Mol Biol Rep (2012) 39:3683–3693
DOI 10.1007/s11033-011-1143-1

Isolation, identification and sequence analysis of a thioredoxin *h* gene, a member of subgroup III of *h*-type Trxs from grape (*Vitis vinifera* L. cv. Askari)

Reza Heidari Japelaghi · Raheem Haddad · Ghasem-Ali Garoosi

Received: 1 March 2011 / Accepted: 24 June 2011 / Published online: 6 July 2011

Abstract Thioredoxins (Trxs) are small ubiquitous proteins which play a regulatory role in a variety of cellular processes. In contrast to other organisms, plants have a great number of Trx types, consisting of six well-defined groups: *f*, *m*, *x*, and *y* in chloroplasts, *o* in mitochondria, and *h* mainly in cytosol. A full-length cDNA, designated Vv*CxxS*2, encoding Trx *h* polypeptide was isolated and cloned from grape (*Vitis vinifera* L. cv. Askari) berries organ by reverse transcription polymerase chain reaction (RT-PCR). The cDNA was 381 bp nucleotides in length with a deduced amino acid of 126 residues, possessing a WCIPS active site, which belongs to the subgroup III of *h*-type Trxs based on phylogenetic analysis. The calculated molecular mass and the predicted isoelectric point of the deduced polypeptide are 14.25 kDa and 4.68, respectively. Nucleotide sequence analysis of genomic DNA fragment of Vv*CxxS*2 gene revealed that this gene possesses two introns at positions identical to the previously sequenced Trx *h* genes. A modeling analysis indicated that Vv*CxxS*2 shares a common structure with other Trxs, and is preferably reduced by Grx rather than NADPH-dependent thioredoxin reductase (NTR). The deduced protein sequence showed a high similarity to Trx *h* from other plants, in particular from castor bean (*Ricinus communis*), *Betula pendula* and sweet orange (*Citrus sinensis*). Semiquantitative RT-PCR experiments indicated that the transcripts of Vv*CxxS*2 gene are present in all plant organs and different developmental stages. In addition, the higher expression of the Vv*CxxS*2 gene was observed in berry organ as compared to the other organs.

Keywords Cloning · Disulfide bond · Grape · Intron position · Thioredoxin

Abbreviations

dpa	Days post anthesis
EST	Expressed sequence tag
FTR	Ferredoxin-dependent thioredoxin reductase
Grx	Glutaredoxin
GSH	Glutathione
NTR	NADPH-dependent thioredoxin reductase
ORF	Open reading frame
RT-PCR	Reverse transcription polymerase chain reaction
Trx	Thioredoxin
UTR	Untranslated region

Introduction

Thioredoxins (Trxs) are small ubiquitous proteins with a catalytically active disulfide group which play a regulatory role in a number of cellular processes including metabolism, photorespiration, translation, protein degradation, vitamin biosynthesis [20], photosynthesis [3], transcription [6], DNA synthesis [3], pathogenic and oxidative stresses [26, 49] and seed germination [14]. They have been found in virtually all organisms and were first identified as one of the hydrogen donors for ribonucleotide reductase in *Escherichia coli* over 40 years ago [24]. The conserved

Electronic supplementary material The online version of this article (doi:10.1007/s11033-011-1143-1) contains supplementary material, which is available to authorized users.

R. H. Japelaghi · R. Haddad (✉) · G.-A. Garoosi
Department of Agricultural Biotechnology, Imam Khomeini International University, P.O. Box 34149-288, Qazvin, Islamic Republic of Iran
e-mail: raheemhaddad@yahoo.co.uk

active site sequence of Trxs forms a disulfide bond in the oxidized form of the protein. The reduced, dithiol form of Trx can modulate the activity of numerous target proteins by reduction of their disulfide bonds [3]. Whereas most organisms have low numbers of Trxs that achieve multiple functions, plants have a great number of Trx types [37, 40], for example at least 48 Trx and Trx-like genes have been detected in the fully sequenced genome of *Arabidopsis* (*Arabidopsis thaliana*) [37]. This outstanding diversity indicates either functional specialization or a high level of redundancy.

Higher plants contain two Trx systems; one is ferredoxin-dependent whereas the other one is NADPH-dependent. The ferredoxin-dependent system, located in the chloroplast, is composed of nuclear-encoded Trxs, *f*, *m*, *x*, and *y*. These Trxs are reduced by ferredoxin via ferredoxin-dependent thioredoxin reductase (FTR) [9]. The NADPH-dependent system is composed of Trx *o* (with mitochondrial localization) and Trx *h* (with cytosolic localization) which are reduced by NADPH via a NADPH-dependent thioredoxin reductase (NTR) [41]. Cell sorting prediction programs suggest that Trxs *h* are cytosolic. This point is in disagreement with previous results showing the purification of Trx *h* from mitochondria [19], plasma membrane [34] and even the nucleus [49]. Thus, subcellular localization of Trxs *h* needs to be investigated. In addition, Ishiwatari et al. [23] described Trx *h* as being one of the major proteins in sieve tubes of rice (*Oryza satia* L. var Kantou). Schobert et al. [47] identified Trxs in the sieve tubes of several monocotyledonous and dicotyledonous plants where they are proposed to maintain sieve tube proteins and provide protection from oxidative damage. Recently, novel types of Trxs, called *s* and *z*, have been identified in *Medicago truncatula* and *Arabidopsis*, respectively. Trx *s* with atypical putative catalytic site and endoplasmic reticulum localization does not belong to any of the types previously described [1], whereas, Trx *z* contains the active site signature typical of Trxs and involved in plastid gene expression [48].

The Trxs *h* share the presence of the conserved catalytic site WC[G/P]PC with other groups of Trxs, except the isoforms labeled CxxS. The catalytic site WCGPC is common to the majority of Trxs *h*, to the mitochondrial Trxs *o* and to the chloroplastic isotypes, whereas the WCPPC active site is only present in some isoforms of subgroup I [18]. They can be divided in three different subgroups I, II and III. Members of subgroup I and II are reduced by NTR, while reduction of Trxs *h* subgroup III is dependent on the GSH (glutathione)/Grx (glutaredoxin) system [16].

In this study, we report the isolation and cloning of a full-length cDNA encoding a Trx *h*, designated Vv*CxxS*2, from grape (*Vitis vinifera* L.) berry organ of an Iranian cultivar, called Askari, and also reveal that Vv*CxxS*2 genomic sequence possesses two introns at positions identical to the previously sequenced Trx *h* genes of other organs. By constructing structural models, we found that Vv*CxxS*2 may be reduced by Grx rather than NTR as shown for its ortholog from *Arabidopsis*. Finally, using a phylogenetic analysis with Trxs from other plants, we show that this gene is *h*-type Trx, belonging to the subgroup III and expresses in all grape organs.

Materials and methods

Plant materials

Berries, leaves, petioles, clusters, stems, roots and seeds from grape (*V. vinifera* L. cv. Askari) were collected from plants grown in the grape collection of the Grape Research Station, Takistan-Qazvin, Iran, during the 2008 field season. Berries at six ripening stages, and leaves, petioles and clusters at three stages of development were harvested. All samples were immediately frozen in liquid nitrogen upon collection and then stored at −80°C. Seeds were removed from the berries at veraison stage by gently breaking open the berries in liquid nitrogen, then pericarp and seed portions were stored separately until use.

