Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Identification and Evolutionary Characteristic Analysis of STARD Gene Family, and Overexpression *VvSTARD5* Responses to Salt Stress in Tomato

Honghong He
Gansu Agricultural University
Shixiong lu
Gansu Agricultural University
Huiming Gou
Gansu Agricultural University
Xuejing Cao
Gansu Agricultural University
Ping Wang
Gansu Agricultural University
Zonghuang Ma
Gansu Agricultural University
Baihong Chen
Gansu Agricultural University
Juan Mao (maojuan@gsau.edu.cn)
Gansu Agricultural University https://orcid.org/0000-0001-9344-6068

Research Article

Keywords: STARD gene family, Expression analysis, salt stress, tomato

Posted Date: January 10th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1109124/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

This study aimed to have a full understanding of the steroidogenic acute regulatory gene family member and evolutionary relationship in grape. 23 *VvSTARD* gene members were identified and divided into five groups in different species. Analyses of the gene codon preference, selective pressure, and tandem duplication of the *VvSTARD*, *AtSTARD*, and *OsSTARD* genes indicated that synteny relationship occurred in grapes, *Arabidopsis thaliana*, and rice genomes. The 8 lipid transporter proteins were found in the tertiary structure of the *STARD* gene family in grape. Expression profiles of the three species microarrays showed that the expression levels of the *STARD* genes in different organs and the response to abiotic stress in the same subgroup had similar characteristics. In addition, analysis of the *VvSTARD* genes expression levels was detected in response to different hormones and abiotic stresses by quantitative real-time polymerase chain reaction (qRT-PCR), and the results were the same as those predicted by the *cis*-elements and the expression profiles. Meanwhile, *VvSTARD5* gene was screened in high concentration NaCl treatment by qRT-PCR. Furthermore, the *VvSTARD5* was located at the nucleus by subcellular location. Through the function analysis of salt tolerance in transgenic tomato, overexpression *VvSTARD5* obviously improved tolerance to salt stress. Taken together, our findings Preliminary identify the functions of *VvSTARD* gene family and vertify *STARD5* that be likely involved in regulating salt tolerance, which may have potential application molecular breeding in grape.

Key Message

The 23 members of the *STARD* gene family were identified in grapes. In addition, Overexpression *VvSTARD5* improved tolerance to salt stress in transgenic tomato.

Introduction

Salt stress is an important constraint factor on the crop quality and yield particularly in grape. Although many studies have reported on the mechanism of plant salt tolerance (Cheong et al. 2003; Shi et al. 2003; Cao et al. 2007), numerous genes have not been excavated and studied yet. The steroidogenic acute regulatory protein-related lipid transfer domain (STARD), which was first discovered in mammals, has a 210-amino-acid conserved sequence, which forms an α/β helix-grip structure, thereby creating a hydrophobic cavity that binds to the ligand and small globular modules (Roderick et al. 2002; Schrick et al. 2004; Clark 2012, 2020; Tillman et al. 2020). Previous studies that diverse ligands, such as phospholipids, oxysterols, sphingolipids, cholesterol, and possibly fatty acids, bind to START domains in mammal and have functions in controlling thioesterase enzyme activity, tumour suppression and non-vesicular lipid transport (Ponting et al. 1999; Suricata et al. 2000; Roccio et al. 2003; Strauss et al. 2003).

The STARD protein family has been identified because many proteins contain the START domains in plants, and the homeodomain leucine zipper (HD–Zip III and HD–Zip IV subfamilies) transcription factor family is part of the *STARD* gene family (Nakamura et al. 2006). A total of 21 HD–Zip START domain transcription factors, which plays an important role in vascular bundle development, meristem formation, and polarity construction in *Arabidopsis* (Schrick et al. 2004). These factors include epidermal hair growth (GL2) (Szymanski et al. 1998), anthocyanin accumulation (ANL2 and FWA) (Thirtyish et al. 1999; Kubo et al. 1999; Ryo et al. 2008; Fujimoto et al. 2008), floral organ development (PDF2), *Arabidopsis* meristem layer 1 (ATML1) (Sessions et al. 1999; Abe et al. 2003), vascular bundle development (ATHB-8) (Mcconnell et al. 2001; Baima et al. 2001), and polarity of near and far axes of leaves and embryos (PHV, PHB, and REV) (Talbert et al. 1995; Emery et al. 2003; Elhiti et al. 2009). Study the function of GL2 showed that the HD–Zip START structure domain is required for the GL2 transcription factor activity (Schrick et al. 2014). HD–Zip III subfamily possesses the START domain, HD–START-associated domain and Me–Glu–Lys–Hi–Leu–Ala (MEKHLA) domain, but the HD–Zip IV subfamily lacks the MEKHLA domain (Williams et al. 2005; Zhang et al. 2020). HD–Zip IV genes are expressed explicitly in the outer cell and epidermal and subepidermal cells of multiple species during biotic and abiotic stresses (Ingram et al. 2000; Nakamura et al. 2006). Ectopic expression of the HD-Zip IV gene HDG11 can improve the drought tolerance and increase the grain yield of transgenic rice plants (Yu et al. 2013). *Cis*-acting element analysis show that the HD–Zip III genes may be involved in responses to light, hormones, abiotic stressors, and stem development, but this analysis fails to verify the function of those genes (Li et al. 2019). Although the HD–Zip gene family had been reported, studies on HD–Zip III and HD–Zip IV containing the START domain has little research focus on plant response to various abiotic stresses. In addition, genes that only contain the START domain in grapes have not been reported. Furthermore, previous research has reported that the START domain combined with the pleckstrin homology (PH) domain at the same site used for the PH domain membrane binding, which confers the complex function regulation of the ceramide transfer (CERT) protein (Prashek et al. 2017; Li et al. 2017). The EDR2 gene is associated with the plant defense stress reaction in *Arabidopsis* (Tang et al. 2005). In *Arabidopsis*, the *AtAPOSTART1* is PH–START domain protein, which is involved in seed germination (Resentini et al. 2014). Nevertheless, studies on genes containing the PH–START domain in grapes abiotic stress are not available.

HD–Zip III and HD–Zip IV containing the START domain proteins in plants have been widely reported, but studies on the resistance of such proteins to abiotic stress in plants are unclear. Moreover, the function mechanism of the PH–START or START domains proteins is still not understood when plants were subjected to abiotic stress. Therefore, this study focused on identifying the *VvSTARD* gene family and verifying the tolerance of the members of this family to salt stress in grapes. The phylogenetic tree, intragenomic tandem duplication events, extra genomic synteny relationship, selective pressure of genes and codon preference are analyzed to predict the evolutionary relationship amongst grapes, *Arabidopsis*, and rice. Meanwhile, qRT-PCR is conducted to verify their expressive difference of *VvSTARD* genes family in grape through response to different hormones and abiotic stresses. The *VvSTARD5* is overexpressed in “Micro Tom” tomato plants to identify its function of salt stress tolerance. These findings will lay a solid foundation for further investigations into the molecular mechanism of the *STARD* gene in grape salt stress resistance.

Materials And Methods

Plant materials and treatments

The *V. vinifera* “Pinot Noir” tube seedling was used as materials and cultured in the Fruit Tree Physiology and Biotechnology Laboratory of Gansu Agricultural University. The single-shoot stem segments of the test tube seedlings were attached to a solid GS (modified B5 solid medium) and cultured under white LED for 35 days. The grape seedings were treated with 0.2 mmol·l^{-1} of abscisic acid (ABA), 150 μmol·l^{-1} of methyl jasmonate (MeJA), 50 mg·l^{-1} of salicylic acid (SA), 100 μmol·l^{-1} of indole acetic acid (IAA), 50 mg·l^{-1} of gibberellin 3 (GA_3), 10% PEG6000, and 400 mmol·l^{-1} of NaCl at low temperature (4 °C) for 12 and 24 h. Three replicates were prepared for each treatment, and an equal volume of distilled water was used as control. All materials were collected, frozen in liquid N_2, and stored at −80 °C for RNA extraction and qRT-PCR.

Cotyledons of “Micro Tom” tomato were used to transform the *VvSTARD5* gene, and young seedings of three weeks were used for the salt treatment. The transgenic tomato was watered every 3 h with 400 mmol l^{-1} of NaCl, and the control was supplemented with the same volume of distilled water. Three biological replicates for each treatment and fresh sample leaves of tomato (0.1 g) were collected. The relative electrical conductivity, proline and malondialdehyde contents of tomato leaves were determined using the commercial ELISA kit (Jiangsu Keming Biotechnology Institute, Suzhou, China) in accordance with the manufacturer’s protocol.

Identification of STARD genes in grape

The *AtSTARD* sequences were downloaded from the *Arabidopsis* genome website (http://www.arabidopsis.org/). Grape and rice genome annotated information were downloaded from the phytozome website (https://phytozome.jgi.doe.gov/pz/portal.html) (Goodstein et al. 2012). The AtSTARD protein sequences (accession numbers: At1g05230, et al) were compared with the grape genome sequences, and the START conserved domain (PF01852) of all proteins were obtained (Table S1). The START conserved domain was used as queries to perform the BLASTP analysis ($E < 10^{-10}$). HMMER (https://www.ebi.ac.uk/Tools/hmmer/), and Pfam (http://pfam.xfam.org/) (Potter et al. 2018; El-Gebali et al. 2019) were used to confirm the sequence accuracy, which contained START domain. Simultaneously, the *STARD* genes of *Arabidopsis* and rice were also named in the same way. The physicochemical properties of the VvSTARD proteins, such as molecular weight (MW), isoelectric point (pI), grand average of hydropathicity (GRAVY), aliphatic index and instability index, were obtained from the ExPASy (https://www.expasy.org/) (Wilkins et al. 1999).

