Microarray analysis of *E2Fa-DPa*-overexpressing plants uncovers a cross-talking genetic network between DNA replication and nitrogen assimilation

Kobe Vlieghe[1], Marnik Vuylsteke[1], Kobe Florquin[1], Stephane Rombauts[1], Sara Maes[1], Sandra Ormenese[2], Paul Van Hummelen[3], Yves Van de Peer[1], Dirk Inzé[1,*] and Lieven De Veylder[1]

[1]Department of Plant Systems Biology, Flanders Interuniversity Institute for Biotechnology (VIB), Ghent University, B-9052 Gent, Belgium
[2]Laboratory of Plant Physiology, University of Liege, B-4000 Liège, Belgium
[3]VIB MicroArray Facility, Gasthuisberg, Onderzoek en Navorsing, B-3000 Leuven, Belgium
*Author for correspondence (e-mail: dirk.inze@psb.ugent.be)

Accepted 19 June 2003
Journal of Cell Science 116, 4249-4259 © 2003 The Company of Biologists Ltd
doi:10.1242/jcs.00715

Summary

Previously we have shown that overexpression of the heterodimeric *E2Fa-DPa* transcription factor in *Arabidopsis thaliana* results in ectopic cell division, increased endoreduplication, and an early arrest in development. To gain a better insight into the phenotypic behavior of *E2Fa-DPa* transgenic plants and to identify E2Fa-DPa target genes, a transcriptomic microarray analysis was performed. Out of 4,390 unique genes, a total of 188 had a twofold or more up- (84) or down-regulated (104) expression level in *E2Fa-DPa* transgenic plants compared to wild-type lines. Detailed promoter analysis allowed the identification of novel E2Fa-DPa target genes, mainly involved in DNA replication. Secondarily induced genes encoded proteins involved in cell wall biosynthesis, transcription and signal transduction or had an unknown function. A large number of metabolic genes were modified as well, among which, surprisingly, many genes were involved in nitrate assimilation. Our data suggest that the growth arrest observed upon *E2Fa-DPa* overexpression results at least partly from a nitrogen drain to the nucleotide synthesis pathway, causing decreased synthesis of other nitrogen compounds, such as amino acids and storage proteins.

Key words: *Arabidopsis thaliana*, Cell cycle, E2F, Microarray, Nitrogen assimilation

Introduction

Progression through the cell cycle is essential for the continued existence of all uni- and multicellular organisms. It is crucial for the survival of a cell that its DNA is correctly replicated. In mammals, the onset of DNA replication is regulated by the activity of the heterodimeric E2F-DP transcription factor. The mammalian E2F family contains six proteins (E2F1, E2F2, E2F3, E2F4, E2F5 and E2F6) (Trimarchi and Lees, 2002). All E2Fs have an N-terminally located DNA-binding domain immediately followed by a dimerization domain, allowing them to pair with a dimerization partner (DP1 or DP2). Dimerization of E2F with DP is a prerequisite for high affinity, sequence-specific binding to the E2F consensus DNA-binding site. E2F activity is negatively regulated by retinoblastoma (Rb), which binds to the transcriptional activation domain of the E2F-DP factor, rendering it inactive. Moreover, the recruitment by Rb of DNA-modifying enzymes, such as histone deacetylases and polycomb proteins, leads to chromatin condensation with suppression of promoter activity of E2F-DP target genes as a result. Phosphorylation of Rb by cyclin-dependent kinases (CDKs) counteracts its inhibitory function, resulting in the release of transcriptionally active E2F-DP and consequential onset of DNA replication.

The mechanism of DNA replication seems to be conserved between mammals and plants, because *E2F* and *DP* genes have been isolated from different plant species, including wheat, tobacco, carrot, *Arabidopsis* and rice (Ramírez-Parra et al., 1999; Sekine et al., 1999; Albani et al., 2000; Magyar et al., 2000; Ramirez-Parra and Gutierrez, 2000; Kosugi and Ohashi, 2002a). In the *Arabidopsis* genome there are three *E2F* (*E2Fa*, *E2Fb* and *E2Fc*) and two *DP* (*DPa* and *DPb*) genes (Vandepoele et al., 2002). Recently, we have analyzed the phenotypes of plants co-overexpressing the *E2Fa-DPa* genes (De Veylder et al., 2002). Transgenic plants were smaller than control plants, had curled leaves and cotyledons, and were arrested in growth at an early stage of development. Microscopic analysis revealed that E2Fa-DPa-overproducing cells underwent ectopic cell division or endoreduplication, depending on the cell type. Whereas extra cell divisions resulted in cells smaller than those seen in the same tissues of control plants, supplementary endoreduplication caused the formation of giant nuclei. By using reverse transcription (RT)-PCR, we demonstrated that the expression levels of genes involved in DNA replication (*CDC6*, *ORC1*, *MCM* and *DNA pol α*) were strongly up-regulated (De Veylder et al., 2002).

Physiologically important targets of the mammalian E2F-DP transcription factors have been identified by microarray hybridization experiments, chromatin immunoprecipitations and computer-assisted prediction (Ishida et al., 2001; Kel et al., 2001; Müller et al., 2001; Weinmann et al., 2001; Ren et al., 2002). *E2F-DP*-responsive genes can be found among genes involved in cell division, DNA repair and replication, mitotic

progression, apoptosis and differentiation. Although little is known about the plant *E2F-DP* target genes, a database search has been published recently, in which the *Arabidopsis* genome was screened for genes harboring the TTTCCCGCC *cis*-acting element in their promoter (Ramirez-Parra et al., 2003). However, it is still unclear whether this specific *cis*-acting element is the only one recognized by the plant E2F-DP complexes, or whether the presence of the TTTCCCGCC element is sufficient to mark a gene as a true E2F-DP target gene. In order to identify the functional classes of genes regulated by E2Fa-DPa and to understand the nature of the phenotype of the *E2Fa-DPa*-overexpressing plants, we designed a microarray experiment that compared the transcript levels of 4,571 genes of wild-type and transgenic lines. We found distinct classes of genes that were up- or down-regulated in the *E2Fa-DPa* plants. Promoter analysis allowed us to distinguish among the downstream expressed genes, the genes that were putatively under direct control of E2Fa-DPa. Furthermore, we found that the increased expression levels of *E2Fa-DPa* have a large impact on the expression levels of genes involved in nitrogen assimilation and metabolism.

Materials and Methods

Plant material

Double transgenic *CaMV35S-E2Fa-DPa* plants were obtained by crossing homozygous *CaMV35S-E2Fa* and *CaMV35S-DPa* plants (De Veylder et al., 2002). Double transformants were grown under a 16-hour light/8-hour dark photoperiod at 22°C on germination medium (Valvekens et al., 1988).

Construction of microarrays

The *Arabidopsis thaliana* (L.) Heynh. microarray consisted of 4,608 cDNA fragments spotted in duplicate, distant from each other, on Type V silane-coated slides (Amersham Biosciences, Little Chalfont, UK). The clone set included 4,571 *Arabidopsis* cDNAs from the unigene clone collection Arabidopsis Gem I (Incyte Genomics, Palo Alto, CA). The functional annotation of the genes related to the spotted cDNAs was retrieved by BLASTN against genomic sequences. To facilitate the analysis, a collection of genomic sequences was built each bearing only one gene. In each of these sequences, the upstream intergenic sequence was followed by the exon-intron structure of the gene and the downstream intergenic sequence, or, in other words, the whole genomic sequence between start and stop codons from neighboring protein-encoding genes. From the BLASTN output, the best hits were extracted and submitted to a BLASTX search against protein databases. From this analysis, the set of 4,571 cDNAs appeared to constitute 4,390 unique clones. To obtain more detailed information concerning the potential function of the genes, protein domains were searched using ProDom. The complete set can be found at http://www.psb.ugent.be/E2F/. The cDNA inserts were amplified by PCR with M13 primers, purified with MultiScreen-PCR plate (Millipore, Bedford, MA), and arrayed on slides using a Generation III printer (Amersham Biosciences). Slides were blocked in 3.5% SSC (1× SSC, 150 mM NaCl, 15 mM sodium citrate, pH 7.0), 0.2% sodium dodecyl sulfate (SDS), 1% bovine serum albumin for 10 minutes at 60°C.