DNA and total RNA extraction

Total RNA was extracted from various grape organs at different developmental stages, whereas genomic DNA was isolated from young leaf organ as described by Heidari Japelaghi et al. [22].

Reverse transcription polymerase chain reaction (RT-PCR), cloning and DNA sequencing

For first strand cDNA synthesis, 5 μg of total RNA treated with DNase I (Fermentas) was used as a template using Oligo $(dT)_{18}$ primer (1 μg/μl, Qiagen) for 5 min at 70°C. Then reaction mixture was incubated with RevertAid™ M-MuLV Reverse Transcriptase (200 u/μl, Fermentas) for 60 min at 42°C. The reaction was stopped by heating the mixture at 70°C for 10 min. The oligonucleotide primers used in this study were designed based on the available expressed sequence tag (EST EE089310), identified with the BLAST program (http://www.ncbi.nlm.nih.gov) for the amplification of the Vv*CxxS*2 gene by RT-PCR. The upstream oligonucleotides were synthesized homologous to the coding strand and the downstream oligonucleotides were complementary to the coding strand. The oligonucleotide primers included an addition of three nucleotides and of the *Bam*HI restriction site at the 5′-ends (Table 1).

Table 1 Nucleotide sequence of oligonucleotide primers used for RT-PCR

Primer name	Sequence (5′–3′)	Melting temperature (°C)	Size of amplicon (bp)
Primers used for DNA segment cloning			
VTrx3F1	tacggatccGGAGGAAGAGGAAGAGGAAA	78.3	519[b]
VTrx3R1	atcggatccGATTAGGAATTTAGGATTTAAATC	75.7	
VTrx3F2	tacggatccATGGAAAATCAGGAGCCG	78.7	866[a]–381[b]
VTrx3R2	atcggatccCTAGGCTACATACACGCGAAA	78.2	
Primers used for sq. RT-PCR			
Sq-VTrx3F	ATGGAAAATCAGGAGCCG	62.6	381[b]
Sq-VTrx3R	CTAGGCTACATACACGCGAAA	61.7	
AtAct2F	GTTAGCAACTGGGATGATATGG	66.8	530[b]
AtAct2R	AGCACCAATCGTGATGACTTG	69.6	

The oligonucleotide primers used for cloning include an addition of three nucleotides (small letters) and of a *Bam*HI restriction site (underlined) at the 5′-ends

[a] The length of amplified DNA segments

[b] The length of amplified ORFs

The RT-PCR reaction was carried out in a thermal cycler (Techene, UK) programmed for 35 cycles; conditions for each cycle being denaturation at 94°C for 30 s, annealing at 58°C for 1 min, and extension at 72°C for 1 min. The final extension was carried out at 72°C for 5 min. Also, the PCR reaction was performed using grape genomic DNA as a template and two specific primers designed based on the sequencing cDNA of Vv*CxxS*2 gene (Table 1) under the same RT-PCR conditions except that the extension time was changed to 2 min.

The amplified products were separated on a 0.8% agarose gel, and then excised from the gel and purified using GF-1 PCR Clean-Up Kit (Vivantis). The purified PCR products were cloned into pUC19 cloning vector to generate the pVTRXh-3 plasmid and were transformed into the competent *E. coli* strain DH5α as described [45]. After screening, target DNA was sequenced by the dideoxynucleotide chain termination method in both directions (Sequence Laboratories Gottingen, Germany).

Sequence analysis

The obtained nucleotide sequence from sequencing was translated using Translate tool (http://www.expasy.ch/tools/dna.html) and the properties of deduced amino acid sequence were estimated using ProtParam (http://www.expasy.ch/tools/protparam.html) [13], ProtScale (http://arbl.cvmbs.colostate.edu/molkit/hydropathy/index.html) and TMHMM (http://www.cbs.dtu.dk/services/) programs. The subcellular localization prediction of Vv*CxxS*2 was performed using a combination of three programs, TargetP (http://www.cbs.dtu.dk/services/TargetP/) [10], iPSORT (http://ipsort.hgc.jp/) [5] and YLOC (http://www-bs.informatik.uni-tuebingen.de/Services/YLoc/) [7], and structural and functional important regions were identified in deduced protein sequence by Conseq (http://conseq.tau.ac.il/) [4] and PatchFinder (http://patchfinder.tau.ac.il/) [39] services. Also, secondary structure was determined by SOPMA (http://npsa-pbil.ibcp.fr) [21] and PSIpred (http://bioinf.cs.ucl.ac.uk/psipred/) programs.

Phylogenetic analysis

Protein sequences were retrieved from the GenBank through the BLASTp algorithm at the National Center for Biotechnology Information (NCBI) using Vv*CxxS*2 sequence as query and several Trx *h* with the highest score from different plants were selected. Sequences were aligned using the ClustalW2 at European Bioinformatics Institute (http://www.abi.ac.uk) using default parameters. Neighbor Joining tree was constructed with MEGA4.1 Beta 2 software [50]. Bootstrap analysis with 1,000 replicates was also conducted in order to obtain confidence levels for the branches [11].

Prediction of three-dimensional structure of Vv*CxxS*2

The three-dimensional structure of Vv*CxxS*2 was predicted using I-TASSER (http://zhang.bioinformatics.ku.edu/I-TASSER) [51] with the crystal structure of HvTrx *h*1 (PDB ID code 2vlvB) [33] as a template. Superimposition analysis of the 3D models of Vv*CxxS*2 and its templates, such as; human (*Homo sapiens*) Trx protein (PDB ID code 1ertA), *Chlamydomonas reinhardtii* Trx *h*1 (PDB ID code 1ep7A), *Saccharomyces cerevisiae* Trx1 protein (PDB ID code 3f3qA) and HvTrx *h*1from barley was done using 3-Dimensional Structural Superposition (3d-SS) service (http://cluster.physics.iisc.ernet.in/3dss/several input.html) [43]. Conserved amino acids at the protein

surface were determined using ConSurf (http://consurf.tau.ac.il/overview.html) [30].

Analysis of docking modeling of Vv*CxxS*2 using of PatchDock and FireDock services

Docking properties of Vv*CxxS*2 with NTR and Grx were analyzed using a two-step procedure of PatchDock and FireDock services. Interaction models of Vv*CxxS*2 and AtNTRB (PDB ID code 1vdc) and of Vv*CxxS*2 and PtGrxC1 (PDB ID code 1z7p) were predicted by the PatchDock web server (http://bioinfo3d.cs.tau.ac.il/PatchDock/) [46] with AtTrx *h*1 as control. Obtained results were redirected for refinement and scoring by the FireDock server (http://bioinfo3d.cs.tau.ac.il/FireDock/) [2].

Semiquantitative RT-PCR analysis

Semiquantitative RT-PCR was performed using 5 µg of DNase I-treated total RNA for first-strand synthesis of cDNA, as described previously. About 1/20 of the reverse transcription reaction was used for RT-PCR with specific primers for distinct Vv*CxxS*2 cDNA. Grape actin gene (At*Act2*) showing invariant expression across the samples was amplified in parallel. PCR amplifications were performed in a thermal cycler programmed with the following temperature parameters: 3 min at 94°C, followed by 30 cycles of 30 s at 94°C, 1 min at 58°C, 30 s at 72°C, and the final extension of 5 min at 72°C. A negative control lacking a template was included for each set of RT-PCR reactions. Reactions were performed in triplicates. Amplification products were separated by agarose gel electrophoresis and quantified using ImageJ software (W.S. Rasband; 1997–2007; National Institutes of Health; http://rsb.info.nih.gov/ij). Signal intensities were normalized with respect to At*Act2* gene from the same sample.