Analysis of phylogenetic tree, gene structures and motifs

The multiple sequence alignment of the *STARD* genes of *Arabidopsis*, rice, and grapes was conducted using the ClustalX 2.0 (Conway Institute, University College Dublin, Dublin, UK) (Larkin et al. 2007). MEGA 7.0 (Pennsylvania State University, State College, PA, USA) was used to perform phylogenetic tree analysis (Kumar et al. 2016) with the NJ, and the “Poisson model” was adopted. The gap was set to “complete deletion”, and the check parameter was bootstrap 1000 times with random seed. GSDS 2.0 (http://gsds.cbi.pku.edu.cn/) was used to analyze gene structures, namely, exon and intron (Hu et al. 2015). MEME online software (http://meme-suite.org/) was used to predict the conserved domain of the protein (Bailey et al. 2009), and the number of motifs in the conserved domain was set to 20.

Analysis of the *STARD* gene synteny and the Ka/Ks in grapes

For the collinearity analysis, the MCScanX was used to detect the collinearity of the *STARD* gene synteny (Wang et al. 2012), and the diagram was drawn via TBtools (Chen et al. 2018). The nonsynonymous/synonymous (Ka/Ks) values of duplicate gene pairs or triplicate gene groups (between any two genes in one triplicate gene group) were calculated through DnaSP 6.0, an application released by Universitat de Barcelona.

Codon usage bias analysis

The codon bias refers to the unequal use of synonymous codons for an amino acid (Hershberg et al. 2008; Larracuente et al. 2008; Plotkin et al. 2011; Guo et al. 2017; Wang et al. 2018). The coding sequences of the STARD were used to determine the codon adaptation index (CAI), codon bias index (CBI), frequency of optimal codons (FOPs), relative synonymous codon usage (RSCU), GC content and GC content at the third site of the synonymous codon (GC3s content) by using the online software CodonW 1.4.2 (http://codonw.sourceforge.net) (Wang et al. 2018). The R programming language was used to analyse the correlation amongst the T3s, C3s, A3s, G3s, GC, GC3s, L_sym, L_aa, GRAVY and Aromo.

Analysis of subcellular localization, secondary and tertiary structure prediction

WoLF PSORT (https://wolfpsort.hgc.jp/) was used to predict the subcellular localization of the *VvSTARD* genes. The NPS@: SOPMA secondary structure (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) was used to predict the secondary structure of *VvSTARD* proteins. SWISS-MODEL (http://www.expasy.org/swissmod/) was used to predict the 3D structure of some atypical HDs, and tertiary structure was prepared using PyMOL software (DeLano 2002, The PyMOL molecular graphics system. http://www.pymol.org).

Cis-acting element and expression analyzes of *STARD* gene in grapes

The promoter sequence of the 2,000 bp upstream of the coding region of *VvSTARD* genes was obtained from the website of grape genomes, and the PlantCARE was used to analyze the gene promoter elements (Lescot et al. 2002; Wang et al. 2016). The diagrams of *cis*-acting elements were constructed via GSDS2.0 (http://gsds.cbi.pku.edu.cn/) (Hu et al. 2015). Expression profile data of different abiotic stresses was revitalized from GEO databases (Affymetrix GeneChip 16K *Vitis vinifera* Genome Array, accession number: GSE31594) (Wang et al. 2018). The expression data of *VvSTARD* genes was extracted from grape. The tissue expression data of grape, *Arabidopsis* and rice were retrieved from the Bio-Analytic Resource for Plant Biology (BAR, https://bar.utoronto.ca/) databases. In addition, stress expression data were retrieved from the BAR databases in *Arabidopsis* and rice. Heat maps were drawn in accordance with TBtools (Chen et al. 2018).

RNA isolation and qRT-PCR

The plant total RNA isolation was performed through kit (Sigma, St. Louis, MO, USA). The M-MLV Reverse Transcriptase (RNase H^-) kit (Takara Bio, Inc., Japan) was utilized to synthesis reverse-strand complementary DNA (cDNA). The purified total RNA (1 μg) was reverse transcribed into the first-strand cDNA and used for qRT-PCR. Subsequently, the TaKaRa SYBR Premix Ex Taq. II (Takara Bio, Inc., Japan) was used for qRT-PCR (Light Cycler 96 Real-Time PCR System, Roche, Basel, Switzerland). The cycling parameters were 95 °C for 30 s, 40 cycles at 95 °C for 5 s, and 60 °C for 30 s. For melting curve analysis, a program consisting of 95 °C for 15 s followed by a constant increase from 60 °C to 95 °C, was included following the PCR cycles. *VvGAPDH* (GenBank accession no. CB973647) and *SlActin* (GenBank accession no. NM_001330119) were used as control genes. The primer sequences are presented in Table S2. The relative expression levels of the genes were calculated using the $2^{-\Delta\Delta Ct}$ method (Willems et al. 2008), and figures were drawn using the Origin 9.0 software.

Subcellular localization and function identification of *VvSTARD5*

The coding sequences of *VvSTARD5* were amplified and inserted into pBI221–EGFP to clarify its expressing site in *Arabidopsis* protoplasts cell. And recombinant vector was constructed by using the NovoRec®PCR One Step Cloning Kit (Novoprotein Scientific Inc., China). The GFP fluorescence was detected using confocal laser-scanning microscopy (Olympus FV1000 Viewer, Tokyo, Japan). *Arabidopsis* protoplasts were transformed in accordance with the method of Yoo et al. (2007).

"Micro Tom" tomato was used for the transformation of the *VvSTARD5*. The complete coding regions of *VvSTARD5* were inserted behind the 35S promoter and constructed pCAMBIA1300-*VvSTARD5* recombinant plasmids that were introduced into the *Agrobacterium* strain GV3101. The *Agrobacterium*-mediated transformation of the "Micro Tom" leaves was performed as previously described (Ruf et al, 2001). The DNA of tomato plants was extracted using the TransDirect Plant Tissue PCR Kit (Beijing Quantising Biotechnology Co., Ltd.), and positive plants were detected using gene-specific primers (35S-F: 5'-TGACGCACAATCCCACTATC-3'; STARD5-R: 5'-CGATGGTAGCGCTTCTTCTT-3').

Statistical analysis

Statistical analysis was performed by one-way ANOVA using the IBM SPSS v.22 (IBM, Armonk, NY, USA). The $p < 0.05$ and < 0.01 indicated a significant difference and extreme significant difference, respectively.

Results

Identification of the STARD gene family in grape

A total of 23 *VvSTARD* candidate genes were retrieved in the grape genome database (Table 1). *VvSTARD1*– *VvSTARD23* were named on the basis of conserved domains and chromosomes sites. The second chromosome was retrieved 4 genes, and the 5th, 6th, 11th, 16th, and 17th chromosomes were only retrieved one gene among 12 chromosomes, respectively. The CDS coding sequences of VvSTART domain in grapes encoded 237–886 amino acids. The MW of VvSTARDs ranged from 26.77 kD (VvSTARD23) to 99.56 kD (VvSTARD7), showing large differences. VvSTARD proteins had hydrophilic values ranging from −0.466 to −0.077. The predicted *pI* values of the VvSTARD proteins ranged from 5.60 (VvSTARD5) to 9.66 (VvSTARD22). Furthermore, 20 VvSTARD proteins (86.95%) had an instability with index greater than 40.

Further analysis showed the VvSTARD proteins were predicted in the nucleus, chloroplast, and cytoplasm (Table S3). Except for VvSTARD19, VvSTARD21 and VvSTARD22, many proteins were predicated and located in the nucleus. Unlike other proteins, the VvSTARD1, VvSTARD2, VvSTARD8, VvSTARD14, and VvSTARD19 were not present in the chloroplast. A total of 13 (VvSTARD2, VvSTARD4, VvSTARD5, VvSTARD9, VvSTARD10, VvSTARD11, VvSTARD12, VvSTARD13, VvSTARD14, VvSTARD17, VvSTARD18, VvSTARD20, and VvSTARD23), 2 (VvSTARD16, and VvSTARD19), 1 (VvSTARD12), 4 (VvSTARD10, VvSTARD16, VvSTARD17, and VvSTARD22), 2 (VvSTARD17, and VvSTARD20), 1 (VvSTARD20) and 6 (VvSTARD9, VvSTARD10, VvSTARD11, VvSTARD13, VvSTARD18, and VvSTARD19) proteins were predicated and located in the cytoplasm, plasma membrane, cytoskeleton, mitochondria, extracellular matrix, golgi apparatus, and vacuole, respectively.

Phylogenetic and structural analyses of the START domain proteins

The phylogenetic tree was constructed using STARD protein sequences of grapes, *Arabidopsis* and rice (Fig. 1A). These START genes were mainly divided into five subgroups (groups 1–5). The members of 20 START domain proteins in group 1 (4, 8, and 8 members from grapes, rice, and *Arabidopsis*, respectively), which contained START and HD domains. 18 members were in group 2 (4, 4, and 10 members from grapes, rice, and *Arabidopsis*, respectively), which contained START and HD domains. 18 members in group 3 (5, 8, and 5 members from grape, rice and *Arabidopsis*, respectively), which contained the START, HD and MEKHLA domains. 13 members in group 4 (4, 2, and 7 members from grapes, rice, and *Arabidopsis*, respectively), which contained the structural START domain. 16 members in group 5 (6, 3, and 5 members from grapes, rice, and *Arabidopsis*, respectively), which contained the structural START, PH and DUF1336 domains (Fig. S1).

Further analysis showed that members from the same subgroups had similar exon/intron structures and motifs. As shown in Fig. 1B, the exon of *VvSTARD* gene members ranged from 5 to 22. Moreover, 6 conserved motifs (motifs 1, 2, 3, 4, 5, and 13) were shared by groups 1, 2, and 3 of the VvSTARD proteins family (Figs. 1C and S2). The 6 motifs (motifs 8, 9, 11, 15, 17, and 18) were shared by groups 1 and 2. The 4 motifs (motifs 6, 7, 16, and 19) were shared by group 3, and 3 motifs (motifs 10, 12, and 14) were shared by group 5. However, no system-conserved motif in the VvSTARD protein family was observed in group 4. In addition, the motif 16 was shared by groups 1 and 2. These results indicated that genes with very similar structures distributed in the same subgroups which might have similar biological functions, whereas the genes distributed in different subgroups likely have different biological functions.

