Theor Appl Genet (2009) 118:891–901
DOI 10.1007/s00122-008-0947-4

ORIGINAL PAPER

cDNA cloning, expression levels and gene mapping of photosynthetic and non-photosynthetic ferredoxin genes in sunflower (*Helianthus annuus* L.)

M. Venegas-Calerón · A. Zambelli · N. Ruiz-López · L. Youssar · A. León · R. Garcés · Enrique Martínez-Force

Received: 24 April 2008 / Accepted: 4 December 2008 / Published online: 8 January 2009

Abstract Fatty acid desaturation in plastids and chloroplasts depends on the electron-donor activity of ferredoxins. Using degenerate oligonucleotides designed from known photosynthetic and heterotrophic plant ferredoxin sequences, two full-length ferredoxin cDNAs were cloned from sunflower (*Helianthus annuus* L.) leaves and developing seeds, *Ha*Fd1 and *Ha*Fd2, homologous to photosynthetic and non-photosynthetic ferredoxins, respectively. Based on these cDNAs, the respective genomic sequences were obtained and the presence of DNA polymorphisms was investigated. Complete sequencing of the *Ha*Fd1 and *Ha*Fd2 genes in different lines indicated the presence of two haplotypes for *Ha*Fd2 and their alignment showed that sequence polymorphisms are restricted to the 5′-NTR intron. In addition, specific DNA markers for the *Ha*Fd1 and *Ha*Fd2 genes were developed that enabled the genes to be mapped. Accordingly, the *Ha*Fd1 locus maps to linkage group 10 of the public sunflower map, while the *Ha*Fd2 locus maps to linkage group 11. Both ferredoxins display different spatial-temporal patterns of expression. While *Ha*Fd2 is expressed at similar levels in all tissues tested (leaves, stem, roots, cotyledons and developing seeds), *Ha*Fd1 is more strongly expressed in green tissues than in all the other tissues tested. Both photosynthetic- and heterotrophic-ferredoxins are present in sunflower seeds and may contribute to fatty acid desaturation during oil accumulation. Nevertheless, the levels of *Ha*Fd2 expression during seed formation are distinct in lines that only varied in the *Ha*Fd2 haplotypes they expressed.

Communicated by A. Bervillé.

M. Venegas-Calerón · L. Youssar · R. Garcés · E. Martínez-Force (✉)
Instituto de la Grasa (CSIC), Av. Padre García Tejero 4, 41012 Sevilla, Spain
e-mail: emforce@cica.es

A. Zambelli · A. León
Centro de Biotecnología, Advanta Semillas, 7620 Balcarce, Argentina

N. Ruiz-López
Crop Performance and Improvement Division, Rothamsted Research, Harpenden, Herts AL5 2JQ, UK

Introduction

Plant ferredoxins (Fds) are soluble, low MW proteins that mediate the transfer of one electron from a donor to an acceptor. The redox active centre is a [2Fe–2S] cluster that confers a highly negative redox potential to the protein (−350 to −450 mV), making Fds a powerful reductant. This cluster is maintained by four highly conserved Cys residues. Although ferredoxins have been most extensively studied in photosynthetic tissue where they serves as a major electron carrier in photosystem I to produce NADPH, multiple Fd isoforms are now known to exist in both photosynthetic and non-photosynthetic tissues (Wada et al. 1989; Morigasaki et al. 1990; Green et al. 1991; Hase et al. 1991a, b; Kamide et al. 1995; Aoki and Wada 1996). Indeed, Fds are involved in a wide spectrum of redox events associated with metabolism in higher plant plastids.

The donation of electrons by Fds to many other plastid enzymes is essential for cellular processes as varied as nitrogen and sulfur assimilation, redox regulation, and amino acid, and fatty acid synthesis (Knaff 1996). Based on their tissue distribution, plant Fds can be classified into two broad categories. Photosynthetic ferredoxins have been shown to be light regulated and are predominantly

expressed in photosynthetic tissues. In contrast, the heterotrophic Fds are not regulated by light and they have a more ubiquitous tissue distribution (Kimata and Hase 1989; Hase et al. 1991a). The sequencing of the *Arabidopsis* genome has enabled four Fd genes expressed in that organism to be studied, showing that variation between *Arabidopsis* Fd sequences confers different functional characteristics on the individual proteins, coupled to different expression patterns (Hanke et al. 2004).

In higher plants, unsaturated fatty acids are synthesized by the sequential insertion of double bonds into palmitic and stearic acid derivatives. These reactions are carried out by desaturases and they require the donation of electrons from an electron-donating system. The first of these desaturation events is carried out in plastids and chloroplasts by the soluble stearoyl-ACP desaturase (SAD). Subsequently, to synthesize the 16:2/18:2 and 16:3/18:3 fatty acids that are essential for plant membranes, a second and a third double bond are introduced by $\Delta 12$ and ω-3 membrane-bound desaturases, respectively. These desaturases can be found in plastids/chloroplasts (the so-called "prokaryotic pathway") where they use Fds as their electron donor, a pathway used predominantly in photosynthetically active tissues (McKeon and Stumpf 1982; Schmidt and Heinz 1990). Alternatively, desaturase activity in the endoplasmic reticulum ("eukaryotic pathway") relies on cytochrome b5 as the electron-donor, a microsomal pathway that predominates in non-green tissues and developing seeds (Smith et al. 1990; Kearns et al. 1991; Browse and Somerville 1991). Because SAD catalyzes the first reaction of desaturation, it plays a key role in determining the ratio of total saturated to unsaturated fatty acids. Plant Fd isoforms, both photosynthetic and heterotrophic, may interact specifically with acyl-ACP desaturases, including SAD, and modify their activity. Indeed, heterotrophic Fd isoforms may be the electron donor for this reaction in vivo (Schultz et al. 2000).

Biochemical characterization of the high-stearic CAS-3 mutant sunflower line (*Helianthus annuus* L.; Cantisán et al. 2000) and its QTL analysis in a segregating population derived from this line indicated that SAD (locus SAD17A) was the principal gene involved in the high-stearic phenotype (Pérez-Vich et al. 2002). RFLP–AFLP linkage maps from two different mapping populations derived from CAS-3 mapped the SAD17A locus to linkage group (LG) 1 of the sunflower genetic map (according to the LG nomenclature of Berry et al. 1997). Other minor QTL that affect stearic acid content in sunflower seeds but that are not associated with candidate genes have been detected (Pérez-Vich et al. 2004), which mapped to LG3, LG7, and LG13.

Taking into account that different Fd isoforms could act as electron donors of the SAD reaction, the genes that encode them might correspond to these minor genes responsible for regulating the stearic acid content. Here, we describe the cloning and sequencing of the Fds expressed in vegetative tissues and developing seeds of sunflower, their gene structure, chromosomal location, and expression levels.

Materials and methods

Biological material

Sunflower plants from the public RHA274 line and CAS3 line were cultivated in growth chambers at 25/15°C (day/night cycles), with a 16-h photoperiod and a photon flux density of 200 μmol m^{-2} s^{-1}.

Escherichia coli DH5α strain (Bethesda Research Laboratories) was used for all plasmid manipulations. The bacteria were grown in Luria broth (1% Bacto tryptone, 0.5% Bacto yeast extract, 1% NaCl, pH 7: Sambrook et al. 1989) with vigorous agitation. The medium was supplemented with antibiotics as required (ampicillin at 50 μg ml^{-1}; kanamycin at 50 μg ml^{-1}).