Accession numbers

The Vv*CxxS*2 gene was submitted in NCBI GenBank database under accession number HM370528. The NCBI, EMBL and SwissProt accession numbers for the sequences described and mentioned in this study are as follows: *Arabidopsis* (*A. thaliana*): AtTrx *h*1 (At3g51030, U35827); AtTrx *h*2 (At5g39950); AtTrx *h*3 (At5g42980); AtTrx *h*4 (At1g19730); AtTrx *h*5 (At1g45145); AtTrx *h*7 (At1g59730); AtTrx *h*8 (At1g69880); AtTrx *h*9 (At3g08710); At*CxxS*1 (At2g40790); At*CxxS*2 (At1g11530); At*CxxC*2 (At3g56420); AtTrx *o*1 (At2g35010); AtTrx *o*2 (At1g31020); AtTrx *m*1 (At1g03680); AtTrx *m*2 (At4g03520); AtTrx *m*3 (At2g15570); AtTrx *m*4 (At3g15360); AtTrx *f*1 (At3g02730); AtTrx *f*2 (At5g16400); AtTrx *x* (At1g50320); AtTrx *y*1 (At1g76760); AtTrx *y*2 (At1g43560); AtTrx *z* (At3g06730); *Betula pendula*: Bp*CxxS* (CD278293); castor oil (*Ricinus communis*): Rc*CxxS*1 (XP_002518569); Rc*CxxS*2 (XP_002522827); *C. reinhardtii*: Chlamy Trx *h*1 (X80887); *Codonopsis lanceolata*: Cl*CxxS* (BAE16559); *E. coli*: E. coli Trx1 (M54881); Grape (*V. vinifera*): VvTrx *h*1 (HM370524); VvTrx *h*2 (HM370525); VvTrx *h*3 (CF518184); VvTrx *h*4 (EE067811); VvTrx *h*5 (CB004453); Vv*CxxS*1 (CF513566); VvTrx *o* (EE085878); VvTrx *m*1 (EC944820); VvTrx *m*2 (BM436574); VvTrx *m*3 (EC979521);VvTrx *m*4 (EC 951344); VvTrx *f* (EE102290); VvTrx *x* (CB342548); VvTrx *y* (GO653033); VvTrx *z* (CB035759); *Hevea brasiliensis*: Hb*CxxS*1 (AAD33596); Hb*CxxS*2 (ACI00350); human (*H. sapiens*): Human Trx (P10599); *M. truncatula*: MtTrx *s*1 (DQ121444); MtTrx *s*2 (DQ121445); Mt*CxxS* (ACJ84725); poplar (*Populus trichocarpa*): Pt*CxxS*1 (CA823821); Pt*CxxS*3 (BU874060); rice (*Oryza sativa*): OsTrx *h*1 (D26547); Os*CxxS* (NP_001053968); *S. cerevisiae*: Yeast Trx2 (M59169); soybean (*Glycine max*): Gm*CxxS*1 (ACU14593); Gm*CxxS*2 (ACU14856); sweet orange (*Citrus sinensis*): Cs*CxxS* (BQ623126); tobacco (*Nicotiana tabacum*): NtTrx *h*2 (Z11803).

Results and discussion

Cloning and nucleotide sequence analysis of grape Trx *h* cDNA

Looking for Trx sequences in the grape EST database led to identify one EST sequence from the berry grape EST library (Pena-Cortes et al., unpublished results) corresponding to a putative full-length Trx *h* cDNA called Vv*CxxS*2 in this study. In order to characterize the full-length cDNA, total RNA isolation and cDNA synthesis were performed using oligonucleotides primers VTrx3F1 and VTrx3R1 (Table 1), designated based on the EST sequence. The PCR product was cloned, sequenced, and oligonucleotide primers VTrx3F2 and VTrx3R2 were synthesized to amplify a complete full-length open reading frame (ORF) and genomic DNA based on the obtained sequence. The full-length ORF and genomic DNA were amplified by PCR, using oligonucleotide primers VTrx3F2 and VTrx3R2 (Table 1). Finally, the PCR products were cloned and then sequenced. The Vv*CxxS*2 cDNA was 519 nucleotides long and contains a single ORF of 381 bp, a 5′-untranslated region (UTR) of 23 bp, and a 3′-UTR of 115 bp containing a putative polyadenylation signal (TATAAT) with one difference to the consensus sequence (A/GATAAA/T) for plant genes [35]. The ORF with 49%

G + C content, encodes a protein of 126 amino acid residues, beginning at the initiation codon ATG (position 24) and ending at the stop codon TAG (position 402) of the cDNA (Fig. S1).

Analysis of the deduced polypeptide sequence

The calculated molecular mass and the predicated isoelectric point of the deduced polypeptide sequence are 14.25 kDa and 4.68, respectively. The aliphatic index, regarded as a positive factor for increased thermostability [28], was calculated 86.59, in contrast to *C. reinhardtii* Trx *m* that was estimated 94.71 using ProtParam [13]. The *C. reinhardtii* Trx *m* is a thermostable Trx, due to the presence of a hydrophobic cluster of four aromatic amino acids consisting of Phe10, Phe25, Tyr68 and Phe79 [14]. In Vv*CxxS*2, the corresponding residues are Trp22, Phe40, Met82 and Met93, based on the alignment with Chlamydomonas Trx *h*1 and *E. coli* Trx1 sequences, indicating that the hydrophobic tetrahedral cluster is not conserved. By calculating the hydropathy value by the method of Kyte and Doolittle [29], it was determined that Vv*CxxS*2 was highly hydrophobic, which could be explained by Vv*CxxS*2 easily interacting and reducing target proteins (Fig. S1).

The deduced primary structure of grape Trx *h* exhibits an unusual active site WCIPS [Trp44-Cys45-Ile46-Pro47-Ser48] that is shown in the atypical Trxs CxxS of the third subgroup of Trxs *h*. Despite *E. coli*, human and some of the cereal sequences, all Trxs *h* characterized so far, consist a conserved Trp residue [18] that it locates in position 22 (W22) of the Vv*CxxS*2 sequence. Prediction of subcellular localization using combination of three programs, TargetP, iPSORT and YLOC suggest a cytosolic localization for Vv*CxxS*2. However, using TMHMM program Vv*CxxS*2 predicted to be an outcellular protein and may be capable of moving from cell-to-cell similar to rice OsTrx *h*1, but this potentially unique function has yet to be demonstrated.

The substrate-binding loop motif composed of hydrophobic and uncharged residues plays a central role in the recognition of substrate proteins [33]. This motif is conserved in the Trx *h* sequences and involves three loop segments also indicated in Vv*CxxS*2; Trp44-Cys45-Ile46-Pro47, Ala86-Met87-Pro88, and Val103-Gly104-Ala105. The conserved hydrophobic residues implicated in the *E. coli* NTR:Trx interaction (Trp31, Ile60, Gly74, Gly92, and Ala93) [31] are also found in Vv*CxxS*2 (Trp44, Val72, Ala86, Gly104, and Ala105). Positions of other amino acids (i.e. Asn3, Val13, Ala42, Trp44, Cys45, Pro47, Ser48, Phe67, Asp71, Val72, Asp73, Met87, Pro88, Thr89, Lys94, Gly96, Val103, Gly104, Ala105, and Ser120) are important for maintaining the three-dimensional structure and function in Vv*CxxS*2 based upon ConSeq and PatchFinder servers.