Analysis of VvSTARD, AtSTARD, and OsSTARD genes codon preference

A total of 23 *VvSTARD*, 35 *AtSTARD*, and 25 *OsSTARD* gene families contained 15 989, 24 209, and 31 815 codons, respectively (including stop codons) (Fig. 2A). And the three species had RSCU > 1 codons of 9 916, 15 413, and 10 459, respectively (Fig. 2B). Among the RSCU > 1 codons, ending in A or U of coding STARD proteins had preferred codons in the grape and *Arabidopsis*. In grape, the total of 2 193, 4 674, and 3 049 codons ending in A, U, and G or C, respectively, accounting for 22.12%, 47.14%, and 30.74%, respectively, of the total number of codons with RSCU > 1. In *Arabidopsis*, codons ending with A, U and G or C accounted for 21.83%, 49.45%, and 28.72%, respectively, of the total codons in RSCU > 1. However, rice contained codons ending in G and C, accounting for 43.24% and 46.17%, respectively, of the total codons in RSCU > 1, whereas codons ending in A or U only accounted for 10.59% of the total codons in RSCU > 1 (Fig. 2B and Table S4).

The grape, *Arabidopsis*, and rice *STARD* genes had average CAI values of 0.193, 0.201, and 0.227, respectively; average CBI values of −0.063, −0.022, and 0.093, respectively; and average FOP values of 0.380, 0.405, and 0.469, respectively (Tables S5-1, S5-2 and S5-3). Grape, *Arabidopsis*, and rice had average Nc values of 54.45, 54.02, and 47.93, respectively; minimum values of 50.69 (*VvSTARD14*), 46.10 (*AtSTARD35*), and 31.76 (*OsSTARD7*), respectively; and maximum values of 57.51 (*VvSTARD2*), 61 (*AtSTARD34*), and 56.66 (*OsSTARD21*), respectively. Among the 23 and 35 *STARD* genes of grapes and *Arabidopsis*, respectively, none had an Nc value of less than 35. However, among the 25 *OsSTARD* genes, six (*OsSTARD5*, *OsSTARD6*, *OsSTARD7*, *OsSTARD8*, *OsSTARD10*, and *OsSTARD25*) showed an Nc value less than 35. The GC3 values in grapes ranged from 0.33 to 0.54, and the distribution was relatively concentrated. The GC3 values in *Arabidopsis* ranged from 0.29 to 0.49, and the distribution was relatively concentrated. The GC3 values in rice ranged from 0.37 to 0.94, and the distribution was relatively scattered. These findings showed that the codon usage preferences of the grape and *AtSTARD* gene families were strong and affected by selective pressure during evolution, whereas those of the *VvSTARD* gene family were weak and affected by the mutation pressure during evolution.

Correlation analysis revealed that the T3s had a negative correlation with C3s, G3s, GC3s, CBI, and Fop and that the C3s had a positive correlation with CBI, Fop, GC, and GC3s in grape, *Arabidopsis*, and rice (Fig. 2C, 2D). These correlations were highly consistent in grapes and *Arabidopsis* but quite different from those in rice (Fig. 2E). For instance, the T3s had a positive correlation with Nc in rice, but the T3s had a negative correlation with Nc in grape and *Arabidopsis*. Nc had a negative correlation with CAI, CBI, and Fop in rice, but Nc had a positive correlation with CAI, CBI, and Fop in grapes and *Arabidopsis*. Collectively, from the above-mentioned results, the genetic relationship between grapes and *Arabidopsis* was inferred to be close.

Chromosomal distribution and gene duplication analysis

As shown in Fig. 3A and Table S5-4, *VvSTARD* genes were unevenly distributed in four linkage groups (chr). The chr6/chr13 linkage group had two *VvSTARD* gene pairs. chr1, chr3, chr14, chr18, and chr19 had no synteny *VvSTARD* gene. In this study, tandem duplication genes, namely, *VvSTARD14*/ *VvSTARD15* and *VvSTARD10*/ *VvSTARD13*, were discovered on chr6 and chr13, respectively. A pair of collinear genes (*VvSTARD6*/ *VvSTARD7*) was observed on chr15 and chr16, and another pair (*VvSTARD9*/ *VvSTARD11*) was found on chr4 and chr9. These results suggested that some *VvSTARD* genes might be manufactured via gene duplication, and the primary driving force of the *VvSTARD* evolution was these duplication events.

Three representative comparative systematic maps of *Arabidopsis*, grapes, and rice were constructed to further forecast the phylogenetic element of the *VvSTARD* family (Fig. 3B and Table S5-5). A total of 13, 14, and 9 *STARD* genes in grapes, *Arabidopsis*, and rice showed a collinearity relationship. Amongst these genes, 15 were homologous pairs of the *STARD* genes in grape and *Arabidopsis*, and 14 were homologous pairs of the *STARD* genes in grapes and rice. Some *VvSTARD* genes particularly the *VvSTARD* and *AtSTARD* genes were linked with three pairs of synonymous genes, such as *VvSTARD7*, which might play a critical role in the evolution of the *STARD* gene family. Some *STARD* collinear gene pairs between grapes and *Arabidopsis* were settled on highly conserved synonymous blocks. The phylogenetic relationship and codon preference analyses demonstrated that the evolutionary relationship between grapes and *Arabidopsis* might be close.

The modes of selection could be estimated using the ratio of the number of nonsynonymous substitutions per nonsynonymous site (Ka) to the number of synonymous substitutions per synonymous site (Ks). The Ka/Ks ratios of the *STARD* gene pairs of grapes, *Arabidopsis*, and rice were calculated to further understand the evolutionary relationship of the *VvSTARD* gene family (Fig. 4 and Table S5-6, S5-7, and S5-8). A total of 202 homologous gene pairs were found in the grape *STARD* gene family (Fig. 4A). A total of 79 pairs had Ka/Ks > 1, and 123 pairs had Ka/Ks < 1. A total of 382 homologous gene pairs were found in the *AtSTARD* gene family (Fig. 4B). A total of 161 pairs had Ka/Ks > 1, and 221 pairs had Ka/Ks < 1. A total of 260 homologous gene pairs were found in the *OsSTARD* gene family (Fig. 4C). A total of 70 pairs had Ka/Ks > 1. One pair (*OsSTARD7*/ *OsSTARD1*) had Ka/Ks = 1, and 189 pairs had Ka/Ks < 1. These results showed that the *VvSTARD*, *AtSTARD*, and *OsSTARD* gene families might be dominated by purification selection during evolution.

Secondary and tertiary structure prediction of VvSTARD proteins

The results of the secondary structure analysis of the VvSTARD proteins family demonstrated that the secondary structures were α helix, β turn, and random coil (Table S6). The percentages of α helix, β turn and random coil were 30.52% (VvSTART17) to 44.11% (VvSTARD13), 3.24% (VvSTARD16) to 6.75% (VvSTARD23) and 32.91% (VvSTARD23) to 49.05% (VvSTARD17), respectively. The 3D structure analysis showed structures peculiar to several STARD proteins (Fig.S3 and Table S7). These proteins included thioesterase adipose-associated isoform brown fat-inducible thioesterase 2 (BFIT2; observed in VvSTARD14, VvSTARD16, and VvSTARD18), CERT (observed in VvSTARD2, VvSTARD3, VvSTARD4, VvSTARD5, VvSTARD6, VvSTARD7, VvSTARD8, VvSTARD15, VvSTARD17, VvSTARD20, and VvSTARD21), metastatic lymph node 64 (MLN64) protein (observed in VvSTARD9, VvSTARD10, VvSTARD11, and VvSTARD12), PCTP (observed in VvSTARD21 and VvSTARD22), START protein3 (observed in VvSTARD4 and VvSTARD7), cholesterol-regulated START protein4 (observed in VvSTARD11, VvSTARD13, and VvSTARD19), START protein5 (observed in VvSTARD10 and VvSTARD13) and START protein3 (observed in VvSTARD1–VvSTARD13 and VvSTARD20).

The secondary and tertiary structure analyses showed that MLN64, PCTP, cholesterol-regulated START protein 4, and START protein 5 contained four α-helixes, of which two α helices (α2 and α3) formed an internal hydrophobic cavity that could hold a ligand molecule (Fig.S3). α4 was visible on the top of the hydrophobic channel, and α helix at the C-terminus formed the lid. In addition, START protein13 had two α-helices (α1 and α2), and the C-terminal α2 helix served as lid, thereby establishing an internal hydrophobic cavity. BFIT2, CERT and START protein3 contained six α-helixes. Further research found that START protein 5 contained only one 8-chain antiparallel β-sheet, whereas MLN64, PCTP, BFIT2, CERT, START protein 3, cholesterol-regulated START protein 4, and START protein 13 contained a 9-chain antiparallel β-sheet. The side view showed that the antiparallel β-sheets, that was β4, β5 and β6 at one end of the hydrophobic cavity formed a basket structure, whereas the β-sheets on the other side, that is β1, β2, β3, β7, β8, and β9, were formed another basket structure. These results suggested that VvSTARD proteins played a significant role in regulating plant lipid metabolism.

Cis -acting element and expression pattern of VvSTARD genes

Cis- acting elements related to the hormone and abiotic stress responses were speculated in the promoter region of the *VvSTARD* genes. There were 9 types *cis-*acting elements of hormone and stress-relation were presented in the promoters of *STARD* genes in grapes (Fig. 5A and Table S8-1). Three stress-related *cis-*acting elements, including TC-rich repeats (defense and stress), MBS (drought), and low-temperature responsive elements were annotated in grape genomic data. Six hormone-related *cis-* acting elements, including TGA element/AuxRR core (auxin), O_2 site (zein metabolism), TCA element (salicylic acid), abscisic acid (ABA)-responsive element, GARE-motif/P-box/TATC-box (gibberellin), and CGTCA/TGACG motif (MeJA responsive element) were identified. All genes of *VvSTARD* contained *cis-* acting elements of abiotic stress or hormonal responses. Among of the *VvSTARD* genes, the promoter of 14 genes included ABA response element, and 14 genes were detected in the drought response element. In addition, the *VvSTARD* genes contained 14 auxins, 10 zeins, 9 GA3, 11 SA, and 13 MeJAresponsive elements. The results showed that the *VvSTARD* genes could regulate the metabolism of various hormones and abiotic stresses in response to different environmental factors. The expression mode and function of the *STARD* gene family in plants were not clear. Moreover, we analyzed the *STARD* gene expression data for organs/tissues and abiotic stress in grapes, rice, and *Arabidopsis* were downloaded from the BAR database.

