found in large number, could be randomly distributed and are more likely to be spurious, whereas it is conceivable that rare and more complex *cis*-elements have higher probability to be relevant for gene regulation.

4.2.1 Analysis in silico of the AtDRTS1 promoter region

As listed in table 4A and summarized in figure 4.9, the analysis of the 1743 bp long *AtDRTS1At/SFH7* intergenic region revealed the presence of 11 uncommon *cis*-element sequences. Most notably, a BS1 site (binding site 1) of sequence **AGCGGG,** required for vascular expression of a Cinnamoyl-CoA reductase gene, is found approximately 150 bp upstream of the *AtDRTS1* transcription start point (TSP). Moreover, a "SV40 core enhancer" of sequence **GTGGWWHG** is located approximately 130 bp further upstream. Both of these *cis*-elements could be responsible for the strong vascular expression of the *AtDRTS1* promoter. Interestingly, also two auxin responsive sites are detected in the *AtDRTS1* 5' region: the first one (CATATGGMSAUR) in the small intron located in the 5' UTR of the gene, and the other (ARFAT) approximately 210 bp upstream of the transcription start. A putative MYB recognition site is located in the middle of the intergenig region whereas only two of the 11 uncommon *cis*-element sequences, a sucrose responsive element and a CarG motif, are found in the proximity of the *AtSFH7* sequence.

Tab. 4A: List of the most significant putative regulatory sites identified by in-silico analysis in the intergenic region spanning between AtDRTS1 and AtSFH7

Site	***Position*** Bp (distance from ATG of At*drts1*)	***Site Sequence***	***Function***
CATATGGMSAUR	**111**	**CATATG**	Sequence found in NDE element in soybean (G.m.) SAUR (Small Auxin-Up RNA) 15A gene promoter; Involved in auxin responsiveness. Studied in soybean (*Glycine max*).
CIACADIANLELHC	**129**	**CAANNNNATC**	Region necessary for circadian expression of tomato (L.e.) Lhc gene.Studied in tomato (*Lycopersicon esculentum).*
CAREOSREP1	**311**	**CAACTC**	"CAREs (CAACTC regulatory elements)" found in the promoter region of a cystein proteinase (REP-1) gene in rice. Studied in *Oryza sativa* (rice)
BS1EGCCR	**339**	**AGCGGG**	"BS1 (binding site 1)" found in E. gunnii Cinnamoyl-CoA reductase (CCR) gene promoter; nuclear protein binding site; Required for vascular expression. Studied in *Eucalyptus gunnii.*
ARFAT	**395**	**TGTCTC**	ARF (auxin response factor) binding site found in the promoter of primary/early auxin response genes of *Arabidopsis thaliana* (A.t.); AuxREBinding site of Arabidopsis ARF1 (Auxin response factor1); Sequence found in NDE element in Soybean (G.m.) SAUR (Small Auxin-Up RNA) 15A gene promoter; Found in D1 or D4 element in Soybean (G.m.) GH3 promoter. Studied in *Arabidopsis thaliana*; Soybean (*Glycine max*); *Oryza sativa* (rice)
ACGTABREMOTIFA20SEM	**392**	**ACGTGKC**	Experimentally determined sequence requirement of ACGT-core of motif A in ABRE of the rice gene OSEMDRE and ABRE are interdependent in the ABA-responsive expression of the rd29A in Arabidopsis. (K=G/T). Studied in rice (*Oryza sativa*); *Arabidopsis thaliana.*

SV40COREENHAN	**467**	**GTGGWWHG**	"SV40 core enhancer"; Similar sequences found in rbcS genes. (W=A/T). Studied in virus; plant; pea (*Pisum sativum*); *Arabidopsis thaliana*. .
NRRBNEXTA	**469**	**TAGTGGAT**	"NRR (negative regulatory region)" in promoter region of Brassica napus (B.n.) extA extensin gene; Removal of this region leads to expression in all tissues within the stem internode, petiole and root. Studied in *Brassica napus.*
MYBPLANT	**731**	**MACCWAMC**	Plant MYB binding site; Consensus sequence related to box P in promoters of phenylpropanoid biosynthetic genes such as PAL, CHS, CHI, DFR, CL, Bz1; Myb305. (M=A/C; W=A/T). The AmMYB308 and AmMYB330 transcription factors from Antirrhinum majus regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. Studied in petunia (*Petunia hybrida*); *Arabidopsis thaliana*; maize (*Zea mays*); parsley (*Petroselinum crispum*).
SURE1STPAT21	**1275**	**AATAGAAAA**	Sucrose Responsive Element (SURE); A motif conserved among genes regulated by sucrose. Found between -184 and -156 bp in the patatin (a major tuber protein) gene promoter of potato (S.t.). Studied in potato (*Solanum tuberosum*).
CARGCW8GAT	**1391**	**CWWWWWWWWG**	A variant of CArG motif with a longer A/T-rich core; Binding site for AGL15 (AGAMOUS-like 15).(W=A/T). Studied in *Arabidopsis thaliana.*

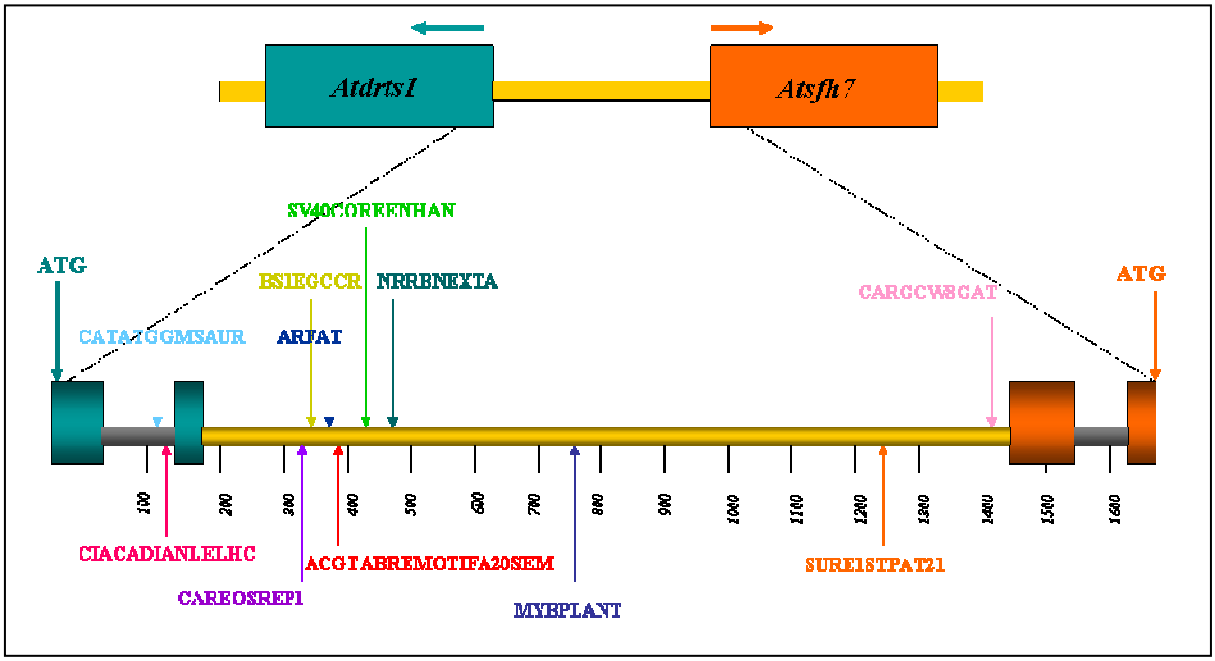


Figure 4.9: Arrangement of the AtDRTS1 (blue) and AtSFH7 (orange) genes. The intergenic region is drawn in yellow and the different sites are shown in different colours.

4.2.2 Analysis in silico of the AtDRTS2 promoter region

Analysis of the *AtDRTS2/AtSFH1* intergenic region of 1968 bp revealed the presence of 13 uncommon *cis*-elements, listed in table 4B and shown in figure 4.10. Only four of the sites are found in the vicinity of the AtSFH1 sequence whereas all the remaining *cis*-elements are closely grouped at the 5'end of the *AtDRTS2* gene. Most remarkably, among the latter motifs, several sites have been shown to regulate gene expression in proliferating cells. The most noticeable *cis*-elements are an E2F binding site, overlapping the transcription start point (TSP), and an adjacent Hexamer motif. Of particular interest are also a "Up1" motif, found about 80 bp upstream of TSP, and a "Up2" motif, also known as "telo-box", located in the first intron of the *AtDRTS2* gene. Both sites have been shown to be overrepresented in the promoters of genes up-regulated after main stem decapitation and involved in the growth of axillary buds in *Arabidopsis* (Tremoysaugue et al., 2003, Tatematsu et al, 2005).

Tab. 4B: List of the most significant putative regulatory sites identified by in-silico analysis in the intergenic region exstending between AtDRTS2 and AtSFH1

Site	***Position*** Bp (distance from ATG of Ats*fh1 or cow1*)	***Site Sequence***	***Function***
CCA1ATLHCB1	**325**	**AAMAATCT**	CCA1 binding site; CCA1 protein (myb-related transcription factor) interact with two imperfect repeats of AAMAATCT in Lhcb1*3 gene of *Arabidopsis thaliana* (A.t.); Related to regulation by phytochrome. Studied in *Arabidopsis thaliana*.
CARGCW8GAT	**331**	**CWWWWWWWWG**	A variant of CArG motif (see S000404), with a longer A/T-rich core; Binding site for AGL15 (AGAMOUS-like 15). (W=A/T). Studied in *Arabidopsis thaliana.*
BOXIINTPATP	**369**	**ATAGAA**	"Box II" found in the tobacco (N.t.) plastid atpB gene promoter; Conserved in several NCII (nonconsensus type II) promoters of plastid genes; Important for the activity of this NCII promoter. Studied in tobacco (*Nicotiana tabacum*)
MYB1LEPR	**423**	**GTTAGTT**	Tomato Pti4(ERF) regulates defence-related gene expression via GCC box and non-GCC box cis elements (Myb1(GTTAGTT), G box (CACGTG). Studied in *Arabidopsis thaliana*; *Lycopersicon esculentum* (tomato).
-10PEHVPSBD	**961**	**TATTCT**	"-10 promoter element" found in the barley (H.v.) chloroplast psbD gene promoter; Involved in the expression of the plastid gene psbD which encodes a photosystem II reaction center chlorophyll-binding protein that is activated by blue, white or UV-A light. Studied in barley (*Hordeum vulgare*).
MYBCORE	**1013**	**CNGTTR**	Binding site for all animal MYB and at least two plant MYB proteins ATMYB1 and ATMYB2, both isolated from Arabidopsis; ATMYB2 is involved in regulation of genes that are responsive to water stress in Arabidopsis; A petunia MYB protein (MYB.Ph3) is involved in regulation of flavonoid biosynthesis. Studied in Arabidopsis thaliana; animal; petunia (Petunia hybrida)

MYB2CONSENSUSAT	**1013**	**YAACKG**	MYB recognition site found in the promoters of the dehydration-responsive gene rd22 and many other genes in Arabidopsis.(Y=C/T; K=G/T). Studied in *Arabidopsis thaliana.*
UP1ATMS	**1020**	**GGCCCAWWW**	"Up1" motif found in 162 of the 1184 up-regulated genes after main stem decapitation in Arabidopsis. (W=A/T). Studied in *Arabidopsis thaliana.*
GCN4OSGL2B1	**1087**	**TGAGTCA**	"GCN4 motif" found in GluB-1 gene in rice (O.s.); Required for endosperm-specific expression; AACA and ACGT motifs was found sufficient to confer a detectable level of endosperm expression; This motif is the recognition site for a basic leucine zipper transcription factor that belongs to the group of maize Opaque-2 (O2)-like proteins; Although all the RISBZ proteins are able to interact with the GCN4 motif, only RISBZ1 is capable of activating the gene expression. Studied in rice (*Oryza sativa)*
HEXAMERATH4	**1098**	**CCGTCG**	Hexamer motif of Arabidopsis thaliana (A.t.) histone H4 promoter. Studied in *Arabidopsis thaliana.*
E2FAT	**1104**	**TYTCCCGCC**	"E2F-binding site" found in many potential E2F target genes; most potential E2F targets identified *in silico* show a cell cycle-regulated expression. (Y=T/C). Studied in *Arabidopsis thaliana.*
TELOBOXATEEF1AA1	**1173**	**AAACCCTAA**	"telo-box" (telomere motif) found in the Arabidopsis (A.t.) eEF1AA1 gene promoter; Conserved in all known plant eEF1A gene promoters; Found in the 5' region of numerous genes encoding components of the translational apparatus; Required for the activation of expression in root primordia; Acts co-operatively with tef-box; Binding site of AtPur alpha-1. Studied in Arabidopsis (*Arabidopsis thaliana).*
UP2ATMSD(-)	**1174**	**AAACCCTA**	"Up2" motif found in 193 of the 1184 up-regulated genes after main stem decapitation in Arabidopsis; W=A/T; Axillary bud outgrowth. Studied in *Arabidopsis thaliana*

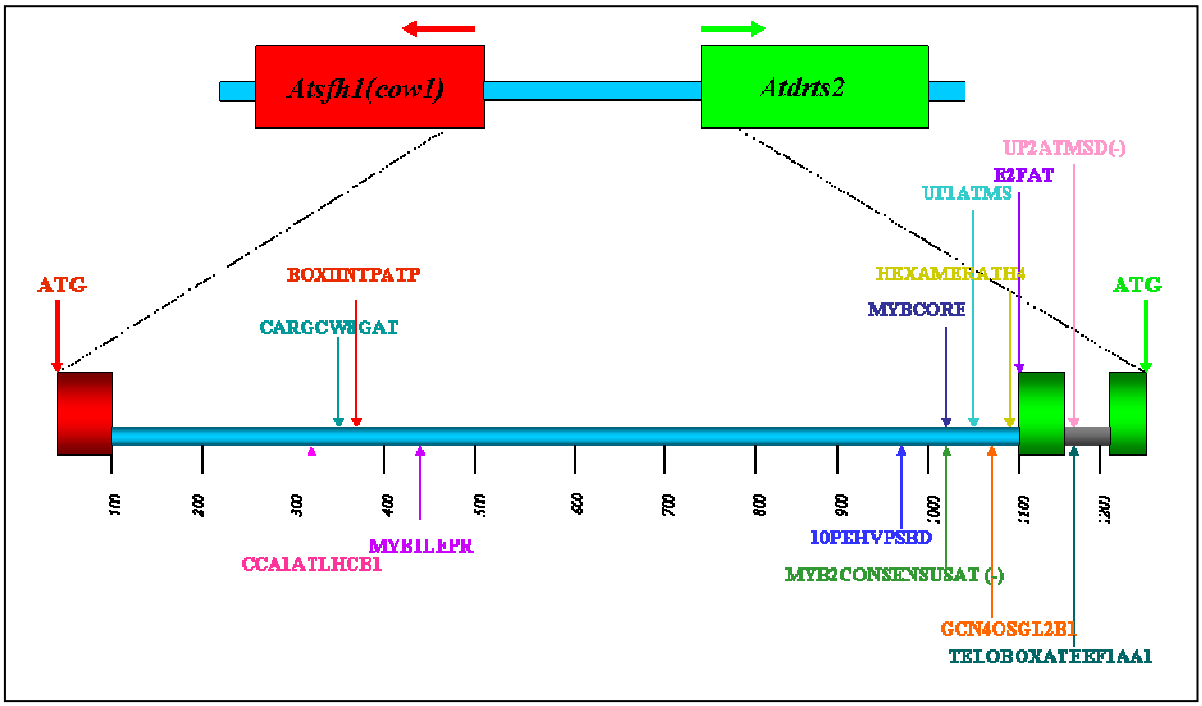


Figure 4.10: Arrangement of the AtDRTS2 (green) and AtSFH1(red) genes. The intergenic region is shown in blue with different putative cis elements shown in different colours.

4.2.3 Analysis in silico of the DRTS3 promoter region

Analysis of the *AtDRTS3/AtSFH3* intergenic region, long 3471 bp, revealed the presence of 15 uncommon *cis*-elements as described in table 4C and in figure 4.11. Most remarkably, an E2F-binding site can be detected also in this intergenic region, although its location is approximately 1400 bp upstream of the *AtDRTS3* start of transcription. Additionally, a "Up1" motif is found about 850 bp upstream of the *AtDRTS3* TSP. Concerning the putative regulatory regions of the *AtSFH* promoter, the short intron in the 5'UTR of the *AtSFH3* gene contains a binding consensus sequence for the product of the floral homeotic gene *Agamous* and could be important for the pollen-specific expression of this gene.