Position of the introns in grape Trx *h* genomic sequence and comparison with other Trxs

The Vv*CxxS*2 genomic sequence with 867 bp length contains three exons, with 107, 122 and 152 nucleotides length, and two introns with 271 and 214 bp, respectively. The exon–intron junctions obey the GT–AG rule [12] and, as in higher plants genes, the introns are relatively AT rich compared to the coding regions. The Vv*CxxS*2 genomic sequence together with other grape Trx *h* sequences [25] reveals introns at the same positions. These nucleotide sequences show the presence of two introns in VvTrx *h*1 [24I (0); 66S (0)], VvTrx *h*2 [24V (0); 66T (0)], VvTrx *h*3 [24M (0); 66D (0)], VvTrx *h*4 [24M (0); 66D (0)], Vv*CxxS*1 [24V (0); 66E (0)], and Vv*CxxS*2 [24V (2); 66A (1)]. In Vv*CxxS*2, Val and Ala residues are at positions 36 and 77, respectively, and the splicing sites are between 2–3 and 1–2 nucleotides [24V (GT/T); 66A (G/CG)], respectively. Whereas, in other grape Trx *h* sequences, the indicated residues are at different positions, but their splicing sites are similar with each other. The analysis of other plant Trx *h* sequences indicated that *Arabidopsis* AtTrx *h*1 [24V (0); 66S (0)] [44], tobacco NtTrx *h*2 [24I (0); 66S (0)] [8] and OsTrx *h*1 from rice [24V (0); 66E (0)] [23] contain introns at the same positions as the grape Trxs *h*. It is confirming that although Trxs *h* is having diverged in their amino acid sequences, they most probably share a common ancestor, having had two introns at positions 24 and 66 [36]. Figures 1 and S2 show the position and size of introns in Vv*CxxS*2 and the different Trx genes.

Most of the results reveal the conserved position and the diverse size of the two introns for higher plants Trxs *h*. In grape, VvTrx *h*1 and *h*2 genes (subgroup I) have far larger introns than Vv*CxxS*2 and other grape Trxs *h*. Also, despite Vv*CxxS*1, intron 1 in VvTrx *h*1, *h*2 and Vv*CxxS*2 is larger than intron 2, and VvTrx *h*3 and *h*4 contain two introns of similar sizes. The size of the introns is related to the species [44]: the size of *Arabidopsis* Trxs *h* introns is far shorter than those of grape (VvTrx *h*1 and *h*2), tobacco (NtTrx *h*2) and rice (OsTrx *h*1). In contrast to VvTrx *h*3, *h*4 and Vv*CxxS*1, VvTrx *h*1, *h*2 and Vv*CxxS*2 are similar to AtTrx *h*1, NtTrx *h*2 and OsTrx *h*1 that intron 1 is larger than intron 2. This suggests a coordinated reduction or extension of the different introns of a sequence during evolution [44].

The intron positions were also analyzed in *E. coli*, *S. cerevisiae*, *C. reinhardtii* and human Trx genes. In *E. coli* and *S. cerevisiae* Trx genes do not contain any introns, whereas the Trx genes of *C. reinhardtii* and human, harbor three and four introns, respectively. In Chlamydomonas Trx *h*1 gene, the first and third introns coincide with the two introns of Vv*CxxS*2 and higher plants Trxs *h*, whereas the second one is positioned in the middle of exon 2 before the active site of the Trx. Interestingly, the third

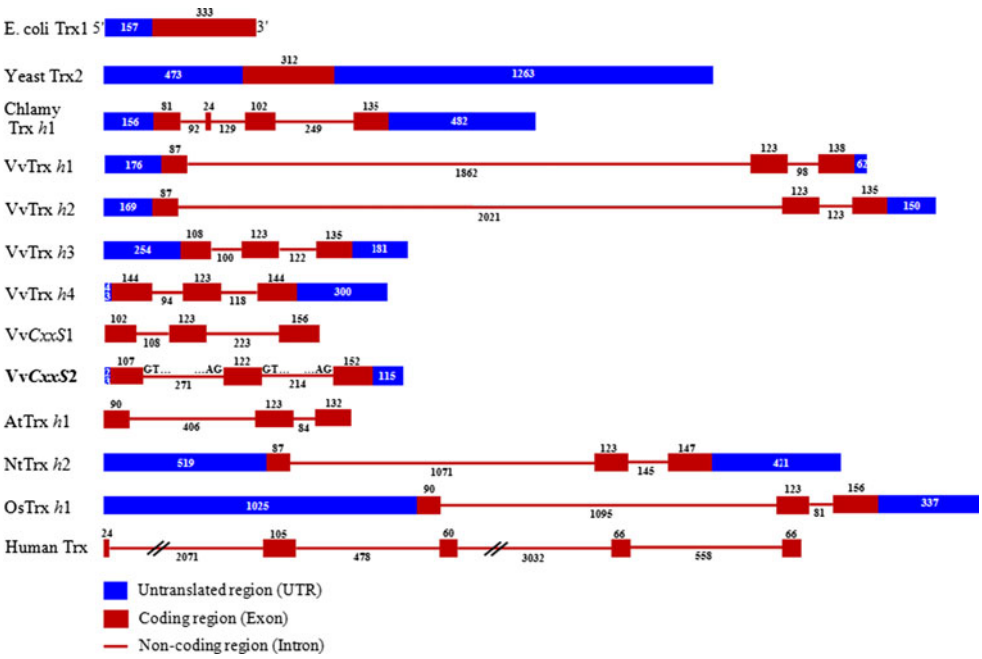

Fig. 1 Structure of Vv*CxxS*2 gene, grape Trxs *h* genes and genes encoding Trx *h* from plant and *C. reinhardtii* with human Trx, *S. cerevisiae* Trx2 and *E. coli* Trx1. The UTRs, coding regions and non-coding regions are indicated in *blue boxes*, *red boxes* and *red lines*, respectively. Accession numbers are given in the "Materials and methods" section. (Color figure online)

intron of human Trx gene is located at the same position as the second intron of plant Trxs *h*.

Analysis of 3D structure of grape Trx *h*

Secondary structure analysis of Vv*CxxS*2 protein by SOMPA and PSIpred programs revealed that Vv*CxxS*2 consists of 54 α-helixes (42.86%), 9 *β*-turns (7.14%) jointed by 26 extended strands (20.63%), and 37 random coils (29.37%) (Fig. S1). Using HvTrx *h*1 as a template for comparative modeling, a predicted 3D structure was determined for Vv*CxxS*2 by applying I-TASSER simulation. Thus, Vv*CxxS*2 similar to all Trxs [24], includes five *β*-sheets surrounded by four α-helices with secondary structural elements: *β*1α1*β*2α2*β*3α3*β*4*β*5α4, and the active site WCIPS is present at the N-terminal end of helix α2 (Fig. 2a). Vv*CxxS*2 3D structure matches almost perfectly with Trx crystal structures from human, Chlamydomonas Trx *h*1, *S. cerevisiae* Trx2 and HvTrx *h*1 (Fig. 2b). Analysis of the evolutionary conservation of its surface amino acids was performed using ConSurf program (Fig. 2c). Several residues with high scores were found to be functional and structural residues in the protein by using PachFinder (Fig. 2d) and ConSeq servers.