RNA amplification and labeling

Antisense RNA was amplified with a modified protocol of in vitro transcription (Puskás et al., 2002). For the first-strand cDNA synthesis, 5 μg of total RNA was mixed with 2 μg of a HPLC-purified anchored oligo(dT) + T7 promoter (5′-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG-T_{24}(ACG)-3′) (Eurogentec, Seraing, Belgium), 40 units of RNAseOUT (Invitrogen, Gaithersburg, MD) and 0.9 M D(+)trehalose (Sigma-Aldrich, St. Louis, MO) in a total volume of 11 μl and heated to 75°C for 5 minutes. To this mixture, 4 μl 5× first-strand buffer (Invitrogen), 2 μl 0.1 M dithiothreitol, 1 μl 10 mM dNTP mix, 1 μl 1.7 M D(+)trehalose (Sigma-Aldrich), and 1 μl SuperScript II (Invitrogen) were added to 20 μl final volume. The sample was incubated in a UnoII thermocycler (Whatman Biometra, Göttingen, Germany) at 37°C for 5 minutes, at 45°C for 10 minutes, 10 cycles at 60°C for 2 minutes and at 55°C for 2 minutes. To the first-strand reaction mix, 103.8 μl water, 33.4 μl 5× second-strand synthesis buffer (Invitrogen), 3.4 μl 10 mM dNTP mix, 1 μl of 10 U/μl DNA ligase (Invitrogen), 4 μl 10 U/μl DNA Polymerase I (Invitrogen), and 1 μl 2 U/μl RNAse H (Invitrogen) were added and incubated at 16°C for 2 hours. The synthesized double-stranded cDNA was purified with Qiaquick (Qiagen, Hilden, Germany). Antisense RNA was synthesized by AmpliScribe T7 high-yield transcription kit (Epicentre Technologies, Madison, WI) in a total volume of 20 μl according to the manufacturer's instructions. The RNA was purified with the RNeasy purification kit (Qiagen). From this RNA, 5 μg was labeled by reverse transcription using random nonamer primers (Genset, Paris, France), 0.1 mM d(G/T/A)TPs, 0.05 mM dCTP (Amersham Biosciences), 0.05 mM Cy3-dCTP or Cy5-dCTP (Amersham Biosciences), 1× first-strand buffer, 10 mM dithiothreitol, and 200 U of SuperScript II (Invitrogen) in 20 μl total volume. The RNA and primers were denatured at 75°C for 5 minutes and cooled on ice before the remaining reaction components were added. After 2 hours incubation at 42°C, mRNA was hydrolyzed in 250 mM NaOH for 15 minutes at 37°C. The sample was neutralized with 10 μl of 2 M 3-(*N*-morpholino)propane-sulfonic acid and purified with Qiaquick (Qiagen).

Array hybridization and post-hybridization processes

The probes were resuspended in 30 μl hybridization solution (50% formamide, 5× SSC, 0.1% SDS, 100 μg/ml salmon sperm DNA) and prehybridized with 1 μl poly(dT) (1 mg/ml) at 42°C for 30 minutes to block hybridization on the polyA/T tails of the cDNA on the arrays. Mouse COT DNA (1 mg/ml) (Invitrogen) was added to the mixture and placed on the array under a glass coverslip. Slides were incubated for 18 hours at 42°C in a humid hybridization cabinet (Amersham Biosciences). Post-hybridization washing was performed for 10 minutes at 56°C in 1× SSC, 0.1% SDS, twice for 10 minutes at 56°C in 0.1× SSC, 0.1% SDS, and for 2 minutes at 37°C in 0.1× SSC.

Scanning and data analysis

Arrays were scanned at 532 nm and 635 nm using a Generation III scanner (Amersham Biosciences). Image analysis was performed with ArrayVision (Imaging Research Inc, St. Catharines, Ontario, Canada). Spot intensities were measured as artifact-removed total intensities (ARVol) without correction for background. We first addressed within-slide normalization by plotting for each single slide a 'MA-plot' (Yang et al., 2002), where $M=\log_2 (R/G)$ and $A=\log_2\sqrt{R\times G}$. Dye intensity differences were corrected with the 'LOWESS' normalization. Subsequently, between-slide normalization and differentially expressed gene identification between the two genotypes were performed by sequential analysis of variances (ANOVAs), according to the method of Wolfinger et al. (Wolfinger et al., 2001). (i) The base-2 logarithm of the 'LOWESS'-transformed measurements for all 73,136 spots (y_{iklm}) were subjected to a normalization model $y_{iklm}=\mu+A_k+A_kD_lR_m+\varepsilon_{iklm}$, where μ is the sample mean, A_k the effect of the kth array (k=1-4), $A_kD_lR_m$ the channel effect (AD) for the mth replication (m=2; left and right) of the total collection of i (i=1, …, 4571) cDNA fragments, and ε_{iklm} the stochastic error. (ii) The residuals from this model were subjected to 4,571 gene-specific models $r_{ijkl}=\mu+G_iA_k+G_iD_l+G_iC_j+\gamma_{ijkl}$, where G_iA_k is the spot effect, G_iD_l the gene-specific dye effect, G_iC_j the signal intensity for genes that can

specifically be attributed to the genotypes (effect of interest), and γ_{ijkl} the stochastic error. All effects were assumed to be fixed, except for ε_{iklm} and γ_{ijkl}. T-tests for differences between the G_iC_j effects were performed, all based on n_1+n_2-6 degrees of freedom, where n_1 and n_2 correspond to the number of wild-type and E2Fa-DPa hybridizations, respectively. Bonferroni adjustment for the 4,571 tests to assure an experiment-wise false positive rate of 0.05 results in a *P*-value cut-off of $1e^{-5.0}$, which is certainly too conservative. Thus, no further adjustments for multiple testing were done. Therefore, we chose to set the *P*-value cut-off arbitrarily at the 0.05 level. We used Genstat for both the normalization and gene model fits.

RT-mediated PCR analysis

RNA was isolated from plants 8 days after sowing with the Trizol reagent (Amersham Biosciences). First-strand cDNA was prepared from 3 µg of total RNA with the Superscript RT II kit (Invitrogen) and oligo$(dT)_{18}$ according to the manufacturer's instructions. A 0.25 µl aliquot of the total RT reaction volume (20 µl) was used as a template in a semi-quantitative RT-mediated PCR amplification, ensuring that the amount of amplified product remained in linear proportion to the initial template present in the reaction. From the PCR reaction, 10 µl was separated on a 0.8% agarose gel and transferred onto Hybond N^+ membranes (Amersham Biosciences) that were hybridized at 65°C with fluorescein-labeled probes (Gene Images random prime module; Amersham Biosciences). The hybridized bands were detected with the CDP Star detection module (Amersham Biosciences). Primers used were 5′-AAAAAGCAGGCTGTGTCGTACGATCTTCTCCCGG-3′ and 5′-AGAAAGCTGGGTCATGTGATAGGAGAACCAGCG-3′ for *E2Fa*, 5′-ATAGAATTCGCTTACATTTTGAAACTGATG-3′ and 5′-ATAGTCGACTCAGCGAGTATCAATGGATCC-3′ for *DPa*, 5′-CAGATCTTGTTAACCTTGACATCTCAG-3′ and 5′-GGGTCAAAAGATACAACCACACCAG-3′ for glutamine synthetase *(GS)*, 5′-GGTTTACGAGCTACATGGCCC-3′ and 5′-GAGCAATCCGTTCAGCCTCC-3′ for glutamate synthase (*GOGAT*), 5′-GCGTTTGACCACTCTTGGAGAC-3′ and 5′-GAACGCCATTGAGAAAGTCCGC-3′ for histone acetylase *HAT B*, 5′-GTTACCGGCTCGACTTGAAGATC-3′ and 5′-GAATCGGAGGGAAAGTCTGACG-3′ for LOB domain protein 41, 5′-GTGTGGTTTCCAAGCTTTCCTACG-3′ and 5′-GGTGAAGGGACTAGCCTTGTGG-3′ for isocitrate lyase, 5′-GGGATCAATCCTCAGGAGAAGG-3′ and 5′-CCGTCCATCTTTATTAGCGGCATG-3′ for nitrite reductase (*NiR*), and 5′-TTACCGAGGCTCCTCTTAACCC-3′ and 5′-ACCACCGATCCAGACACTGTAC-3′ for actin 2 (*ACT2*).