Tab. 4C: List of the most significant putative regulatory sites identified by in-silico analysis in the intergenic region spanning between AtDRTS3 and AtSFH3

Site	***Position*** Bp (distance from ATG of *Atsfh3*)	***Site Sequence***	***Function***
GAREAT	**843 1602**	**TAACAAR**	GARE (GA-responsive element); Occurrence of GARE in GA-inducible,GA-responsible, and GA-nonresponsive genes found in Arabidopsis seed germination was 20, 18, and 12%, respectively. Studied in *Arabidopsis thaliana.*
BOXLCOREDCPAL	**355 3133**	**ACCWWCC**	Consensus of the putative "core" sequences of box-L-like sequences in carrot (D.c.) PAL1 promoter region; DCMYB1 bound to these sequences in vitro (W=A/T). Studied in *Daucus carota* (carrot)
AGATCONSENSUS	**413**	**TTWCCWWWWNNGGWW**	Binding consensus sequence for the product of the Arabidopsis (A.t.) floral homeotic gene AGAMOUS (AG); AG protein contains a region similar to the DNA binding domain of SRF and MCM1; The consensus sequence contains a CArG box; AG protein is a putative transcription factor for floral genes; H=A/T/C; W=A/T.
ELRECOREPCRP1	**673 1694**	**TTGACC**	ElRE (Elicitor Responsive Element) core of parsley (P.c.) PR1genes; consensus sequence of elements W1 and W2 of parsley PR1-1 and PR1-2 promoters; Box W1 and W2 are the binding site of WRKY1 and WRKY2, respectively; ERE; "WA box"; One of the W boxes foun in the Parsley (P.c.) WRKY1 gene promoter; Required for elicitor responsiveness constitute a palindrome; WRKY1 protein binding site; W-box found in thioredoxin h5 gene in Arabidopsis. Studied in *Petroselinum crispum* (parsley); *Nicotiana tabacum* (tobacco).
D1GMAUX28	**902**	**ACAGTTACTA**	DNase I protected sequence found in the soybean (G.m.) auxin responsive gene, Aux28, promoter; D1 and D4 share a very similar core sequence TAGTXXCTGT and TAGTXCTGT, respectively; D1/D4-like sequence were identified in several other auxin responsive genes; Binding site of GmGT-2 which is the GT-2 family of transcription factors; GmGT-2 are down-regulated by light in a phytochrome-dependent manner soybean. Studied in *Glycine max.*

HDZIP2ATATHB2	**1226 2619**	**TAATMATTA**	Binding site of the Arabidopsis (A.T.) homeobox gene (ATHB-2) found in its own promoter; Located between -72 and -80; Similar to the HD-ZIP-2 binding consensus sequence; ATHB-2 is regulated by light signals which function as a negative autoregulator of its own gene. (M=C/A). Studied in *Arabidopsis thaliana*
ERELEE4	**1257 1575**	**AWTTCAAA**	"ERE (ethylene responsive element)" of tomato (L.e.) E4 and carnation GST1 genes; GST1 is related to senescence; Found in the 5'-LTR region of TLC1.1 retrotransposon family in Lycopersico chilense; ERE motifs mediate ethylene-induced activation of the U3 promoter region. Stuided in tomato (*Lycopersicon esculentum*); carnation (*Dianthus caryophillus*); *Lycopersicon chilense.*
E2FAT	**1872**	**TYTCCCGCC**	"E2F-binding site" found in many potential E2F target genes; most potential E2F targets identified *in silico* show a cell cycle-regulated expression. (Y=T/C). Studied in *Arabidopsis thaliana.*
UP1ATMSD	**2424**	**GGCCCAWWW**	"Up1" motif found in 162 of the 1184 up-regulated genes after main stem decapitation in Arabidopsis. (W=A/T). Studied in *Arabidopsis thaliana.*
SORLIP5AT	**2641**	**GAGTGAG**	One of "Sequences Over-Represented in Light-Induced Promoters (SORLIPs) in Arabidopsis; Computationally identified phyA-induced motifs; Over-represented in both light-induced cotyledon-specific and root-specific genes of *Arabidopsis thaliana.* Studied in *Arabidopsis thaliana.*
ABRELATERD1	**2670**	**ACGTG**	ABRE-like sequence (from -199 to -195) required for etiolation-induced expression of erd1 (early responsive to dehydration) in Arabidopsis. Studied in *Arabidopsis thaliana*.
CATATGGMSAUR	**2794 3021**	**CATATG**	Sequence found in NDE element in soybean (G.m.) SAUR (Small Auxin-Up RNA) 15A gene promoter; Involved in auxin Responsiveness. Studied in soybean (*Glycine max*)
CANBNNAPA	**2923**	**CNAACAC**	Core of "(CA)n element" in storage protein genes in *Brasica napus* (B.n.); embryo- and endosperm-specific transcription of napin (storage protein) gene, napA; seed specificity; activator and repressor. Studied in *Brassica napus*.

SV40COREENHAN	**3030**	**GTGGWWHG**	SV40 core enhancer"; Similar sequences found in rbcS genes. (W=A/T). Studied in virus; plant; pea (*Pisum sativum*); *Arabidopsis thaliana.*
300-CORE	**3080**	**TGTAAAG**	"TGTAAAG core motif" in "300 elements" of alpha-zein genes of maize: 300 element core or "prolamin box" P-box that binds with P-box binding factor (PBF). PBF is a DNA-binding protein of the DOF class of transcription factors. Studied in maize (*Zea mays*); wheat (*Triticum aestivum*); barley (*Hordeum vulgare*); tobacco (*Nicotiana tabacum*)

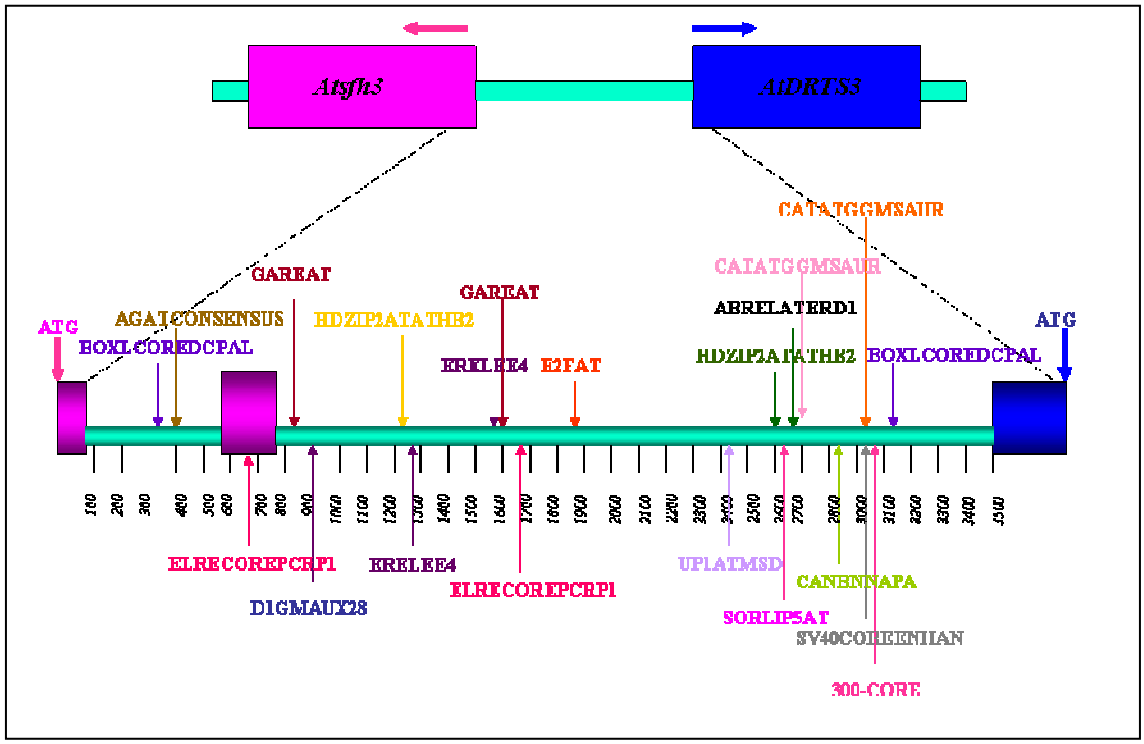


Figure 4.11: Arrangement of the AtDRTS3 (blue) and AtSFH3 (fucsia) genes. The intergenic region is shown in green with different putative cis elements shown in different colours.

4.3 The E2F sites detected in the AtDRTS promoters are recognized by E2F factors and regulate promoter activity

4.3.1 ChIP analysis of the AtDRTS promoters containing E2F cis-elements

The *in silico* analysis of the promoters of the three *AtDRTS* genes revealed the presence of putative E2F cis-elements only in the promoters, *AtDRTS2* and *AtDRTS3*, which are active in the shoot apical meristems. These sites are likely to be important for meristematic expression because E2F sites have been found in one or two copies in close proximity of the transcription start sites of several plant cell cycle-specific promoters.

To verify whether the E2F sites detected in the two *AtDRTS* promoters could be functional and could be recognized by activating AtE2F factors, ChIP analyses were performed using cross-linked chromatin isolated from *Arabidopsis* suspension cells. The ChIP protocol was optimized using polyclonal antibodies raised against the carrot E2F factor DcE2F; these antibodies recognize only activating AtE2Fs with a far greater preference for AtE2Fa. In order to assess the efficiency of the ChIP analysis, PCR reactions on the immunoprecipitated genomic fragments were performed as positive control using primers specific for the promoter of *AtRNR1b*, a well characterized E2F responsive gene (Chaboutè et al, 2002; Egelkrout et al, 2002). PCR reactions using primers for the *AtDRTS* promoters on mock reactions incubated without the addition of the antibodies were performed as negative control. As shown in figure 4.12, the ChIP experiments revealed that DNA fragments of both *AtDRTS2* and *AtDRTS3* promoters can be selectively immunoprecipitated using antibodies against *Dc*E2F. Thus, these results indicate that the E2F sites found in the *AtDRTS* promoters are recognized *in vivo* by activating AtE2F factors and are expected to be important for the expression of *AtDRTS2* and *AtDRTS3* in meristematic tissues.

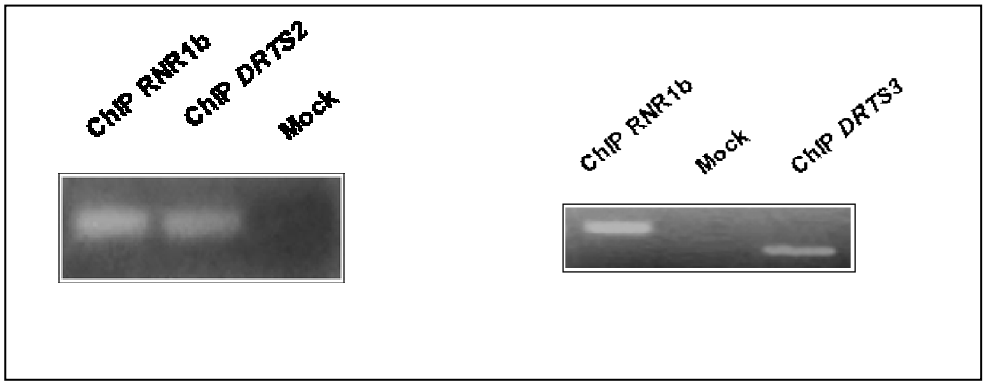


Figure 4.12. ChIP analysis of the AtDRTS2 and AtDRTS3 promoters

4.3.2 Functional analysis of the E2F sites in the two AtDRTS promoters

The binding of activating AtE2Fs to the *AtDRTS* promoters detected by ChIP analysis suggests that the E2F *cis*-elements found in the *AtDRTS2* and *AtDRTS3* promoter may exert positive roles on promoter activity. However, E2F sites have been shown to regulate cell cycle-dependent genes in both positive or negative ways. Therefore, to verify the specific role of the E2F sites in the regulation of the *AtDRTS* genes, the *AtDRTS2* and *AtDRTS3* promoters were mutated at their E2F sites and were assembled in GFP/GUS reporter constructs (pDRTS2-ΔE2F and pDRTS3-ΔE2F) as shown in figure 4.13

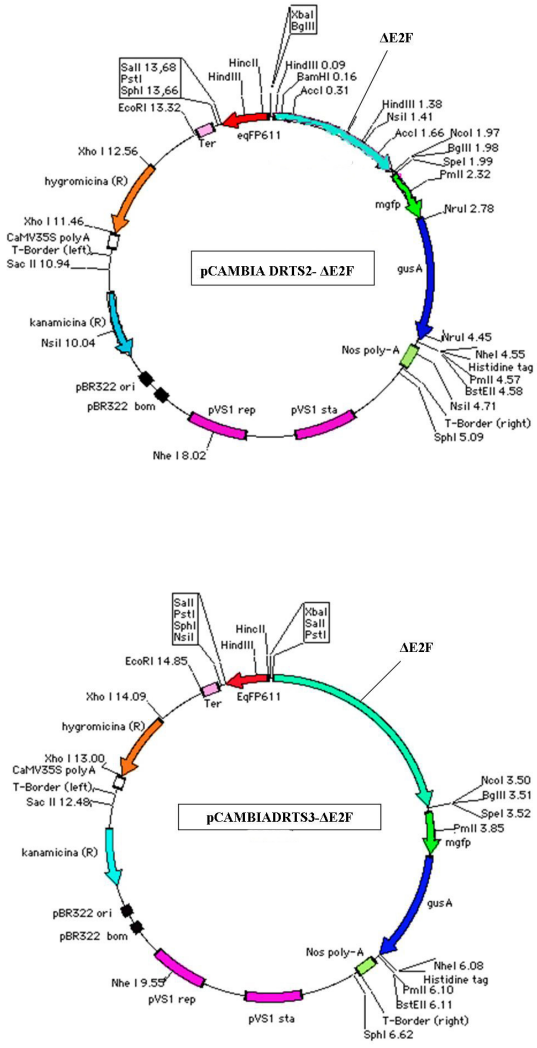


Figure 4.13. Structure of the AtDRTS promoter construct: AtDRTS2 and AtDRTS3 with mutated E2F sites

Following Arabidopsis transformation, several primary (T1) transformants were obtained and selfed to produce T2 seedlings. The activity of the promoters mutated in the E2F sites was then studied by histochemical GUS assays. As shown in figure 4.14 and 4.15, the spatial pattern of expression of the two *AtDRTS-ΔE2F* promoters appeared to be similar to the pattern obtained with the original promoter constructs. However, the level of GUS activity driven by the two *AtDRTS-ΔE2F* constructs appeared to be strongly increased in the vascular tissues and upregulated also in the meristematic tissues which were already positive with the wild-type promoter constructs (figures 4.14 and 4.15).

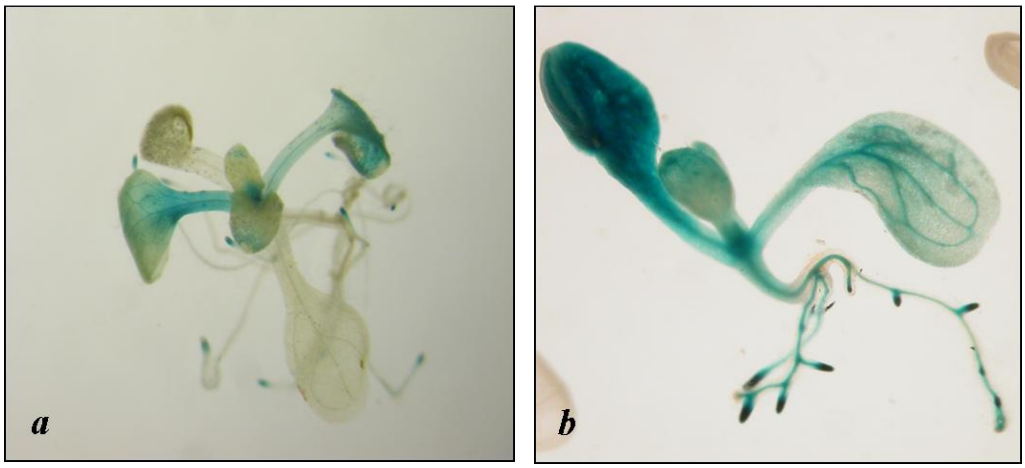


Figure 4.14 Patterns of expression of the original AtDRTS2 promoter (panel a) versus the mutated AtDRTS2-ΔE2F promoter (panel b) in 10 day-old transgenic plants.

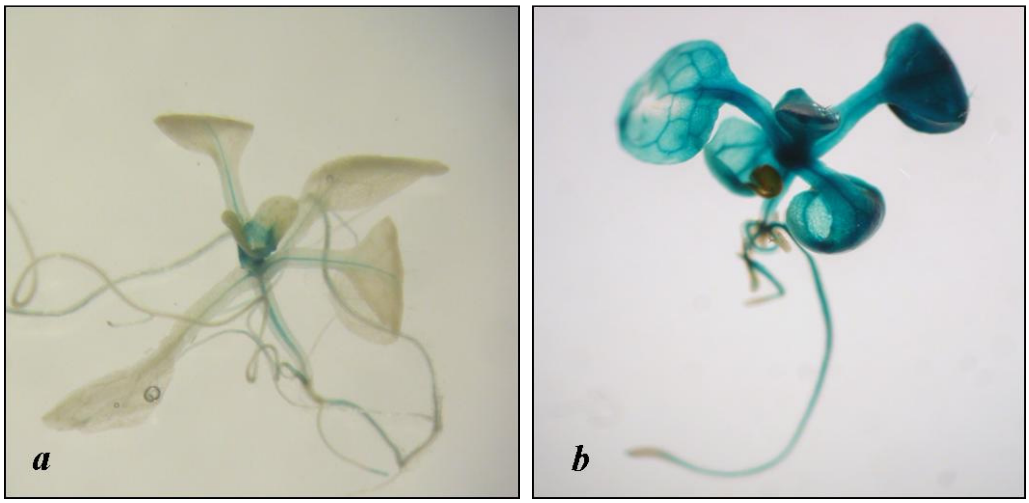


Figure 4.15. Comparison of pattern expression of the original AtDRTS3 promoter (panel a) versus the mutated AtDRTS3-ΔE2F promoter (panel b) in 10 day-old transgenic plants.

To quantify the level of activity, fluorimetric assays were performed on seedlings of the same developmental stage (emergence of the third pair of leaves). As shown in figures 4.16 and 4.17, the mutation of the E2F *cis* elements increased considerably the activity of both promoters indicating that the E2F factors recognizing these sites are able to repress *AtDRTS* expression.

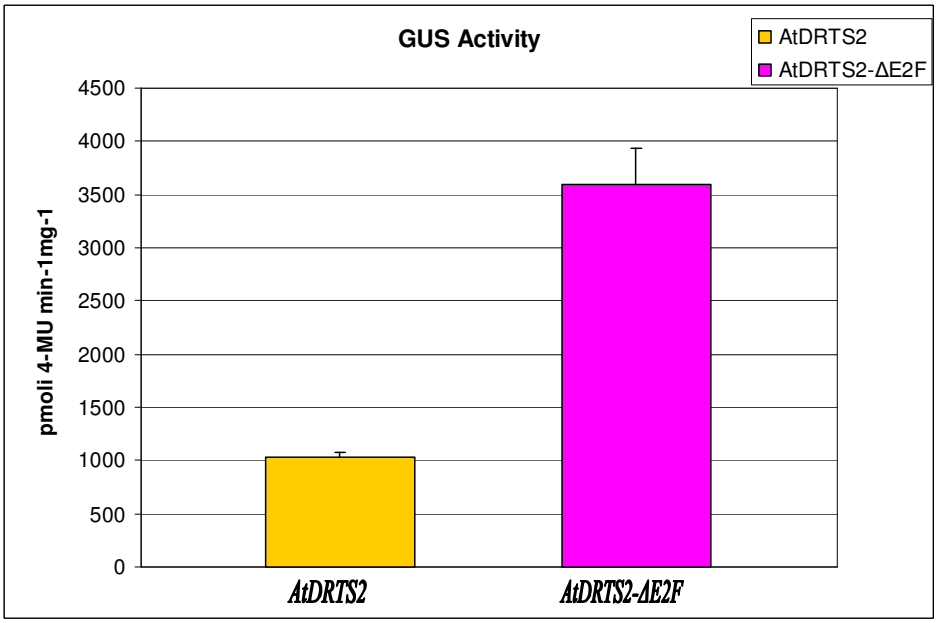


Figure 4 16 Increased fluorimetrical activity of the mutated AtDRTS2-ΔE2F promoter in 2 week-old plants.