Cloning of sunflower photosynthetic and heterotrophic ferredoxin cDNAs and genomic sequences

Approximately 0.4 g of developing sunflower endosperm, 15 days after flowering (DAF), was ground in liquid N_2 in a precooled sterile pestle and mortar. The mRNA was isolated using a MicroFastTrack Kit (Invitrogen, Groningen, The Netherlands), the mRNA pellet was resuspended in 33 μl TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8), and cDNAs were obtained using a Ready-To-Go T-Primed First-Strand Kit (Amersham Bioscience, Roosendaal, The Netherlands).

Plant photosynthetic and heterotrophic ferredoxin protein sequences from public databases were aligned to identify regions of homology using the ClustalX v1.8 program (Thompson et al. 1997). The degenerate oligonucleotides, FdphoF1 (5′-GHNGGNMWYGANTTRC C-3′) and FdhetF1 (5′-GCNGGNBTYGANCTNCC-3′), were used in combination with the FA2Z oligonucleotide (5′-AACTGGAAGAATTCGCGG-3′), complementary to the sequences incorporated during the initial cDNA synthesis, to obtain PCR fragments (330 bp long) that corresponded to the 3′-ends of the sunflower photosynthetic and heterotrophic ferredoxins, respectively. The 5′- ends were obtained using the Smart™-RACE cDNA amplification kit (Clontech) and the internal oligonucleotides: FdphoR1 (5′-GTCATCTCCTCCTCCTTGTGGG-3′) and FdphoR2 (5′-ATCTGGTCATCATCAAGAAAACTCTG G-3′); or FdhetR1 (5′-GTGTGTGAACAACACAATCA

CCAGTCGGG-3′) and FdhetR2 (5′-CGACCCATCAG ATTGGTCAACAGCACC-3′). All primers were synthesized by MWG Biotech AG (Ebersberg, Germany). The PCR fragments were cloned into the pGEM®-T Easy Vector (Promega, Madison, WI, USA), sequenced by GATC Gmbh (Konstanz, Germany) and assembled to obtain DNA sequences of about 429 bp and 474 bp. Once their identity was confirmed using the Blast software (Altschul et al. 1990), the photosynthetic sunflower ferredoxin was named *Ha*Fd1 and the heterotrophic sunflower ferredoxin was named *Ha*Fd2. The complete cDNA sequences of the sunflower photosynthetic and heterotrophic ferredoxins, *Ha*Fd1 and *Ha*Fd2, were deposited in GenBank with the accession numbers AY189833 and DQ012385, respectively.

Total genomic DNA was isolated from leaf tissue from the sunflower RHA274 line by a modified CTAB miniprep method described elsewhere (Lenardon et al. 2005). Based on the cDNA sequences, oligonucleotides were designed to amplify the coding regions of *Ha*Fd1 and *Ha*Fd2 by PCR using total genomic DNA as the template. For *Ha*Fd1 the primer combination used was Fd1-F_1 (5′-ACCATGGC CAGCACCTCCTT-3′) and Fd1-R_2 (5′-AACTGACT TAATGTAAAGGC-3′; Fig. 1) with the following PCR cycles: an initial denaturing step at 94°C for 1 min; a program of 35 cycles of 94°C for 30 s, 58°C for 30 s, 72°C for 30 s; and a final elongation step of 72°C for 10 min. PCR products were purified using Wizard SV Gel and the PCR clean-up system (Promega, Madison, USA) for direct sequencing. To obtain the genomic sequence of *Ha*Fd2 two primer combinations were designed based on the cDNA sequence and that produced overlapping fragments: Fd2-F_1 (5′-GACCTCTTGATTTCTCCGCT-3′) and Fd2-R_1 (5′-GT ATGGCAGCTCAATTCCCG-3′); and Fd2-F_2 (5′-CGT GACACCTGATGGTGAAC-3′) and Fd2-R_2 (5′-CAACTC TAAACGACAAACCG-3′; Fig. 1). The PCR conditions, product purification and sequencing were as indicated for *Ha*Fd1. The *Ha*Fd2 PCR products were bigger than expected from the cDNA, indicating the possible presence of introns or non-transcribed regions (NTR). We define NTR as non-transcribed DNA regions not flanked by exons, differentiating them thus from introns, which are flanked by exons. Besides, NTR could be classified as 5′- or 3′-NTR, depending on its gene-location. The DNA bands were gel purified, cloned into the pGEM®-T Easy Vector (Promega, Madison, WI, USA) and sequenced by GATC Gmbh (Konstanz, Germany). Based on the primary sequencing results, additional specific primers were designed in order to obtain the whole genomic sequence of the gene: Fd2-F_3 (5′-GGA ATTGGAAATCCTAGTGT-3′) and Fd2-F_4 (5′-TTGTG TCCGTTTGAGTTTTA-3′); and Fd2-R_3 (5′-CACAATA AATCACAAAGCCA-3′; Fig. 1). Whole genomic sequences of *Ha*Fd1 and *Ha*Fd2 from a set of sunflower lines were obtained and compared by alignment using ClustalX v.1.8 program and the default settings (Thompson et al. 1997).

DNA marker development and genetic mapping

Comparison of the genomic sequence of *Ha*Fd1 and *Ha*Fd2 from different sunflower lines showed DNA polymorphisms that allowed specific molecular markers to be developed for each gene. An allele specific amplification (ASA) marker for *Ha*Fd1 was developed that detected a T/G single nucleotide polymorphism (SNP) at nucleotide position (np) 192 (see "Results"). Two forward primers that only differ at their 3′-end (according to the two allelic variants of the SNP 192) were designed: Fd1A_1-F (5′-TTCTTGACCATTGTGAAGAT-3′) and Fd1A_2-F (5′-TTCTTGACCATTGTGAAGAG-3′). Both primers were used separately in combination with the Fd1-R_2 primer to produce a 290 bp fragment. Two independent PCR reactions per individual were performed to differentiate the two allelic variants of the *Ha*Fd1.

For *Ha*Fd2 a specific DNA marker was developed on the basis of an insertion-deletion (INDEL 1) found in the 5′-NTR region (see "Results"). Two forward primers were specifically designed for each of the two allelic variants of the gene: Fd2A1-F (5′-GGTTTTATTACTGTAAGTTC-3′) (deleted allele); and Fd2A2-F (5′-AAGGTTTTTAGGT TTTATAC-3′) (inserted allele). Both primers were used separately in combination with Fd2-R (5′-TATCGTAA-CAGCATCAGATC-3′) Fd1-R_2 to produce a 214-bp and 222-bp fragment, respectively. Two independent PCR reactions per individual were performed to differentiate the two allelic variants of the *Ha*Fd2.