Grx fits for the reduction of Vv*CxxS*2 rather than NTR

Computational docking methods used to predict protein–protein interactions can help define parameters valuable to understanding biochemical mechanisms [34]. Vv*CxxS*2 seemed not to interact with NTR as revealed by computational docking analysis, using a two-step procedure of PatchDock and FireDock services. Moreover, predicted structure indicated that it was preferentially reduced by GSH/Grx system, such as At*CxxS*1 from *Arabidopsis* belonging to Trxs *h* subgroup III. NTR (AtNTRB) was too large to interact with Vv*CxxS*2 and produced a weak and unstable complex so that binding site was in place contrary to what was expected for NTR:Trx interaction (Fig. 3b, c). Whereas, Grx (PtGrxC1) was perfectly accommodating with Vv*CxxS*2, yielding a tight and stable complex (Fig. 3e, f). It seems that the conserved hydrophobic residues near the active site are implicated in Grx:Vv*CxxS*2 interaction in three dimensional conformations, i.e. Trp44, Val72, Ala86, Gly104, and Ala105 as described in the NTR:Trx interaction [31]. Predicted specificity of Vv*CxxS*2 was supported by AtTrx *h*1 used as a control, which is reduced by NTR, not Grx, and fits perfectly with the AtNTRB enzyme (Fig. 3a, d).

Comparison of grape Trx *h* with other Trxs

A phylogenetic tree was constructed using the Trxs described for *Arabidopsis* [20], grape (Heidari et al., unpublished) and novel type of Trx, called *s* from *M. truncatula* [1] (Fig. S3). The tree is divided into two major clusters. One cluster contains prokaryotic Trxs (*m*, *s*, *x*, *y*, and *z*) and the other contains Trxs *f*, *h*, and *o* that are specific to eukaryotic organisms. Trxs *h* are also classified into three different subgroups I, II, and III, based on the primary structure analysis [17]. The third subgroup of Trxs *h* comprises the atypical Trxs CxxS/C (At*CxxS*1, At*CxxS*2,

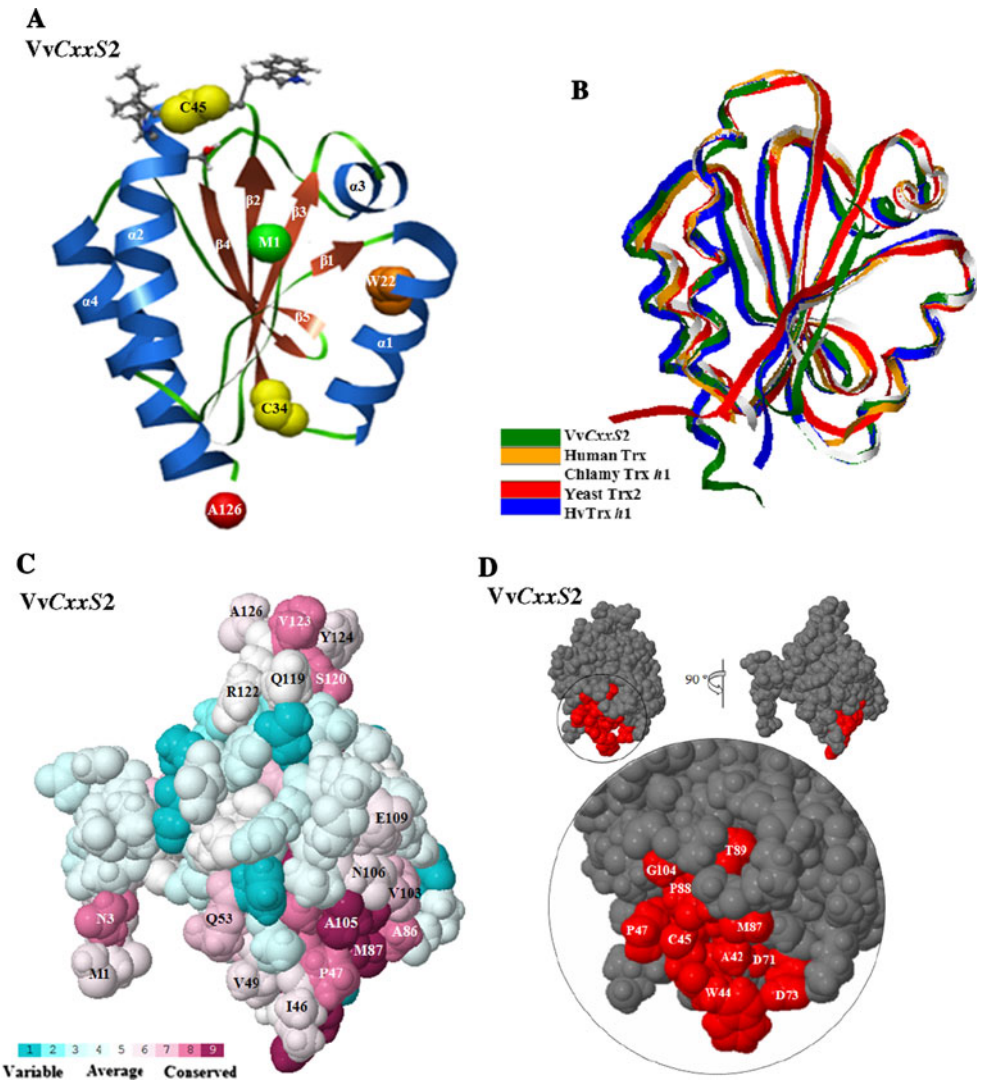

Fig. 2 Three-dimensional models and conserved residue prediction for Vv*CxxS*2. **a** Cartoon display of the three-dimensional structure of Vv*CxxS*2. The α-helices, β-sheets, and coiled coil regions are colored in *blue*, *red*, and *green*, respectively. The cysteine, characteristic of tryptophan, initiation methionine, and end alanine residues are indicated by *yellow*, *orange*, *green* and *red spheres*, respectively. **b** Superimposition of 3D model of Vv*CxxS*2 (*green*) and the top four templates of: human Trx protein (*orange*; PDB ID code 1ertA); *C. reinhardtii* Trx *h*1 (*white*; PDB ID code 1ep7A); *S. cerevisiae* Trx2 protein (*red*; PDB ID code 3f3qA); and HvTrx *h*1 protein from barley (*blue*; PDB ID code 2vlvB), using 3d-SS (3-Dimensional Structural Superposition) service. **c** Conserved residue analysis of Vv*CxxS*2. Residue conservation from variable to conserved is shown in *green* to *dark red*, respectively. **d** Identification of functional important residues using PatchFinder server. The identified residues are involved in the active site sequence and the substrate-binding loop motif are colored in *red*. (Color figure online)

At*CxxC*2, Vv*CxxS*1, and Vv*CxxS*2) and Trxs exhibiting the usual catalytic site WCGPC (AtTrx *h*9 and VvTrx *h*5) whose main feature is interactions with the GSH/Grx system [16]. Due to the absorbance of the second cysteinyl residue in the catalytic site [42] Vv*CxxS*2 could not be accounted as true Trxs compared to grape Vv*CxxS*1 and At*CxxS*1 from *Arabidopsis*. Interestingly, the atypical At*CxxS*2 and At*CxxC*2 cluster that are separate from At*CxxS*1, Vv*CxxS*1 and Vv*CxxS*2, are related to AtTrx *h*9 and VvTrx *h*5. Trxs *s* that may be specific to legumes do not belong to any of the types previously described. Nevertheless, on the basis of protein sequence and gene structure, they are both related to Trxs *m* and probably have evolved from Trxs *m* after the divergence of the higher plant families [1]. Also, Trxs *z* with typical putative catalytic site (WCGPC) and plastid localization, are distantly related to other known Trxs and clustered separately [48].