The results of the analysis of the grape abiotic stress expression data (Fig. 5B and Table S8-3) showed that six genes (*VvSTARD1*, *VvSTARD2*, *VvSTARD3*, *VvSTARD5*, *VvSTARD6*, and *VvSTARD8*) belonged to groups 1 and 2, whereas five genes (*VvSTARD9*, *VvSTARD10*, *VvSTARD11*, *VvSTARD12*, and *VvSTARD13*) belonged to group 3, and such genes were related to salt stress. The expression profiles indicated that most *VvSTARD* genes were highly expressed at different times of NaCl, PEG and low temperature (5°C) treatments. Genes belonging to groups 5 (*VvSTARD15*, *VvSTARD16*, *VvSTARD19*, and *VvSTARD23*), 4 (*VvSTARD20* and *VvSTARD22*) and 3 (*VvSTARD9*, *VvSTARD10*, *VvSTARD11*, and *VvSTARD13*) were related to drought stress. *VvSTARD* genes related to low-temperature stress were distributed in different groups, and two genes were found in groups 1 and 2 (*VvSTARD6* and *VvSTARD8*).

The expression patterns of various tissues and organs of the *AtSTARD* gene family demonstrated that the expression of genes in different subfamilies had similarities (Fig. S4A and Table S8-4). Most *STARD* genes distributed in group 1, such as *AtSTARD15*, *AtSTARD10*, *AtSTARD1*, *AtSTARD6*, and *AtSTARD9*, were expressed in seeds. Two *AtSTARD* genes (*AtSTARD5* and *AtSTARD19*) belonged to group 2, and such genes were expressed in seeds. Most *AtSTARD* genes in group 3, such as *AtSTARD17*, *AtSTARD18*, *AtSTARD19*, *AtSTARD20*, and *AtSTARD21*, were not expressed in the pollen but normally expressed in other tissues and organs. Two *AtSTARD* genes in group 5 (*AtSTARD24* and *AtSTARD25*) were expressed in all organs and tissues. Except in seeds, *AtSTARD22* belonged to group 5 and expressed in all tissues and organs. *AtSTARD26* belonged to group 5, but it was expressed only in the roots and stamens. Most of the *AtSTARD* genes in group 4, such as *AtSTARD28* and *AtSTARD30*, were not expressed in the pollen, seed, shoot and root but normally expressed in other tissues. *AtSTARD27* and *AtSTARD30* were not expressed in the shoot, and *AtSTARD27* was not expressed in the root. Only *AtSTARD31* could be expressed in various tissues and organs.

The results of abiotic stress expression analysis demonstrated that the *AtSTARD* genes clustered in the same group had similar resistance and different expression patterns (Fig. S4B and Table S8-5). In group 4, one gene (*AtSTARD28*) was highly expressed in the shoot and root under control, cold, salt, drought, wound, and heat stresses. Group 3 had three genes (*AtSTARD18*, *AtSTARD19*, and *AtSTARD21*) under the control, cold, salt, drought, wound, and heat stresses that were expressed higher in the root than in the shoot. In addition, under the control, cold, salt, drought, wound, and heat stresses, some genes showed a higher expression level in root than in shoot, with one gene belonging to group 5 (*AtSTARD25*) and another gene belonging to group 4 (*AtSTARD31*). Moreover, under the control, cold, salt, drought, wound, and heat stresses, the expression level in the shoot was higher than that in the root, and the genes were distributed in groups 1 (*AtSTARD10* and *AtSTARD12*) and 4 (*AtSTARD27*, *AtSTARD29*, and *AtSTARD30*).

The expression patterns of the *OsSTARD* gene family in various tissues and organs showed that the expression of genes in different subfamilies had similarities (Fig. S4C and Table S8-6). Most of the *STARD* genes in groups 1 and 2, such as *OsSTARD5*, *OsSTARD9*, *OsSTARD10*, *OsSTARD1*, *OsSTARD11*, and *OsSTARD6*, were expressed in rice seeds, shoot apical meristem (SAM) and inflorescence. Some *OsSTARD* genes (*OsSTARD15* and *OsSTARD13*) were placed in group 3 and expressed in SAM, in florescence and seedling root. Furthermore, *OsSTARD14* and *OsSTARD12* were expressed in SAM and inflorescence. Group 4 only contained one gene, that is, *OsSTARD21*, which was expressed in mature leaves, inflorescence P2, and seeds S2–S5. Group 5 contained three *OsSTARD* genes, namely, *OsSTARD18*, *OsSTARD19*, and *OsSTARD20. OsSTARD19* was highly expressed in inflorescence P6 and seed S5. *OsSTARD20* was highly expressed in SAM and young inflorescence. *OsSTARD18* was highly expressed in mature and young leaves.

The analysis of rice abiotic stress expression data demonstrated that 17 genes were expressed in the normal growing shoot and root and evenly distributed in five subgroups (Fig. S4D and Table S8-7). Nine genes belonged to groups 1 (*OsSTARD5*, *OsSTARD10*, *OsSTARD4*, and *OsSTARD2*), 3 (*OsSTARD16*, *OsSTARD13*, and *OsSTARD14*), and 5 (*OsSTARD19* and *OsSTARD18*), and such genes were highly expressed in the root and shoot under salt stress and evenly distributed amongst four subgroups. Groups 2, 1, 3, and 5 with 1 (*OsSTARD24*), 1 (*OsSTARD7*), 1 (*OsSTARD2*), 2 (*OsSTARD12* and *OsSTARD15*), and 1 (*OsSTARD20*) gene were expressed in the root and shoot under cold stress and evenly distributed in six subgroups.

Analysis of *VvSTARD* gene family tissues (Fig. 5C and Table S8-2) demonstrated that the tissue expression of the *VvSTARD* genes in the same group was similar, but the tissue expression sites differed because of evolutionary differences. *VvSTARD4*, *VvSTARD5*, *VvSTARD6*, and *VvSTARD7* were members of the

group 1, which contained the HD–START domain. Interestingly, *VvSTARD4*, *VvSTARD5*, and *VvSTARD6* were expressed in the leaves, seedling, stems, flowers, buds, fruits, skin, seed, stamen, petals, pericarp, and carpel. However, the *VvSTARD7* was only expressed in the leaves and seed-post fruits. *VvSTARD1* and *VvSTARD8*, which were classified into group 2 and contained the HD–START domain, were expressed in the leaves, buds, flowers, pollens, and seeds. *VvSTARD9*, *VvSTARD10*, *VvSTARD11*, *VvSTARD12*, and *VvSTARD13* belonged to group 3 and contained the HD–START–MEKHLA domain. *VvSTARD10* and *VvSTARD11* were not expressed in the pollen, seed, flesh, rachis, pericarp, and other tissues and organs. *VvSTARD9* and *VvSTARD12* were detected in the tendrils, leaves, seedling, stems, roots, flowers, buds, fruits, and carpels. Nevertheless, *VvSTARD13* was extremely lowly expressed or not expressed in many tissues. *VvSTARD20*, *VvSTARD21*, *VvSTARD22*, and *VvSTARD23*, which were classified into group 4 and contained the START domain only, were expressed at different developmental stages of each organ and tissue. *VvSTARD14*, *VvSTARD15*, *VvSTARD16*, *VvSTARD17*, *VvSTARD18*, and *VvSTARD19* belonged to group 5. *VvSTARD14*, *VvSTARD15*, and *VvSTARD18* were expressed in other tissues except for seed, petal, seedling, and bud winter. The *VvSTARD16* was expressed at different developmental stages of each organ and tissue, and *VvSTARD17* was downregulated or not expressed in many organs. The *VvSTARD19* was upregulated in the pollen, flesh midripening, flesh ripening, flesh, pericarp, and skin. *VvSTARD23* was also upregulated in the tendrils, young leaves, seedlings, stalks, flowers, carpel, stamen, petals, pollen, seed veraison, flesh veraison, skin veraison, and pericarp veraison. Tissue expression analysis indicated that the expression levels of the *VvSTARD* genes in different tissues at different developmental stages of grapes had obvious difference.

qRT-PCR of the VvSTARD gene family

qRT-PCR was utilized to verify the expression profile data and further verify the physiological characteristics of the *VvSTARD* gene family. The results showed that most of the *VvSTARD* gene families could be expressed in grape leaves in response to hormones and abiotic stresses (Fig. 6). The expression levels of different hormones and abiotic stresses at 24 h were more evident than those at 12 h. A considerable degree of agreement was found among the predicted results. As shown in the chip expression profile, the *VvSTARD* gene family was expressed in grape leaves, which could respond to the exogenous hormone treatment and presented a high expression level. The expression levels of MeJA, SA, IAA, and GA3 were the same as those of *VvSTARD1– VvSTARD4*, *VvSTARD14– VvSTARD15*, *VvSTARD7– VvSTARD10*, *VvSTARD16– VvSTARD21*, *VvSTARD10*, *VvSTARD13*, and *VvSTARD23*. under the 400 mmol l^{-1} NaCl treatments for 24 h, expression levels of 17 genes (*VvSTARD1– VvSTARD15*, *VvSTARD17*, and *VvSTARD19*) were obvious upregulation compared with control. The members of *VvSTARD* gene family could obviously respond to high-salt stress conditions in grape.