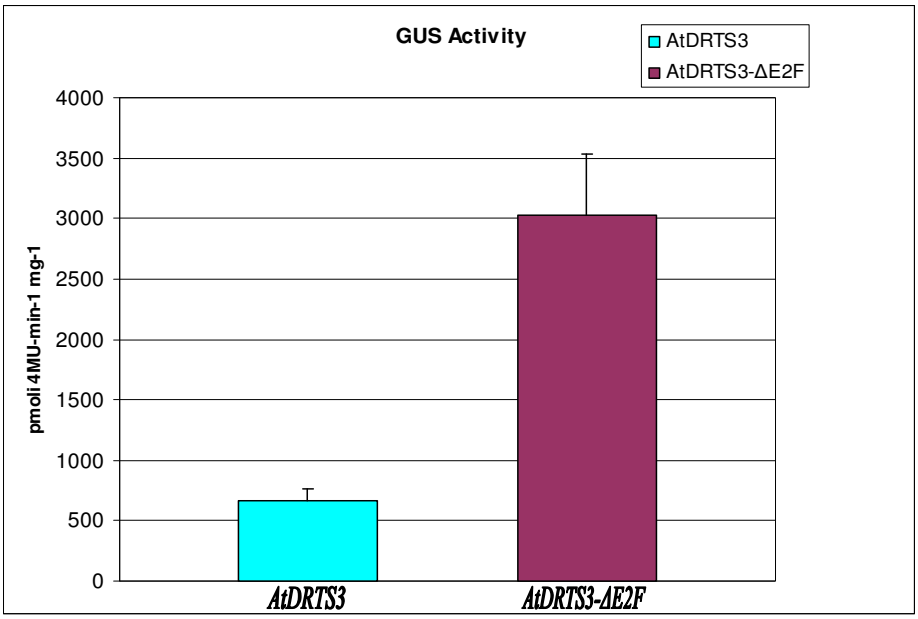


Figure 4.17. Increased fluorimetrical activity of the mutated AtDRTS3-ΔE2F promoter in 2 week-old plants

4.4 The activity of the AtDRTS2 promoter is controlled by multiple regulatory regions

The discovery that in both *AtDRTS* promoters the E2F site plays a repressive role implies that other *cis*-acting elements must be involved in the activation of expression in meristematic cells. In particular we focused our studies on the *AtDRTS2* gene which is specifically meristematic. Its pattern espression agrees with results found in *Daucus carota* where *DHFR-TS* transcripts have been shown to accumulate to high level in dividing meristematic cells (Albani et al, 2005).

As revealed by the *in silico* analysis, the 5' region of the *AtDRTS2* gene contains two remarkable *cis*-elements called Up1 and Up2. The first site is located 80 bp upstream of the transcription start site whereas the second is located at the start of the 100 bp intron found in the 5' untranslated region of the gene. These sites are also known in literature as Motif Site II and Telo Box and previous studies have suggested their conservative association in the promoters of several genes that are highly expressed in dividing cells (Tremousaygue et al, 2003).

4.4.1 Study of the role of the Up1 and Up2 sites in the AtDRTS2 promoter

To analyse the possible role of the Up1 and/or Up2 sites in the control of *AtDRTS2* expression, single mutations (AtDRTS2-ΔUp1 and AtDRTS2-ΔUp2) as well as mutations of both sites (AtDRTS2-ΔΔ) were introduced in the *AtDRTS2* promoter construct driving expression of the GFP/GUS reporter gene as shown in figure 4.18.

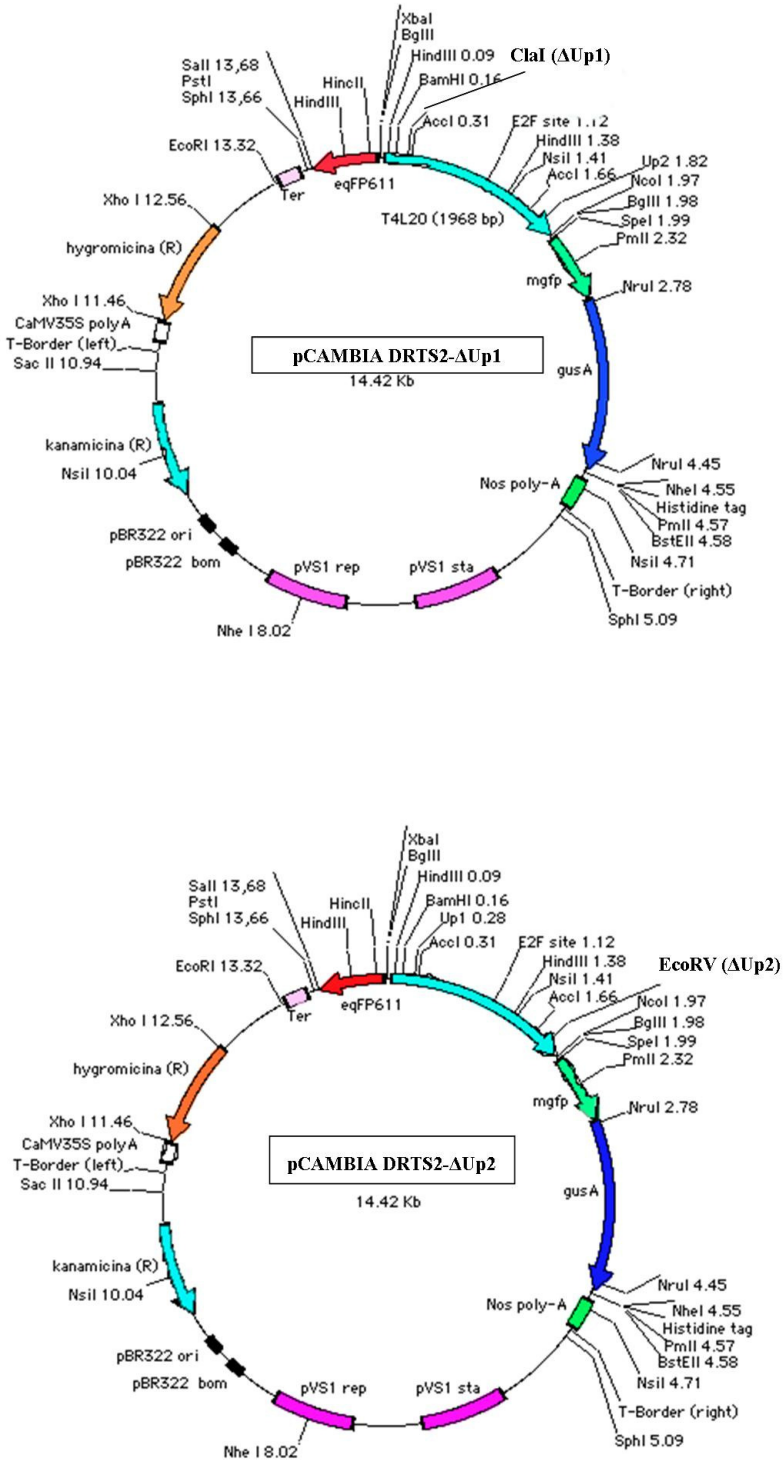


Figure 4.18. Structure of the promoter construcs with mutated UP1 or UP2 sites

As shown in figures 4.19, 4.20 and 4.21 the spatial patterns of activity of the AtDRTS2-ΔUp1, AtDRTS2-ΔUp2 and AtDRTS2- ΔUps promoter constructs appear to be very similar to the one conferred by the wild type *AtDRTS2* promoter. All the constructs are expressed in shoot and root apical meristems as well as in the developing leaves and also in vascular tissues. A clear GUS activity is seen also in the developing ovaries but no expression is seen in mature seeds of the siliques (panels d, e, f). The activity of the mutated promoters is fairly constant during development (panels a, b, c) but appears to be slightly decreased compared to the wild type *AtDRTS2* promoter. To quantify the level of activity, fluorimetric assays were performed on seedlings of the transgenic lines of two weeks of age and at the same developmental stage. As shown in figure 4.22, the mutation of either the Up1 or Up2 site lowered considerably, and at the same level, the activity of the promoter compared to the wild type construct. Also the mutation of both Up1 and Up2 sites in the double mutant decreased similarly, but not at lower levels, the activity of the *AtDRTS2* promoter indicating that the proteins recognizing the two sites must act together, but not synergistically, to drive high level of expression of the *AtDRTS2* gene.

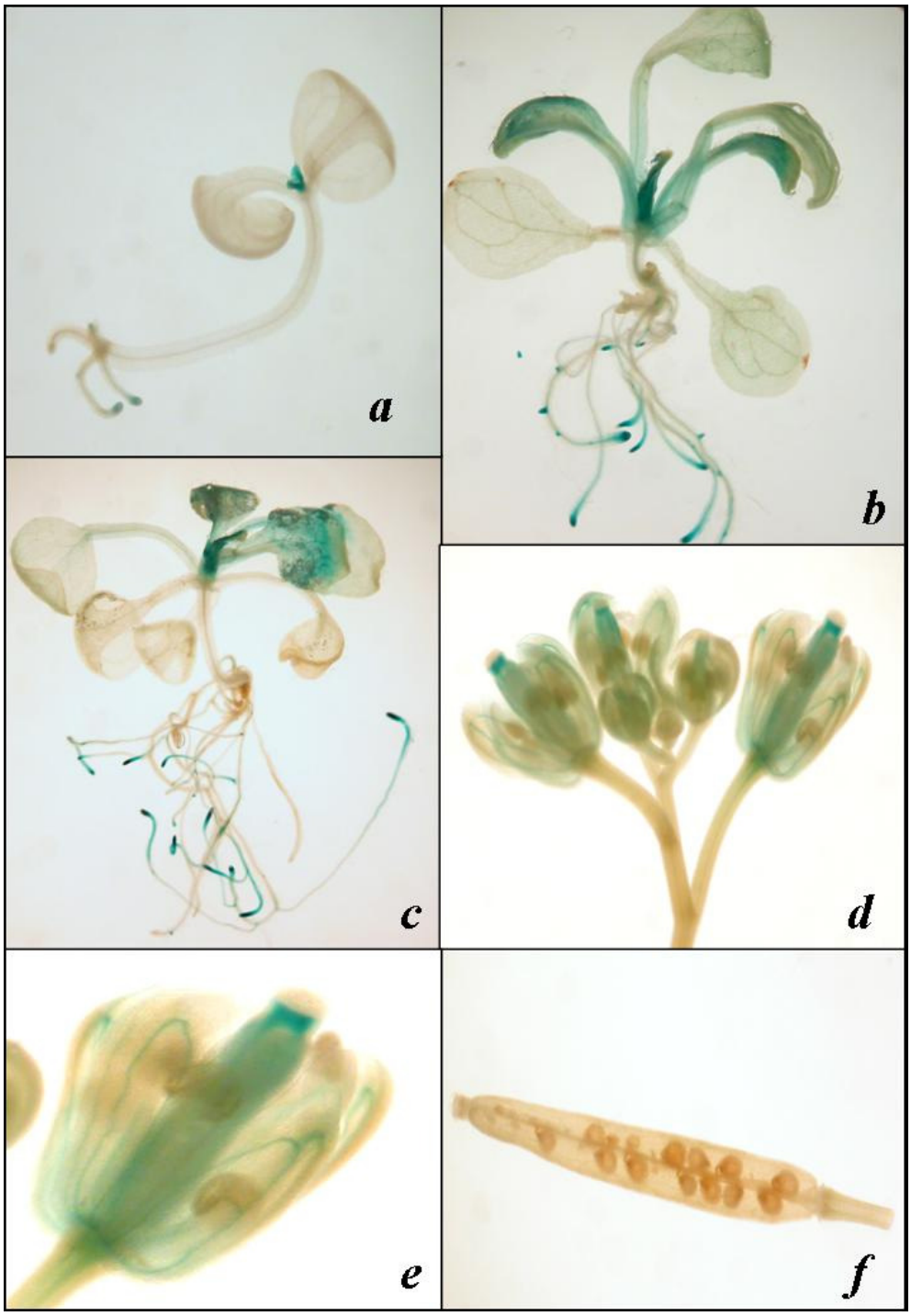


Figure 4.19. Activity of the DRTS2-ΔUP1 promoter. Panels a, b, c show plantlets of one, two and three weeks, respectively; panels d and e inflorescences and flowers; panel f siliques.

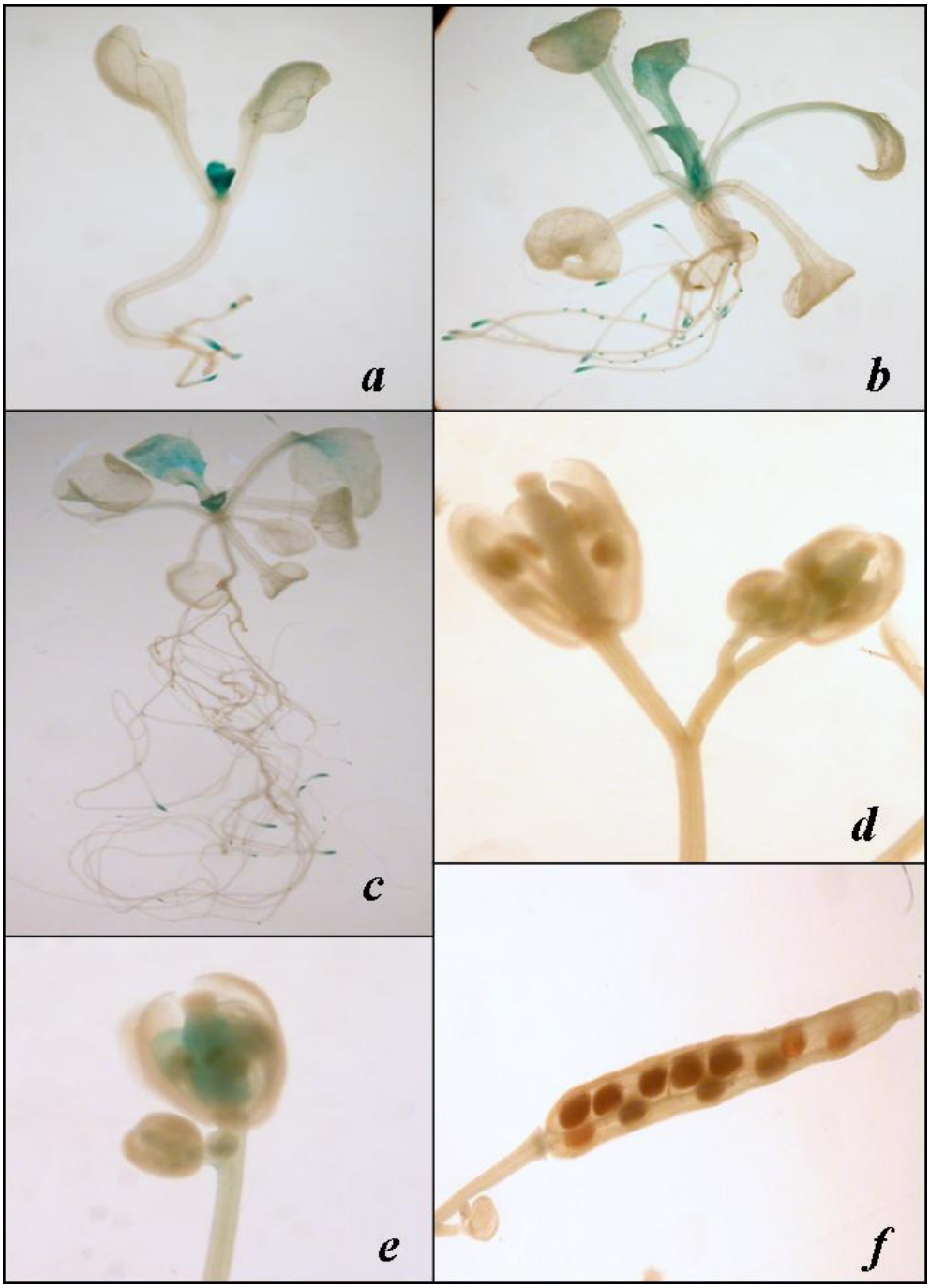


Figure 4.20. Activity of the DRTS2-ΔUP2 promoter Panels a, b, c show plantlets of one, two and three weeks, respectively; panels d and e inflorescences and flowers; panel f siliques.

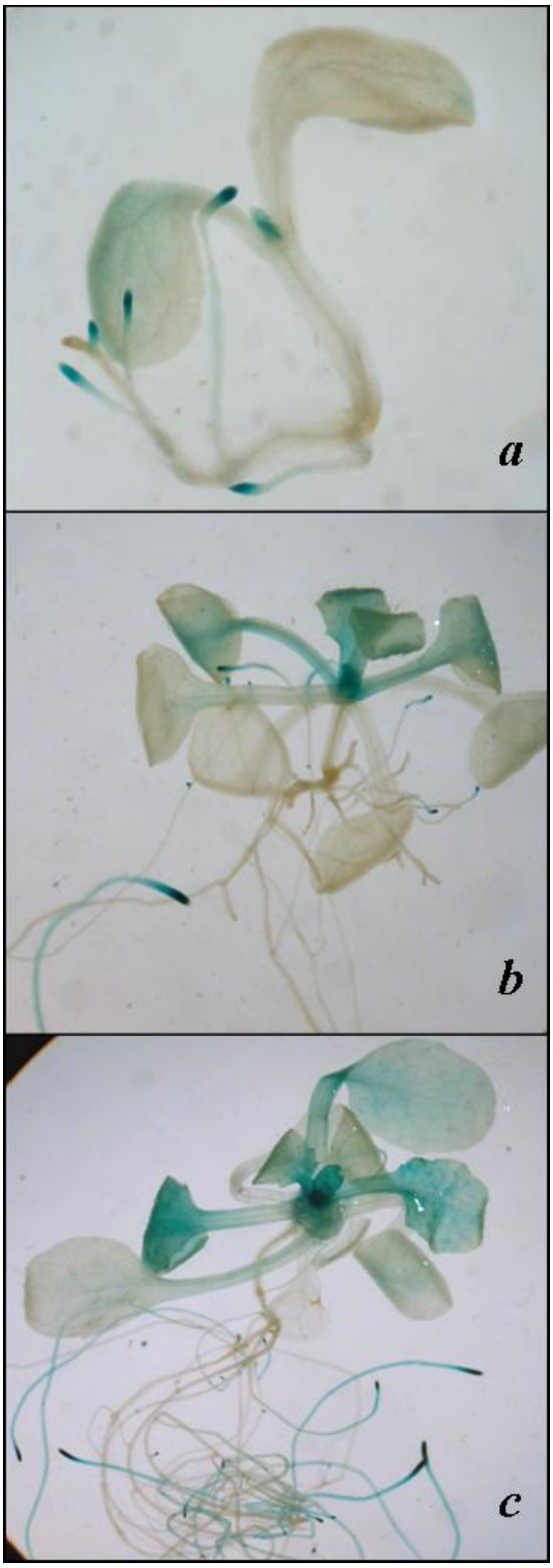


Figure 4.21. Activity of the DRTS2-ΔUPs promoter Panels a, b, c show plantlets of one, two and three weeks, respectively.