Homology analysis of Vv*CxxS*2 with Trxs *h*III from other plants revealed that the putative polypeptide sequence encoded by Vv*CxxS*2 shares a high degree of homology with castor oil (*R. communis*: Rc*CxxS*1; 75.4% identity and 89.7% similarity), *B. pendula* (Bp*CxxS*; 75.8% identity and 89.1% similarity), and sweet orange (*C. sinensis*: Cs*CxxS*; 75.0% identity and 88.3% similarity). In contrast, Vv*CxxS*2 shares lower degrees of identity to Trx of human, *C. reinhardtii*, *S. cervisiae* and *E. coli* with 44.6, 40.2, 46.6, and 28.1%, respectively. The overall sequence identity between Vv*CxxS*2 with other grape Trxs ranges from 20.9 to 72.9%. The Vv*CxxS*2 shows a high identity with Vv*CxxS*1 (72.9% identity and 83.6% similarity) belonging to Trxs *h*III. However, there is also a low identity of sequences between Vv*CxxS*2 and grape Trx *o*; Vv*CxxS*2 shows 20.9% identity and 50.4% similarity with VvTrx *o*. Interestingly, Vv*CxxS*2 with less than 73% identity with all grape Trxs, shares a high degree of identity

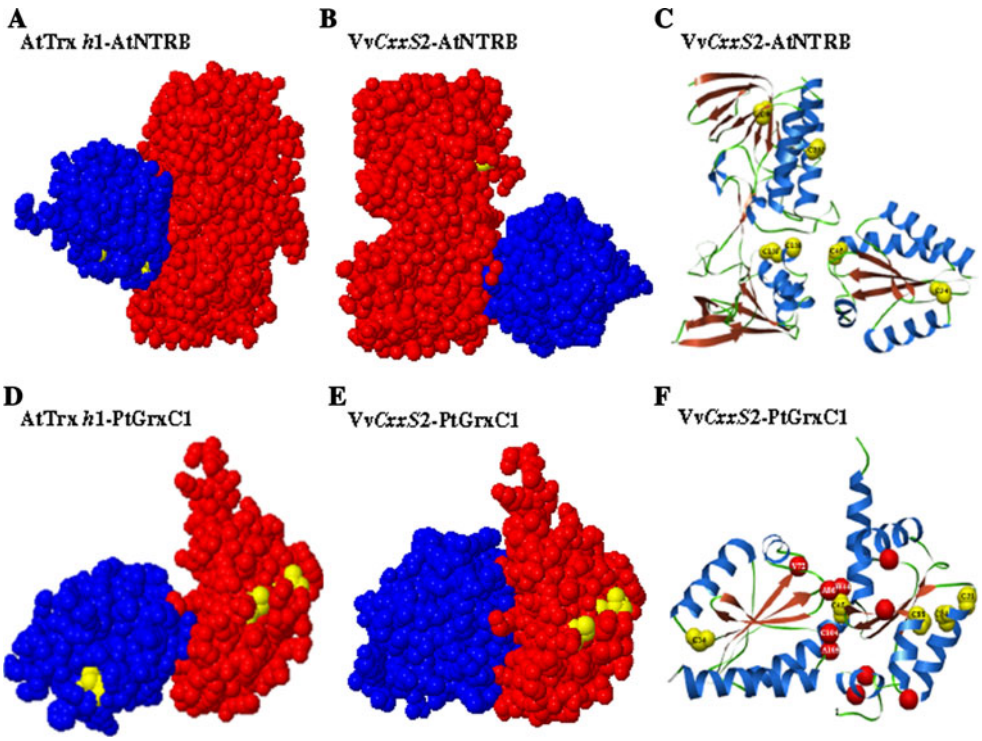

Fig. 3 Docking property analysis of Vv*CxxS*2 and AtTrx *h*1. **a** Docking model of AtTrx *h*1 (PDB ID code 1xflA) with AtNTRB (PDB ID code 1vdc). **b**, **c** Docking models of Vv*CxxS*2 with AtNTRB. **d** Docking model of AtTrx *h*1 with PtGrxC1 (PDB ID code 1z7p). **e**, **f** Docking models of Vv*CxxS*2 with PtGrxC1. AtNTRB is shown in *red* in **a** and **b**. PtGrxC1 is shown in *red* in **d** and **e**. AtTrx *h*1 is shown in *blue* in **a** and **d**, and Vv*CxxS*2 in *blue* in **b**, and **e**. Cysteines are shown in *yellow* spheres in **c** and **f** with their numbered amino acid position. Also, the residues implicated in NTR/Grx:Trx interaction are indicated in **f** as *red* and *numbered spheres*. (Color figure online)

with other plants indicating that Trxs diverged early in plant evolution [27]. The presence of other highly conserved regions between Vv*CxxS*2 and Trxs from other organisms is shown in Fig. S3.

Semiquantitative RT-PCR

The number of ESTs coding for Vv*CxxS*2 in plant organs was counted among the entire grape ESTs recorded on the NCBI GenBank database. Out of 570 ESTs detected for Trxs families, around 397 ESTs belong to the whole Trx *h* group. The Vv*CxxS*2 gene with 103 ESTs at different grape organs (berry, bud, leaf, petiole, flower, stem, root, and cluster) plays a critical role in this plant. However, no ESTs coding for Vv*CxxS*2 were found in seed organ in the NCBI GenBank database. Based on the number of identified ESTs in different organs, Vv*CxxS*2 gene thus appears to be transcribed at a high level in berries (46 hits), and at low levels in clusters (1 hit), roots (2 hits), and stems (3 hits).

The expression of Vv*CxxS*2 gene was analyzed in different grape organs at various developmental stages by semiquantitative RT-PCR (Fig. 4). Transcripts of Vv*CxxS*2 were present in all organs at different developmental stages, and its higher expression was related to the berry organs. Similarity, Trx *h*1 and *h*2 genes from poplar and *HbTRX1* gene from *H. brasiliensis* were detected in all compartments of the plant, especially leaves [15, 32]. Also, the pea Trx *h*3 and *h*4 genes are clearly expressed in all organs, while Trx *h*1 and *h*2 genes are only detected in green leaves in a very low amount [38]. In berry organ, expression analysis was studied at six developmental stages consisting of 14, 28, 42, 56, 82, and 110 days post anthesis (dpa). The amount of Vv*CxxS*2 transcripts was approximately constant from 14 to 42 dpa, then an increase was observed from 42 dpa to veraison stage (56 dpa) and then remained constant to 82 dpa, and finally it followed by a relatively dramatic decrease to ripeness (110 dpa) (Fig. 4a). Therefore, it seems that the highest expression level of Vv*CxxS*2 gene is related to veraison stage in berry organ. The expression of Vv*CxxS*2 gene was also analyzed in leaf, petiole and cluster organs at young, mid and old stages. The gene expression pattern was highly similar in leaf, petiole and cluster organs, with the higher expression at young stages (Fig. 4b). The Vv*CxxS*2 gene expression was also confirmed in green stems, young roots, and seeds. The lowest amounts of Vv*CxxS*2 transcripts were observed in root and seed organs.