Promoter analysis

The intergenic sequence corresponding to the promoter area of each gene spotted on the microarray was deduced from genomic sequences. From these intergenic sequences, up to 500 bp upstream of the ATG start codon were extracted and subjected to motif searches to retrieve potential E2F elements. Of the 4,571 expressed sequence tags (ESTs) spotted on the microarray, we could retrieve the genomic sequence of 4,390. This difference is due to the presence of duplicate genes and mitochondrial or chloroplast DNA on the microarray. Both the position and frequency of occurrence were determined with the publicly available MatInspector (version 2.2) by using matrices extracted from PlantCARE and matrices made especially for this particular analysis (Lescot et al., 2002). The relevance of each motif was evaluated against a background consisting of all the sequences from the dataset by using the Fisher exact test.

Results and Discussion

Experimental setup and statistical analysis

A microarray containing in replicate 4,571 *Arabidopsis* expressed sequence tags (ESTs) was used to compare the transcriptome of the wild-type with that of *E2Fa-DPa*-overexpressing plants. cDNA was synthesized from total RNA of plants harvested 8 days after sowing. At that stage, transgenic plants can be distinguished from control plants by the appearance of curled cotyledons that display ectopic cell divisions and enhanced endoreduplication (De Veylder et al., 2002). In the first two hybridizations, including a biological repeat, fluorescently Cy3- and Cy5-labeled probe pairs of control and *E2Fa-DPa* cDNAs were used. Subsequently, a dye-swap replication was performed for both hybridizations, resulting in a total of four cDNA microarray hybridizations. Because each cDNA was printed in duplicate on the array, eight data points for every gene were obtained.

Fluorescence levels were analyzed to establish whether the expression level of each gene varied according to the overexpression of the *E2Fa-DPa* transcription factor. Two sequential ANOVA models were used, as proposed by Wolfinger et al. (Wolfinger et al., 2001). First, the model called 'normalization model' accounts for experiment-wise systematic effects, such as array and channel effects, which could bias inferences made on the data from the individual genes. The residuals from this model represent normalized values and are the input data for the second model, called the 'gene' model. The gene models are fitted separately to the normalized data from each gene (see Materials and Methods). In this procedure, normalized expression levels rather than ratios are used as units.

For each of the 4,571 genes on the arrays the genotype-specific signal intensity was determined and *t*-tested for significant differences ($P<0.05$). Fig. 1 presents the *P* values obtained (as the negative log10 of the *P* value) against the

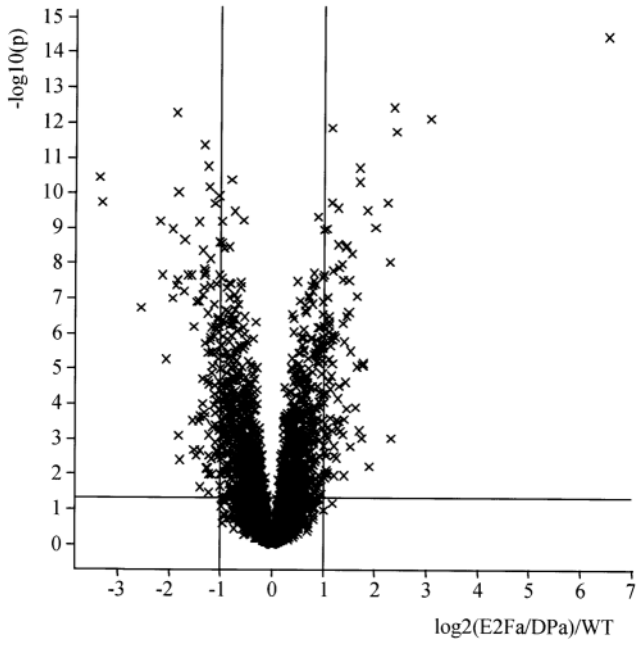

Fig. 1. Volcano plot of significance against effect. Each x represents one of the 4,571 genes, with the negative log10 of the *P* value from the gene model plotted against the difference between least-square means for the genotype effect. The horizontal line represents the test-wise threshold of $P=0.05$. The two vertical reference lines indicate a twofold cut-off for either repression or induction.

Table 1. *Arabidopsis* genes up-regulated twofold or more in *E2Fa-DPa* plants sorted according to functional category

Gene identification	Accession number	ORF name	Fold induction	E2F motif	Position*	Strand
DNA replication and modification (14)						
Putative thymidine kinase	AI997851	At3g07800	8.44			
DNA methyltransferase	AI994691	At5g49160	5.37			
Msi3	AW004204	At4g35050	4.89	TTTCCCGC	–75	–
Putative linker histone protein	AI994590	At3g18035	3.31			
Putative replication factor c	AI997934	At1g21690	3.30	TTTCCCGC	–96	–
Topoisomerase 6 subunit A	AI995290	At5g02820	2.62	TTTCCCGC	–66	+
				TTTGGCGG	–369	+
Histone H4-like protein	AI999171	At3g46320	2.55	TTTGGCGC	–310	+
Histone acetylase HAT B	AI998229	At5g56740	2.36	TTTCCCGC	–50	+
Putative histone H1	AI996137	At1g06760	2.27			
Histone H2A-like protein	AI995882	At4g27230	2.23			
Putative DNA gyrase subunit A	AI995400	At3g10690	2.20	ATTCGCGC	–91	+
				TTTGGCGG	–107	–
Histone H2B-like protein	AI999101	At5g59910	2.16	ATTCCCGC	–329	–
				ATTGGCGC	–303	–
Putative mismatch binding protein	AI993280	At3g24320	2.10			
Adenosyl homocysteinase	AI996953	At4g13940	2.07			
Cell cycle (2)						
E2Fa	AJ294534	At2g36010	94.88			
CDKB1;1	D10851	At3g54180	2.60	TTTCCCGC	–151	–
Cell wall biogenesis (11)						
Xyloglucan endo-1,4-β-D-glucanase (meri-5)	AI994459	At4g30270	3.74			
Putative glycosyl transferase	AI999244	At1g70090	3.38			
α-Galactosyltransferase-like protein	AI998223	At3g62720	3.26			
Putative xyloglucan endotransglycosylase	AI999683	At3g23730	2.85			
Xyloglucan endo-1,4-β-D-glucanase-like protein	AI998301	At4g30280	2.74			
Putative xyloglucan endotransglycosylase	AI994477	At1g14720	2.51			
Putative glycosyl transferase	AI999770	At1g24170	2.39			
Putative UDP-glucose glucosyltransferase	AI997288	At1g22400	2.34	TTTCGCGC	–20	+
Putative glucosyltransferase	AI998872	At2g15480	2.15			
Peroxidase	AI994622	At2g38380	2.11	TTTCGCGC	–314	–
β-1,3-glucanase-like protein	AI994681	At3g55430	2.05			
Chloroplastic genes (7)						
Large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase	N96785	rbcL	4.71			
Ribosomal protein L33	AI994194	rpl33	3.54	TTTCCCCC	–315	–
Photosystem II protein	AW004203	PsbI	2.81			
Ribosomal protein L2	AW004266	rpl2	2.61			
ATP-dependent protease subunit	AI997947	clpP	2.60			
Cytochrome B6	AI997102	PetB	2.55	TTTCGCGG	–36	–
ATPase ε subunit	AW004251	atpE	2.17	TTTCCCGG	–160	–
Mitochondrial genes (1)						
26S ribosomal RNA protein	AW004275	orf107a	2.87			
Transcription factors (6)						
LOB domain protein 41	AI996685	At3g02550	4.01			
WRKY transcription factor 21	AI992739	At2g30590	2.78	TTTCCCCC	–23	–
GATA Zn-finger protein	AI995731	At3g16870	2.75			
Anthocyaninless2	AI993655	At4g00730	2.73			
Leucine zipper-containing protein	AI995691	At1g07000	2.43	TTTCCCCG	–33	+
Homeodomain transcription factor (Athb-6)	AI999190	At2g22430	2.30			
Metabolism and biogenesis (11)						
Alcohol dehydrogenase	AI998773	At1g77120	5.09			
Putative isocitrate lyase	AI999168	At3g21720	3.08			
Protochlorophyllide reductase precursor	AI993342	At4g27440	2.39			
Sugar transporter-like protein	AI997793	At4g36670	2.27			
NADH-dependent glutamate synthase (GOGAT)	AI997600	At5g53460	2.25	TTTCGCCG	–225	+
Nitrate reductase (NIA2)	AI996208	At1g37130	2.15			
Pectate lyase-like protein	AJ508995	At3g54920	2.13			
Putative sterol dehydrogenase	AI996340	At2g43420	2.10			
Glutamine synthetase root isozyme 1 (GS)	161G19T7	At1g66200	2.06			
Monosaccharide transporter STP3	AI997045	At5g61520	2.05			
Signal transduction (6)						
Calcium-dependent protein kinase	AI996555	At5g66210	2.96			
WD-40 repeat protein	AI993055	At5g14530	2.70	TTTCGCGG	–104	–
Receptor-protein kinase-like protein	AI994727	At5g54380	2.59			
Putative phytochrome A	AI998146	At1g09570	2.45			
Putative leucine-rich receptor-like protein kinase	AI999651	At1g72180	2.13			
Putative receptor-like kinase	AI993298	At3g23750	2.06			