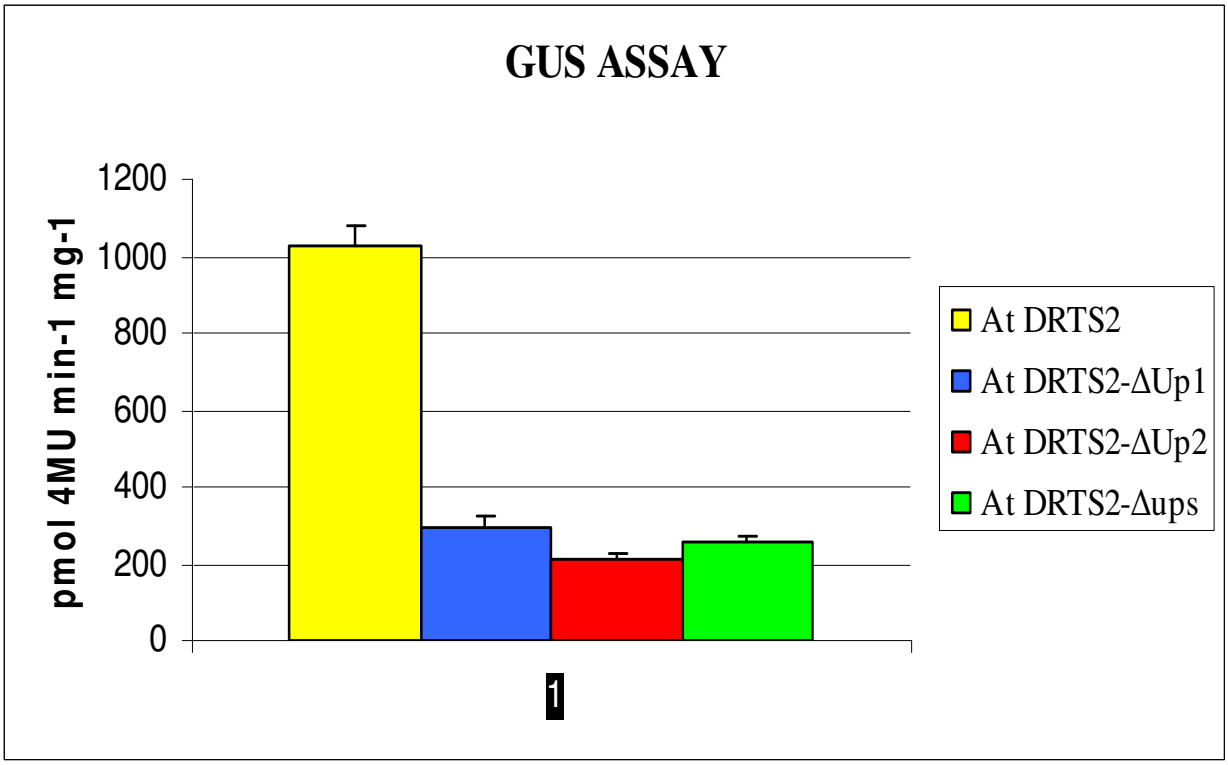


Figure 4.22. Activity of AtDRTS2-ΔUp1, AtDRTS2-ΔUp2 and AtDRTS2-ΔUps promoters compared to the wild type AtDRTS2 promoter in two week-old seedlings.

4.4.2 The intron located in the 5'UTR region of the AtDRTS2 gene is crucial for its meristematic expression

To further investigate regions of the *AtDRTS2* promoter which are necessary for its activation in the meristems a new construct lacking the promoter region of the *AtSFH1* gene, extending form position -830 to position +175 relative to the proposed transcription start point of the gene (AtDRTS2 -830/+175), was devised. Moreover, because the first intron in the 5'-UTR region of the *AtDRTS2* gene contains an Up2 site which is conferring high levels of expression but is not essential for the meristematic expression of the gene, an additional construct lacking also the first intron was prepared (AtDRTS2 -830/+51). As shown in figure 4.23, these promoter fragments were cloned upstream of the *GUS* reporter gene into the binary vector pCAMBIA1300 and the resulting plasmids were used for the transformation of *Arabidopis thaliana*.

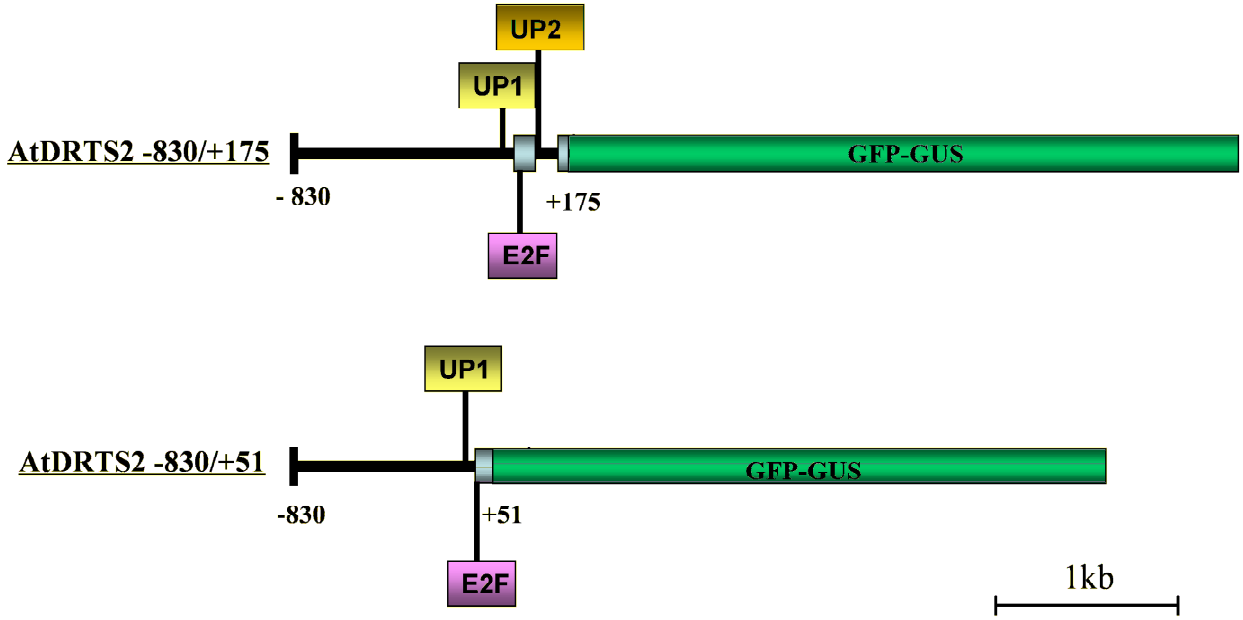


Figure 4.23. Structure of the new AtDRTS2 promoter constructs

Analyses of the resulting transgenic plants revealed that the promoter construct containing the first intron of *AtDRTS2* is strongly active in the meristems (figure 4.24) and shows a very high expression also in the vascular tissues, which is possibly linked to the nearby presence of the strong constitutive 35S promoter of the selectable marker gene (hygromycin resistance). Conversely, the removal of the 100 bp long intron in the -830/+51 construct drastically changed the pattern of activity of the *AtDRTS2* promoter, as shown in figure 4.25, abolishing completely the expression in root and shoot meristems. This surprising result suggests that at least one additional regulatory site, not detected by *in silico* analysis, is contained in the first intron and is strictly necessary for the meristematic expression of the *AtDRTS2* gene.

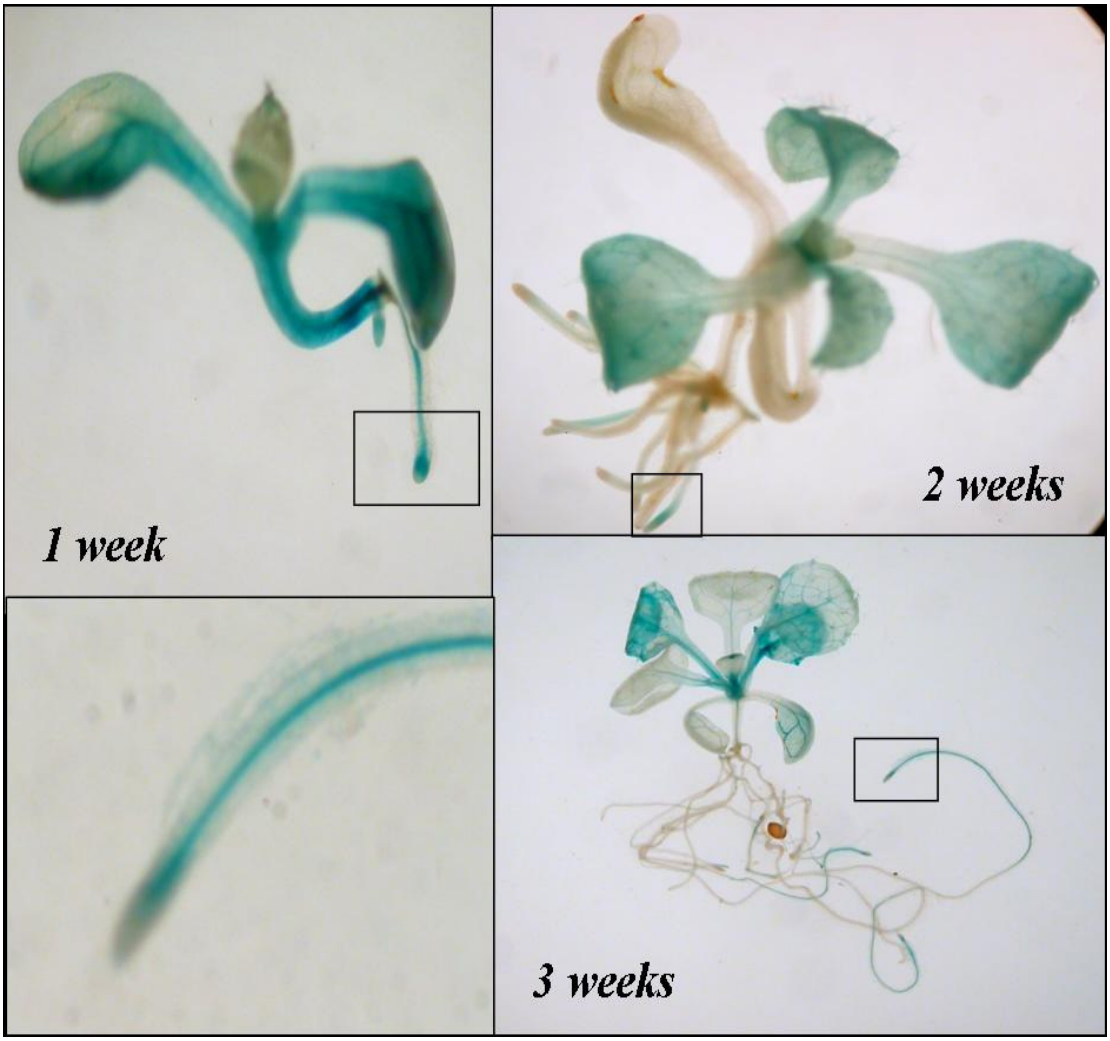


Figure 4.24. Activity of the AtDRTS2 promoter construct (AtDRTS2 -830/+175) with the first intron of the gene. In the figure are shown plantlets of one, two and three weeks of age and the enlarged view of a root tip.

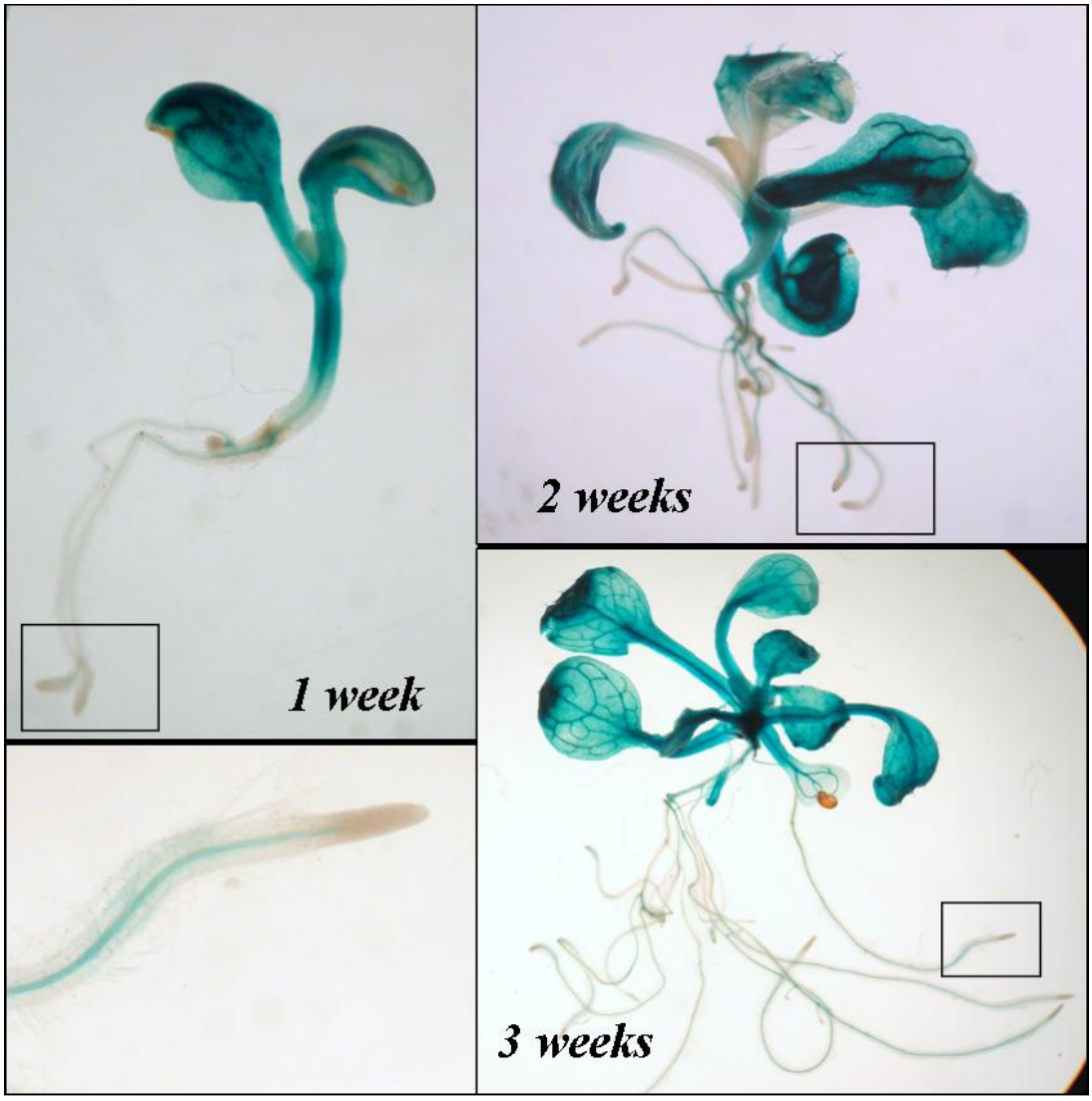


Figure 4.25. Activity of the AtDRTS2 promoter construct (AtDRTS2 -830/+51) lacking the first intron of the gene. In the figure are shown plantlets of one, two and three weeks of age and the enlarged view of a root tip.

4.4.3 A site located in the intron in the 5'UTR region of the AtDRTS2 gene is crucial for its meristematic expression

As shown by preliminary results of construct lacking the first intron (-830/+51) the pattern of activity of the *AtDRTS2* promoter is drastically changed with the complete abolition of the GUS expression in root and shoot meristems of the plant. This result suggests that at least one additional regulatory site, not detected by *in silico* analysis, is contained in the first intron and is strictly necessary for the meristematic expression of the *AtDRTS2* gene. To identify this site, sequential mutations of 8 bp were introduced along the intronic region in the original *AtDRTS2* promoter construct driving expression of the *GFP/GUS* reporter gene as shown in figure 4.26. Preliminary analyses by histochemical GUS staining were performed on root tips of T1 generation transgenic plants. Only one of the construct, the *AtDRTS2-ΔIII* mutant, showed altered expression with complete lack of GUS activity while the other nine constructs gave high expression in root meristems comparable to the wild type promoter *AtDRTS2*. As shown in figure 4.27, the lack of GUS activity of the mutated promoter is constant during development (panels a, b, c). To quantify the level of activity, fluorimetric assays were performed on seedlings of the transgenic lines of approximately two weeks of age and at the same developmental stage. As shown in figure 4.28, mutation in the ΔIII mutant of the *tttgtaaa* sequence located 28 bp downstream of the donor splicing site led to nearly complete abolition of promoter activity.

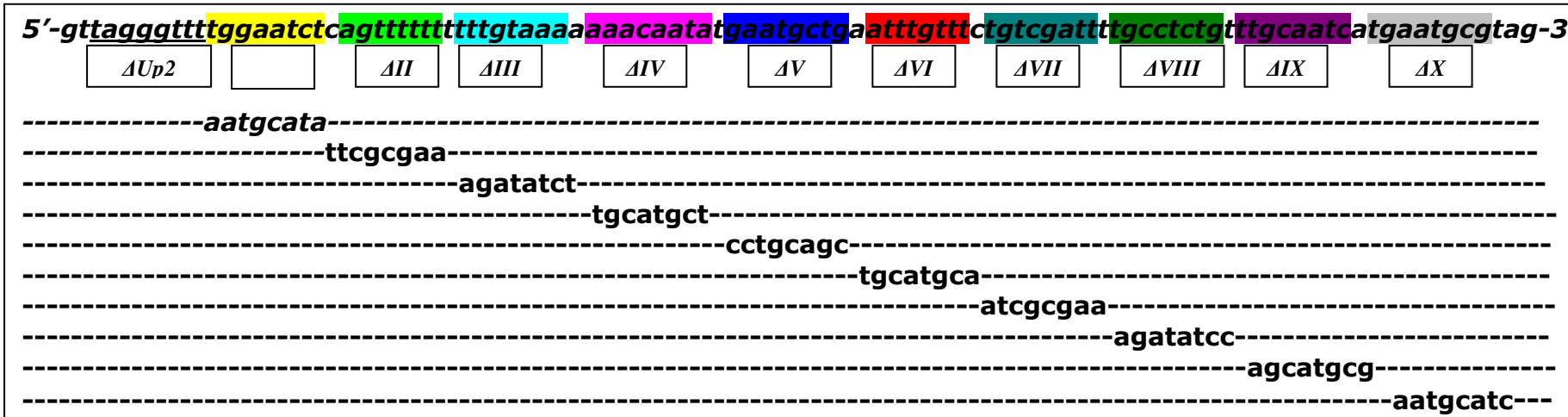


Figure 4.26. : Sequence of the first intron and the linker scanning mutations: sites with inserted mutations are underlined with colours while down of the sequence where designed the site mutations.