The specific *Ha*Fd1 and *Ha*Fd2 markers were polymorphic in the proprietary Advanta lines P504 and R232. Therefore, both markers were genotyped in an F2 segregating sunflower population generated by crossing P504 × R232 and integrated into the genetic linkage map obtained using simple sequence repeat (SSR) and insertion-deletion (INDEL) markers from previously published genetic linkage maps (Tang et al. 2002; Yu et al. 2003). Genetic mapping analyses were performed using MAPMAKER (Lander et al. 1987), essentially as described by Tang et al. (2002).

cDNA and protein sequence analysis

Alignment of the amino acid sequences, including the transit peptides, for ferredoxin proteins deposited at GENBANK was performed using the ClustalX v.1.8 program with the default settings (Thompson et al. 1997). These entire alignments were used to generate a phylogenetic tree based on the neighbor-joining algorithm (Saitou and Nei 1987), and the resulting 'phenogram' was drawn

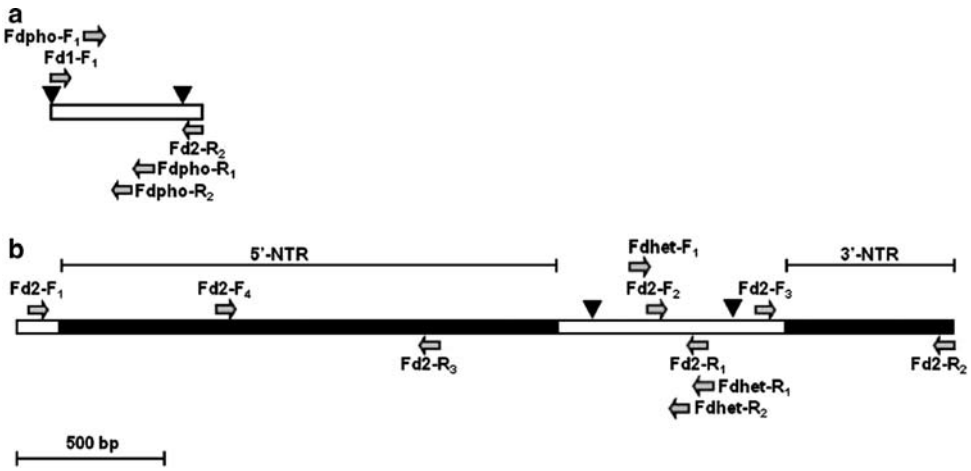

Fig. 1 Schematic representation of the genomic organization of the *Ha*Fd1 (**a**) and *Ha*Fd2 (**b**) genes. cDNA segments are depicted in *white* while the 5′- and 3′ non-transcribed regions are shown in *black* (5′-NTR and 3′-NTR, respectively). *Arrows* show the position of all primers used to amplify the entire *Ha*Fd1 and *Ha*Fd2 gene sequences. *Arrowheads* (*inverted triangle*) flank the coding region of the gene

using the TreeView program (Page 1996). The chloroplast transit peptides were identified through alignment with known ferredoxin sequences and using the network-based method TargetP V1.0 (Emanuelsson et al. 2000).

Determination of mRNA levels by real time-PCR

The cDNAs from developing seeds, roots, hypocotyls, green cotyledons, and leaves were obtained as described earlier. These cDNAs were subjected to real time-PCR with specific primer pairs and SYBR Green I according to the manufacturer's instructions (QuantiTect™ SYBR® Green PCR Kit, Quiagen, Crawley, UK) using a SmartCycler system (Cepheid, Sunnyvale, CA, USA). The primers used for amplification were as follows: for *Ha*Fd1 QPCRFd1F1 (5′-GTTCGACTGCGCTGATGATA-3′) and QPCRFd1R1 (5′-TGTGCATGAGGAGCAAGAAC-3′); and for *Ha*Fd2 QPCRFd2F1 (5′-AAGTACCCATCCACCCTTCC-3′) and QPCRFd2R1 (5′-CGTGTTCACCATCAGGTGTC-3′). The reaction mixture was heated to 95°C for 120 s before subjecting it to PCR cycles of 95°C for 15 s, 56°C for 30 s, and 72°C for 30 s, while monitoring the resulting fluorescence. The specificity of the PCR amplification was checked by the heat dissociation curve (from 65 to 95°C) after following the final cycle of the PCR. Calibration curves were drawn using sequential dilutions of the pGEM-T: *Ha*Fd1 and pGEM-T:*Ha*Fd2 plasmids and they were used to estimate the transcript content of each sample.

Results

Isolation and sequence analysis of sunflower ferredoxin cDNAs

Conserved regions from known photosynthetic ferredoxin sequences were used to design the oligonucleotide primer FdphoF1, with a 3,072-fold degree of degeneration. Using this primer and FA2Z, a primer complementary to the sequences incorporated during the cDNA synthesis, 330 bp fragments were amplified from sunflower leaf cDNA by PCR, corresponding to fragments of ferredoxin mRNAs. Subsequently, the full-length *HaFD1* cDNA clone of 429 bp was obtained by RACE using the primers shown in "Materials and methods" and Fig. 1. This PCR fragment was cloned and sequenced, and its identity as a photosynthetic ferredoxin was confirmed using the Blast software (Altschul et al. 1990). The full-length cDNA was predicted to generate a preprotein of 142 amino acids (Fig. 2), with a molecular mass of 15.17 kDa and a pI of 4.15.

In a similar way, conserved regions from known heterotrophic ferredoxin sequences were used to design the oligonucleotide primer FdhetF1, with a 1,536-fold degree of degeneration. Using this primer and FA2Z as above, we amplified 330 bp fragments from developing sunflower seed cDNA by PCR, corresponding to fragments of ferredoxin mRNA. Subsequently, we obtained the full-length *Ha*Fd2 cDNA clone of 474 bp by RACE using the primers shown as earlier, in "Materials and methods" and Fig. 1. This PCR fragment was cloned and sequenced, and its identity as a non-photosynthetic or heterotrophic ferredoxin confirmed using the Blast software (Altschul et al. 1990). The full-length cDNA was predicted to generate a preprotein of 157 amino acids (Fig. 2), with a molecular mass of 16.92 kDa and a pI of 4.95.

Through alignment with known ferredoxin sequences and using the network-based method TargetP V1.0 to identify chloroplast transit peptides (Emanuelsson et al. 2000), Ala46 of *Ha*FD1 and Ala62 of *Ha*FD2 sequences were the best candidates to be the N-terminal amino acid of the mature proteins (Fig. 2).

Using our data and other known ferredoxin sequences we generated a phylogenetic tree (Fig. 3), which demonstrated that the protein encoded by *Ha*Fd1 is highly

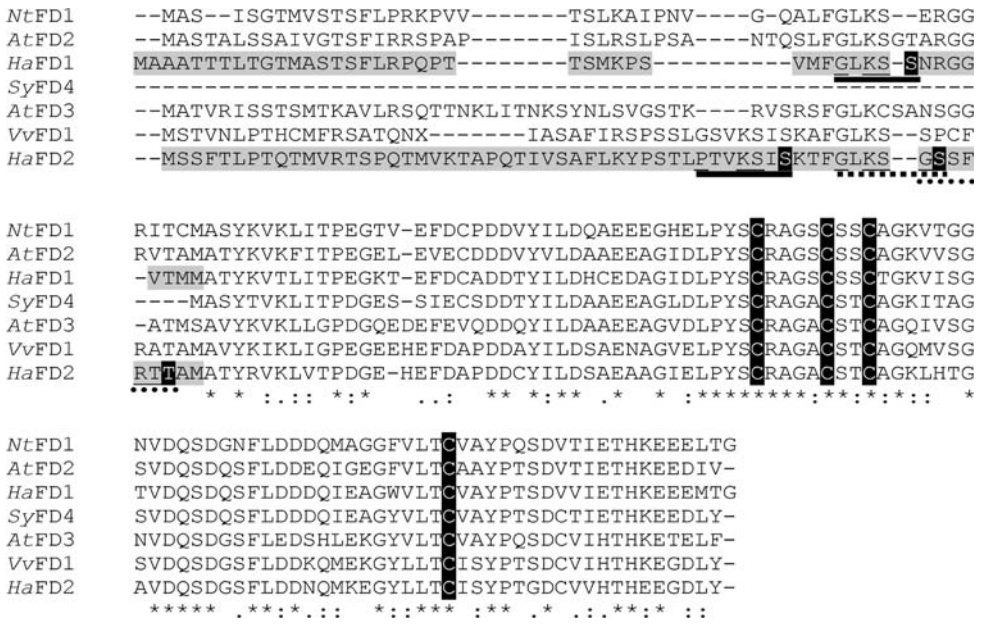

Fig. 2 Alignment of the deduced amino acid sequences of sunflower photosynthetic *Ha*FD1 and heterotrophic *Ha*FD2 ferredoxins, with the closely related sequences from *Arabidopsis thaliana* (*At*FD2, gi15219837; *At*FD3, gi15225888), *Nicotiana tabacum* (*Nt*FD1, gi45357074), *Vitis vinifera* (*Vv*FD1, gi147819070) and *Synechocystis* sp. PCC 6714 (SyFD4, gi118573556). The putative transit peptides are shaded in *gray*, the predicted phosphorylation sites in the putative signal peptide are *underlined*, and the four conserved cysteine residues from the plant ferredoxin cluster-binding motif ($CX_4CX_2CX_{29}C$) are shaded in *black*. *Asterisks* identical residues, *colon* conservative changes, and *dot* weakly conservative changes between the sequences