Table 1. ***Continued***

Gene identification	Accession number	ORF name	Fold induction	E2F motif	Position*	Strand
Others (13)						
Putative pollen allergen	AI996548	At3g45970	3.22			
Cold-regulated protein COR6,6	AW004198	At5g15970	3.03			
Phi-1-like protein	AI994601	At5g64260	2.60			
Lipid-transfer protein-like	AI998609	At5g01870	2.33			
DnaJ homologue	AI994551	At5g06910	2.32	ATTGGCGC	–103	+
Blue copper-binding protein	AI996535	At5g20230	2.30			
Src-2-like protein	AI998679	At1g09070	2.19			
Ring finger protein	AI999491	At3g61460	2.14			
Putative Ticc22	AI993361	At3g23710	2.14			
Nodulin-like protein	AI996322	At1g80530	2.07			
Putative resistance protein	AI997549	At1g61100	2.06			
Seed imbitition protein-like	AI993446	At5g20250	2.05			
Putative disease resistance protein	AI998978	At1g72900	2.04	TTTGGCGG	–175	+
Unknown function (14)						
Putative protein	AI994686	At3g45730	5.14			
Putative protein	AI994734	At5g66580	3.18			
Unknown protein	AI999397	At2g38310	2.79	TTTGCCCC	–280	–
Hypothetical protein	AI998042	At1g57680	2.66			
Unknown protein	AI995465	At2g47440	2.50			
Unknown protein	AI994871	At1g76970	2.34			
Hypothetical protein	AI998366	At1g27500	2.21			
Putative protein	AI996967	At4g33050	2.20			
Putative protein	AI995917	At3g43690	2.18			
Unknown protein	AI993084	At2g25970	2.15			
Unknown protein	AI993077	At1g68580	2.13			
Putative protein	AI993019	At5g14420	2.05	TTTCGCCG	–443	–
Hypothetical protein	AI997428	At1g57990	2.02			
Unknown protein	AI997827	At5g53740	2.01	TTTGGCGG	–66	+

*Relative position upstream from the translation initiation site.

magnitude of the effect (log2 of estimated fold change). This so-called volcano plot illustrates the substantial difference of significance testing as opposed to cut-offs strictly based on the fold change. The two vertical reference lines indicate a twofold cut-off for either repression or induction, whereas the horizontal reference line refers to the *P*-value cut-off at 0.05. These reference lines divide the plot into six meaningful sectors. The 3,126 genes in the lower middle sector have low significance and low fold change, and both methods are in agreement that the corresponding changes are not significant. The 188 genes in the upper left and right sectors have high significance ($P<0.05$) and high fold change (≥2); 84 of these genes show a significant two-or-more-fold induction of expression, whereas the remaining 104 genes show a significant two-or-more-fold repression of expression in the *E2Fa-DPa* plants. The identity of these genes was confirmed by sequencing, and the induction of a random set of selected genes was confirmed by RT-PCR analysis (Fig. 2). Finally, the 1,257 genes in the upper middle sector represent significant ($P<0.05$) up- or down-regulated genes, but with a low (≤2) fold change. The full dataset of genes can be viewed at http://www.psb.ugent.be/E2F/.

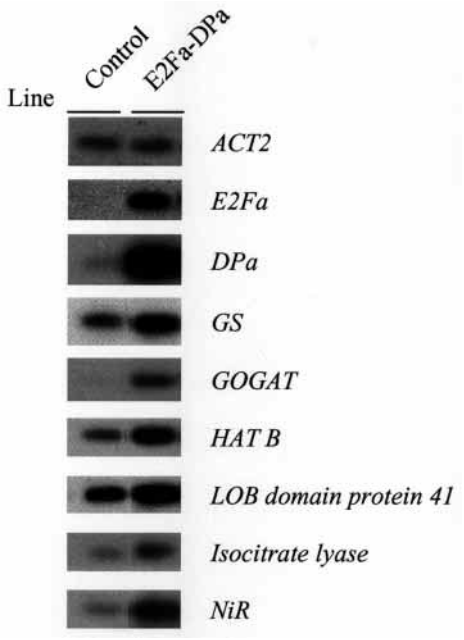

Fig. 2. Verification of microarray analysis by RT-PCR. RT-PCR analysis was carried out under linear amplification conditions. The actin 2 gene (*ACT2*) was used as loading control. GS, glutamine synthetase; GOGAT, glutamate synthase; NiR, nitrite reductase.

DNA replication and cell cycle genes

Genes up- or down-regulated in the *E2Fa-DPa* transgenic plants can be classified into clear groups according to their function (Tables 1 and 2). Among the genes that are twofold or more up-regulated, 14 belong to the class of DNA replication and modification, correlating with the observation that *E2Fa-DPa*-overexpressing plants undergo extensive endoreduplication. Most of these genes have previously been

Table 2. *Arabidopsis* genes repressed twofold or more in *E2Fa-DPa* plants sorted according to functional category