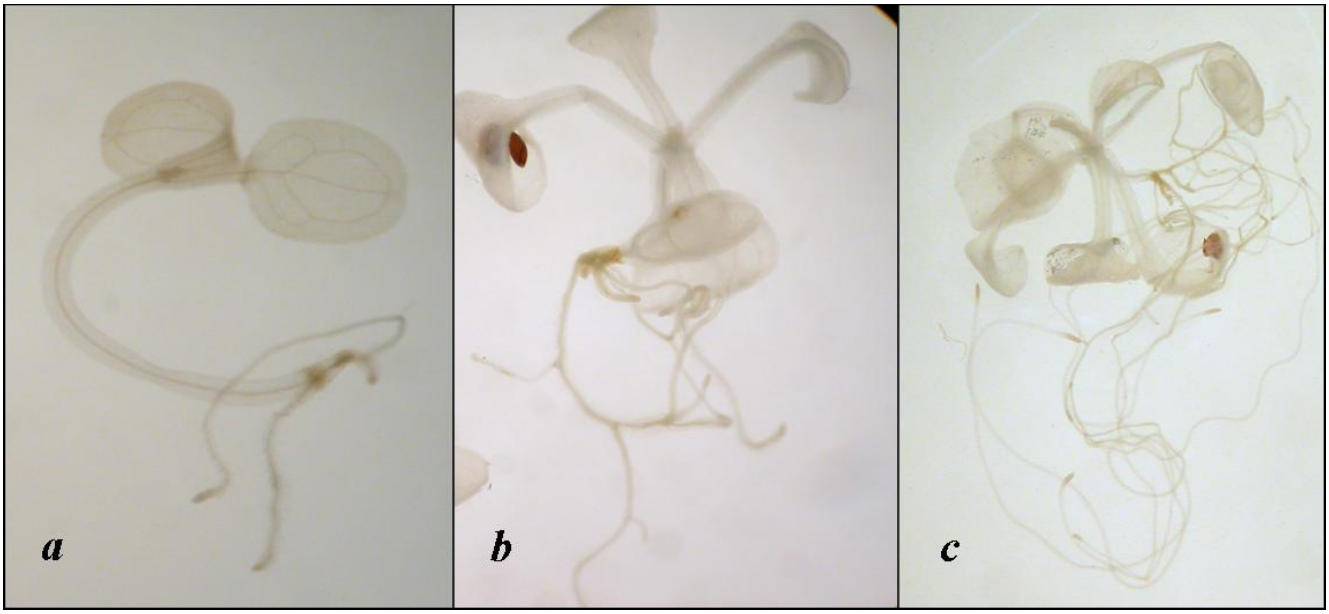


Figure 4.27. Histochemical analysis of the activity of the DRTS2-ΔIII promoter. Panels a, b, c show plantlets of one, two and three weeks of age, respectively.

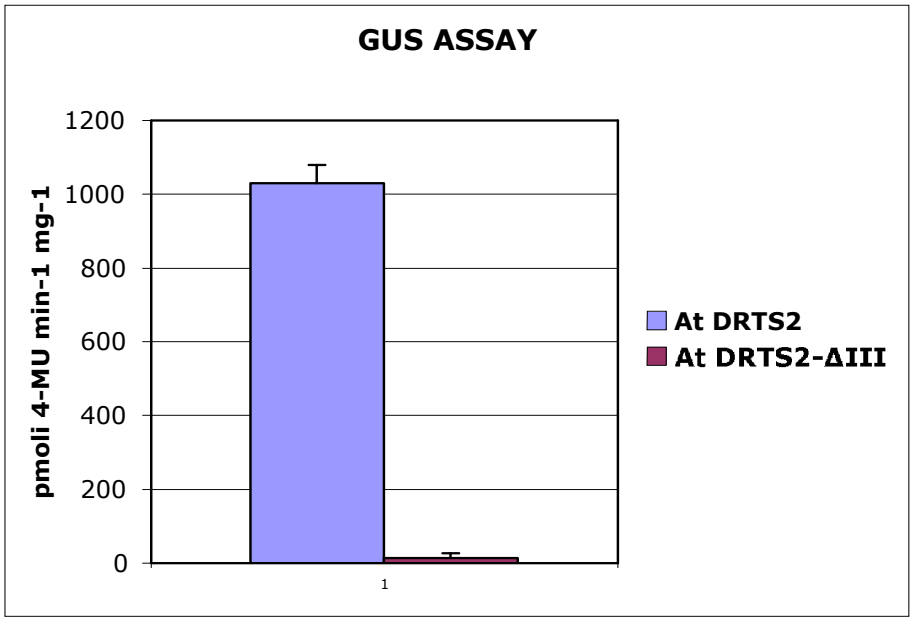
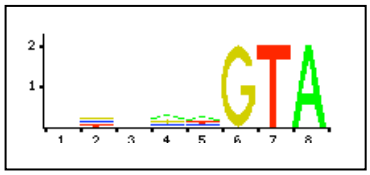
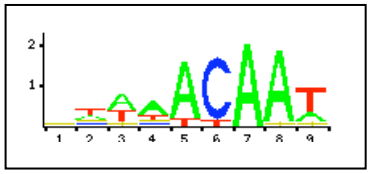
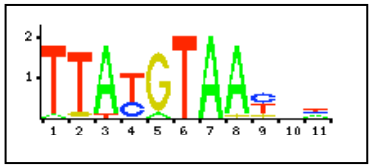
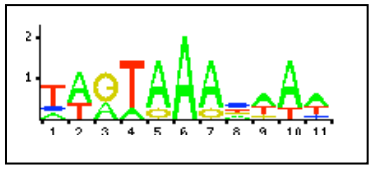


Figure 4.28. Flurimetric analysis of the activity of AtDRTS2-ΔIII promoter compared to the wild type AtDRTS2 promoter in two week-old seedlings.

To verify which type of transcription factor could binds the site mutated in the DRTS2-ΔIII construct we analysed the wild type sequence against the Jaspar high-quality transcription factor binding profile database (jaspar.genereg.net (Sandelin et al, 2004)). The result of this analysis revealed that five diffent animal transcription factors recognize a binding site that can match the *tttgtaaa* sequence. The sites recognized by these transcription factors are listed in table 4D. The existance of plant hortologues of these transcription factors is not clear at present.

Tab. 4D: List of the most significant putative regulatory sites identified by Jaspar Database in the region of mutated site ΔIII

Site	*Sequence recognizable at the ΔIII site of DRTS2*	BINDING SEQUENCE MATRX
FOXC1	**tttttgta**	
SRY	**tttgtaaaaa**	
NFIL3	**ttttgtaaaaa**	
Broad-Complex IV and Hunchback	**ttgtaaaaaaa**	

5. DISCUSSION

Plant growth is finely regulated by endogenous and environmental signals and relies on both meristematic cell proliferation and the expansion of differentiating undividing cells. Thus, the control of cell proliferation is vital for plant development and productivity. Plant genes that play important roles during cell proliferation are expected to be highly expressed in meristematic cells and in specific phases of the cell cycle. However, the molecular mechanisms involved in the control of gene expression during the cell cycle in plants are still far from being clear. Studies performed in recent years have led to the discovery that important cell cycle transcriptional regulators are conserved in plant and animal cells. In particular, it has been shown that the retinoblastoma/E2F pathway, a critical element in the control of the G1/S phase gene expression in animal cells, disrupted in several types of cancers, is also conserved in plant cells.

To better understand the control of gene expression during cell proliferation in plants we have studied the regulation of genes coding for enzymes that are known to be important for S phase progression. In this respect, the genes coding for dihydrofolate reductase (DHFR) and thymidilate synthase (TS) play central roles in DNA precursor biosynthesis and are expected to be regulated during G1/S phase of the cell cycle. Unlike animal cells, where DHFR and TS are coded by distinct genes, plant and protozoan genomes contain genes coding for bifunctional DHFR-TS enzymes. In plants, DHFR-TS enzymes, like the animals counterparts, were found to be highly accumulated in proliferating cells and virtually undetectable in quiescent cells (Vandiver and Fites, 1979), implying that these enzymes are active in tissues where DNA synthesis is occurring (Cox et al, 1999).

Also *in situ* hybridization studies from *Daucus carota* confirmed that the *DHFR-TS* transcripts accumulate in meristematic tissues and northern blot experiments have confirmed a higher accumulation of the *DHFR-TS* transcripts in proliferating suspension cells compared to cells in stationary phase (Albani et al, 2005). However, the data reported by Neuburger et al (1996) which described strong accumulation of DHFR-TS in the mitochondria of mature pea leaves are in contrast with the expression analyses performed in carrot. This discrepancy could derive from the presence of paralog genes that might be differentially expressed during development and in different tissues.

The genome of *Arabidopsis thaliana* has been fully sequenced and contains three *DHFR-TS* genes called *AtDRTS1, AtDRTS2* and *ATDRTS3*. Interestingly, only the first two genes appear to code for bifunctional DHFR-TS enzymes whereas *ATDRTS3* appears to code for an isoform truncated in the C-terminal TS domain. In fact, *ATDRTS3* contains an ATREP transposable element in its third intron which impairs the correct splicing and causes the retention of the intronic region which leads to a premature termination of the messenger RNA and of the *AtDRTS*3 open reading frame. Moreover, analysis of the genomic organization of the *AtDRTS* loci has revealed that three members of the *SFH* (Sec Fourteen Homologues) gene family are located upstream of the three *DRTS* genes and oriented in divergent direction with respect to them. The SFHs are a diverse group of proteins with distinct subcellular localizations and varied physiological functions related to lipid metabolism, phosphoinositide mediated signalling and membrane trafficking. The conserved organization of the three *AtDRTS/AtSFH* loci suggests that they must derive from a single ancestral locus that underwent at least two duplication events.

In view of the conserved arrangement of the *AtDRTS/AtSFH* loci, we decided to study the patterns of expression of the three *AtDRTS* genes and to verify eventual correlations with the expression of the neighbouring *AtSFH* genes. For this analysis we prepared bidirectional expression constructs in which the genes coding for a chimeric GFP-GUS protein and for a red fluorescent protein (eqFP611) were placed at the extremities of the intergenic regions spanning between the putative *AtDRTS* and *AtSFH* ATG codons. These divergent reporter gene constructs were assembled into a binary vector suitable for Agrobacterium-mediated plant transformation. Histochemical and enzymatic GUS assays were then performed on the resulting transgenic plants to reveal the activity of the *AtDRTS* promoters, whereas fluorescence microscopy analyses were done to monitor the activity of the *AtSFH* promoters. These analyses revealed that *AtDRTS1* is not expressed in the meristems but is highly expressed in the vascular tissues of plantlets at one, two and three weeks of age. Conversely, the *AtDRTS2* promoter appears to drive a very strong and constant cell proliferation-specific gene expression in root and shoot apical meristems. Surprisingly, the *AtDRTS3* promoter showed strong activity in the shoot meristem but not in the meristem of the root where GUS expression was confined to the root columella and the central cylinder. Thus, the three *AtDRTS* genes appear to have very different patterns of expression, and only *AtDRTS2* is specific to division cells and that confirm its activity seen by other studies. Surprisingly the

other two genes are differentially expressed during development in different tissues suggesting that this gene family can be expressed differentially for the presence of other functions not analysed precedently.

The activity of the *AtSFH* promoters, that drive expression of the *eqFP611* red fluorescent protein gene, was detected by fluorescence microscopy analyses of primary transformants and T2 progeny. Concerning the activity of the *AtSFH7* promoter, located upstream of *AtDRTS1*, analyses of the corresponding transgenic plants did not allow the detection of any clear accumulation of the eqFP611 red fluorescent protein. Thus, is likely that the level of activity of the *AtSFH7* promoter is too weak and general to allow its detection using a fluorescent protein reporter gene. This evidence is in accordance with microarrays data included in the Genevestigator database (https://www.genevestigator.com/) that suggest a weak but widespread expression of the *AtSFH7* gene.

Conversely, confirming results published in other studies, specific patterns of expression were detected for the *AtSFH1* and *AtSFH3* promoters, located upstream of *AtDRT2* and *AtDRTS3* respectively. The *AtSFH1/cow1* gene has been shown to be strongly and specifically expressed in roots where its activity is required for polarized membrane growth during the development of root hairs (Bohme et al, 2004; Vincent et al, 2005) whereas the *AtSFH3* gene has been shown to be highly expressed in mature and germinating pollen (Mo et al, 2007). Accordingly, strong red fluorescence was detected only in the roots or in the pollen of our transgenic plants. Thus, in all three cases, the pattern of expression of the *AtSFH* genes is not correlated with the expression of the *AtDRTS* genes.

In silico analyses of the *AtDRTS* promoters revealed the presence of several putative *cis*-acting elements which in the case of the *AtDRTS2* and *AtDRTS3* promoters, in accordance with the expression of these two genes in shoot meristem, include sites conserved in the promoters of several cell cycle-dependent genes. In particular, an E2F binding site is contained in both promoters. Remarkably, the E2F site in *AtDRTS2* is overlapping the transcription start point. The E2F transcription factors are components of the E2F/retinoblastoma pathway which appears to be conserved in animal and plant cells. The diversity of the E2F family members in both plants and animals (Vandepoele et al, 2002) could reflect the complexity of the E2F-mediated regulation of various target genes. The six E2F proteins found in Arabidopsis can be classified as activators or repressors of gene expression (Mariconti et al, 2002, Kosugi and Ohashi 2002). Indeed, several studies have

shown that the E2F sites can be involved in both activation or repression of gene expression.
In 1997 Yamaguchi studied the promoter of the DNA polymerase α 180 KDa catalytic subunit gene of Drosophila which contains three E2F recognition sequences. Mutations were introduced in each of the three E2F sites and their effects on E2F binding and promoter activity in cultured cells and in living flies were evaluated. From these analyses it is resulted that only E2F site 3, the closest to the start of transcription, functions positively to stimulate DNA polα gene promoter activity, while the other two sites (E2F sites 1 and 2) have a negative function throughout Drosophila development. Also in plants, E2F sites are known to be able to mediate transcriptional activation or repression depending on the different stage of development (Egelkrout et al, 2002). Moreover, it has been shown that E2F sites can act as positive elements in the G1/S phase of cell cycle or as negative elements during the progression of cell cycle (Chaboutè et al, 2000).
In both *AtDRTS2* and *AtDRTS3* promoters the E2F site appear to be functional because ChIP analyses have revealed that both sites are recognized *in vivo* by AtE2Fa or AtE2Fb which have been characterized as activating AtE2F factors. Thus, the binding of activating AtE2Fs to the *AtDRTS* promoters suggested that the E2F *cis*-elements in the *AtDRTS2* and *AtDRTS3* promoter may exert positive roles on promoter activity. In most cases, in both animal and plant promoters the E2F *cis*-elements have been shown to be involved in activation of gene expression. Lincker et al. identified two E2F sites on the promoter of the *RNR1* gene of *Nicotiana benthamiana*. Functional analyses revealed that both E2F elements are important for driving *RNR1* promoter induction at the G1/S transition in synchronized tobacco BY-2 cells as well as in response to UV irradiation in dividing non-synchronized cells and S-phase cells. Mutation of both elements resulted in decreased *RNR1* promoter activity in growing cells. However, other *cis*-elements appeared to act as positive regulators in concert with the E2F sites.
To better define the role of the E2F sites located in the *AtDRTS2* and *AtDRTS3* promoters we analysed the effect of their mutation on the regulation of these *AtDRTS* genes. The mutations were introduced in the same bidirectional reporter constructs used for the analysis of the *AtDRTS* promoters givin rise to the pDRTS2-ΔE2F and pDRTS3-ΔE2F vectors. The activity of the promoters with the mutated E2F sites was then analysed by histochemical GUS assays which revealed a pattern of expression very similar to the one obtained with the original promoter constructs. However, the level of GUS activity driven by the two promoters mutated

appeared to be strongly increased in the vascular tissues and upregulated also in the meristematic tissues which were already positive with the wild-type promoter constructs. The level of activity was quantified performing fluorimetric assays on seedlings of the same developmental stage (emergence of the third pair of leaves). These analyses confirmed the fact that the inactivation of the E2F *cis*-elements increased considerably the activity of both promoters, indicating that the E2F factors recognizing these sites are able to repress *AtDRTS* expression. Thus, for the first time, it appears that plant E2Fs so far identified as transcriptional activators by conventional models could also mediate repression of gene expression. Interestingly, the evidence that E2F sites in plant promoters can negatively regulate gene expression has been demonstrated only in the case of promoters which contain more than one E2F site. For example, Egelkrout et al. demonstrated that one of the two E2F elements in the *PCNA* promoter activates the transcription in proliferating cells whereas the other E2F site is a repressor in differentiated tissues (Egelkrout et al, 2002).

The discovery that in both *AtDRTS* promoters the E2F site plays a repressive role implies that other *cis*-acting elements must be involved in the activation of expression in meristematic cells. In particular we focused our studies on the *AtDRTS2* gene which is specifically meristematic. Its pattern of expression agrees with results found in *Daucus carota* where *DHFR-TS* transcripts have been shown to accumulate to high level in dividing meristematic cells (Albani et al. 2005). In silico analyses of the sequence of this promoter have revealed the presence of two remarkable *cis*-elements, called Up1 and Up2, in the 5' region of the *AtDRTS2* gene. The first site is located 80 bp upstream of the transcription start site whereas the second site is located at the start of the 100 bp intron found in the 5' untranslated region of the gene. The Up1 and Up2 sites are *cis*-acting elements that are found in the promoters of several genes involved in cellular division. These sequences are also known as site II motifs and telo-box motifs respectively (Tremousaygue et al, 2003).