homologous to photosynthetic ferredoxins, and more specifically with those of the Solanaceae family like *Capsicum annuum* (Dayakar et al. 2003/gi34921349) and *Nicotiana tabacum* (Ham et al. 2005/gi78192120), both showing 65% identity to *Ha*Fd1. In contrast, the protein encoded by *Ha*Fd2 is highly homologous to heterotrophic ferredoxins, and more specifically with those from *Vitis vinifera* (Velasco et al. 2006/gi147819070) with 67% identity and those from *Citrus sinensis* (Alonso et al. 1995/gi1360725) with 63% identity. Significantly, *Ha*FD1 was clearly different to *Ha*FD2, showing 49% identity, confirming the presence of two types of ferredoxins in sunflower.

Tissue expression of the sunflower ferredoxins *Ha*Fd1 and *Ha*Fd2

While the expression of photosynthetic ferredoxin is closely related to active photosynthesis in green tissues of all plant species studied to date, heterotrophic ferredoxins have mainly been described in roots (Kimata and Hase 1989; Hase et al. 1991a; Aoki and Wada 1996). By RT-PCR, we analyzed the expression of *Ha*Fd1 and *Ha*Fd2 in roots, stems, cotyledons, and leaves from seedlings (15 days after sowing, DAS), as well as in developing seeds of RHA274 (Fig. 4a). As a result, we found the highest levels of *Ha*Fd1 expression in leaves, strong expression in other green tissues like the cotyledons and stem, and the weakest expression in the developing seeds. In contrast, *Ha*Fd2 showed more constant levels of expression in all the tissues assayed except in roots where this gene was most strongly expressed. When comparing the expression of both genes, *Ha*Fd1 was predominant in green tissues, *Ha*Fd2 in developing seeds and they were expressed at similar levels in roots. These results contrast with the ones observed in *Arabidopsis* seeds (Fig. 4b) where the expression levels of the photosynthetic ferredoxins are maintained. This is probably due to the fact that *Arabidopsis* seeds are photosynthetically actives.

Genomic sequences of *Ha*Fd1 and *Ha*Fd2

The genomic sequence of sunflower *Ha*Fd1 was isolated by PCR using the primer combination Fd1-F1/Fd1-R2 and the PCR product was sequenced with the same primers. The nucleotide analysis indicated that the gene has no intron, with a coding region 429 bp long. A comparison with sequences obtained from different sunflower lines allowed us to define two haplotypes that differ at nucleotide positions 225 (T–G transversion) and 300 (T–C transition, Fig. 5a). The deduced amino acidic sequence of the protein (142 aa) indicated that the sole nucleotide polymorphism (SNP) observed at np225 produced an Asp to Glu substitution in the *Ha*FD1 protein (Fig. 5b).