Conclusion

An *h*-type Trx, designated Vv*CxxS*2, was isolated from grape berry organ of an Iranian cultivar, called Askari. It

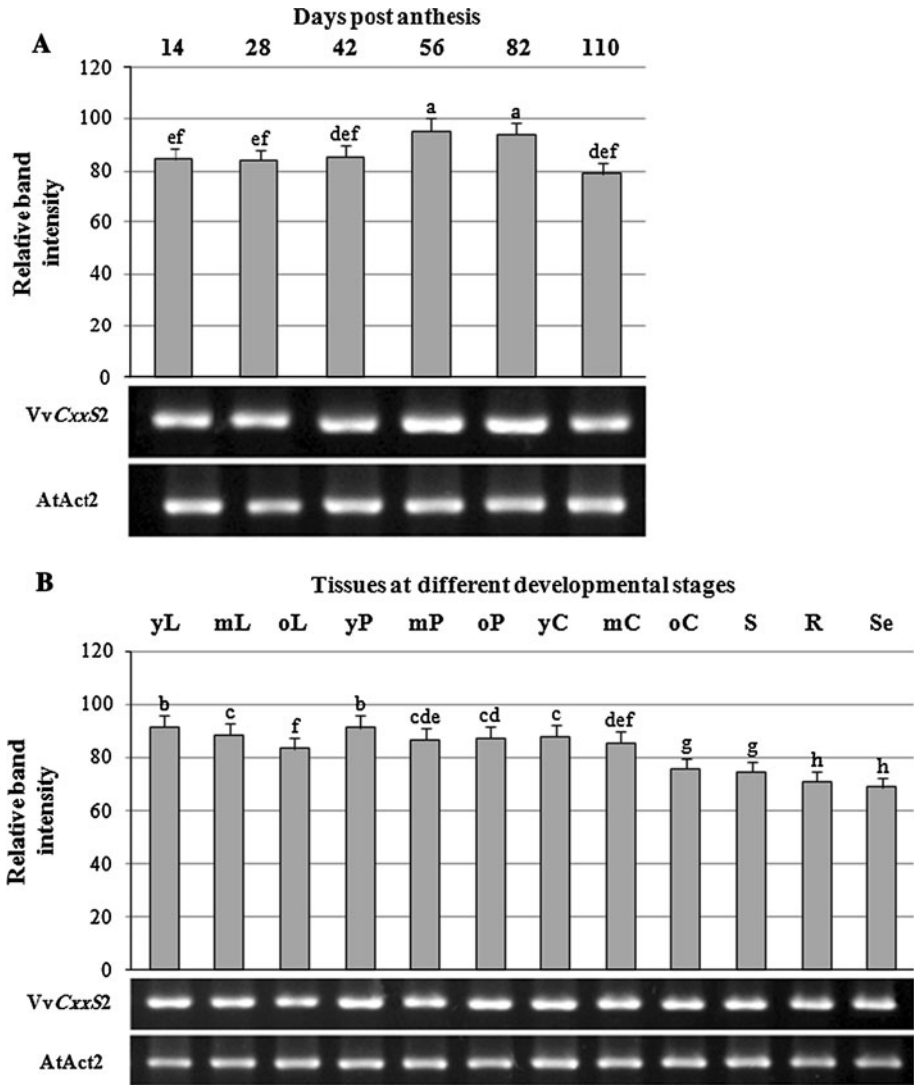

Fig. 4 Gene expression of Vv*CxxS*2 in different grape organs at various developmental stages. **a** RT-PCR analysis of Vv*CxxS*2 gene in grape berry organ at six times of sampling (14–110 dpa). **b** RT-PCR analysis of Vv*CxxS*2 gene in leaf, petiole, cluster, stem, root, and seed organs. One representative gel is shown from three independent replicates. Relative band intensities were normalized to the At*Act2* band intensity (100%). *Bars* with the *same lowercase letter* are significantly different ($P < 0.01$). Each *bar* represents the mean ± SD obtained from three independent RT-PCR reactions. *yL* young leaf, *mL* mid leaf, *oL* old leaf, *yP* young petiole, *mP* mid petiole, *oP* old petiole, *yC* young cluster, *mC* mid cluster, *oC* old cluster, *S* stem, *R* root, *Se* seed

was shown that this gene harbors two introns at positions identical to the previously sequenced Trx *h* genes. A modeling analysis indicated that Vv*CxxS*2 shares a common structure with other Trxs, and is preferably reduced by Grx rather than NTR. The constructed phylogenetic tree revealed that Vv*CxxS2* belongs to *h*-type Trx group, subgroup III, and contains a high degree of identity with other plant Trxs *h*. Also, the results of the semiquantitative RT-PCR indicated that this gene is transcribed in all organs at different developmental stages, although it appears to be expressed in berry more than other organs. It seems that this gene plays a critical role in the plant development and involves in important reactions. Therefore, further studies are needed to understand the physiological function of Vv*CxxS*2 in grape.

Acknowledgments We wish to thank Dr. Ramin Hosseini and Dr. A. H. Beiki (Imam Khomeini International University, Qazvin, IR of Iran) for helpful and technical assistance.

References

1. Alkhalfioui F, Renard M, Frendo P, Keichinger C, Meyer Y, Gelhaye E et al (2008) A novel type of thioredoxin dedicated to symbiosis in legumes. Plant Physiol 148:424–435
2. Andrusier N, Nussinov R, Wolfson HJ (2007) FireDock: fast interaction refinement in molecular docking. Proteins 69:139–159
3. Arner ES, Holmgren A (2000) Physiological functions of thioredoxin and thioredoxin reductase. Eur J Biochem 267:6102–6109
4. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. http://nar.oxfordjournals.org/content/38/suppl_2/W529.abstract
5. Bannai H, Tamada Y, Maruyama O, Nakai K, Miyano S (2002) Extensive feature detection of N-terminal protein sorting signals. Bioinformatics 18:298–305
6. Besse I, Buchanan BB (1997) Thioredoxin-linked plant and animal processes: the new generation. Bot Bull Acad Sin 38:1–11
7. Briesemeister S, Rahnenführer J, Kohlbacher O (2010) Going from where to why-interpretable prediction of protein subcellular localization. Bioinformatics 26:1232–1238