The telo-box was first observed within the promoters of Arabidopsis genes encoding the translation elongation factor EF1A. This site, of sequence AAACCCTAA, is conserved at the same location within all the known plant EF1A gene promoters (Axelos et al. 1989; Curie et al. 1991; Liboz et al. 1991) and is the target of nuclear proteins (Regad et al, 1994). Moreover, database searches indicated the presence of telo-boxes in the 5' region of numerous genes encoding components of the translational apparatus. The putative role of the telo-box in the regulation of the Arabidopsis EF1A gene has been previously studied (Curie et

al. 1993). The promoter of this gene exhibits a modular organization, with *cis*-acting elements located both upstream and downstream of the transcription initiation site. Within this promoter, the telo-box is located 63 bp upstream of the transcription initiation site, between the TATA box (-33) and the tef *cis*-acting element (-105), another regulatory sequence involved in the activation of EF1A genes in cycling cells (Regad et al. 1994; Manevski et al. 1999). The telo-box is absolutely required for the activation of expression in root primordia. This activation occurs in a co-operative manner with the tef-cis acting element and was observed when the telo-box was inserted either upstream or downstream of the transcription initiation site. These results indicate that telomere motifs are involved in control of gene expression in plants. Telo-boxes are also observed within the promoter of other plant genes expressed in late G1, such as ribonucleotide reductase (*RNR*) or proliferating cell nuclear antigen (*PCNA*) raising the possibility that this element could be involved in a common regulatory process which connects expression of a set of genes at the G1-S transition (Tremoysaugue et al. 2003).

Both maize and Arabidopsis extracts contain a DNA-binding activity able to specifically interact with the double-stranded telomeric sequence AAACCCTAA and a mutation within this motif (AcgtCCTAA) prevented this specific interaction (Regad et al. 1994). Other studies on *Arabidopsis thaliana* have identified and characterized a protein, called AtPuralpha, that interact with the telomere motif (Tremoysaugue 1999). AtPuralpha protein could be involved in the telo-box dependent regulation of gene expression in dividing cells. In animal cells, the puralpha protein was originally described as a HeLa cell nuclear protein with affinity for a purine-rich element in the 5' region of a *c-myc* gene (Bergemann and Johnson 1992). Although the exact role of Puralpha in mammalian cell replication is still unknown, recent studies have implicated puralpha in the control of the promoters of several different cellular and viral genes (Herault et al. 1992; Kelm et al. 1997; Safak et al. 1999). Data suggest a role for puralpha in cell division through cross-family protein interactions. For example the human puralpha has been shown to interact with the Rb protein, and this interaction suppresses the DNA binding activity of puralpha (Johnson at al, 1995). Morover the association of puralpha and Rb is modulated during the cell cycle, and depends on the phoshorylation state of Rb, with Puralfa binding only to the hypophosphorylated form of Rb and Rb phosphorylation beginning in the late G1. The fact that Rb proteins are present also in plants (Ach et al, 1997)

raises the possibility that AtPurAlpha could interact with this protein also in Arabidopsis cells.

In addition to a telo-box located downstream of the transcription initiation site, the rice PCNA promoter contains site II motifs (the site IIa and IIb motifs) which regulate gene expression in the meristems of transgenic tobacco plants and specifically interact with two rice bHLH transcription factors (PCF1 and PCF2) (Kosugi et al. 1995; Kosugi and Ohashi 2002). Other studies in *Arabidopsis thaliana* demonstrated the interaction of TCP 20 protein, that belongs to a subfamily of 13 members related to rice PCF1 and PCF2 factors, with site II motifs (Cubas, 2002). These elements activate gene expression in Arabidopsis root primordia in synergy with a telo-box (Manevski et al, 2000). Trémousaygue et al. (2003) analysed two PCNA promoters of Arabidopsis that have two elements (TTGGGCC) similar to that IIa motif of the rice PCNA promoter (TGGGCC). To understand the possible roles of these sites a 33bp DNA fragment containing the two putative site II motifs of the Arabidopsis PCNA-2 promoter was inserted, with or without a telo-box, upstream of an Arabidopsis EF1A A1 minimal promoter. The GUS reporter gene was placed downstream of the resulting chimaeric promoters. The insertion of a 33bp DNA fragment containing the two putative site two motifs led to a detectable GUS expression in root primordia and in young leaves. Mutation of these two site II motifs totally abolished the expression of the reporter gene. The presence of a telobox within the chimaeric promoter strongly stimulated the expression of the GUS gene in both root primordia and young leaves. These observations indicate that the telo-box increases the activation driven by other *cis*-acting elements but is not sufficient by itself to induce gene expression in cycling cells. (Trémousaygue et al. 2003).

Structural analysis of the Arabidopsis PCNA, RNR, TS promoters revealed the presence of both site II motifs (Up1 sites) and telo-boxes (Up2 sites). The gene ontology of genes with Up elements in their promoter includes a higher proportion of genes involved in cell cycle, DNA processing and protein synthesis. Studies have defined the function of the Up1 and Up2 site in the expression of genes during axillary bud outgrowth in *Arabidopsis* (Tatematsu et al. 2005). Analysis of genic expression demonstrated that the two sites acts in a synergistic manner as positive elements for the activation of gene expression after decapitation of buds. Up1 and Up2 sites are also significantly over-represented in the upstream regions of genes upregulated during germination. (Tatematsu et al, 2008).

To assess the function of Up elements, Tatematsu et al. (2008) analyzed transgenic lines harbouring the GUS reporter gene driven by the Up1/Up2 containing native promoter of RPL15B, a ribosomal protein gene, or by the same promoter with mutations of Up1 or Up2, or with both Up elements mutated. The analysis of mature seeds revealed that the lines transformed with the wild type promoter showed weak GUS staining in the entire seed, including embryo and endosperm. Following seed germination, strong GUS staining was evident in the radicle and vascular tissues of these lines. The lines with the mutated Up1 or with both Up sites mutated had no detectable GUS staining in either the dry seed or germinated seedling whereas the lines with the mutated Up2 site showed a similar GUS staining pattern as the lines with the wild type promoter. These results indicate that Up1, but not Up2, is necessary for the induction of *RPL15B* expression during germination. Moreover, GUS staining of the Up2-mutated lines was weaker than that of wt lines, suggesting that Up2 may act as an enhancer of Up1-mediated gene induction. Regulation of Up1-mediated gene expression appears to be a common mechanism for the activation of growth in germinating seeds and axillary shoots, both of which are quiescent when necessary.

Based on these evidences, it was important to analyse the possible role of the Up1 and Up2 sites in the control of *AtDRTS2* expression. For this analysis, we introduced single mutations (AtDRTS2-ΔUp1 and AtDRTS2-ΔUp2) as well as mutations of both sites (AtDRTS2-ΔΔ) in the *AtDRTS2* promoter construct driving expression of the GFP/GUS reporter gene. The spatial patterns of activity of the AtDRTS2-ΔUp1, AtDRTS2-ΔUp2 and AtDRTS2-ΔUps promoter constructs appeared to be very similar to the one conferred by the wild type *AtDRTS2* promoter. Both constructs are expressed in shoot and root apical meristems as well as in the developing leaves and also in vascular tissues. However the quantitative analysis of the GUS activity in AtDRTS2-ΔUp1, AtDRTS2-ΔUp2 and AtDRTS2-ΔUps lines revealed clear differences with respect to the lines with the wild type construct and the inactivation of either the Up1 or Up2 sites lowered considerably the activity of the mutated promoter compared to the wild type promoter. Interestingly, the activity of the promoter with both Up sites mutated was similar to the activity of the promoters with the single Up mutations. Therefore, it appears that the Up1 and Up2 sites in the *AtDRTS2* promoter are not necessary for the expression in proliferating cells of Arabidopsis and act together, but not sinergistically, to drive the high level of activity of this promoter.

Because the Up2 site in the first intron of the *AtDRTS2* gene is conferring high levels of expression but is not essential for the meristematic expression of the gene, we analysed the effect of the removal of the entire intron. In fact, introns located in the 5' untranslated region (5'-UTR), as in the case of *DRTS2*, often have a role in the regulation of genic expression. An increase in gene expression associated to the presence of intronic regions has been reported in both monocotyledonous (Callis et al., 1987; McElroy et al., 1990; Christensen et al., 1992; Xu et al., 1994; Jeon et al., 2000; Morello et al., 2002) and dicotyledonous plants (Norris et al., 1993; Gidekel et al., 1996; Rose and Last, 1997; Plesse et al., 2001; Mun et al., 2002). The exact mechanisms responsible for this intron-mediated enhancement of expression are not clear. Some introns contains promoter regulatory elements (Gidekel et al, 1996), while others seems to affect transcript stability or translation efficiency (Rose and Last, 1997), suggesting the existence of different mechanisms. Some introns can also affect tissue-specific gene expression (Terada *et al.*, 1995; Bolle et al, 1996; Jeon et al, 2000). In 2001, Chaubet and Gigot demonstrated that introns in the 5' UTR of the replacement H3 genes, functionally combined with the endogenous promoters, are necessary the high and constitutive expression observed *in planta*. This effect was studied using the promoter of the histone H4 gene H4A748 of Arabidopsis, which drives expression in an S phase- and meristem-specific pattern (Atanassova *et al.*, 1992; Chaubet *et al.*, 1996). The insertion downstream to this cell cycle-dependent promoter of the 5'-UTR intron of either replacement histone H3 genes of Arabidopsis resulted in a meristem-independent expression (loss of preferential meristem expression). Moreover, this study demonstrated that the constitutive expression of the replacement H3 genes of *Arabidopsis* results from both the presence of the intron and the nature of the promoters, which lack some *cis*-elements found in replication-dependent promoters. Other studies in *Arabidopsis thaliana* analysed the involvement of the regions located upstream of the translation start site in the expression of two nuclear COX5c genes encoding subunit 5c of mitochondrial cytochrome c oxidase (Curi et al., 2005). It was observed that these regions, which include a leader intron, direct a tissue-specific expression mainly in root and shoot meristems, actively growing tissues and vascular strands. Removal of the leader intron produced a significant decrease in expression to values only slightly higher than those observed with a promoterless reporter gene. (Curi et al., 2005).

To verify the importance first intron of *AtDRTS2*, located at the 5'-UTR region of the gene, we analysed the activity of promoter constructs containing or lacking the first intron

(AtDRTS2 -830/+175 and AtDRTS2 -830/+51) placed upstream of the *GUS* reporter gene. The lack of the intron in the -830/+51 construct drastically changed the pattern of activity of the *AtDRTS2* promoter, abolishing completely the expression in root and shoot meristems. This surprising result suggests that at least one additional regulatory site different from Up2 is contained in the first intron and is strictly necessary for the meristematic expression of the *AtDRTS2* gene. Linker scanning mutations of 8 bp along the entire first intron revealed that the expression of the *AtDRTS2* promoter requires a small intronic DNA region. Mutation of this intronic sequence in the *AtDRTS2-ΔIII* construct completely abolished the expression of the reporter gene in Arabidopsis plants. This result led us to analyse this particular region, to understand what kind of DNA binding factor could recognize this DNA sequence. A research in Jaspar database yielded a list of transcription factors that could bind the DNA sequence of interest. This list is composed by FOXC1, SRY, NFIL3, Broad Complex IV and Hunchback.

FoxC1 (Forkhead box C1), is a protein which in humans is encoded by the *FOXC1* gene. This gene belongs to the forkhead family of transcription factors which is characterized by a distinct DNA-binding fork head domain. (der Horst et al, 2007). However, hortologues of his protein are not found in plants.

SRY (sex-determining region Y) belongs to the SOX/TCF (SRY-related HMG box/T cell factor) group of transcription factors which is part of the HMG box (high mobility group) superfamily of proteins (Riechmann et al, 2000). Most HMG box proteins contain two or more HMG boxes and act as architectural components of chromatin, binding DNA in a relatively sequence-aspecific manner. Conversely, the SOX/TCF factors, which contain a single HMG box, show sequence-specific DNA binding and transactivation activities. There are 14 genes in the *Arabidopsis* genome encoding HMG box-containing proteins, but phylogenetic analyses indicate that none of these proteins belong to the SOX/TCF group (Riechmann et al, 2000). However, the YABBY family of putative plant transcription factors contains a putative Cys_2–Cys_2 zinc-finger domain near the amino terminus, and a high mobility group (HMG)-like domain, designated the YABBY domain (Bowman and Smyth 1999, Villanueva et al, 1999), which shows sequence similarity to DNA-binding domains of HMG transcription factors such as the human SRY protein. The HMG-like domain of SRY consists of three α-helices separated by turns, with the first two helices binding directly to DNA (Werner et al. 1995). The YABBY domains align with the first two helices of this motif and are predicted to form two α-helices separated by a turn (Villanueva et al. 1999).

NFIL3 is a mammalian nuclear factor, interleukin 3 regulated, also known as E4BP4, which is restricted to activated T cells, natural killer (NK) cells, and mast cell lines. It is a bZIP protein with no clear hortologues in plants, but is worth noticing that several bZIP proteins are also present in plants.

Broad Complex is a superfamily of proteins that has been described also in plants and contains a conserved protein–protein interaction motif named broad complex, tramtrack, bric-a-brac (BTB), also known as the poxvirus and zinc finger (POZ) domain. BTB/POZ domains can be found in combination with various other protein–protein interaction motifs, indicating involvement in various biological processes and the BTB/POZ proteins are often transcriptional regulators containing a C2H2 domain for DNA binding. The *Arabidopsis* genome contains eighty genes encoding BTB/POZ domain-containing proteins. However, the actin-binding Kelch/BTB proteins as well as the DNA-binding zinc finger BTBs are absent from the *Arabidopsis* superfamily (Gingerich et al, 2007).

Hunchback proteins are proteins showing sequence-specific DNA binding activity, which indicates that they might regulate gene expression at the level of transcription. These proteins are present in animals but there is no evidence of their presence in plants.

It is clear therefore that the transcription factors that could recognize the ΔIII region in the first intron of *AtDRTS2* is unlikely to be an hortologue of the animal proteins identified in the Jaspar database. Thus, the identification of this novel plant transcription factor, essential for the meristematic expression of *AtDRTS2*, will require additional experiments.

6. BIBLIOGRAPHY

Ach RA, Durfee T, Miller AB, Taranto P, Hanley-Bowdoin L, Zambryski PC, Gruissem W. 1997. RRB1 and RRB2 encode maize retinoblastoma-related proteins that interact with a plant D-type cyclin and geminivirus replication protein. *Mol Cell Biol*. **17**, 5077-86.

Albani D, Giorgetti L, Pitto L, Luo M, Cantoni RM, Erra Pujada M, Rotino GL, Cella R. 2005.Proliferation-dependent pattern of expression of a dihydrofolate reductase-thymidylate synthase gene from Daucus carota *Eur J Histochem*. **49**, 107-15.

Albani D, Mariconti L, Ricagno S, Pitto L, Moroni C, Helin K, Cella R. 2000. DcE2F, a functional plant E2F-like transcriptional activator from Daucus carota. .*J Biol Chem*. **275**, 19258-67.

Appling DR. 1991. Compartmentation of folate-mediated one-carbon metabolism in eukaryotes.
Arabidopsis thaliana. *Plant Mol Biol*. **43**, 583-593.

Arch RH, Thompson CB. 1998. 4-1BB and Ox40 are members of a tumor necrosis factor (TNF)-nerve growth factor receptor subfamily that bind TNF receptor-associated factors and activate nuclear factor kappaB. *Mol Cell Biol*. **18**, 558-65.

Atanassova R, Chaubet N, Gigot C.A.1992. 126 bp fragment of a plant histone gene promoter confers preferential expression in meristems of transgenic Arabidopsis. *Plant J*.**2**, 291-300.

Attwooll C, Lazzerini Denchi E, Helin K. 2004. The E2F family: specific functions and overlapping interests. *EMBO J.* **23**, 4709-16.

Axelos M, Bardet C, Liboz T, Le Van Thai A, Curie C, Lescure B. 1989. The gene family encoding the Arabidopsis thaliana translation elongation factor EF-1 alpha: molecular cloning, characterization and expression. *Mol Gen Genet*. **219**, 106-12.

Ayusawa D, Koyama H, Shimizu K, Kaneda S, Takeishi K, Seno T. 1986. Induction, by thymidylate stress, of genetic recombination as evidenced by deletion of a transferred genetic marker in mouse FM3A cells. *Mol Cell Biol.* **6**, 3463-9.

Bachmann B, Follmann H.1987. Deoxyribonucleotide biosynthesis in green algae: characterization of thymidylate synthase-dihydrofolate reductase in Scenedesmus obliquus.*Arch Biochem Biophys*. **256**, 244-52.

Bailey LB, Gregory JF 3rd. 1999. Polymorphisms of methylenetetrahydrofolate reductase and other enzymes: metabolic significance, risks and impact on folate requirement. *J Nutr*. **129**, 919-22.

Baker SJ, Reddy EP. 1996. Transducers of life and death: TNF receptor superfamily and associated proteins. *Oncogene*. **12**, 1-9.

Baldauf SL.A. 1999. Search for the Origins of Animals and Fungi: Comparing and Combining Molecular Data. *Am Nat*. **154**, 178-188.

Baserga R.1991. Growth regulation of the PCNA gene. *J Cell Sci*. **98**, 433-6.

Basset GJ, Ravanel S, Quinlivan EP, White R, Giovannoni JJ, Rébeillé F, Nichols BP, Shinozaki K, Seki M, Gregory JF 3rd, Hanson AD. 2004. Folate synthesis in plants: the last step of the p-aminobenzoate branch is catalyzed by a plastidial aminodeoxychorismate lyase. *Plant J*. **40**, 453-61.

Bergemann AD, Johnson EM. 1992.The HeLa Pur factor binds single-stranded DNA at a specific element conserved in gene flanking regions and origins of DNA replication. *Mol Cell Biol.* **12**, 1257-65.

Bianchi ME, Falciola L, Ferrari S, Lilley DM. 1992. The DNA binding site of HMG1 protein is composed of two similar segments (HMG boxes), both of which have counterparts in other eukaryotic regulatory proteins. *EMBO J.* **11**, 1055-63.

Böhme K, Li Y, Charlot F, Grierson C, Marrocco K, Okada K, Laloue M, Nogué F. 2004. The Arabidopsis COW1 gene encodes a phosphatidylinositol transfer protein essential for root hair tip growth. *Plant J*. **40**, 686-98.

Bolle C, Herrmann RG, Oelmüller R. 1996. Intron sequences are involved in the plastid- and light-dependent expression of the spinach PsaD gene. *Plant J*. **10**, 919-24.