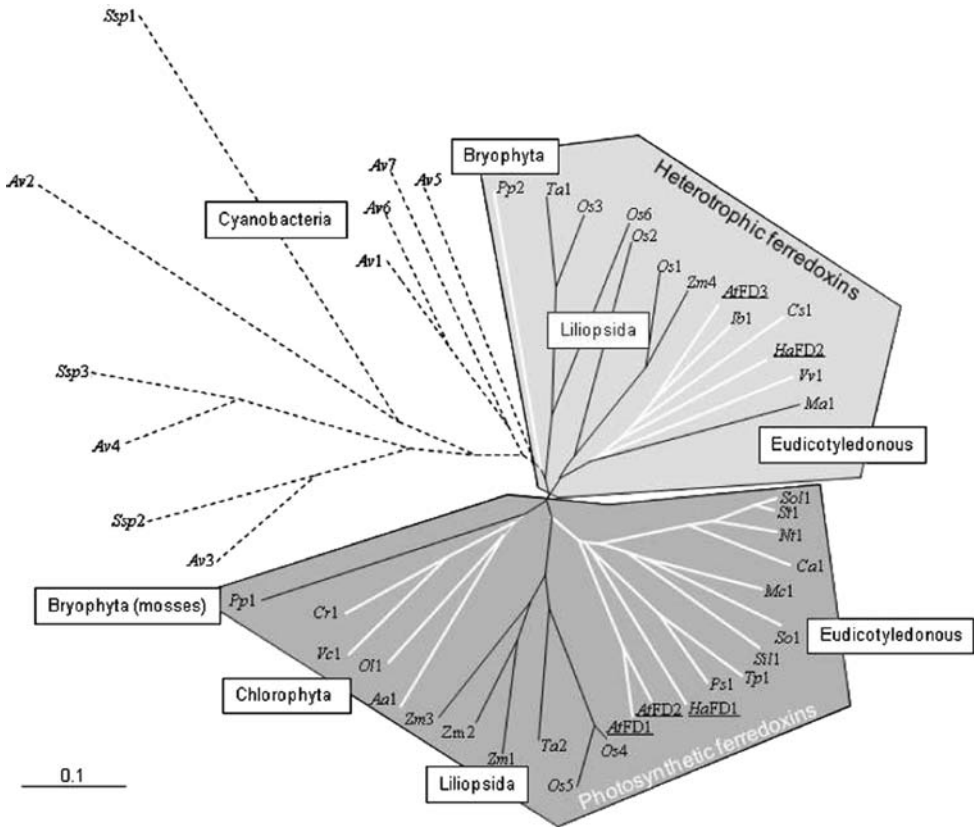

Fig. 3 Unrooted phylogenetic tree of plant photosynthetic and heterotrophic ferredoxin proteins and their homologues in cyanobacteria. Photosynthetic and heterotrophic ferredoxins are shown in *boxes* with a *dark* or *light gray* background, respectively. *Cyanobacteria branches* are shown as *dashed lines*, and the *Helianthus* and *Arabidopsis* sequences are *underlined*. Accession numbers of the different ferredoxins included are as follows: Cyanobacteria: *Anabaena variabilis* (*Av*1, gi157831119; *Av*2, gi75906422; *Av*3, gi75906791; *Av*4, gi75907204; *Av*5, gi75907755, *Av*6, gi75910144; *Av*7, gi75910456) and *Synechocystis* sp. PCC 6803 (*Ssp*1, gi16329789; *Ssp*2, gi16330020; *Ssp*3, gi16330840); Bryophyta: *Physcomitrella patens* (*Pp*1, gi17366013; *Pp*2, gi168017501); Chlorophyta: *Acetabularia acetabulum* (*Aa*1, gi58613453), *Chlamydomonas reinhardtii* (*Cr*1, gi462079), *Ostreococcus lucimarinus* (*Ol*1, gi145355410) and *Volvox carteri* f. *nagariensis* (*Vc*1, gi121077583); Liliopsida: *Musa acuminata* (*Ma*1, gi109390460), *Oryza sativa* (*Os*1, gi115456441; *Os*2, gi115458276; *Os*3, gi115464151; *Os*4, gi18698985; *Os*5, gi2305115; *Os*6 gi56784805), *Triticum aestivum* (*Ta*1, gi19569591; *Ta*2, gi462081) and *Zea mays* (*Zm*1, gi119928; *Zm*2, gi119961; *Zm*3, gi3417455; *Zm*4, gi119958); Eudicotyledons: *Arabidopsis thaliana* (*At*FD1, gi15220256; *At*FD2, gi15219837; *At*FD3, gi15225888), *Capsicum annuum* (*Ca*1, gi34921349), *Citrus sinensis* (*Cs*1, gi1360725), *H. annuus* (*Ha*Fd1, gi37779195; *Ha*Fd2, gi68137465), *Impatiens balsamina* (*Ib*1, gi13182955), *Mesembryanthemum crystallinum* (*Mc*1, gi3023743), *Nicotiana tabacum* (*Nt*1, gi45357074), *Pisum sativum* (*Ps*1, gi119931), *Silene latifolia* (*Sil*1, gi120026), *Solanum lycopersicum* (*Sol*1, gi3023752), *Spinacia oleracea* (*So*1, gi119937), *Solanum tuberosum* (*St*1, gi14041724), *Trifolium pratense* (*Tp*1, gi33520415), and *Vitis vinifera* (*Vv*1, gi147819070)

To obtain the genomic sequence of *Ha*Fd2, two overlapping primer combinations were used for PCR amplification: Fd2-F_1/Fd2-R_1 and Fd2-F_2/Fd2-R_2. In both cases, the size of the PCR products was larger than that expected from the cDNA sequence, indicating the presence of NTR (Fig. 1). The first combination produced a fragment of about 2.2 Kb, while the second produced two fragments of 320 and 990 bp indicating that one of the priming sites was duplicated. All PCR fragments were cloned and fully sequenced. The 2.2 Kb fragment (corresponding to the 5′-end of the gene) was initially sequenced with the universal primers in order to obtain the partial sequence of the insert. Based on the sequence obtained, internal primers Fd2-F_4 and Fd2-R_3 were designed to complete the sequencing of the insert. Overlapping all of these sequences enabled the complete genomic sequence of the *Ha*Fd2 gene to be reconstructed, demonstrating that the gene has a 5′-NTR of about 1.8 Kb with no introns in the coding region (Fig. 1). The *Ha*Fd2 gene was sequenced in a set of sunflower lines and aligned in order to identify DNA polymorphisms. No nucleotide differences were found among the lines analyzed when the coding region and the 3′-end of the gene were compared. However, when the 5′-NTR was analyzed two INDELs and several SNPs were observed that defined two distinct haplotypes (Fig. 6a), and there were repeated sequences throughout the gene. In the 5′-NTR, there were two repeat sequences named A and B, and while there were two copies of the

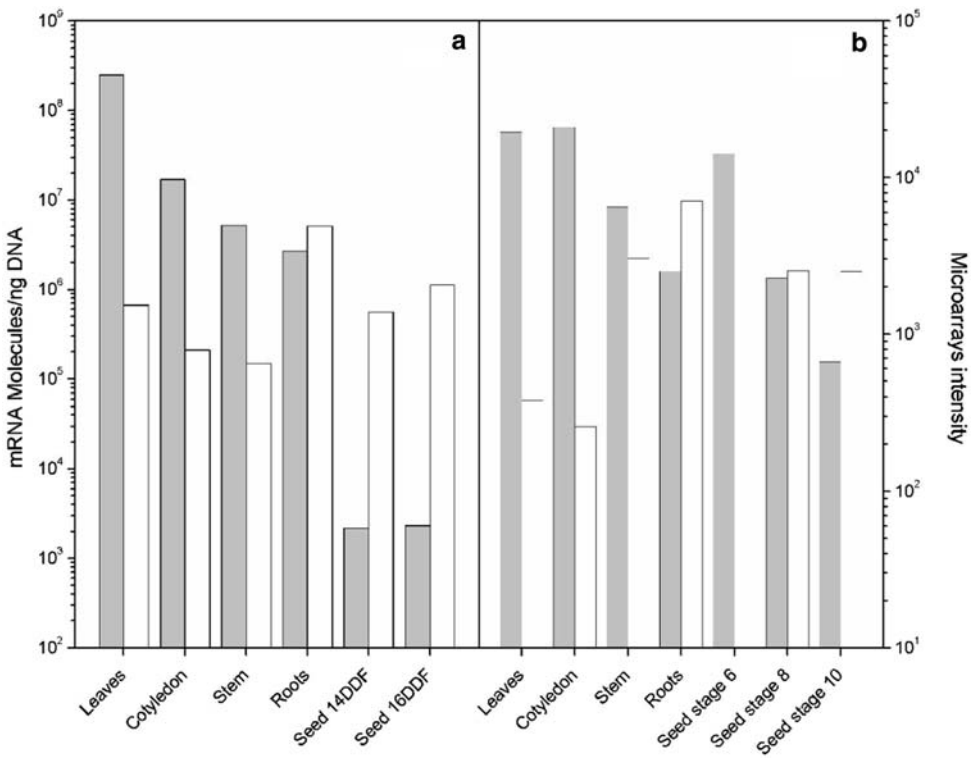

Fig. 4 Expression of photosynthetic (*gray columns*) and heterotrophic (*white columns*) ferredoxins in *Helianthus annuus* (**a**) and *Arabidopsis thaliana* (**b**, data estimated from the results of Schmid et al. 2005). **a** *Ha*Fd1 and *Ha*Fd2 expression determined by real time-PCR, **b** *Arabidopsis* photosynthetic (sum of *At*FD1 and *At*FD2 levels) and heterotrophic (*At*FD3, At2g27510) ferredoxin levels were estimated from microarrays. Values are the means from the mRNA samples from three plants

Fig. 5 Alignment of the genomic DNA (**a**) and the deduced amino acid (**b**) sequences of the *Ha*Fd1 gene. A comparison of the genomic DNA sequences identified two haplotypes that differ at nucleotide positions 192 and 267, and a D–E substitution when the deduced amino acidic sequences were aligned

first (A1–2), four copies of the second repeat were detected (B1–4). At the 3′-end there were three repeated sequences, each present in two copies (C1–2, D1–2 and E1–2, Fig. 6a, b). In addition, one of the INDELs (INDEL 1) was located in the A1 repeat, producing an insertion of eight nucleotides in haplotype 2 (Fig. 6b).

Expression of the sunflower *Ha*Fd1 and *Ha*Fd2 ferredoxins in lines with different *Ha*Fd2 haplotyes

To determine the possible phenotypic effect of the differences found between the *Ha*Fd2 haplotypes, the levels of Fd gene expression were determined by Q-PCR in developing sunflower seeds from lines with a similar background but which differ in the *Ha*Fd2 haplotype (CAS3, Fig. 7). While the *Ha*Fd1 levels were very similar in both lines, different patterns of *Ha*Fd2 expression were observed during seed formation. In addition, if we compare *Ha*Fd2 expression levels from these lines and those obtained with RHA274 (Fig. 4), the differences found were mostly due to the genetic backgrounds.