Subcellular localization and the heterologous expression of STARD5

A fusion protein of *VvSTARD5* and GFP were introduced into *Arabidopsis* protoplasts to determine expression site of *VvSTARD5* (Fig. 7A). The *VvSTARD5* gene was amplified and recombinational *Agrobacterium* identification (Fig.S5A-S5C). Confocal microscopy revealed that the 35S::*VvSTARD5*-EGFP fluorescence signal was localized at the nucleus (Figs. 7B). Transgenic tomato plants were obtained by *Agrobacterium* medicating leaf disc method and PCR identification (Figs.S6A-S6H). The overexpression recombine vector map and salt-tolerant phenotype of wild-type (WT) and transgenic tomatoes are shown in Figs. 7C and 7D. Combined with the result of the qRT-PCR analysis, *VvSTARD5* showed a high level of expression under 24 h salt treatment (Fig. 7E). In addition, the relative electrical conductivity, malondialdehyde and proline contents of WT and transgenic tomato under salt stress were measured. The results showed that the relative electrical conductivity of transgenic tomatoes were significantly (p<0.01) lower than WT (Fig. 7F). Moreover, the contents of proline were significantly (p<0.01) higher than WT. However, the content of malondialdehyde were lower than WT (Fig. 7H). These results showed that the ectopic overexpression of *STARD5* could significantly enhance the salt tolerance of tomatoes plants.

Discussion

Prior studies that have highlighted the importance of the HD–Zip transcription factor family (Ding et al. 2017; Chen et al. 2017; Zhang et al. 2020). However, a few studies have explored the HD–Zip III and the HD–Zip IV subfamily members containing the START domain proteins on plant salt stress. In addition, studies on proteins containing START and PH–START domains in grapes have not been conducted. Therefore, this study focused on the *STARD* gene identification and salt stress tolerance in grape.

Previous studies have reported that there have 35 and 25 members in *Arabidopsis* and rice, respectively. (Schrick et al. 2004). In the current study, 23 *VvSTARD* genes were found from the grape genome database, and the number of *START* genes is less in grape than *Arabidopsis* and rice. These results demonstrated that the number of *STARD* genes does not correlated with the genome size of the plant species, which may partly result from tandem and segmental duplication in grapes. On the basis of previous studies, the members of *STARD* gene families in *Arabidopsis*, rice and grape, namely, *AtSTARD1– AtSTARD35*, *OsSTARD1– OsSTARD25*, and *VvSTARD1– VvSTARD23*, are renamed in accordance with the sequence of the gene containing the conserved domains and the position of the gene on the chromosome. The *STARD* gene family is divided into five subfamilies (groups 1–5) on the basis of the inclusion of HD–START, HD–START–MEKHLA, START, and PH–START–DUF1336 structural domains (Soccio et al. 2003; Schrick et al. 2004; Clark 2012). The results showed that the positions of 23 *VvSTARD* genes on chromosomes are different, and the most relevant members in the same subfamily have similar exons/introns. Moreover, some differences in physical and chemical properties are observed in different groups. These results are consistent with those in previous studies (Schrick et al. 2004; Hill et al. 2016; Zhang et al. 2020). The analysis of the tertiary structure shows that eight major functional proteins, namely, MLN64 (Soccio et al. 2003; Murcia et al. 2006), PCTP (Alpy et al. 2005; Krisko et al. 2017), BFIT2 (Adams et al. 2001; Chen et al. 2012), CERT (Kudo et al. 2008; Agaisse et al. 2014), cholesterol-regulated START protein4 (Tan et al. 2019), START protein3 (Vassilev et al. 2015), START protein5 (Lorin et al. 2013), and START protein13 (Zhou et al. 2017), which are verified by the *VvSTARD* gene family, and these findings are similar to the results of previous studies (Schrick et al. 2004).

Various abiotic stresses are related to the expansion of some genes because of tandem and segmental duplication events (Cannon et al. 2004; Lynch et al. 2000; Raes et al. 2003; Otto et al. 2002; Duarte et al. 2006; Wang et al. 2010; Finet et al. 2013). For example, the expression of *VvSTARD15* is 20-fold higher than that of the control when the plant is exposed to low temperature stress, whereas *VvSTARD14* is not difference to low temperature stress. In addition, gene duplication, through either segmental or tandem duplication, played important roles in the expansion of new members during the evolution of a gene family

(Holub 2001). Synteny analysis of the *VvSTARD* gene family reveals four pairs of tandem duplication genes distributed in a common subfamily, the results probably because certain fragments of the gene have been copied, exchanged, inverted, and changed during evolution and other events (Shen et al. 2014; Li et al. 2017). In addition, the synteny analysis of grapes and *Arabidopsis* shows that 14 pairs of synteny gene are distributed in same subfamily, and only one pair of genes (*VvSTARD12*/ *AtSTARD33*) doesn't belong to the same group, *VvSTARD12* belongs to group 3, and the *AtSTARD33* belongs to group 2. Synteny analysis of grapes and rice has revealed nine pairs of synteny genes distributed in the same subgroup (Fig. 3B and Table S6). The Ka/Ks analysis suggests that the evolution of the grapes, *Arabidopsis*, and rice *STARD* gene families is primarily a purification choice (Yang, 2007; Wang et al. 2018).

Previous studies showed that the proteins of possess the START domain include the HD–Zip III, HD–Zip IV, PH–START and the START subfamilies. However, studies on these proteins under abiotic stress are relatively few. The members of the *STARD* gene family are analysed using the evolution and the tertiary structure analyses. The analyses of transgenic *Arabidopsis* plants carrying the gene-specific promoter fused to the bacterial β-glucuronidase reporter gene have revealed that some of the promoters have high activities in the epidermal layer of SAM and developing shoot organs, whereas others are temporarily active during the development of the reproductive organ (Nakamura et al. 2006; Khosla et al. 2014). However, the main functions of *STARD* genes in plants remain unclear. The HD–Zip genes of subfamilies III and IV encode an additional conserved domain call the START domain (Ponting et al. 1999), which has a putative function in sterol binding (Schrick et al. 2004).

In this study, the members of groups 1–3 belong to the *HD–Zip* genes of subfamilies III and IV, according to the accession numbers of grape, such as group 1 members *HDZ8* (*GSVIVT01035612001*), *HDZ19* (*GSVIVT01012643001*). *HDZ20* (*GSVIVT01030605001*) and *HDZ26* (*GSVIVT01027508001*); group 2 members *HDZ6* (*GSVIVT01013073001*), *HDZ10* (*GSVIVT01035238001*), *HDZ16* (*GSVIVT01017073001*), and *HDZ31* (*GSVIVT01029396001*); and group 3 members *HDZ11* (*GSVIVT01025193001*), *HDZ15* (*GSVIVT010170701001*), *HDZ18* (*GSVIVT01021625001*), *HDZ21* (*GSVIVT01016272001*) and *HDZ29* (*GSVIVT01010600001*). This result suggests that the HD–Zip IV has a potential role in defense in response ethylene (Li et al. 2017). The members of HD–Zip I and HD–Zip are reported to be related to salt stresses in *Eucalyptus* (Zhang et al. 2020). *EgHD–Zip27* from the HD–Zip II subfamily and *EgHD–Zip37* from the HD–Zip I subfamily play an essential role in cope with salt stress (Zhang et al. 2020), but the members of HD–Zip III and HD–Zip IV with salt stress are not mentioned.

In this study, the *VvSTARD5* (*HDZ20*) from the HD–Zip IV subfamily plays key roles in salt stress. In addition, the present study has described the functional characterization of the PH–START protein *AtAPO1* (*Arabidopsis APOSTART1*), indicating that the *AtAPO1* is involved in the control of seed germination (Resentini et al. 2014), whereas plants withstand drought and low temperature conditions. However, in the present study, the expression of PH–START proteins VvSTARD14 and VvSTARD15 are upregulated under salt and cold stresses, and HD–START proteins can also exhibit high expression levels under high-salt stress conditions. For instance, the HD–Zip IV subfamily member *VvSTARD5* has high expression level under salt stress (Li et al. 2017). Moreover, members with only one START domain have low or even no expression under high-salt stress conditions (Fig. 6). The relative electrolyte leakage serves as an indicator for the damage caused by salt stress (Cao et al. 2007), and the proline and the malondialdehyde contents can change under the salt stress in plants (Fedina et al. 2002). Therefore, the relative electrolyte leakage, proline and malondialdehyde contents are determined from tomato leaves of overexpression *VvSTARD5*, the results showed that the relative electrolyte leakage and malondialdehyde were lower in transgenic tomato plants than WT, but the content of proline significant increase. These results demonstrated that the overexpression of *VvSTARD5* can reduce salinity leading to cell membrane damage of leaves, and increasing transgenic tomato plants salt tolerance. The data from the present study strongly indicates the important functions of *VvSTARD* genes in response to salt stress.

Conclusion

In this study, 23 *VvSTARD* genes are identified in grape. Subsequently, these genes are divided into five subgroups and disseminated broadly on 12 chromosomes of grape genomes. Different expression pattern in the function of *STARD* genes are found amongst grape, *Arabidopsis* and rice. and the majority of *VvSTARD* genes can response salt stress. In addition, the *VvSTARD5* can increase salt tolerance in transgenic tomato. Therefore, the *VvSTARD* genes were identified and further explore its function for genetic improvement of agronomic traits of grapes.

Abbreviations

Gene ID, gene identification number; NJ, neighbour-joining method; CDS, coding sequence length; NC, number of codons; FOP, frequency of optimal codons; CAI, codon adaptation index; CBI, codon bias index; RSCU, relative synonymous codon usage; Ks, synonymous; Ka, nonsynonymous; MW, molecular weight; pI, isoelectric point; GRAVY, grand average of hydropathicity; II, instability index; AI, aliphatic index; qRT-PCR, quantitative reverse-transcription polymerase chain reaction; WT, wild-type tomato plants; #1, #2, #4, overexpression plants; CDS, coding sequence; GEO, gene expression omnibus; ABA, abscisic acid; MeJA, methyl jasmonate; SA, salicylic acid; IAA, indole acetic acid; GA3, gibberellin 3; PEG6000, polyethylene glycol 6000; GS, modified B5 solid medium; SAM, shoot apical meristem.