Boudolf V, Vlieghe K, Beemster GT, Magyar Z, Torres Acosta JA, Maes S, Van Der Schueren E, Inzé D, De Veylder L. 2004. The plant-specific cyclin-dependent kinase CDKB1;1 and transcription factor E2Fa-DPa control the balance of mitotically dividing and endoreduplicating cells in Arabidopsis. *Plant Cell*. **16**, 2683-92.

Brosius J, Gould SJ. 1992. On "genomenclature": a comprehensive (and respectful) taxonomy for pseudogenes and other "junk DNA".*Proc Natl Acad Sci U S A*. **89**, 10706-10.

Brosius J. 2003. The contribution of RNAs and retroposition to evolutionary novelties.*Genetica*. **118**, 99-116.

Brosius J. 2005. Waste not, want not--transcript excess in multicellular eukaryotes.Trends *Genet*. **21**, 287-8.

Callis J, Fromm M, Walbot V. 1987. Introns increase gene expression in cultured maize cells. *Genes Dev*. **1**, 1183-200.

Cavalier-Smith T. 2000. Membrane heredity and early chloroplast evolution.*Trends Plant Sci***.** **5**, 174-82.

Cella, R. and Parisi, B. 1993. Dihydrofolate reductase and thymidylate synthase plants: an open problem. *Physiologia Plantarum.* **88,** 509-521.

Chabouté ME, Clément B, Philipps G. 2002. S phase and meristem-specific expression of the tobacco RNR1b gene is mediated by an E2F element located in the 5' leader sequence. *J Biol Chem***. 277**, 17845-51.

Chabouté ME, Clément B, Sekine M, Philipps G, Chaubet-Gigot N. 2000.Cell cycle regulation of the tobacco ribonucleotide reductase small subunit gene is mediated by E2F-like elements.*Plant Cell.* **12**, 1987-2000.

Chen L, Chan SY, Cossins EA.1997. Distribution of Folate Derivatives and Enzymes for Synthesis of 10-Formyltetrahydrofolate in Cytosolic and Mitochondrial Fractions of Pea Leaves. *Plant Physiol.* **115**, 299-309.

Cheng Y, Qin G, Dai X, Zhao Y. 2007. NPY1, a BTB-NPH3-like protein, plays a critical role in auxin-regulated organogenesis in Arabidopsis. *Proc Natl Acad Sci U S A*. **104**, 18825-9.

Christensen J, Cloos P, Toftegaard U, Klinkenberg D, Bracken AP, Trinh E, HeeranM, Di Stefano L, Helin K. 2005. Characterization of E2F8, a novel E2F-like cell-cycle regulated repressor of E2F-activated transcription. *Nucleic Acids Res*. **33**, 5458-5470.

Christensen AH, Sharrock RA, Quail PH. 1992. Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol Biol*. **18**, 675-89.

Cossins E.A. 2000. The fascinating world of folate and one-carbon metabolism. *Can. J. Bot*. **78**, 691-708.

Cossins E.A.1984. Folates in biological materials. In: Folates and pterins. (Blakley R.L., Benkovic S.K. eds.) *Wiley (intersciences)*, New York. Vol**. I**: 1-59.

Cowell IG, Skinner A, Hurst HC.1992. Transcriptional repression by a novel member of the bZIP family of transcription factors. *Mol Cell Biol*. **12**, 3070-7.

Cox K, Robertson D, Fites R. 1999. Mapping and expression of a bifunctional thymidylate synthase, dihydrofolate reductase gene from maize. *Plant Mol Biol*. **41**, 733-9.

Cubas P. 2002. Role of TCP genes in the evolution of morphological characters in angiosperms. In Developmental Genetics and Plant Evolution (Cronk QCB, BAterman RM and Hawkins JA; eds), London: Taylor and Francis, pp: 247-266

Curie C, Liboz T, Bardet C, Gander E, Médale C, Axelos M, Lescure B. 1991. Cis and trans-acting elements involved in the activation of Arabidopsis thaliana A1 gene encoding the translation elongation factor EF-1 alpha. *Nucleic Acids Res*. **19**, 1305-10.

Curie C, Axelos M, Bardet C, Atanassova R, Chaubet N, Lescure B. 1993. Modular organization and development activity of an Arabidopsis thaliana EF-1 alpha gene promoter. *Mol Gen Genet*. **238**, 428-36.

de Bruin A, Maiti B, Jakoi L, Timmers C, Buerki R, Leone G. 2003. Identification and characterization of E2F7, a novel mammalian E2F family member capable of blocking cellular proliferation. *J. Biol. Chem*. **278**, 42041-42049.
de Jager SM, Menges M, Bauer UM, Murray JA. 2001. *Arabidopsis* E2F1 binds a sequence present in the promoter of S-phase-regulated gene AtCDC6 and is a member of a multigene family with differential activities. *Plant Mol. Biol*. **47**, 555-568.

De Veylder L, Beeckman T, Beemster GT, de Almeida Engler J, Ormenese S, Maes S, Naudts M, Van Der Schueren E, Jacqmard A, Engler G, Inze D. 2002. Control of proliferation, endoreduplication and differentiation by the *Arabidopsis* E2Fa-DPa transcription factor. *EMBO J*. **21,** 1360-1368.

del Pozo JC, Boniotti MB, Gutierrez C. 2002. *Arabidopsis* E2Fc functions in cell division and is degraded by the ubiquitin-SCF(AtSKP2) pathway in response to light. *Plant Cell*. **14**, 3057-3071.

del Pozo JC, Diaz-Trivino S, Cisneros N, Gutierrez C. 2006. The balance between cell division and endoreplication depends on E2FC-DPB, transcription factors regulated by the ubiquitin-SCFSKP2A pathway in Arabidopsis. *Plant Cell*. **18**, 2224-35.

den Boer BG, and Murray JA. 2000. Control of plant growth and development through manipulation of cell-cycle genes. *Curr. Opin. Biotechnol*. **11**, 138-145.

Denny P, Swift S, Brand N, Dabhade N, Barton P, Ashworth A.A. 1992. Conserved family of genes related to the testis determining gene, SRY. *Nucleic Acids Res*. **20**, 2887.

van der Horst A, Burgering BM. 2007.Stressing the role of FoxO proteins in lifespan and disease. *Nat. Rev. Mol. Cell Biol.* **8**, 440–50.

Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, Creelman R, Pilgrim M, Broun P, Zhang JZ, Ghandehari D, Sherman BK, Yu G. 2000. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science*. **290**, 2105-10.

Villanueva JM, Broadhvest J, Hauser BA, Meister RJ, Schneitz K, Gasser CS. 1999. INNER NO OUTER regulates abaxial- adaxial patterning in Arabidopsis ovules. *Genes Dev.* **13**, 3160-9.

Bowman JL, Smyth. DR. 1999. *CRABS CLAW*, a gene that regulates carpel and nectary development in Arabidopsis, encodes a novel protein with zinc finger and helix-loop-helix domains. *Development* **126:** 2387–2396.

Werner MH, Huth JR, Gronenborn AM, and Clore GM.1995. The molecular basis of human 46x,Y sex reversal revealed from the three-dimensional solution structure of the human Sry-DNA complex. *Cell* **81:** 705–714.

Gingerich DJ, Gagne JM, Salter DW, Hellmann H, Estelle M, Ma L, Vierstra RD. 2005. Cullins 3a and 3b assemble with members of the broad complex/tramtrack/bric-a-brac (BTB)

protein family to form essential ubiquitin-protein ligases (E3s) in Arabidopsis. *J Biol Chem*. **280**, 18810-21.

Gingerich DJ, Hanada K, Shiu SH, Vierstra RD. 2007. Large-scale, lineage-specific expansion of a bric-a-brac/tramtrack/broad complex ubiquitin-ligase gene family in rice. *Plant Cell.* **19**, 2329-48.

Desvoyes B, Ramirez-Parra E, Xie Q, Chua NH, Gutierrez C. 2006. Cell type-specific role of the retinoblastoma/E2F pathway during Arabidopsis leaf development. *Plant Physiol*. **140**, 67-80.

Di Fiore B, Palena A, Felsani A, Palitti F, Caruso M, Lavia P. 1999. Cytosine methylation transforms an E2F site in the retinoblastoma gene promoter into a binding site for the general repressor methylcytosine-binding protein 2 (MeCP2).*Nucleic Acids Res*. **27**, 2852-9.

Di Stefano L, Jensen MR, Helin K. 2003. E2F7, a novel E2F featuring DP-independent repression of a subset of E2F-regulated genes. *EMBO J*. **22**, 6289-6298.

Diffley JF, Stillman B.A.1991. Close relative of the nuclear, chromosomal high-mobility group protein HMG1 in yeast mitochondria. *Proc Natl Acad Sci U S A*. **88**, 7864-8.

Dimova DK, Dyson NJ. 2005. The E2F transcriptional network: old acquaintances with new faces. *Oncogene*. **24**, 2810-26.

Dong X, Li X, Zhang Y, Fan W, Kinkema M, Clarke J. 2001. Regulation of systemic acquired resistance by NPR1 and its partners. *Novartis Found Symp*. **236**:165-73.

Dyson N. 1998. The regulation of E2F by pRB-family proteins. E2F3 activity is regulated during the cell cycle and is required for the induction of S phase. *Genes Dev*. **12**, 2120-2130.

Egelkrout EM, Mariconti L, Settlage SB, Cella R, Robertson D, Hanley-Bowdoin L. 2002. Two E2F elements regulate the proliferating cell nuclear antigen promoter differently during leaf development. *Plant Cell*. **14**, 3225-3236.

Egelkrout EM, Robertson D, Hanley-Bowdoin L. 2001. Proliferating cell nuclear antigen transcription is repressed through an E2F consensus element and activated by geminivirus infection in mature leaves. *Plant Cell*. **13**, 1437-1452.

Ewen ME, Xing YG, Lawrence JB, Livingston DM. 1991. Molecular cloning, chromosomal mapping, and expression of the cDNA for p107, a retinoblastoma gene product-related protein. *Cell*. **66**, 1155-64

Ferrari S, Harley VR, Pontiggia A, Goodfellow PN, Lovell-Badge R, Bianchi ME. 1992. SRY, like HMG1, recognizes sharp angles in DNA. *EMBO J*. **11**, 4497-506.

Ferrari S, Harley VR, Pontiggia A, Goodfellow PN, Lovell-Badge R, Bianchi ME. 1992. SRY, like HMG1, recognizes sharp angles in DNA. *EMBO J.* **11**, 4497-506.

Flemington EK, Speck SH, Kaelin WG Jr. 1993. E2F-1-mediated transactivation is inhibited by complex formation with the retinoblastoma susceptibility gene product.*Proc Natl Acad Sci U S A*. **90**, 6914-8.

Fry CJ, Farnham PJ. 1999. Context-dependent transcriptional regulation. *J Biol Chem*. **274**, 29583-6.

Fuxa M, Skok JA. 2007. Transcriptional regulation in early B cell development. *Curr.Opin Immunol.* **19**, 129-36.
Gambonnet B., Jabrin S., Ravanel S., Karan M., Douce R., Rébeillé F. 2001. Folate distribution during higher plant development. *J. Sci. Food Agrie*. **81**, 835-841.

Gidekel M, Jimenez B, Herrera-Estrella L. 1996. The first intron of the Arabidopsis thaliana gene coding for elongation factor 1 beta contains an enhancer-like element. *Gene*.**170**, 201-6.

Giese K, Cox J, Grosschedl R. 1992. The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures.*Cell*. **69**, 185-95.

Gingerich DJ, Gagne JM, Salter DW, Hellmann H, Estelle M, Ma L, Vierstra RD. 2005. Cullins 3a and 3b assemble with members of the broad complex/tramtrack/bric-a-brac (BTB) protein family to form essential ubiquitin-protein ligases (E3s) in Arabidopsis. *J Biol Chem*. **280**, 18810-21.

Gingerich DJ, Hanada K, Shiu SH, Vierstra RD. 2007. Large-scale, lineage-specific expansion of a bric-a-brac/tramtrack/broad complex ubiquitin-ligase gene family in rice. *Plant Cell.* **19**, 2329-48.

Grafi G, Larkins BA. 1995. Endoreduplication in Maize Endosperm: Involvement of M Phase--Promoting Factor Inhibition and Induction of S Phase--Related Kinases. *Science.* **269**, 1262-1264.

Gutierrez C. 1998. The retinoblastoma pathway in plant cell cycle and development. *Curr. Opin. Plant Biol*. **1**, 492-497.

Ha CM, Jun JH, Nam HG, Fletcher JC. 2004. BLADE-ON-PETIOLE1 encodes a BTB/POZ domain protein required for leaf morphogenesis in Arabidopsis thaliana.*Plant Cell Physiol*. **45**, 1361-70.

Hanson AD, Gregory JF 3rd. 2002. Synthesis and turnover of folates in plants. *Curr Opin Plant Biol*. **5**, 244-9.

Hanson AD, Roje S. 2001. One-Carbon Metabolism in Higher Plants. *Annu Rev Plant Physiol Plant Mol Biol.* **52**, 119-137.

Haqq CM, King CY, Donahoe PK, Weiss MA. 1993. SRY recognizes conserved DNA sites in sex-specific promoters. *Proc Natl Acad Sci U S A.* **90**, 1097-101.

Harbour JW, Dean DC. 2000. The Rb/E2F pathway: expanding roles and emerging paradigms. *Genes Dev.* **14**, 2393-409.

Harley VR, Jackson DI, Hextall PJ, Hawkins JR, Berkovitz GD, Sockanathan S, Lovell-Badge R, Goodfellow PN. 1992. DNA binding activity of recombinant SRY from normal males and XY females. *Science.* **255**, 453-6.

Herault Y, Chatelain G, Brun G, Michael D. 1992. v.src induced-transription of the avian clusterin gene. *Nucl. Acids Res*. **20**, 6377-6383.

Hülskamp M, Lukowitz W, Beermann A, Glaser G, Tautz D. 1994. Differential regulation of target genes by different alleles of the segmentation gene hunchback in Drosophila. *Genetics*. **138**, 125-34.

Huynh KD, Bardwell VJ. 1998. The BCL-6 POZ domain and other POZ domains interact with the co-repressors N-CoR and SMRT. *Oncogene.* **17**, 2473-84.

Ivanetich KM, Santi DV. 1990. Bifunctional thymidylate synthase-dihydrofolate reductase in protozoa. *FASEB J*. **4**, 1591-7.

Jabrin S, Ravanel S, Gambonnet B, Douce R, Rébeillé F. 2003. One-carbon metabolism in plants. Regulation of tetrahydrofolate synthesis during germination and seedling development. *Plant Physiol.* **131**,1431-9.

Jantzen HM, Admon A, Bell SP, Tina R. 1990. Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. *Nature*. **344**, 830-6.

Jenh CH, Geyer PK, Johnson LF. 1985. Control of thymidylate synthase mRNA content and gene transcription in an overproducing mouse cell line. *Mol Cell Biol*. **5**, 2527-32.

Jeon JS, Lee S, Jung KH, Jun SH, Kim C, An G. 2000. Tissue-preferential expression of a rice alpha-tubulin gene, OsTubA1, mediated by the first intron. *Plant Physiol.* **123**, 1005-14.

Johnson EM, Chen PL, Krachmarov CP, Barr SM, Kanovsky M, Ma ZW, Lee WH. 1995. Association of human Pur alpha with the retinoblastoma protein, Rb, regulates binding to the single-stranded DNA Pur alpha recognition element. *J Biol Chem*. **270**, 24352-60.

Kaestner KH, Knochel W, Martinez DE. 2000. Unified nomenclature for the winged helix/forkhead transcription factors. *Genes Dev*. **14**, 142-6.

Kelley KM, Wang H, Ratnam M. 2003. Dual regulation of ets-activated gene expression by SP1. *Gene*. **307**, 87-97.

Kelm RJ Jr, Elder PK, Strauch AR, Getz MJ. 1997. Sequence of cDNAs encoding components of vascular actin single-stranded DNA-binding factor 2 establish identity to Puralpha and Purbeta. *J Biol Chem*. **272**, 26727-33.

Kim S, Choi HI, Ryu HJ, Park JH, Kim MD, Kim SY 2004. ARIA, an Arabidopsis arm repeat protein interacting with a transcriptional regulator of abscisic acid-responsive gene expression, is a novel abscisic acid signaling component. *Plant Physiol.* **136**, 3639-48.

Kosugi S, Ohashi Y. 2002. E2F sites that can interact with E2F proteins cloned from rice are required for meristematic tissue-specific expression of rice and tobacco proliferating cell nuclear antigen promoters. *Plant J.* **29**, 45-59.

Kosugi S, Ohashi Y. 2003. Constitutive E2F expression in tobacco plants exhibits altered cell cycle control and morphological change in a cell type-specific manner. *Plant Physiol*. **132**, 2012-2022.

Kovesdi I, Reichel R, Nevins JR. 1987. Role of an adenovirus E2 promoter binding factor in E1A-mediated coordinate gene control. *Proc Natl Acad Sci U S A*. **84**, 2180-4.

Kowles RV, Yerk GL, Haas KM, Phillips RL. 1997.Maternal effects influencing DNA endoreduplication in developing endosperm of Zea mays. *Genome*. **40**, 798-805.

Kowles RV. and Phillips R.L. 1988. Endosperm development in maize. *Int. Rev. Cytol.* **112**, 97-136.

Lazar G, Zhang H, Goodman HM. 1993. The origin of the bifunctional dihydrofolate reductase-thymidylate synthase isogenes of *Arabidopsis thaliana*. *Plant J*. **3**, 6576-6678.

Liboz T, Bardet C, Le Van Thai A, Axelos M, Lescure B. 1990. The four members of the gene family encoding the Arabidopsis thaliana translation elongation factor EF-1 alpha are actively transcribed. *Plant Mol Biol*. **14**, 107-10.

Luo M, Orsi R, Patrucco E, Pancaldi S, Cella R. 1997. Multiple transcription start sites of the carrot dihydrofolate reductase-thymidylate synthase gene, and sub-cellular localization of the bifunctional protein. *Plant Mol Biol*. **33**, 709-22.

Luo M, Piffanelli P, Rastelli L, Cella R. 1993. Molecular cloning and analysis of a cDNA coding for the bifunctional dihydrofolate reductase-thymidylate synthase of Daucus carota. *Plant Mol Biol*. **22**, 427-35.

Luo M.Z., Cella R.1998. Structure of dihydrofolate reductase- thymidylate synthase gene (Accession N. AJ003139) from *Daucus carota* Plant Gene Register PGR98-055. *Plant Physiology* **116**, 1603.

Magyar Z, Atanassova A, De Veylder L, Rombauts S, Inze D. 2000. Characterization of two distinct DP-related genes from *Arabidopsis thaliana*. *FEBS Lett*. **486**, 79-87.