Chromosome mapping of *Ha*Fd1 and *Ha*Fd2

The sequence polymorphism found for both genes allowed specific co-dominant DNA markers to be generated for each of these (Fig. 8). The *Ha*Fd1 marker was based on the SNP at nucleotide 192, generating an ASA marker capable of identifying the T or G alleles. For *Ha*Fd2, a specific molecular marker was developed based on INDEL 1 that

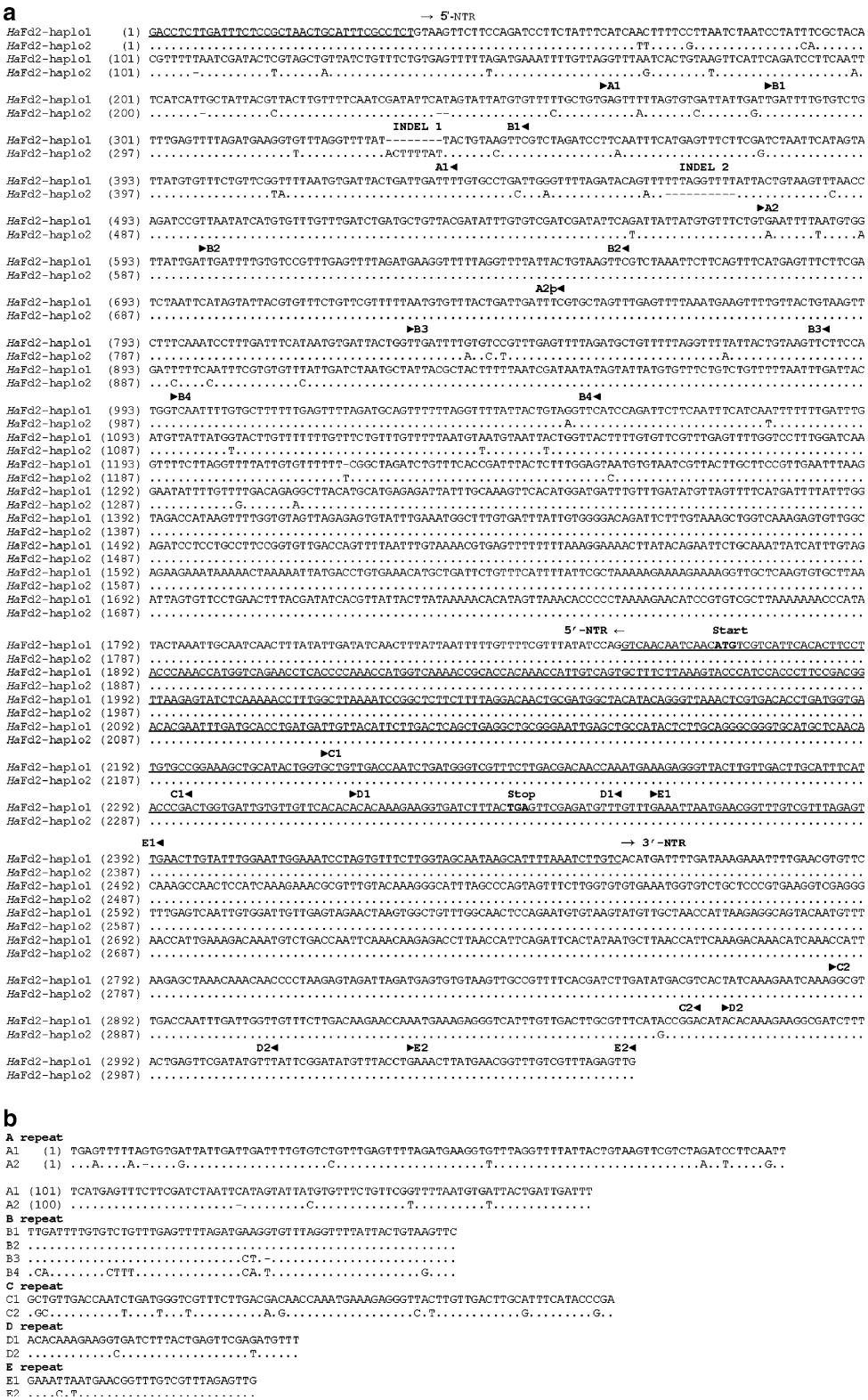

Fig. 6 **a** Alignment of the complete sequence of the two haplotypes of the sunflower *Ha*Fd2 gene. The haplotypes differ at several SNPs and two INDELs are present in the 5′-NTR. The start and stop codons in the coding region are indicated and the transcribed regions are *underlined*. *Arrows* flank the 5′-NTR and 3′-NTR and *arrowheads* flank the different repeated motifs present in *Ha*Fd2 gene: A1–2, and B1–4 (in 5′-NTR) and C1–2, D1–2 and E1–2 (in 3′-NTR). **b** Alignment of the different internal repeated motifs found in the *Ha*Fd2 genomic sequence from haplotype 1

was found in the genomic sequence of the gene (Fig. 8). Thus, the *Ha*Fd1 and *Ha*Fd2 genes were genotyped and genetically mapped in the F2 sunflower population P504 × R232, since both these genes were polymorphic in the parental lines. Accordingly, the *Ha*Fd1 locus mapped to linkage group 10 of the public sunflower map, while *Ha*Fd2 locus mapped to linkage group 11 (Tang et al. 2002; Yu et al. 2003; Fig. 9).

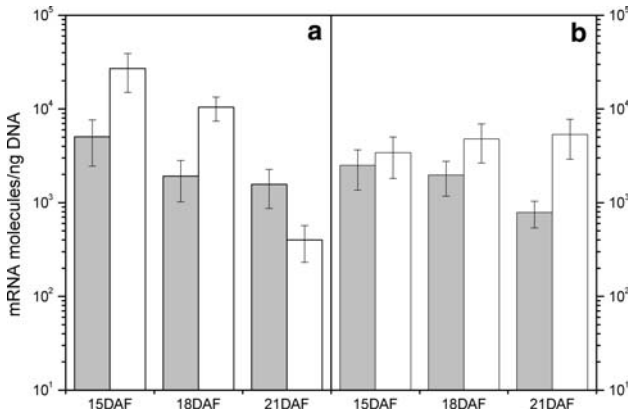

Fig. 7 Photosynthetic (*gray columns*) and heterotrophic (*white columns*) ferredoxin expression in developing *Helianthus annuus* seeds in *lines* that harbor the two different haplotypes for *Ha*Fd2: haplotype 1 (**a**) and haplotype 2 (**b**). Values are the means from the mRNA samples from three plants

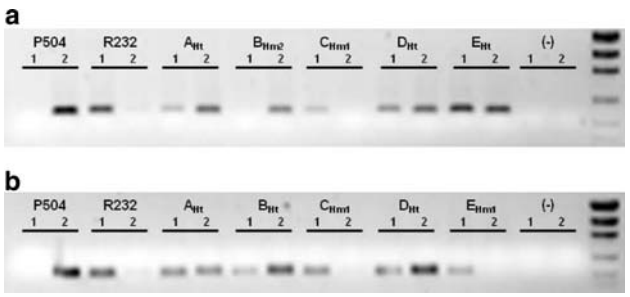

Fig. 8 Representative agarose gel electrophoresis showing the ASA molecular markers for the *Ha*Fd1 (**a**) and *Ha*Fd2 (**b**) genes. For each genotyped individual two independent PCR reactions were assayed: 1, specific amplification for haplotype 1; and 2, specific amplification for haplotype 2. P504 and R232 are the parental lines of the segregating F2 population used to map both genes, and A–E are the different F2 individuals. The genotype for each marker is indicated as a *subscript*: *Hm1* homozygous for haplotype 1, *Hm2* homozygous for haplotype 2; and *Ht* heterozygous