Declarations

Acknowledgements

This work was supported by the FuXi Foundation of Gansu Agricultural University (No. Ganfx-03J02), Youth Innovation and Entrepreneurship Talent Project of Longyuan (2018LYQN01), Discipline Construction Fund Project of Gansu Agricultural University (GSAU-XKJS-2018-226) and the Science and Technology Major Project of Gansu Province (18ZD2NA006).

Author Contributions

HHH, SXL, and JM designed the experiments, coordinated, and organized the whole research activities. HMG, QZ, XJC, PW, SXL, ZHM, participated in most of the experiments and data collection. HMG, QZ, XJC, PW, SXL, ZHM, provided technical assistance to HHH. HHH wrote the manuscript with contributions from all the authors. BHC and JM revised the manuscript. All authors read, reviewed, and approved the final manuscript.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

References

1. Adams SH, Chui C, Schilbach SL, Yu XX, Lewin DA (2001) Bfit, a unique acyl-CoA thioesterase induced in thermogenic brown adipose tissue: cloning, organization of the human gene and assessment of a potential link to obesity. Biochemical Journal 360: 135–142
2. Abe M, Katsumata H, Komeda Y, Takahashi T (2003) Regulation of shoot epidermal cell differentiation by a pair of homeodomain proteins in Arabidopsis. Development 130: 635–643
3. Alpy F, Tomasetto C (2005) Give lipids a START: the StAR-related lipid transfer (START) domain in mammals. Journal of Cell Science 118: 2791–2801
4. Agaisse H, Derre I (2014) Expression of the effector protein IncD in Chlamydia trachomatis mediate recruitment of the lipid transfer protein CERT and the endoplasmic reticulum-resident protein VAPB to the inclusion membrane. Infection and Immunity 82: 2037–2047
5. Baima S, Possenti M, Matteucci A, Wisman E, Altamura MM, Ruberti I, Morelli G (2001) The Arabidopsis ATHB-8 HD-zip protein acts as a differentiation-promoting transcription factor of the vascular meristems. Plant Physiology 126: 643–655
6. Bailey TL, Bodén M, Fabian A, Buske FA, Frith M, Grant CE, Clementi L, Ren JY, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Research 37: W202–W208
7. Cheong YH, Kim KN, Pandey GK, Gupta R, Grant JJ, Luan S (2003) CBL1, a calcium sensor that differentially regulates salt, drought, and cold responses in Arabidopsis. Plant Cell 15: 1833–1845
8. Cannon SB, Mitra A, Baumgarten A, Young ND, May G (2004) The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. BMC Plant Biol 4: 10
9. Cao WH, Liu J, He XJ, Mu RL, Zhou HL, Chen SY, Zhang JS (2007) Modulation of ethylene responses affects plant salt-stress responses. Plant Physiology 143: 707-719
10. Clark BJ (2012) The mammalian START domain protein family in lipid transport in health and disease. Journal of Endocrinology 212:257–275
11. Chen D, Latham J, Zhao H, Bisoffi M, Farelli J, Dunaway-Mariano D (2012) Human brown fat inducible thioesterase variant 2 cellular localization and catalytic function. Biochemistry 51: 6990–6999
12. Chen DM, Chen Z, Wu M, Wang Y, Wang YJ, Yan HW, Xiang Y (2017) Genome-Wide Identification and Expression Analysis of the HD-Zip Gene Family in Moso Bamboo (Phyllostachys edulis). Journal of Plant Growth Regulation 36:323-337
13. Chen CJ, Xia R, Chen H, He YH (2018) TBtools, a Toolkit for Biologists integrating various HTS-data 2 handling tools with a user-friendly interface. bioRxiv preprint first posted online. DOI: https://doi.org/10.1101/289660
14. Clark BJ (2020) The START-domain proteins in intracellular lipid transport and beyond. Mol Cell Endocrinol 504: 110704
15. DeLano WL (2002) The PyMOL Molecular Graphics System. DeLano Scientific, Palo Alto, CA, USA. http://www.pymol.org
16. Duarte JM, Cui L, Wall PK, Zhang Q, Zhang X, Leebensmack J, Ma H, Altman N, Depamphilis CW (2006) Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of Arabidopsis. Mol. Biol. Evol 23: 469–478
17. Ding ZH, Fu LL, Yan Y, Tie WW, Xia ZQ, Wang WQ, Peng M, Hu W, Zhang JM (2017) Genome-wide characterization and expression profiling of HD-Zip gene family related to abiotic stress in cassava. PLoS One 12
18. Emery JF, Floyd SK, Alvarez J, Eshed Y, Hawker NP, Izhaki A, Baum SF, Bowman JL (2003) Radial patterning of Arabidopsis shoots by class III HD-ZIP and KANADI genes. Current Biology 13: 1768–1774
19. Elhiti M, Stasolla C (2009) Structure and function of homeodomain-leucine zipper (HD-Zip) proteins. Plant Signaling and Behavior 4: 86–88
20. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE, Finn RD (2019) The Pfam protein families database in 2019. Nucleic Acids Res 47: D427–D432
21. Fedina IS, Georgieva K, Grigorova I (2002) Light-Dark Changes in Proline Content of Barley Leaves under Salt Stress. Biologia Plantarum 45:59-63
22. Fujimoto R, Kinoshita Y, Kawabe A, Kinoshita T, Takashima K, Nordborg M, Nasrallah ME, Shimizu KK, Kudoh H, Kakutani T (2008) Evolution and control of imprinted FWA genes in the genus *Arabidopsis*. PLoS Genetics 4: e1000048
23. Finet C, Bernededieu A, Scutt CP, Marlétaz F (2013) Evolution of the ARF Gene Family in Land Plants: Old Domains, New Tricks. Mol. Biol. Evol 30:45–56
24. Goodstein DM, Shu SQ, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 40: D1178–D1186
25. Guo Y, Liu J, Zhang J, Liu S, Du J (2017) Selective modes determine evolutionary rates, gene compactness and expression patterns in Brassica. Plant J 91: 34–44
26. Hiroyoshi K, Anton JM, Peeter S, Mark GMA, Andy P, Maarten K (1999) ANTHOCYANINLESS2, a homeobox gene affecting anthocyanin distribution and root development in Arabidopsis. The Plant Cell 11: 1217–1226
27. Hershberg R, Petrov DA (2008) Selection on codon bias, Annu. Rev. Genet 42: 287–299

28. Hu B, Jin JP, Guo AY, Zhang H, Luo JC, Gao G (2015) GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31: 1296–1297
29. Hill R J, Ringel A, Knuepfer E, Moon RW, Blackman MJ, van Ooij C (2016) Regulation and Essentiality of the StAR-related Lipid Transfer (START) Domain-containing Phospholipid Transfer Protein PFA0210c in Malaria Parasites. Journal of Biological Chemistry 291: 24280-24292
30. Ingram GC, Boisnard-Lorig C, Dumas C, Rogowsky PM (2000) Expression patterns of genes encoding HD-ZipIV homeo domain proteins define specific domains in maize embryos and meristems. Plant J 22401–14
31. Kubo H, Peeters AJM, Aarts MGM, Koornneef PM (1999) Anthocyaninless2, a homeobox gene affecting anthocyanin distribution and root development in Arabidopsis. The Plant Cell 11: 1217–1226
32. Kudo N, Kumagai K, Tomishige N, Yamaji T, Wakatsuki S, Nishijima M, Hanada K, Kato R (2008) Structural basis for specific lipid recognition by CERT responsible for nonvesicular trafficking of ceramide. Proceedings of the National Academy of Science of the United States of America 105: 488–493
33. Khosla A, Paper JM, Boehler AP, Bradley AM, Neumann TR, Schrick K (2014) HD-Zip Proteins GL2 and HDG11 Have Redundant Functions in Arabidopsis Trichomes, and GL2 Activates a Positive Feedback Loop via MYB23. The Plant Cell 26: 2184–2200
34. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol 33: 1870–1874
35. Krisko TI, LeClair KB, Cohen DE (2017) Genetic ablation of phosphatidylcholine transfer protein/StarD2 in ob/ob mice improves glucose tolerance without increasing energy expenditure. Metabolism 68: 145–149
36. Lynch M, Conery JS (2000) The Evolutionary Fate and Consequences of Duplicate Genes. Science 290: 1151–1155
37. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, 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
38. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947–2948
39. Larracuente AM, Sackton TB, Greenberg AJ, Wong A, Singh ND, Sturgill D, Zhang Y, Oliver B, Clark AG (2008) Evolution of protein-coding genes in Drosophila. Trends Genet 24: 114–123
40. Lorin A, Letourneau D, Lefebvre A, LeHoux J G, Lavigne P (2013) (1)H, (13)C, and (15)N backbone chemical shift assignments of StAR-related lipid transfer domain protein 5 (STARD5). Biomol NMR Assign 7: 21–24
41. Li X, Liu G, Geng Y, Wu M, Pei W, Zhai H, Zang X, Li X, Zhang J, Yu S, Yu J (2017) A genome-wide analysis of the small auxin-up RNA (*SAUR*) gene family in cotton. BMC Genomics 18: 815
42. Li Z, Zhang C, Guo Y, Niu W, Wang Y, Xu Y (2017) Evolution and expression analysis reveal the potential role of the HD-Zip gene family in regulation of embryo abortion in grapes (*Vitis vinifera* L.) BMC Genomics 18: 744
43. Li LL, Zheng TC, Zhuo XK, Li SZ, Qiu LK, Wang J, Cheng TR, Zhang QX (2019) Genome-wide identification, characterization and expression analysis of the HD-Zip gene family in the stem development of the woody plant Prunus mume. Peerj 7.
44. Mcconnell JR, Emery J, Eshed Y, Bao N, Bowman J, Barton MK (2001) Role of Phabulosa and Phavoluta in determining radial patterning in shoots. Nature 411: 709–713
45. Murcia M, Faraldo-Gomez JD, Maxfield FR, Roux B (2006) Modeling the structure of the StART domains of MLN64 and StAR proteins in complex with cholesterol. Journal of Lipid Research 47: 2614–2630
46. Nakamura M, Katsumata H, Abe M, Yabe N, Komeda Y, Yamamoto Kt, Takahashii T (2006) Characterization of the class IV homeodomain-leucine zipper gene family in Arabidopsis. Physiologia Plantarum 141:1363–1375
47. Otto SP, Yong P (2002) The evolution of gene duplicates. Adv. Genet 46: 451–483
48. Ponting CP, Aravind L (1999) START: a lipid-binding domain in StAR, HD-ZIP and signaling proteins. Trends in Biochemical Sciences 24:130–132
49. Plotkin JB, Kudla G (2011) Synonymous but not the same: The causes and consequences of codon bias, Nature reviews. Genetics 12:32–42
50. Prashek J, Bouyain S, Fu MG, Li Y, Berkes D, Yao XL (2017) Interaction between the PH and START domains of ceramide transfer protein competes with phosphatidylinositol 4-phosphate binding by the PH domain. Journal of Biological Chemistry 292:14217-14228
51. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD (2018) HMMER web server: 2018 update. Nucleic Acids Research 46: W200–W204
52. Ruf S, Hermann M, Berger IJ, Carrer H, Bock R (2001) Stable genetic transformation of tomato plastids and expression of a foreign protein in fruit. Nature Biotechnology,19:870–875
53. Roderick SL, Chan WW, Agate DS, Olsen LR, Vetting MW, Rajashankar KR, Cohen DE (2002) Structure of human phosphatidylcholine transfer protein in complex with its ligand. Nature Structural & Molecular Biology 9:507–511
54. Raes J, Vandepoele K, Simillion C, Saeys Y, Van de Peer Y (2003) Investigating ancient duplication events in the Arabidopsis genome. J. Struct. Funct. Genom 3: 117–129
55. Ryo F, Yuki K, Akira K, Tetsu K, Kazuya T, Magnus N, Mikhail E, Nasrallah A, Kentaro KS, Hiroshi K, Tetsuji K (2008) Evolution and control of imprinted FWA genes in the genus Arabidopsis. PLoS Genetics 4: e1000048
56. Resentini F, Vanzulli S, Marconi G, Colombo L, Albertini E, Masiero S (2014) AtAPOSTART1, an *Arabidopsis thaliana* PH-START domain protein involved in seed germination. Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology 148(6): 1178-1186.
57. Szymanski DB, Jilk RA, Pollock SM, Marks MD (1998) Control of GL2 expression in Arabidopsis leaves and trichomes. Development 125: 1161–1171
58. Sessions A, Weigel D, Yanofsky MF (1999) The *Arabidopsis thaliana* MERISTEM LAYER1 promoter specifies epidermal expression in meristems and young primordia. The Plant Journal 20: 259–263
59. Stocco DM (2001) StAR protein and the regulation of steroid hormone biosynthesis. Annu Rev Physiol 63: 193–213