Magyar Z, De Veylder L, Atanassova A, Bako L, Inze D, Bogre L. 2005. The Role of the *Arabidopsis* E2FB Transcription Factor in Regulating Auxin-Dependent Cell Division. *Plant Cell*.; **17**, 2527-2541.

Maiti B, Li J, de Bruin A, Gordon F, Timmers C, Opavsky R, Patil K, Tuttle J, Cleghorn W, Leone G. 2005. Cloning and characterization of mouse E2F8, a novel mammalian E2F family member capable of blocking cellular proliferation. *J. Biol. Chem*. **280**, 18211-18220.

Manevski A, Bertoni G, Bardet C, Tremousaygue D, Lescure B. 2000. In synergy with various cis-acting elements, plant insterstitial telomere motifs regulate gene expression in Arabidopsis root meristems. *FEBS Lett*. **483**, 43-6.

Mariconti L, Pellegrini B, Cantoni R, Stevens R, Bergounioux C, Cella R, Albani D. 2002. The E2F family of transcription factors from *Arabidopsis thaliana*. Novel and conserved components of the retinoblastoma/E2F pathway in plants. *J. Biol. Chem*. **277**, 9911-9919.

McCarty AS, Kleiger G, Eisenberg D, Smale ST. 2003. Selective dimerization of a C2H2 zinc finger subfamily. *Mol Cell*. **11**, 459-70.

McElroy D, Zhang W, Cao J, Wu R. 1990. Isolation of an efficient actin promoter for use in rice transformation. *Plant Cell*. **2**, 163-71.

Meek T.D., Garvey E.P., Santi D.Y. 1985. Purification and characterization of a bifunctional thymidylate synthase-dihydrofolic reductasi from methotrexate resistant *Leishmania tropica*. *Biochemistry*. **24**, 678-686.

Menges M, and Murray JA. 2002. Synchronous *Arabidopsis* suspension cultures for analysis of cell-cycle gene activity. *Plant J*. **30**, 203-212.

Miyazawa H, Izumi M, Tada S, Takada R, Masutani M, Ui M, Hanaoka F. 1993. Molecular cloning of the cDNAs for the four subunits of mouse DNA polymerase alpha-primase complex and their gene expression during cell proliferation and the cell cycle. *J Biol Chem*. **268**, 8111-22.

Mo P, Zhu Y, Liu X, Zhang A, Yan C, Wang D. 2007.Identification of two phosphatidylinositol/phosphatidylcholine transfer protein genes that are predominately transcribed in the flowers of Arabidopsis thaliana. *J Plant Physiol*. **164**, 478-86.

Morello L, Bardini M, Sala F, Breviario D.A. 2002.long leader intron of the Ostub16 rice beta-tubulin gene is required for high-level gene expression and can autonomously promote transcription both in vivo and in vitro. *Plant J.* **29**, 33-44.

Mun JH, Lee SY, Yu HJ, Jeong YM, Shin MY, Kim H, Lee I, Kim SG. 2002. Petunia actin-depolymerizing factor is mainly accumulated in vascular tissue and its gene expression is enhanced by the first intron. *Gene*. **292**, 233-43.

Neuburger M, Rébeillé F, Jourdain A, Nakamura S, Douce R. 1996. Mitochondria are a major site for folate and thymidylate synthesis in plants. *J Biol Chem*. **271**, 9466-72.

Nishimura DY, Swiderski RE, Alward WL, Searby CC, Patil SR, Bennet SR, Kanis AB, Gastier JM, Stone EM, Sheffield VC. 1998. The forkhead transcription factor gene FKHL7 is responsible for glaucoma phenotypes which map to 6p25. *Nat Genet.* **19**, 140-7.

Norris SR, Meyer SE, Callis J. 1993. The intron of Arabidopsis thaliana polyubiquitin genes is conserved in location and is a quantitative determinant of chimeric gene expression. *Plant Mol Biol.* **21**, 895-906.

Ohtani K, DeGregori J, Leone G, Herendeen DR, Kelly TJ, Nevins JR. 1996. Expression of the HsOrc1 gene, a human ORC1 homolog, is regulated by cell proliferation via the E2F transcription factor. *Mol Cell Biol*. **16**, 6977-6984.

Oosterwegel MA, van de Wetering ML, Holstege FC, Prosser HM, Owen MJ, Clevers HC. 1991. TCF-1, a T cell-specific transcription factor of the HMG box family, interacts with sequence motifs in the TCR beta and TCR delta enhancers. *Int Immunol.* **3**, 1189-92.

Parisi MA, Clayton DA. 1991. Similarity of human mitochondrial transcription factor 1 to high mobility group proteins. *Science*. **252**, 965-9.

Pedmale UV, Liscum E. 2007. Regulation of phototropic signaling in Arabidopsis via phosphorylation state changes in the phototropin 1-interacting protein NPH3. *J Biol Chem*. **282**, 19992-20001.

Philippe H, Germot A, Moreira D. 2000. The new phylogeny of eukaryotes. *Curr Opin Genet Dev.* **10**, :596-601.

Pierrou S, Hellqvist M, Samuelsson L, Enerbäck S, Carlsson P.1994. Cloning and characterization of seven human forkhead proteins: binding site specificity and DNA bending. *EMBO J*. **13**, 5002-12.

Plesse B, Criqui MC, Durr A, Parmentier Y, Fleck J, Genschik P. 2001. Effects of the polyubiquitin gene Ubi. U4 leader intron and first ubiquitin monomer on reporter gene expression in Nicotiana tabacum. *Plant Mol Biol*. **45**, 655-67.

Prabhu V, Chatson KB, Abrams GD, King J. 1996. 13C nuclear magnetic resonance detection of interactions of serine hydroxymethyltransferase with C1-tetrahydrofolate synthase and glycine decarboxylase complex activities in Arabidopsis. *Plant Physiol*. **112**, 207-16.

Ramirez-Parra E, Xie Q, Boniotti MB, Gutierrez C. 1999. The cloning of plant E2F, aretinoblastoma-binding protein, reveals unique and conserved features with animal G(1)/S regulators. *Nucleic Acids Res*. **27**, 3527-3533.

Ravanel S, Cherest H, Jabrin S, Grunwald D, Surdin-Kerjan Y, Douce R, Rébeillé F. 2001. Tetrahydrofolate biosynthesis in plants: molecular and functional characterization of dihydrofolate synthetase and three isoforms of folylpolyglutamate synthetase in Arabidopsis thaliana. *Proc Natl Acad Sci U S A*. **98**,15360-5.

Rébeillé F, Macherel D, Mouillon JM, Garin J, Douce R. 1997. Folate biosynthesis in higher plants: purification and molecular cloning of a bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/7,8-dihydropteroate synthase localized in mitochondria. *EMBO J.* **16**, 947-57.

Regad F, Lebas M, Lescure B. 1994. Interstitial telomeric repeats within the Arabidopsis thaliana genome. *J Mol Biol*. **239**, 163-9.

Roos AJ, Cossins EA.1971. Pteroylglutamate derivatives in Pisum sativum L. Biosynthesis of cotyledonary tetrahydropteroylglutamates during germination. *Biochem J*. **125**, 17-26.

Rose AB, Last RL. 1997. Introns act post-transcriptionally to increase expression of the Arabidopsis thaliana tryptophan pathway gene PAT1. *Plant J.* **11**, 455-64.

Rossignol P, Stevens R, Perennes C, Jasinski S, Cella R, Tremousaygue D, Bergounioux C. 2002. AtE2F-a and AtDP-a, members of the E2F family of transcription factors, induce *Arabidopsis* leaf cells to re-enter S phase. *Mol. Genet. Genomics*. **266,** 995-1003.

Safak M, Gallia GL, Khalili K. 1999. Reciprocal interaction between two cellular proteins, Puralpha and YB-1, modulates transcriptional activity of JCVCY in glial cells. *Mol Cell Biol.* **19**, 2712-23.

Sahr T., Ravanel S., Rébeillé F. 2005.Tetrahydrofolate byosynthesis and distribution in higher plants. *Biochemical Society Transactions*. **33** part 4.

Sandelin A, Alkema W, Engström P, Wasserman WW, Lenhard B. 2004. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res*. **32**, 91-4.

Santi DV, Danenberg PV.1984. Folates in pyrimidine nucleotide biosynthesis. In: Blakley RL, Benkovic SJ (eds) Folates and Pterins, vol. 1, pp. 343-396. *John Wiley*, NewYork.

Schweitzer BI, Dicker AP, Bertino JR. 1990. Dihydrofolate reductase as a therapeutic target. *FASEB J.* **4**, 2441-52.

Scott J., Rébeillé F., Fletcher J. 2000. Folic acid and folates: the feasibility for nutritional enhancement in plant foods. *J. Sci. Food Agrie*. **80**, 795-824.

Sekine M, Ito M, Uemukai K, Maeda Y, Nakagami H, Shinmyo A. 1999. Isolation and characterization of the E2F-like gene in plants. *FEBS Lett*. **460**, 117-122.

Shen WH. 2002. The plant E2F-Rb pathway and epigenetic control. *Trends Plant Sci*. **7**, 505-11.

Sozzani R, Maggio C, Varotto S, Canova S, Bergounioux C, Albani D, Cella R. 2006. Interplay between Arabidopsis activating factors E2Fb and E2Fa in cell cycle progression and development. *Plant Physiol*. **140**, 1355-66.

Stals H, Casteels P, Van Montagu M, Inze D. 2000. Regulation of cyclin-dependent kinases in Stechmann A. and Cavalier-Smith T. 2002. Rooting the eucaryote tree by using a derived gene fusion. *Science.* **297**.

Stevens R, Mariconti L, Rossignol P, Perennes C, Cella R, Bergounioux C. 2002. Two E2F sites in the *Arabidopsis* MCM3 promoter have different roles in cell cycle activation and meristematic expression. *J. Biol. Chem*. **277**, 32978-32984.

Stogios PJ, Downs GS, Jauhal JJ, Nandra SK, Privé GG. 2005. Sequence and structural analysis of BTB domain proteins. *Genome Biol*. **6**, R82.

Storms RK, Ord RW, Greenwood MT, Mirdamadi B, Chu FK, Belfort M.1984. Cell cycle-dependent expression of thymidylate synthase in Saccharomyces cerevisiae. *Mol Cell Biol*. **4**, 2858-64

Tatematsu K, Ward S, Leyser O, Kamiya Y, Nambara E. 2005. Identification of cis-elements that regulate gene expression during initiation of axillary bud outgrowth in Arabidopsis. *Plant Physiol*. **138**, 757-66.

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

Terada R, Nakayama T, Iwabuchi M, Shimamoto K. 1995. A type I element composed of the hexamer (ACGTCA) and octamer (CGCGGATC) motifs plays a role(s) in meristematic expression of a wheat histone H3 gene in transgenic rice plants. *Plant Mol Biol*. **27**, 17-26.

Timmermans MC, Maliga P, Vieira J, Messing J. 1990. The pFF plasmids: cassettes utilising CaMV sequences for expression of foreign genes in plants. *J Biotechnol*. **14**, 333-44.

Trémousaygue D, Garnier L, Bardet C, Dabos P, Hervé C, Lescure B. 2003. Internal telomeric repeats and 'TCP domain' protein-binding sites co-operate to regulate gene expression in Arabidopsis thaliana cycling cells. *Plant J*. **33**, 957-66.

Trimarchi JM. And Lees JA. 2002. Sibling rivalry in the E2F family. *Nat. Rev. Mol. Cell Biol*. **3**, 11-20.

Tsukiyama T, Becker PB, Wu C. 1994. ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. *Nature.* **367**, 525-32.

Tuteja G, Kaestner KH. 2007. Forkhead transcription factors *Cell.* **131**, 192.

Tuteja G., Kaestner KH. 2007. SnapShot: forkhead transcription factors .*Cell*. **130**, 1160.

Uren AG, Vaux DL.1996. TRAF proteins and meprins share a conserved domain.*Trends Biochem Sci.* **21**, 244-5.

van de Wetering M, Clevers H. 1992. Sequence-specific interaction of the HMG box proteins TCF-1 and SRY occurs within the minor groove of a Watson-Crick double helix. *EMBO J.* **11**, 3039-44.

van der Horst A, Burgering BM. 2007. Stressing the role of FoxO proteins in lifespan and disease. *Nat Rev Mol Cell Biol.* **8**, 440-50.

Vandepoele K, Raes J, De Veylder L, Rouze P, Rombauts S, Inze D. 2002. Genome-wide analysis of core cell cycle genes in *Arabidopsis*. *Plant Cell*. **14**, 903-916.

Vandepoele K, Vlieghe K, Florquin K, Hennig L, Beemster GT, Gruissem W, Van de Peer Y, Inze D, De Veylder L. 2005. Genome-Wide Identification of Potential Plant E2F Target Genes. *Plant Physiol.* **139**, 316-328.

Vandiver VV, Fites RC. 1979. Thymidylate Synthase Activity from Chlamydomonas Cells and Cultured Tissues of Nicotiana, Pinus, and Daucus. *Plant Physiol*. **64**, 668-670.

Vauclare P, Diallo N, Bourguignon J, Macherel D, Douce R. 1996. Regulation of the Expression of the Glycine Decarboxylase Complex during Pea Leaf Development. *Plant Physiol.* **112**, 1523-1530.

Vlieghe K, Boudolf V, Beemster GT, Maes S, Magyar Z, Atanassova A, de Almeida Engler J, De Groodt R, Inze D, De Veylder L. 2005. The DP-E2F-like gene DEL1 controls the endocycle in *Arabidopsis thaliana*. *Curr. Biol*. **15**, 59-63.

Wang KL, Yoshida H, Lurin C, Ecker JR. 2004. Regulation of ethylene gas biosynthesis by the Arabidopsis ETO1 protein. *Nature.* **428**, 945-50.

Wang XY, Ohtaka-Maruyama C, Pisano MM, Jaworski CJ, Chepelinsky AB. 1995. Isolation and characterization of the 5'-flanking sequence of the human ocular lens MIP gene. *Gene*. **167**, 321-5.

Weber H, Bernhardt A, Dieterle M, Hano P, Mutlu A, Estelle M, Genschik P, Hellmann H. 2005. Arabidopsis AtCUL3a and AtCUL3b form complexes with members of the BTB/POZ-MATH protein family. *Plant Physiol*. **137**, 83-93.

Weber H, Hano P and Hellmann H. 2007. The charming complexity of CUL3. *Int J Plant Dev Biol.* **1**, 178–184.

Werner-Felmayer G, Golderer G, Werner ER. 2002. Tetrahydrobiopterin biosynthesis, utilization and pharmacological effects. *Curr Drug Metab*. **3**, 159-73.

Xu Y, Yu H, Hall TC. 1994. Rice Triosephosphate Isomerase Gene 5[prime] Sequence Directs [beta]-Glucuronidase Activity in Transgenic Tobacco but Requires an Intron for Expression in Rice. *Plant Physiol*.**106**, 459-467.

Yamaguchi M, Hayashi Y, Hirose F, Nishimoto Y, Matsukage A. 1997. Distinct roles of E2F recognition sites as positive or negative elements in regulation of the DNA polymerase alpha 180 kDa catalytic subunit gene promoter during Drosophila development. *Nucleic Acids Res.* **25**, 3847-54.

Yamaguchi K, Shirakabe K, Shibuya H, Irie K, Oishi I, Ueno N, Taniguchi T, Nishida E, Matsumoto K. 1995. Identification of a member of the MAPKKK family as a potential mediator of TGF-beta signal transduction. *Science*. **270**, 2008-11.

Zhang W, Zhang J, Kornuc M, Kwan K, Frank R, Nimer SD. 1995. Molecular cloning and characterization of NF-IL3A, a transcriptional activator of the human interleukin-3 promoter. *Mol Cell Biol.* **15**, 6055-63.

Zhang W, Zhang J, Kornuc M, Kwan K, Frank R, Nimer SD. 1995. Molecular cloning and characterization of NF-IL3A, a transcriptional activator of the human interleukin-3 promoter. *Mol Cell Biol*. **15**, 6055-63.

Zhu L, Zhu L, Xie E, Chang LS. 1995.Differential roles of two tandem E2F sites in repression of the human p107 promoter by retinoblastoma and p107 proteins. *Mol Cell Biol*. **15**, 3552-62.

Zhu, L., S. van den Heuvel, K. Helin, A. Fattaey, M. Ewen, D. Livingston, N. Dyson, and E. Harlow. 1993. Inhibition of cell proliferation by p107, a relative of the retinoblastoma protein. *Gene Dev*. **7:**1111–1125.

7. SUMMARY

Arabidopsis thaliana contains three *dhfr/ts* (dihydrofolate reductase/thymidylate synthase) genes. Analyses of promoter activity in transgenic plants suggest that the three *AtDRTS* genes have very different patterns of expression, and only *AtDRTS2* appears to be specifically expressed in dividing cells. Surprisingly the other two *AtDRTS* promoters are highly active in specific differentiated tissues suggesting important unknown functions of this enzyme in mature organs. E2F binding sites, which are often involved in cell cycle-dependent expression, are contained in the *AtDRTS2* and *AtDRTS3* promoters. ChIP analyses have revealed that these sites are recognized *in vivo* by E2F factors. To better define the role of the E2F sites located in the two *AtDRTS* promoters we analysed the effect of their mutation on the regulation of these *AtDRTS* genes. Functional analyses of transgenic plants revealed that the inactivation of the E2F *cis*-elements increased considerably the activity of both promoters, indicating that the E2F factors recognizing these sites can repress *AtDRTS* expression. The discovery that in both *AtDRTS* promoters the E2F site plays a repressive role implies that other *cis*-acting elements must be involved in the activation of expression in meristematic cells. In particular we focused our studies on the *AtDRTS2* gene which is specifically meristematic. *In silico* analyses of its promoter region revealed the presence of additional *cis*-acting elements, called Up1 and Up2. Mutational analyses revealed that these sites are not necessary for the expression in proliferating cells of Arabidopsis but act together, and not sinergistically, to drive high level of activity of this promoter. Because the Up2 site in the first intron of the *AtDRTS2* gene confers high levels of expression but is not essential for the meristematic expression of the gene, we analysed the effect of the removal of the entire intron. Surprisingly, the absence of the intron abolished completely the expression in root and shoot meristems suggesting that at least one additional regulatory site in the first intron is strictly necessary for the meristematic expression of the *AtDRTS2* gene. In this respect, an analysis by linker scanning mutations of 8 bps along the entire intron revealed that the expression of the *AtDRTS2* promoter requires a small intronic region which shows no homology to any plant *cis*-element already described.