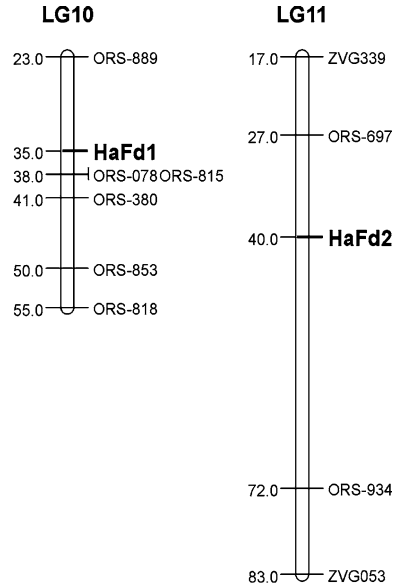

Fig. 9 Location of the *Ha*Fd1 and *Ha*Fd2 genes in the public map of the sunflower genome with respect to the simple sequence repeat (ORS) and INDEL (ZVG) marker loci in linkage groups 10 and 11, respectively

Discussion

Using the sequences available for photosynthetic and heterotrophic plant ferredoxins, we have identified two of the genes encoding these proteins in the leaves and developing seeds of sunflowers, *Ha*Fd1 and *Ha*Fd2.

The proteins deduced from the sequences of these genes contained signal peptides for plastid/chloroplast targeting, a feature common to all plant ferredoxins described to date. Such signals are present in most plastid/chloroplast enzymes with an endosymbiotic origin and in this case, the putative transit peptides also include consensus phosphorylation motifs [(P/G)X(n)(R/K)X(n)(S/T)X(n)(S*/T*), where $n = a$ 0–3 amino acid spacer and S*/T* represents the phosphate acceptor] as described by Waegemann and Soll (1996). While we found just one such motif in *Ha*Fd1, the *Ha*Fd2 sequence contains up to three such consensus motifs. In many plastid/chloroplast enzymes, phosphorylation of the transit peptide leads to the binding of the 14-3-3 protein, which can form a cytosolic guidance complex together with HSP70 (May and Soll 2000). Preproteins that are bound to the guidance complex are more rapidly imported into chloroplasts than monomeric preproteins, suggesting that sunflower ferredoxins are subject to fast regulatory responses. In the case of the photosynthetic *Ha*FD1, the proteolytic processing of the prepeptide would produce a 97-amino-acid protein with a predicted molecular mass of around 10.5 kDa and a predicted pI of 3.80. With regards the heterotrophic *Ha*FD2, a 96-amino-acid mature protein would be produced with a predicted molecular mass of 10.3 kDa and a predicted pI of 4.06. In both ferredoxins, the [2Fe–2S] cluster-binding cysteine residues of the plant ferredoxin binding motif ($CX_4CX_2CX_{29}C$) were completely conserved.

A phylogenetic tree was generated using plant ferredoxin preprotein sequences, including the signal peptide, together with cyanobacterial ferredoxin sequences from genus *Anabaena* and *Synechocystis*. In this phylogenetic tree, clear groups were evident that include the photosynthetic and heterotrophic ferredoxins, and the two sunflower ferredoxins cloned were attributed to the appropriate photosynthetic (*Ha*FD1) and heterotrophic groups (*Ha*FD2). Sequences in each of the ferredoxin types were clearly grouped in Bryophytas (mosses), Chlorophytas, Liliopsidas (monocotyledonous) and Eudicotyledonous. In addition,

the common origin of plant ferredoxins and those in cyanobacteria was evident.

While *Ha*Fd1 and *Ha*Fd2 are expressed in roots, stems, cotyledons, and leaves from 15 DAS RHA274 seedlings, as well as in developing seeds, *Ha*Fd1 is most prominent in green tissues and *Ha*Fd2 in developing seeds, while they are expressed at similar levels in roots. These results are similar to those from seedlings grown in agar (Schmid et al. 2005), with photosynthetic Fds predominant in green tissues and heterotrophic Fds most strongly expressed in roots and developing seeds. In contrast, the photosynthetic isoforms are expressed more intensely in younger *Arabidopsis* seeds, although it should be remembered that they are mostly composed of green tissues at these stages and when they start accumulating oil, the expression of the heterotrophic isoforms increases, as in sunflower seeds.

The complete sequencing of the *Ha*Fd2 gene in different lines indicated the presence of two haplotypes, and their alignment showed that all the sequence polymorphisms are restricted to the 5′-NTR. Apart from several SNPs, the two haplotypes can be distinguished by two INDELs. It is noteworthy that the repeat A1 in haplotype 2 is interrupted by an insertion of eight bases (INDEL 1). Although the physiological importance of the repeated sequences in maize ferredoxin (fdIII) is unclear, an mRNA transcript of about 4 kb was detected by northern-blot in some cases, which was much longer than the mature mRNA (Nakano et al. 1997). Post-transcriptional regulation, such as processing of a precursor may regulate the expression of fdIII. In the case of sunflower, so far no mRNA transcripts significantly longer than that predicted by the coding regions have been identified, suggesting that the 5′-NTR is not transcribed in any tissue or developmental stage. However, the INDEL 1 (present in the A1 repeat) could influence the transcriptional regulation of *Ha*Fd2 since such repeat regions may be involved in regulation. It is know that 5′-UTR introns (also known as 5′-NTR) may be related with intron-mediated enhancement (IME) of gene expression (Rose and Beliakoff 2000; Kim et al. 2006; Lu et al. 2008). To test this hypothesis the expression of the *Ha*Fd1 and *Ha*Fd2 ferredoxin genes was examined in seeds of lines with similar background but that differ in the *Ha*Fd2 allele. The results suggest that the 5′-NTR INDEL 1 may affect the transcriptional regulation of *Ha*Fd2, influencing the developmental expression of these transcripts and differentiating individuals with haplotype 1 and 2.

The *Ha*Fd1 and *Ha*Fd2 genes are located in linkage groups 10 and 11, respectively. Minor QTLs affecting stearic acid content in sunflower seeds were previously described that mapped to LG3, LG7, and LG13 (Pérez-Vich et al. 2004). Since the chromosomal location of the Fd genes analyzed here did not coincide with any of those minor QTLs, it seems unlikely that *Ha*Fd1 and *Ha*Fd2 will affect the stearic acid content in known high-stearic sunflower seeds. The identification of these two markers will allow the influence of ferredoxins in the phenotypes associated with metabolic pathways to be studied, as well as the processes in which they might be involved (e.g., photosynthesis).

Acknowledgments Our thanks are due to Marita Martino for her skillful technical assistance. This work was supported by the MEC and FEDER, project AGL2005-00100, and the seed company Advanta.