Gene identification	Accession number	ORF name	Fold repression	E2F motif	Position*	Strand
Cell wall biogenesis (4)						
Similar to polygalacturonase-like protein	AI993509	At1g10640	3.62			
Putative xyloglucan endo-transglycosylase	AI997647	At2g36870	2.51			
Pectate lyase 1-like protein	AI994801	At1g67750	2.40			
Xyloglucan endo-transglycosylase	AI998832	At3g44990	2.35			
Metabolism and biogenesis (24)						
Fructose-bisphosphate aldolase-like protein	AI994456	At4g26530	5.99	ATTGGCCC	–426	–
Sucrose-phosphate synthase-like protein	AI995432	At4g10120	4.64			
Putative branched-chain amino acid aminotransferase	AI997263	At3g19710	3.31			
Vitamin C-2	AI997404	At4g26850	3.04	TTTGCCGC	–222	+
Nicotianamine synthase	AI993200	At5g04950	2.86			
β-fructosidase	AI994670	At1g62660	2.66	TTTCCCCC	–344	–
Neoxanthin cleavage enzyme-like protein	AI997269	At4g19170	2.66			
Putative starch synthase	AI997174	At1g32900	2.63			
Cytochrome P450 monooxygenase (CYP83A1)	AI994017	At4g13770	2.57			
β-amylase-like protein	AI999322	At5g18670	2.53			
FRO1-like protein; NADPH oxidase-like	AI995987	At5g49740	2.46			
Putative hydrolase	AI997149	At3g48420	2.39			
Furamate hydratase	AI997067	At5g50950	2.31			
5′-adenylylsulfate reductase	AI992757	At1g62180	2.30			
5′-adenylylsulfate reductase	AI996614	At4g04610	2.30			
UDP rhamnose-anthocyanidin-3-glucoside rhamnosyltransferase-like protein	AI996803	At4g27560	2.24			
Cytochrome P450-like protein	AI993171	At5g48000	2.23			
Lactoylglutathione lyase-like protein	AI994552	At1g11840	2.20			
Putative β-glucosidase	AI995306	At4g27820	2.20	ATTGGCCC	–327	–
Adenine phosphoribosyltransferase-like protein	AI994567	At4g22570	2.18			
Catalase	AI995830	At4g35090	2.17	ATTCCCCC	–168	+
Putative glutathione peroxidase	AW004143	At2g25080	2.15			
Putative adenosine phosphosulfate kinase	AW004219	At2g14750	2.13			
Tyrosine transaminase-like protein	AI996914	At4g23600	2.13			
Transcription factors (5)						
Homeobox-leucine zipper protein ATHB-12	AI994027	At3g61890	4.20	ATTGGCCG	–113	–
NAC domain protein NAC2	AI992865	At1g69490	3.68			
Myb-related transcription factor	AI995298	At1g71030	2.78			
Dof zinc finger protein	AI994875	At1g51700	2.30	TTTCCCGG	–18	+
				TTTCCCCG	–13	–
MYB-related transcription factor (CCA1)	AI992931	At2g46830	2.19			
Signal transduction (9)						
Serine/threonine protein kinase-like protein	AI995557	At5g10930	3.91			
Subtilisin proteinase-like	AI993428	At4g21650	3.19	TTTCGCGG	–85	+
Putative oligopeptide transporter	AI996160	At4g10770	2.68			
Putative lectin	AI998542	At3g16400	2.52			
Ca^{2+}-dependent membrane-binding protein annexin	AI998553	At1g35720	2.45			
Putative WD repeat protein	AI997238	At3g15880	2.38			
Putative lectin	AI999016	At3g16390	2.35			
Putative lectin	AI993358	At3g16530	2.31			
SNF1-related protein kinase (ATSRPK1)	AI993111	At3g23000	2.06			
Others (25)						
Putative protease inhibitor Dr4	AI995265	At1g73330	10.30			
Major latex protein homolog-like	AI998305	At2g01520	4.27			
Pollen allergen-like protein	AI993041	At1g24020	3.56	TTTGGCCG	–377	+
Putative heat shock protein	AI997846	At1g06460	3.55			
Putative fibrillin	AI997199	At4g04020	3.55	TTTGGCCG	–435	–
Major latex protein homolog-like	AI997255	At1g70890	3.50			
Putative nematode resistance protein	AI993740	At2g40000	2.95			
Putative auxin-regulated protein	AJ508998	At2g46690	2.86			
Putative myrosinase-binding protein	AI997583	At2g39310	2.61			
Ubiquitin-conjugating enzyme-like protein	AI997782	At5g56150	2.41			
Ubiquitin-conjugating enzyme E2-17 kDa 8	AI994771	At5g41700	2.40			
Vegetative storage protein Vsp2	AI999152	At5g24770	2.35			
Heat shock protein 70	AI994044	At3g12580	2.24			
Chloroplast outer envelope membrane protein	AI997015	At3g63160	2.20			
Translation initiation factor-like protein	AI992786	At5g54940	2.15			
Pseudogene	AI995323	At2g04110	2.07			
Vegetative storage protein Vsp1	AI999546	At5g24780	2.06			
Dehydrin ERD10	AI997518	At1g20450	2.06			
MTN3-like protein	AI997159	At3g48740	2.05			
Putative chlorophyll a/b-binding protein	AI994859	At3g27690	2.05			
Photosystem I reaction center subunit psaN	AI997939	At5g64040	2.03			

Table 2. ***Continued***

Gene identification	Accession number	ORF name	Fold repression	E2F motif	Position*	Strand
Others – *continued*						
AR781 similar to yeast pheromone receptor	AI998194	At2g26530	2.03			
Putative lipid transfer protein	AI997024	At2g15050	2.03			
Peroxidase ATP3a	AI998372	At5g64100	2.03	TTTGGCCG	–492	+
Myosin heavy chain-like protein	AI999224	At3g16000	2.01			
Unknown function (35)						
Unknown protein	AI993767	At1g45200	3.91			
Putative protein	AI993468	At3g56290	3.38			
Hypothetical protein	AI996374	At1g61890	2.78			
Unknown protein	AI994573	At3g15950	2.71			
Putative protein	AI994726	At3g52360	2.65			
Hypothetical protein	AI997393	At4g02920	2.60	TTTGCCCC	–419	–
Unknown protein	AJ508997	At5g43580	2.58			
Unknown protein	AI997866	At1g70760	2.52			
Unknown protein	AI997085	At5g43750	2.51			
Putative protein	AI995724	At5g50100	2.48			
Unknown protein	AI995337	At1g74880	2.42			
Unknown protein	AI998296	At3g19370	2.40			
Unknown protein	AI993346	At3g10420	2.40			
Putative protein	AI999485	At3g61080	2.38			
Unknown protein	AI996923	At1g67860	2.38			
Unknown protein	AI994841	At1g52870	2.35	ATTCCCCC	–74	+
Unknown protein	AI999581	At1g64370	2.35			
Unknown protein	AI997584	At1g05870	2.25			
Putative protein	AI992938	At5g03540	2.21			
Hypothetical protein	AI997712	At2g15020	2.21			
Unknown protein	AI998338	At1g68440	2.20			
Unknown protein	AI996872	At2g21960	2.19			
Putative protein	AI996295	At4g27280	2.18			
Putative protein	AI995642	At3g48200	2.16			
Unknown protein	AI997470	At2g32870	2.14			
Hypothetical protein	AI998460	At1g69510	2.11	ATTCGCGG	–120	+
				TTTGGCCC	–492	+
Putative protein	AI993356	At5g22460	2.10			
Putative protein	AI995956	At5g52060	2.08			
Unknown protein	AI996100	At2g35830	2.06			
Hypothetical protein	AI996039	At3g27050	2.05	ATTGCCCC	–5	–
Unknown protein	AI996020	At5g51720	2.04			
Putative protein	AW004101	At4g39730	2.03			
Hypothetical protein	AI998372	At2g01260	2.03			
Unknown protein	AI999573	At3g61060	2.00			
Unknown protein	AI998562	At2g35760	2.00			
No hit (2)						
No hit on genome	AI995690		2.54			
No hit on genome	AI999010		2.23			

*Relative position upstream from the translation initiation site.

shown to be up-regulated by *E2F-DP* overexpression in mammalian cells, including a putative thymidine kinase, replication factor c, adenosylhomocysteinase, DNA (cytosine-5)-methyltransferase, and histone genes (Ishida et al., 2001; Müller et al., 2001; Ren et al., 2002). Other E2Fa-DPa-induced S phase genes include a linker histone protein, the topoisomerase 6 subunit A, and two subunits of the histone acetyltransferase HAT B complex, namely HAT B and Msi3. The HAT B complex is responsible for the specific diacetylation of newly synthesized histone H4 during nucleosome assembly on newly synthesized DNA (Lusser et al., 1999).

In addition to the overexpressed *E2Fa* gene (90-fold more abundant in transgenic than in control plants), only one cell cycle gene (*CDKB1;1*) has a twofold or more change in expression level upon *E2Fa-DPa* overexpression. *CDKB1;1* had already been predicted to be a candidate E2F-DP target by the presence of a consensus E2F-DP-binding site in its promoter (de Jager et al., 2001). Whereas CDKB1;1 activity is highest at the G2 to M transition, its transcript levels start to increase during S phase (Porceddu et al., 1999; Menges and Murray, 2002). Therefore, up-regulation of *CDKB1;1* might be a mechanism linking DNA replication with the following mitosis. That other cell cycle genes modulated in the *E2Fa-DPa* plants are not detected can be explained by the lack of many important E2F-DP target genes on the microarray and the putative difficulty in identifying changes in expression levels of lowly expressed genes in microarray hybridizations.

Cell wall biogenesis genes

Four members of the xyloglucan endotransglucosylase (XET) gene family are found to be twofold or more up-regulated in the *E2Fa-DPa* plants, one of them identical to the previously

described *Meri-5* gene (Medford et al., 1991). XETs are enzymes that modify cell wall components and are presumed to play a role in altering size, shape and physical properties of plant cells. Reversal breakage of the xyloglucan tethers by XETs has been proposed as a mechanism for allowing cell wall loosening in turgor-driven cell expansion (Campbell and Braam, 1999). However, there are several reasons for believing that E2Fa-DPa-induced XETs are not required for cell expansion. First, cells divide more frequently in the *E2Fa-DPa* plants, but the overall cell size is smaller in transgenic than in control plants; so, no overall increase in expansion rates is needed. Second, no induction is seen of genes with a known role in cell expansion, such as expansins. Therefore, the hydrolytic activity of the XETs might rather be required to incorporate the newly synthesized cell walls formed during cytokinesis into the existing cell wall structure. Alternatively, because XET activity has been shown to be involved in the postgerminative mobilization of xyloglucan storage reserves in *Nasturtium* cotyledons (Farkas et al., 1992; Fanutti et al., 1993), induction of XETs in *E2Fa-DPa* plants might be related to polysaccharide breakdown to serve the metabolic and energy needs that are required to synthesize new nucleotides (see below).