60. Soccio RE, Breslow JL (2003) StAR-related Lipid Transfer (START) Proteins: Mediators of Intracellular Lipid Metabolism. Journal of Biological Chemistry 278: 22183–22186
61. Strauss JF, Kishida T, Christenson LK, Fujimoto T, Hiroi H (2003) START domain proteins and the intracellular trafficking of cholesterol in steroidogenic cells. Molecular and Cellular Endocrinology 202: 59–65
62. Shi H, Lee BH, Wu SJ, Zhu JK (2003) Overexpression of a plasma membrane Na1/H1 antiporter gene improves salt tolerance in *Arabidopsis thaliana*. Nat Biotechnol 21: 81–85
63. Schrick K, Nguyen D, Karlowski WM, Mayer KF (2004) START lipid/sterol-binding domains are amplified in plants and are predominantly associated with homeodomain transcription factors. Genome Biology 5: 50–67
64. Schrick K, Bruno M, Khosla A, Cox PN, Marlatt SA, Roque RA, Nguyen HC, He CW, Snyder MP, Singh D, Yadav G (2014) Shared functions of plant and mammalian StAR-related lipid transfer (START) domains in modulating transcription factor activity. BMC Biology 12: 70–90
65. Shen C, Yue R, Yang Y, Zhang L, Sun T, Xu L, Tie S, Wang H (2014) Genome-wide identification and expression profiling analysis of the Aux/IAA gene family in Medicago truncatula during the early phase of Sinorhizobium meliloti infection. PLoS One 9: e107495
66. Talbert PB, Adler HT, Parks DW, Comai L (1995) The REVOLUTA gene is necessary for apical meristem development and for limiting cell divisions in the leaves and stems of *Arabidopsis thaliana*. Development 121: 2723–2735
67. Tsujishita Y, Hurley JH (2000) Structure and lipid transport mechanism of a StAR-related domain. Nature Structural and Molecular Biology, 7: 408–414
68. Tang D, Ade J, Frye CA, Innes RW (2005) Regulation of plant defense responses in Arabidopsis by EDR2, a PH and START domain-containing protein. Plant J 44:245-257
69. Tan L, Tong J, Chun C, Im YJ (2019) Structural analysis of human sterol transfer protein STARD4. Biochem Biophys Res Commun 520: 466–472
70. Tillman MC, Imai N, Li Y, Khadka M, Okafor CD, Juneja P, Adhiyaman A, Hagen SJ, Cohen DE, Ortlund EA (2020) Allosteric regulation of thioesterase superfamily member 1 by lipid sensor domain binding fatty acids and lysophosphatidylcholine. Proc Natl Acad Sci USA 117: 22080-22089.
71. Vassilev B, Sihto H, Li S, Holtta-Vuori M, Ilola J, Lundin J, Isola J, Kellokumpu-Lehtinen PL, Joensuu H, Ikonen E (2015) Elevated levels of StAR-related lipid transfer protein 3 alter cholesterol balance and adhesiveness of breast cancer cells: potential mechanisms contributing to progression of HER2-positive breast cancers. Am J Pathol 185: 987–1000
72. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Hochstrasser DF (1999) Protein Identification and Analysis Tools in the ExPASy Server. Methods Mol Biol 112: 531–552
73. Williams L, Grigg SP, Xie M, Christensen S, Fletcher JC (2005) Regulation of Arabidopsis shoot apical meristem and lateral organ formation by microRNA miR166g and its AtHD-ZIP target genes. Development. 132:3657–68
74. Willems E, Leyns L, Vandesompele J (2008) Standardization of real-time PCR gene expression data from independent biological replicates. Anal Biochem 37: 127–129
75. Wang Y, Deng D, Bian Y, Lv Y, Xie Q (2010) Genome-wide analysis of primary auxin-responsive Aux/IAA gene family in maize (Zea mays L.). Mol. Biol. Rep 37: 3991–4001
76. Wang Y, Tang H, DeBarry JD, Tan X., Li J, Wang X, Lee T-h, Jin H, Marler B, Guo H (2012) MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res 40: e49
77. Wang P, Gao C, Bian X, Zhao S, Zhao C, Xia H, Song H, Hou L, Wan S, Wang X (2016) Genome-wide identification and comparative analysis of cytosine-5 DNA methyltransferase and demethylase families in wild and cultivated peanut. Front Plant Sci 7: 7
78. Wang PF, Su L, Gao HH, Jiang X, Wu XY, Li Y, Zhang QQ, Wang YM, Ren FS (2018) Genome-Wide Characterization of bHLH Genes in Grape and Analysis of their Potential Relevance to Abiotic Stress Tolerance and Secondary Metabolite Biosynthesis. Front Plant Sci 9: 64
79. Yoo SD, Cho YH, Sheen J (2007) Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. Nature 2: 1565–1572
80. Yang, Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591
81. Yu L, Chen X, Wang Z, Wang S, Wang Y, Zhu Q, Li S, Xiang C (2013) Arabidopsis enhanced drought tolerance1/HOMEODOMAIN GLABROUS11 confers drought tolerance in transgenic rice without yield penalty. Plant Physiol, 162:1378–1391
82. Zhou GQ, Liu XM, Xiong B, Sun YF (2017) Homeobox B4 inhibits breast cancer cell migration by directly binding to StAR-related lipid transfer domain protein 13. Oncology Letters 14: 4625–4632
83. Zhang JS, Wu JZ, Guo ML, Aslam M, Wang Q, Ma HY, Li SB, Zhang XT, Cao SJ (2020) Genome-wide characterization and expression profiling of Eucalyptus grandis HD-Zip gene family in response to salt and temperature stress. Bmc Plant Biology 20: 451