References

Alonso JM, Chamarro J, Granell A (1995) A non-photosynthetic ferredoxin gene is induced by ethylene in Citrus organs. Plant Mol Biol 29:1211–1221

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Aoki K, Wada K (1996) Temporal and spatial distribution of ferredoxin isoproteins in tomato fruit. Plant Physiol 112:651–657

Berry ST, Leon AJ, Peerbolte R, Challis P, Livini C, Jones R, Feingold S (1997) Presentation of the Advanta sunflower RFLP linkage map for public research. In: Proceedings of the 19th sunflower research workshop, Fargo, USA, pp 113–118

Browse J, Somerville C (1991) Glycerolipid synthesis: biochemistry and regulation. Annu Rev Plant Physiol Plant Mol Biol 42:467–506

Cantisán S, Martínez-Force E, Garcés R (2000) Enzymatic studies of high stearic acid sunflower seed mutants. Plant Physiol Biochem 38:377–382

Dayakar BV, Lin H-J, Chen C-H, Ger M-J, Lee B-H, Pai C-H, Chow D, Huang H-E, Hwang S-Y, Chung M-C, Feng T-Y (2003) Ferredoxin from sweet pepper (*Capsicum annuum* L.) intensifying harpin(pss)-mediated hypersensitive response shows an enhanced production of active oxygen species (AOS). Plant Mol Biol 51:913–924

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol 300:1005–1016

Green LS, Yee BC, Buchanan BB, Kamide K, Sanada Y, Wada K (1991) Ferredoxin and ferredoxin-NADP reductase from photosynthetic and nonphotosynthetic tissues of tomato. Plant Physiol 96:1207–1213

Ham B-K, Lee I-J, Kim MJ, Kim HR, Paek K-H (2005) Identification and characterization of ferredoxin for a putative interactor of Tsip1 (Tsi1-Interacting Protein 1) protein (Unpublished)

Hanke GT, Kimata-Ariga Y, Taniguchi I, Hase T (2004) A post genomic characterization of *Arabidopsis* ferredoxins. Plant Physiol 134:255–264

Hase T, Kimata Y, Matsumura T, Sakakibara H (1991a) Molecular cloning and differential expression of the maize ferredoxin family. Plant Physiol 96:77–83

Hase T, Mizutani S, Mukohata Y (1991b) Expression of maize ferredoxin cDNA in *Escherichia coli*: comparison of photosynthetic and nonphotosynthetic ferredoxin isoproteins and their chimeric molecule. Plant Physiol 97:1395–1401

Kamide K, Sakai H, Aoki K, Sanada Y, Wada K, Green LS, Yee BC, Buchanan BB (1995) Amino acid sequences of heterotrophic and photosynthetic ferredoxins from the tomato plant (*Lycopersicon esculentum* Mill). Photosynth Res 46:301–308

Kearns EV, Hugly S, Somerville CR (1991) The role of cytochrome b5 in Δ12 desaturation of oleic acid by microsomes of safflower (*Carthamus tinctorius* L.). Arch Biochem Biophys 284:431–436

Kim MJ, Kim H, Shin JS, Chung CH, Ohlrogge JB, Suh MC (2006) Seed-specific expression of sesame microsomal oleic acid desaturase is controlled by combinatorial properties between negative cis-regulatory elements in the SeFAD2 promoter and enhancers in the 5′-UTR intron. Mol Genet Genomics 276:351–368

Kimata Y, Hase T (1989) Localization of ferredoxin isoproteins in mesophyll and bundle sheath cells in maize leaf. Plant Physiol 89:1193–1197

Knaff DB (1996) Ferredoxin and ferredoxin dependent enzymes. In: Ort DR, Yocum CF (eds) Oxygenic photosynthesis: the light reactions. Kluwer, Dordrecht, pp 333–361

Lander E, Green P, Abrahanson J, Barlow A, Daley M, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Lenardon SL, Bazzalo ME, Abratti G, Cimmino CJ, Galella MT, Grondona M, Giollitti F, Leon AJ (2005) Screening sunflower for resistance to sunflower chlorotic mottle virus and mapping the Rcmo-1 resistance gene. Crop Sci 45:735–739

Lu J, Sivamani E, Azhakanandam K, Samadder P, Li X and Qu R (2008) Gene expression enhancement mediated by the 5′ UTR intron of the rice rubi3 gene varied remarkably among tissues in transgenic rice plants. Mol Genet Genomics. doi: 10.1007/s00438-008-0333-6

May T, Soll J (2000) 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. Plant Cell 12:53–63

McKeon TA, Stumpf PK (1982) Purification and characterization of the stearoyl-acyl carrier protein desaturase and the acyl-acyl carrier protein thioesterase from maturing seeds of safflower. J Biol Chem 257:12141–12147

Morigasaki S, Takata K, Sanada Y, Wada K, Yee BC, Shin S, Buchanan BB (1990) Novel forms of ferredoxin and ferredoxin-NADP reductase from spinach roots. Arch Biochem Biophys 283:75–80

Nakano R, Matsumura T, Sakakibara H, Sugiyama T, Hase T (1997) Cloning of maize ferredoxin III gene: presence of a unique repetitive nucleotide sequence within an intron found in the 5'-untranslated region. Plant Cell Physiol 38:1167–1170

Page RDM (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12:357–358

Pérez-Vich B, Fernández-Martínez JM, Grondona M, Knapp SJ, Berry ST (2002) Stearoyl-ACP and oleoyl-PC desaturase genes cosegregate with quantitative trait loci underlying high stearic and high oleic acid mutant phenotypes in sunflower. Theor Appl Genet 104:338–349

Pérez-Vich B, Knapp SJ, Leon AJ, Fernández-Martínez JM, Berry ST (2004) Mapping minor QTL for increased stearic acid content in sunflower seed oil. Mol Breed 13:313–322

Rose AB, Beliakoff JA (2000) Intron-mediated enhancement of gene expression independent of unique intron sequences and splicing. Plant Physiol 22:535–542

Saitou N, Nei M (1987) The neighbour-joining method—a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Schölkopf B, Weigel D, Lohmann J (2005) A gene expression map of *Arabidopsis* development. Nat Genet 37:501–506

Schmidt H, Heinz E (1990) Involvement of ferredoxin in desaturation of lipid-bound oleate in chloroplasts. Plant Physiol 94:214–220

Schultz DJ, Suh MC, Ohlrogge JB (2000) Stearoyl-acyl carrier protein and unusual acyl-acyl carrier protein desaturase activities are differentially influenced by ferredoxin. Plant Physiol 124:681–692

Smith MA, Cross AR, Jones TG, Griffiths WT, Stymne S, Stobart K (1990) Electron-transport components of the 1-acyl–2-oleoyl-sn-glycero-3-phosphocholine Δ12-desaturase (Δ12-desaturase) in microsomal preparations from developing safflower (*Carthamus tinctorius* L.) cotyledons. Biochem J 272:23–29

Tang S, Yu J-K, Slabaugh M, Shintani D, Knapp S (2002) Simple sequence repeat map of the sunflower genome. Theor Appl Genet 105:1124–1136

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882

Velasco R, Zharkikh A, Troggio M, Bhatnagar S, Pindo M, Cartwright D, Coppola G, Eldredge G, Vezzulli S, Malacarne G, Mitchell J, Stefanini M, Segala C, Gutin N, Grando MS, Pruss D, Dematte' L, Cestaro A, Toppo S, Fontana P, Skolnick M, Gutin A, Salamini F, Viola R (2006) The first genome sequence of an elite grapevine cultivar (Pinot noir *Vitis vinifera* L.): coping with a highly heterozygous genome (Unpublished)

Wada K, Onda M, Matsubara H (1989) Amino-acid sequences of ferredoxin isoproteins from radish roots. J Biochem 105:619–625

Waegemann K, Soll J (1996) Phosphorylation of the transit sequence of chloroplast precursor proteins. J Biol Chem 271:6545–6554

Yu J-K, Tang S, Slabaugh MB, Heesacker A, Cole G, Herring M, Soper J, Han F, Chu W-C, Webb DM, Thompson L, Edwards KJ, Berry S, Leon AJ, Grondona M, Olungu C, Maes N, Knapp SJ (2003) Towards a saturated molecular genetic linkage map for cultivated sunflower. Crop Sci 43:367–387