Interestingly, two XETs can be identified in the set of twofold-or-more down-regulated genes. These XETs are more related to each other than to the induced *XET* genes. This differential response of XETs toward the E2Fa-DPa-induced phenotypes suggests that plant XETs can be classified into at least two different functional classes.

Genes involved in metabolism and biogenesis

A relatively large number of genes involved in metabolism and biogenesis were found in both the up-regulated and down-regulated gene groups. Most remarkable is the induction of genes involved in nitrogen assimilation, such as nitrate reductase (*NIA2*), glutamine synthetase (*GS*), and glutamate synthase (*GOGAT*) (Fig. 3). Although not present on the microarray, the nitrite reductase (NiR) gene was found to be induced as well in the transgenic lines, as demonstrated by RT-mediated PCR analysis (Fig. 2). Nitrogen and nitrite reductase catalyze the first two steps in the nitrogen assimilation pathway, whereas GS and GOGAT are involved both in the primary assimilation of nitrogen and the reassimilation of free ammonium. This mechanism supplies the plant with all nitrogen needed for the biosynthesis of amino acids and other nitrogen-containing compounds.

There are other indications that the nitrogen metabolism is altered in the *E2Fa-DPa* plants; these include the modification of genes homologous to genes expressed during the formation of nitrogen-fixing nodules in *Medicago sativa* (*MTN3* and a nodulin-like gene), and the down-regulation of genes involved in sulfur assimilation (two different genes encoding adenylylsulfate reductase [APR] and a putative adenine phosphosulfate kinase). Genes involved in sulfur assimilation have been shown before to be transcriptionally down-regulated during nitrogen deficiency (Koprikova et al., 2000).

The altered expression of genes involved in nitrogen assimilation and metabolism in the *E2Fa-DPa* transgenic plants might reflect the need for nitrogen for the nucleotide biosynthesis, because purine and pyrimidine bases are rich in nitrogen. If nitrogen assimilation were indeed stimulated by *E2Fa-DPa* overexpression, two requirement should be fulfilled. Firstly, there should be enough α-ketoglutarate to act as an acceptor molecule for ammonium (Lancien et al., 2000)

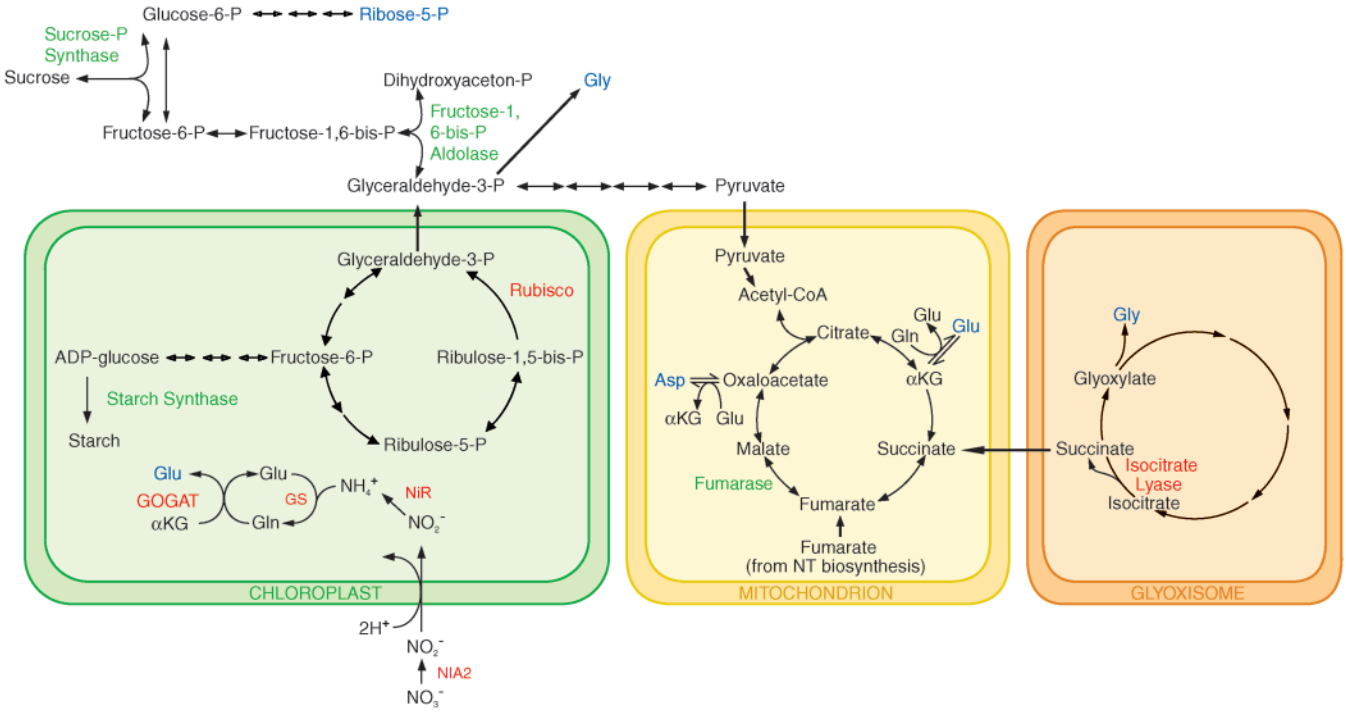

Fig. 3. Sources of α-ketoglutarate in the E2Fa-DPa-overproducing cells. Genes encoding enzymes shown in red and green are up-regulated or down-regulated in the *E2Fa-DPa* versus wild-type plants, respectively. Products indicated in blue act as precursors for nucleotide biosynthesis. α-KG, α-ketoglutarate; GOGAT, glutamate synthase; GS, glutamine synthetase; NIA2, nitrate reductase; NiR, nitrite reductase.

and secondly, because assimilation of nitrogen is energy consuming, the rate of reductant production should be higher in the *E2Fa-DPa* transgenic than in the wild-type plants.

Our microarray data suggest that in the accumulation of α-ketoglutarate in *E2Fa-DPa*-overexpressing plants is stimulated in different ways. First, α-ketoglutarate production is improved by increased photosynthetic activity, as indicated by the 4.7-fold up-regulation of the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (Fig. 3), with accumulation of glyceraldehyde-3-phosphate as a result. Glyceraldehyde-3-phosphate can be converted into fructose-1,6-bisphosphate by fructose bisphosphate aldolase. However, a sixfold down-regulation of the fructose bisphosphate aldolase gene rather suggests the conversion of glyceraldehyde-3-phosphate into pyruvate, which can be converted into α-ketoglutarate in the citrate cycle. The preferential conversion of glyceraldehyde-3-phosphate to pyruvate fits the increased need for amino acids rather than for sugars to drive nucleotide biosynthesis (Fig. 3).

A second source of α-ketoglutarate can be provided by the glyoxylate cycle. In E2Fa-DPa-overproducing plants we observed a 3.1-fold increase in expression of isocitrate lyase, suggesting an increased lipid turnover. Isocitrate lyase activity cleaves isocitrate into glyoxylate and succinate (Fig. 3). Whereas the produced glyoxylate can be converted into glycine, which is also required for de novo nucleotide biosynthesis, succinate can be converted into α-ketoglutarate in the citrate cycle. A 2.3-fold decrease in the expression of the fumarase gene presumably stimulates the subsequent conversion of α-ketoglutarate to glutamate by triggering an accumulation of succinate and fumarate, which are also side products formed during de novo nucleotide biosynthesis (Fig. 3).

Reductant in plants mainly originates from photosynthetic electron transport in leaves. Corresponding with the increased need for reductant, several components of the chloroplast electron transport chain and associated ATP-synthesizing apparatus, such as cytochrome B6, a photosystem II subunit, and the ATPase ε subunit, are up-regulated in the *E2Fa-DPa* transgenic plants. Increased expression of the protochlorophyllide reductase precursor even indicates an increase in chlorophyll biosynthesis.