8. Brugidou C, Marty I, Chartier Y, Meyer Y (1992) The *Nicotiana tabacum* genome encodes two cytoplasmic thioredoxin genes which are differently expressed. Mol Gen Genet 238:285–293
9. Buchanan BB, Schürmann P, Wolosiuk RA, Jacquot JP (2002) The ferredoxin/thioredoxin system: from discovery to molecular structures and beyond. Photosynth Res 73:215–222
10. Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP, and related tools. Nat Protoc 2:953–971. doi:10.1038/nprot.2007.131
11. Felsenstein J (1985) Confidence limit on phylogenies: an approach using the bootstrap. Evolution 39:783–791
12. Gallie DR (1993) Post-transcriptional regulation of gene expression in plants. Ann Rev Plant Physiol Plant Mol Biol 44:77–105
13. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD et al (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) The proteomics protocols handbook. Humana Press, Totowa, pp 571–607
14. Gautier MF, Lullien-Pellerin V, de Lamotte-Guery F, Guirao A, Joudrier P (1998) Characterization of wheat thioredoxin *h* cDNA and production of an active *Triticum aestivum* protein in *Escherichia coli*. Eur J Biochem 252:314–2324
15. Gelhaye E, Rouhier N, Laurent P, Sautiere PE, Martin F, Jacquot JP (2002) Isolation and characterization of an extended thioredoxin *h* from poplar. Physiol Plant 114:165–171
16. Gelhaye E, Rouhier N, Jacquot JP (2003) Evidence for a subgroup of thioredoxin *h* that requires GSH/Grx for its reduction. FEBS Lett 555:443–448
17. Gelhaye E, Rouhier N, Vlamis-Gardikas A, Girardet JM, Sautière PE, Sayzet M et al (2003) Identification and characterization of a third thioredoxin *h* in poplar. Plant Physiol Biochem 41:629–635
18. Gelhaye E, Rouhier N, Jacquot JP (2004) The thioredoxin *h* system of higher plants. Plant Physiol Biochem 42:265–271
19. Gelhaye E, Rouhier N, Gerard J, Jolivet Y, Gualberto J, Navrot N et al (2004) A specific form of thioredoxin *h* occurs in plant mitochondria and regulates the alternative oxidase. Proc Natl Acad Sci USA 101:14545–14550
20. Gelhaye E, Rouhier N, Navrot N, Jacquot JP (2005) The plant thioredoxin system. Cell Mol Life Sci 62:24–35
21. Geourjon C, Deleage G (1995) SOPMA: significant improvement in protein secondary structure prediction by consensus prediction from multiple alignments. Comput Appl Biosci 11:681–684
22. Heidari Japelaghi R, Haddad R, Garoosi GA (2011) Rapid and efficient isolation of high quality nucleic acids from plant tissues rich in polyphenols and polysaccharides. Mol Biotechnol. doi:10.1007/s12033-011-9384-8
23. Ishiwatari Y, Honda C, Kawashima I, Nakamura S, Hirano H, Mori S et al (1995) Thioredoxin *h* is one of the major proteins in rice phloem sap. Planta 195:456–463
24. Jacquot JP, Lancelin JM, Meyer Y (1997) Thioredoxin: structure and function in plant cells. New Phytol 136:543–570
25. Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A et al (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449:463–467
26. Jiang H, Song W, Li A, Yang X, Sun D (2011) Identification of genes differentially expressed in cauliflower associated with resistance to *Xanthomonas campestris* pv. Campestris. Mol Biol Rep 38:621–629
27. Juttner J, Olde D, Langridge P, Baumann U (2000) Cloning and expression of a distinct cluster of plant thioredoxins. Eur J Biochem 267:7109–7117
28. Kim YJ, Shim JS, Krishna PR, Kim SY, In JG, Kim MK et al (2008) Isolation and characterization of a glutaredoxin gene from *Panax ginseng* C. A. Meyer. Plant Mol Biol Rep 26:335–349
29. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105–132
30. Landau M, Mayrose I, Rosenberg Y, Glaser F, Martz E, Pupko T et al (2005) ConSurf: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acid Res. http://www.ncbi.nlm.nih.gov/pubmed/15980475
31. Lennon BW, Williams CH, Ludwig ML (2000) Twists in catalysis: alternating conformations of *Escherichia coli* thioredoxin reductase. Science 289:1190–1194
32. Li HL, Lu HZ, Guo D, Tian WM, Peng SQ (2010) Molecular characterization of a thioredoxin *h* gene (HbTRX1) from *Hevea brasiliensis* showing differential expression in latex between self-rooting juvenile clones and donor clones. Mol Biol Rep 38:1989–1994
33. Maeda K, Hagglund P, Finnie C, Svensson B, Henriksen A (2008) Crystal structures of barley thioredoxin *h* isoforms HvTrxh1 and HvTrxh2 reveal features involved in protein recognition and possibly in discriminating the isoform specificity. Protein Sci 17:1015–1024
34. Meng L, Wong JH, Feldman LJ, Lemaux PG, Buchanan BB (2010) A membrane-associated thioredoxin required for plant growth moves from cell to cell, suggestive of a role in intercellular communication. Proc Natl Acad Sci USA 107:3900–3905
35. Messing J, Geraghty D, Heidecker G, Hu NT, Kridl J, Rubenstein I (1983) Plant gene structure. In: Kosuge T, Meredith CP, Hollaender A (eds) Genetic engineering of plants. Plenum Press, New York, pp 211–227
36. Meyer Y, Vignols F, Reichheld JP (2002) Classification of plant thioredoxins by sequence similarity and intron position. Method Enzymol 347:394–402
37. Meyer Y, Siala W, Bashandy T, Riondet C, Vignols F, Reichheld JP (2008) Glutaredoxins and thioredoxins in plants. Biochim Biophys Acta 1783:589–600
38. Montrichard F, Renard M, Alkhalfioui F, Duval FD, Macherel D (2003) Identification and differential expression of two thioredoxin *h* isoforms in germinating seeds from pea. Plant Physiol 132:1707–1715
39. Nimrod G, Schushan M, Steinberg DM, Ben-Tal N (2008) Detection of functionally important regions in hypothetical proteins of known structure. Structure 16:1755–1763
40. Qiu L, Ma Z, Jiang S, Wang W, Zhou F, Huang J et al (2009) Molecular cloning and mRNA expression of peroxiredoxin gene in black tiger shrimp (*Penaeus monodon*). Mol Biol Rep 37:2821–2827
41. Rouhier N, Gelhaye E, Jacquot JP (2002) Redox control by dithiol-disulfide exchange in plants. Part II. Cytosolic and mitochondrial systems. Ann NY Acad Sci 973:520–528
42. Rouhier N, Gelhaye E, Jacquot JP (2004) Plant glutaredoxins: still mysterious reducing systems. Cell Mol Life Sci 61:1266–1277
43. Russell RB, Breed J, Barton GJ (1992) Conservation analysis and secondary structure prediction of the SH_2 family of phosphotyrosine binding domains. FEBS Lett 304:15–20
44. Sahrawy M, Hecht V, Lopez-Jaramillo J, Chueca A, Chartier Y, Meyer Y (1996) Intron position as an evolutionary marker of thioredoxins and thioredoxin domains. J Mol Evol 42:422–431
45. Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
46. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ (2005) PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acid Res. http://www.ncbi.nlm.nih.gov/pubmed/15980490
47. Schobert C, Baker L, Szederkenyi J, Grossmann P, Komor E, Hayashi H (1998) Identification of immunologically related proteins in sieve-tube exudate collected from monocotyledonous and dicotyledonous plants. Planta 206:245–252
48. Schröter Y, Steiner S, Matthäi K, Pfannschmidt T (2010) Analysis of oligomeric protein complexes in the chloroplast

sub-proteome of nucleic acid-binding proteins from mustard reveals potential redox regulators of plastid gene expression. Proteomics 10:2191–2204
49. Serrato AJ, Cejudo FJ (2003) Type-*h* thioredoxins accumulate in the nucleus of developing wheat seed tissues suffering oxidative stress. Planta 217:392–399
50. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599
51. Zhang Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9:40