E2Fa-DPa plants may suffer from nitrogen starvation that has an impact on amino acid biosynthesis. Three different amino acid aminotransferases are down-regulated in the *E2Fa-DPa* plants. Shortage of nitrogen-rich amino acids is also evident from the reduced expression of genes encoding vegetative storage proteins (*VSP1* and *VSP2*) and *ERD10*, a protein with a compositional bias toward glutamate (Kiyosue et al., 1994). Additional evidence for amino acid shortage comes from the down-regulation of a myrosinase-binding protein and the cytochrome P450 monooxygenase CYP83A1. Both proteins are involved in the biosynthesis of glucosinolates, nitrogen- and sulfur-containing products derived from amino acids (Wittstock and Halkier, 2002).

Promoter analysis of E2Fa-DPa-regulated genes

The DNA-binding domains of the E2F and DP proteins are highly conserved between plants and mammals and, correspondingly, plant E2F-DP proteins have been shown by the technique of electrophoresis mobility shift assay to bind to the same canonical DNA-binding site as their mammalian counterparts (Albani et al., 2000; Ramirez-Parra and Gutierrez, 2000; de Jager et al., 2001). Furthermore, these E2F-binding sites regulate the expression of several plant genes involved in DNA synthesis (Kosugi and Ohashi, 2002a; Chabouté et al., 2000; Castellano et al., 2001; Egelkrout et al., 2001; Stevens et al., 2002).

To distinguish between the putatively direct target genes of E2Fa-DPa and the secondarily induced genes, the first 500 bp upstream of the ATG start codon of the genes with 2-fold or higher change in expression were scanned for the presence of an E2F-like-binding site matching the (A/T)TT(G/C)(G/C)C(G/C)(G/C) sequence, which corresponds to all the different E2F-DP-binding motifs that have been described in plants. Of all the different permutations only the TTTCCCGC and TTTGGCGG elements were enriched significantly ($P<0.01$) in the set of E2Fa-DPa-upregulated genes, suggesting these are the preferred binding site of the E2Fa-DPa complex (Table 3). Moreover, six out of eight target genes containing one or more of these elements belong to the group of genes involved in DNA replication and modification. The observation that not all genes that enclase this DNA sequence in their promoter are induced upon *E2Fa-DPa* overexpression suggests that the presence of the TTTCCCGC or TTTGGCGG motif is not the only element to

Table 3. Number of E2F elements in the different data sets

E2F motif	All genes (4390)*	Upregulated genes (84)	Downregulated genes (104)
TTTCCCCC	49	2	1
TTTCCCCG	31	1	1
TTTCCCGC	46	5	0
TTTCCCGG	61	1	1
TTTCGCCC	16	0	0
TTTCGCCG	76	2	0
TTTCGCGC	19	2	0
TTTCGCGG	30	2	1
TTTGCCCC	35	1	1
TTTGCCCG	13	0	0
TTTGCCGC	24	0	1
TTTGCCGG	34	0	0
TTTGGCCC	54	0	1
TTTGGCCG	38	0	3
TTTGGCGC	18	1	0
TTTGGCGG	47	4	0
ATTCCCCC	14	0	2
ATTCCCCG	21	0	0
ATTCCCGC	11	1	0
ATTCCCGG	23	0	0
ATTCGCCC	10	0	0
ATTCGCCG	42	0	0
ATTCGCGC	13	1	0
ATTCGCGG	9	0	1
ATTGCCCC	14	0	1
ATTGCCCG	6	0	0
ATTGCCGC	15	0	0
ATTGCCGG	0	0	0
ATTGGCCC	42	0	2
ATTGGCCG	13	0	1
ATTGGCGC	12	2	0
ATTGGCGG	28	0	0
Total	864	25	17

*Promoters of mitochondrial and chloroplastic genes were omitted from this analysis

make a gene responsive toward E2Fa-DPa, and that E2Fa-DPa may cooperate with other factors to activate transcription. Alternatively, the promoters of non-responsive genes might be shielded with other transcription factor complexes. A putative candidate is the E2Fc protein which, in analogy with the mammalian E2F6 protein, lacks a strong transactivation domain (del Pozo et al., 2002). Alternative candidates are the recently discovered DEL proteins, proven to bind as monomers to the canonical E2F-binding site (Kosugi and Ohashi, 2002b; Mariconti et al., 2002). Because of a lack of transcriptional activation domain, the DEL proteins are postulated to act as repressors of E2F-DP-regulated genes by competing for the same binding site.

It is not excluded that genes without an E2F-like-binding site are not directly activated by E2Fa-DPa. Chromatin immunoprecipitation experiments have shown that mammalian E2F factors can bind to promoters without a clear E2F recognition motif (Kiyosue et al., 1994), suggesting that E2F-DP might recognize non-canonical binding sites, or might be recruited by promoters through the association of other factors. In this respect, the *Chlorella vulgaris* nitrate reductase gene, of which the *Arabidopsis* homologue was shown here to be induced by E2F-DPa, binds an E2F-DP complex, although a clear consensus binding site is lacking (Cannons and Shiflett, 2001).

E2Fs can activate as well as repress promoter activity (Trimarchi and Lees, 2002). In the PCNA, MCM3 and RNR2 promoters, E2F sequences have been identified that act as a negative regulatory element during development (Chabouté et al., 2000; Egelkrout et al., 2001; Stevens et al., 2002). In the set of down-regulated genes, no particular enrichment of a specific E2F sequence could be seen (Table 3). Therefore, the data suggest that the E2Fa-DPa complex works as a transcriptional activator and that other E2F-DP complexes are involved in E2F-mediated transcriptional repression.

Conclusions

Microarray analysis of *E2Fa-DPa*-overexpressing lines identified a cross-talking genetic network between DNA replication, nitrogen assimilation and photosynthesis. The putatively direct E2Fa-DPa target genes as identified by the presence of an E2F-DP-binding site, belong to the group of genes involved in DNA synthesis, whereas the secondarily induced genes are mainly linked to nitrogen assimilation. In a recently published microarray experiment in which the periodic expression of genes during the cell cycle was monitored, genes with a role in nitrogen assimilation (aspartate aminotransferase and a nitrate transporter) were found to be specifically expressed during the S phase (Menges et al., 2002). Because purine and pyrimidine bases are nitrogen rich, we postulate that induction of nitrogen assimilation genes during DNA synthesis in wild-type and *E2Fa-DPa* transgenic plants is required to supply enough nitrogen for nucleotide biosynthesis. However, in the *EFa-DPa* transgenic plants, increased nitrogen assimilation most probably does not meet all the nucleotide biosynthesis needs, as seen by the expression modulation of many genes involved in nitrogen and carbohydrate metabolism. The drain of nitrogen from essential biosynthetic pathways to the nucleotide biosynthesis pathway is expected to affect other aspects of plant metabolism, as can be seen from the reduced expression of vegetative storage protein genes and genes involved in amino acid biosynthesis. This altered metabolism might, at least in part, contribute to the growth arrest observed in *E2Fa-DPa* transgenic plants.

The exact regulatory pathways and factors controlling the nitrogen assimilation pathway in plants are still unknown. In addition to the genes involved in DNA replication and metabolism, our data contain a relatively large number of genes with unspecified function (Tables 1 and 2). For instance, a GATA zinc-finger-encoded gene with a still unknown function is found between the up-regulated regulatory genes. This gene might encode the ortholog of the *Neurospora crassa* nit-2 protein that has been shown to positively regulate expression of the nitrate reductase gene (Fu and Marzluf, 1990). Other regulatory genes modified in the *E2Fa-DPa* plants encode protein kinases and several putative receptor kinases. These genes might include some novel key regulatory components in the process of nitrogen assimilation or regulation of efficient nitrogen usage. It will be of great interest to analyze their role in nitrogen assimilation, metabolism and plant growth.

The authors thank the members of the cell cycle group for fruitful discussions and useful suggestions and Martine De Cock for help in preparing the manuscript. This work was supported by grants from the Interuniversity Poles of Attraction Programme (Belgian State, Prime Minister's Office – Federal Office for Scientific, Technical and Cultural Affairs; P5/13), the European Union (ECCO QLG2-CT1999-00454), CropDesign NV (0235), and the Fund for Scientific Research (Flanders) (G.0025.02). K.V. and K.F. are indebted to the Instituut voor de aanmoediging van Innovatie door Wetenschap en Technologie in Vlaanderen for predoctoral fellowships. L.D.V. is a postdoctoral fellow of the Fund for Scientific Research (Flanders).

References

Albani, D., Mariconti, L., Ricagno, S., Pitto, L., Moroni, C., Helin, K. and Cella, R. (2000). DcE2F, a functional plant E2F-like transcriptional activator from *Daucus carota*. *J. Biol. Chem*. **275**, 19258-19267.

Campbell, P. and Braam, J. (1999). Xyloglucan endotransglycosylases: diversity of genes, enzymes and potential wall-modifying functions. *Trends Plant Sci*. **4**, 361-366.

Cannons, A. C. and Shiflett, S. D. (2001). Transcriptional regulation of the nitrate reductase gene in *Chlorella vulgaris*: identification of regulatory elements controlling expression. *Curr. Genet*. **40**, 128-135.

Castellano, M. M., del Pozo, J. C., Ramirez-Parra, E., Brown, S. and Gutierrez, C. (2001). Expression and stability of Arabidopsis CDC6 are associated with endoreplication. *Plant Cell* **13**, 2671-2686.

Chabouté, M.-E., Clément, B., Sekine, M., Philipps, G. and Chaubet-Gigot, N. (2000). Cell cycle regulation of the tobacco ribonucleotide reductase small subunit gene is mediated by E2F-like elements. *Plant Cell* **12**, 1987-1999.

de Jager, S. M., Menges, M., Bauer, U.-M. and Murray, J. A. H. (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, G. T. S., de Almeida Engler, J., Ormenese, S., Maes, S., Naudts, M., Van Der Schueren, E., Jacqmard, A., Engler, G. and Inzé, D. (2002). Control of proliferation, endoreduplication and differentiation by the *Arabidopsis* E2Fa/DPa transcription factor. *EMBO J*. **21**, 1360-1368.

del Pozo, J. C., Boniotti, M. B. and 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.

Egelkrout, E. M., Robertson, D. and 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.

Fanutti, C., Gidley, M. J. and Reid, J. S. G. (1993). Action of a pure

xyloglucan *endo*-transglycosylase (formerly called xyloglucan-specific *endo*-(1→4)-β-D-glucanase) from the cotyledons of germinated nasturtium seeds. *Plant J*. **3**, 691-700.

Farkas, V., Sulova, Z., Stratilova, E., Hanna, R. and Maclachlan, G. (1992). Cleavage of xyloglucan by nasturtium seed xyloglucanase and transglycosylation to xyloglucan subunit oligosaccharides. *Arch. Biochem. Biophys*. **298**, 365-370.

Fu, Y.-H. and Marzluf, G. A. (1990). *nit-2*, the major positive-acting nitrogen regulatory gene of *Neurospora crassa*, encodes a sequence-specific DNA-binding protein. *Proc. Natl. Acad. Sci. USA* **87**, 5331-5335.

Ishida, S., Huang, E., Zuzan, H., Spang, R., Leone, G., West, M. and Nevins, J. R. (2001). Role for E2F in control of both DNA replication and mitotic functions as revealed from DNA microarray analysis. *Mol. Cell. Biol*. **21**, 4684-4699.

Kel, A. E., Kel-Margoulis, O. V., Farnham, P. J., Bartley, S. M., Wingender, E. and Zhang, M. Q. (2001). Computer-assisted identification of cell cycle-related genes: new targets for E2F transcription factors. *J. Mol. Biol*. **309**, 99-120.

Kiyosue, T., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1994). Cloning of cDNAs for genes that are early-responsive to dehydration stress (ERDs) in *Arabidopsis thaliana* L.: identification of three ERDs as HSP cognate genes. *Plant Mol. Biol*. **25**, 791-798.

Koprivova, A., Suter, M., Op den Camp, R., Brunold, C. and Kopriva, S. (2000). Regulation of sulfate assimilation by nitrogen in Arabidopsis. *Plant Physiol*. **122**, 737-746.

Kosugi, S. and Ohashi, Y. (2002a). 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. and Ohashi, Y. (2002b). E2Ls, E2F-like repressors of *Arabidopsis* that bind to E2F sites in a monomeric form. *J. Biol. Chem*. **277**, 16553-16558.

Lancien, M., Gadal, P. and Hodges, M. (2000). Enzyme redundancy and the importance of 2-oxoglutarate in higher plant ammonium assimilation. *Plant Physiol*. **123**, 817-824.

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouzé, P. and Rombauts, S. (2002). PlantCare: a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Res*. **30**, 325-327.

Lusser, A., Eberharter, A., Loidl, A., Goralik-Schramel, M., Horngacher, M., Haas, H. and Loidl, P. (1999). Analysis of the histone acetyltransferase B complex of maize embryos. *Nucleic Acids Res*. **27**, 4427-4435.

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

Mariconti, L., Pellegrini, B., Cantoni, R., Stevens, R., Bergounioux, C., Cella, R. and 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.

Medford, J. I., Elmer, J. S. and Klee, H. J. (1991). Molecular cloning and characterization of genes expressed in shoot apical meristems. *Plant Cell* **3**, 359-370.

Menges, M. and Murray, J. A. H. (2002). Synchronous *Arabidopsis* suspension cultures for analysis of cell-cycle gene activity. *Plant J*. **20**, 203-212.

Menges, M., Hennig, L., Gruissem, W. and Murray, J. A. H. (2002). Cell cycle-regulated gene expression in *Arabidopsis*. *J. Biol. Chem*. **277**, 41987-42002.

Müller, H., Bracken, A. P., Vernell, R., Moroni, M. C., Christians, F., Grassilli, E., Prosperini, E., Vigo, E., Oliner, J. D. and Helin, K. (2001). E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis. *Genes Dev*. **15**, 267-285.

Porceddu, A., De Veylder, L., Hayles, J., Van Montagu, M., Inzé, D. and Mironov, V. (1999). Mutational analysis of two *Arabidopsis thaliana* cyclin-dependent kinases in fission yeast. *FEBS Lett*. **446**, 182-188.

Puskás, L. G., Zvara, A., Hackler, L. J. and Van Hummelen, P. (2002). RNA amplification results in reproducible microarray data with slight ratio bias. *Biotechniques* **32**, 1330-1341.

Ramirez-Parra, E. and Gutierrez, C. (2000). Characterization of wheat DP, a heterodimerization partner of the plant E2F transcription factor which stimulates E2F-DNA binding. *FEBS Lett*. **486**, 73-78.

Ramirez-Parra, E., Fründt, C. and Gutierrez, C. (2003). A genome-wide identification of E2F-regulated genes in *Arabidopsis*. *Plant J*. **33**, 801-811.

Ramírez-Parra, E., Xie, Q., Boniotti, M. B. and Gutierrez, C. (1999). The cloning of plant E2F, a retinoblastoma-binding protein, reveals unique and conserved features with animal G_1/S regulators. *Nucleic Acids Res*. **27**, 3527-3533.

Ren, B., Cam, H., Takahashi, Y., Volkert, T., Terragni, J., Young, R. A. and Dynlacht, B. D. (2002). E2F integrates cell cycle progression with DNA repair, replication, and G_2/M checkpoints. *Genes Dev*. **16**, 245-256.

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

Stevens, R., Mariconti, L., Rossignol, P., Perennes, C., Cella, R. and 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.

Trimarchi, J. M. and Lees, J. A. (2002). Sibling rivalry in the E2F family. *Nature Rev. Mol. Cell. Biol*. **3**, 11-20.

Valvekens, D., Van Montagu, M. and Van Lijsebettens, M. (1988). *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. *Proc. Natl. Acad. Sci. USA* **85**, 5536-5540.

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

Weinmann, A., Bartley, S. M., Zhang, T., Zhang, M. Q. and Farnham, P. J. (2001). Use of chromatin immunoprecipitation to clone novel E2F target promoters. *Mol. Cell. Biol*. **21**, 6820-6832.

Wittstock, U. and Halkier, B. A. (2002). Glucosinolate research in the Arabidopsis era. *Trends Plant Sci*. **7**, 263-270.

Wolfinger, R. D., Gibson, G., Wolfinger, E. D., Bennett, L., Hamadeh, H., Bushel, P., Afshari, C. and Paules, R. S. (2001). Assessing gene significance from cDNA microarray expression data via mixed models. *J. Comput. Biol*. **8**, 625-637.

Yang, Y. H., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J. and Speed, T. P. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res*. **30**, e15.