Table

Table 1. The characteristic of START domain-encoding genes

Gene name	GenBank accession numbers	Gene accession No.	Position	Location	Structure	CDS (bp)	Peptide (aa)	Mw (kD)	GRAVY	pI	II	A
VvSTARD1	XM_002277637	GSVIVT01013073001	8739164-8746005	2	HD-START	2397	799	88.50	-0.341	5.65	54.35	8
VvSTARD2	XM_002268236	GSVIVT01035238001	10997910-11025564	4	HD-START	2145	714	78.77	-0.258	5.95	43.56	8
VvSTARD3	XM_002284466	GSVIVT01017073001	3967669-3972062	9	HD-START	2253	750	82.96	-0.466	6.09	48.69	7
VvSTARD4	XM_002266652	GSVIVT01012643001	300629-304962	10	HD-START	2181	726	79.67	-0.31	5.91	43.11	8
VvSTARD5	XM_010659009	GSVIVT01030605001	7101002-7106874	12	HD-START	2274	757	82.70	-0.289	5.6	41.78	7
VvSTARD6	XM_002272228	GSVIVT01027508001	16132617-16138896	15	HD-START	2316	771	83.91	-0.231	5.86	49.68	8
VvSTARD7	CBI31820	GSVIVT01010600001	16145022-16162952	16	HD-START	2658	886	99.56	-0.306	6.48	46.32	8
VvSTARD8	XM_002270976	GSVIVT01029396001	16227031-16230509	17	HD-START	2025	674	74.79	-0.367	6.16	48.05	8
VvSTARD9	XM_002283681	GSVIVT01035612001	2700607-2710044	4	HD-START-MEKHLA	2520	839	92.47	-0.155	5.8	51.04	8
VvSTARD10	XM_010652862	GSVIVT01025193001	3507793-3517339	6	HD-START-MEKHLA	2535	844	92.25	-0.093	5.87	48.36	8
VvSTARD11	XM_002283967	GSVIVT01017010001	3414695-3425101	9	HD-START-MEKHLA	2508	835	91.99	-0.142	6.06	48.4	8
VvSTARD12	XM_002281832	GSVIVT01021625001	8333352-8341424	10	HD-START-MEKHLA	2538	845	92.84	-0.161	5.93	51.31	8
VvSTARD13	XM_002274158	GSVIVT01016272001	5640475-5655021	13	HD-START-MEKHLA	2523	841	91.83	-0.131	6.28	51.49	8
VvSTARD14	XM_010652877	GSVIVT01025201001	3373992-3381809	6	PH-START-DUF1336	2205	734	82.96	-0.417	7.56	48.19	7
VvSTARD15	XM_002274017	GSVIVT01016264001	5771590-5801747	13	PH-START-DUF1336	2187	728	83.17	-0.464	7.02	46.5	7
VvSTARD16	XM_010666453	GSVIVT01022620001	13478753-13525358	2	START-DUF1336	2133	710	81.23	-0.472	6.32	40.81	8
VvSTARD17	XM_002262825	GSVIVT01022623001	13595184-13673525	2	START-DUF1336	2205	734	84.15	-0.47	6.32	36.41	7
VvSTARD18	XM_010658253	GSVIVT01001043001	6787163-6911536	11	START-DUF1336	2289	762	85.80	-0.426	6.45	37.46	8
VvSTARD19	CBI38216	GSVIVT01011334001	8665113-8695586	15	START-DUF1336	1923	641	72.54	-0.077	5.81	37.64	9
VvSTARD20	XM_002278741	GSVIVT01019684001	2435453-2439133	2	START	1155	384	42.47	-0.263	7.46	52.29	5
VvSTARD21	XM_010651704	GSVIVT01018120001	6714620-6724790	5	START	1215	404	45.93	-0.388	9.51	42.1	7
VvSTARD22	XM_010656935	GSVIVT01029461001	22597050-22602216	9	START	1068	355	40.15	-0.374	9.66	46.55	8
VvSTARD23	XM_002272438	GSVIVT01023115001	22403922-22405945	12	START	714	237	26.77	-0.22	6.75	67.4	9

Notes: isoelectric point (pI), molecular weight (Mw), instability index (I.I), aliphatic index (A.I) and grand average of hydropathicity (GRAVY)

Figures

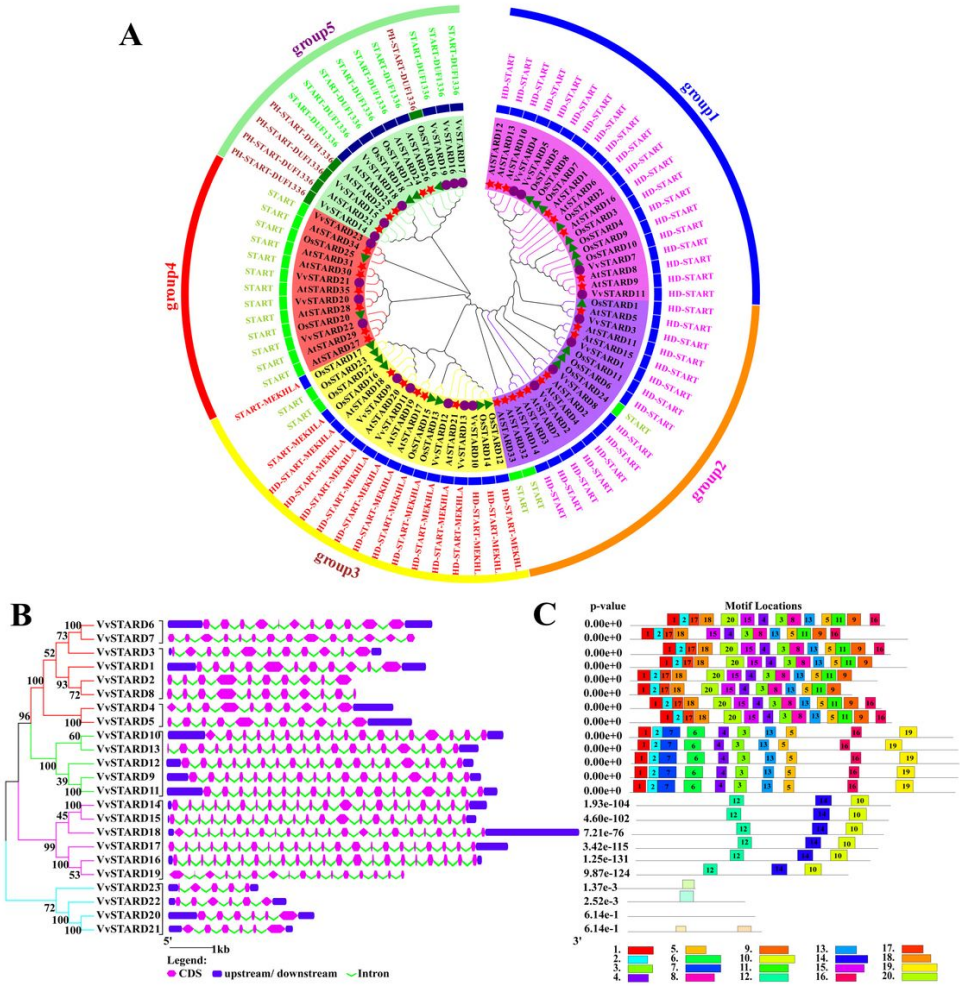

Figure 1

Phylogenetic relationships, gene structure and architecture of conserved protein motifs in VvSTARD proteins. **A** Phylogenetic tree analysis of STARD proteins in grapes (Vv), *Arabidopsis* (At) and rice (Os). Red five-pointed star represents *Arabidopsis*. Purple circle represents grape. Green triangle represents rice. **B** Exon–intron structure of *VvSTARD* genes. Blue boxes indicate untranslated 5'- and 3'-regions; pink boxes indicate exons; green lines indicate introns. **C** Motif composition of VvSTARD proteins. The motifs, numbers 1–20, are displayed in different colored boxes.

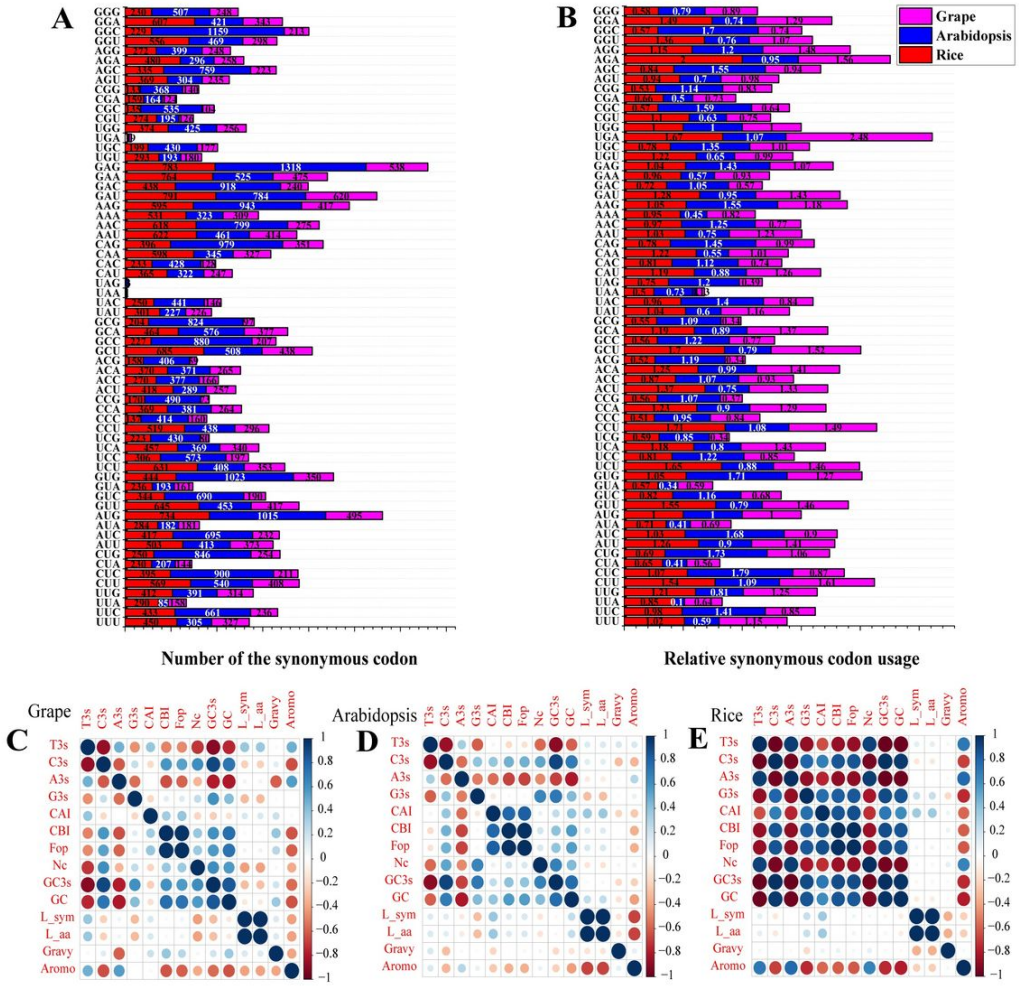

Figure 2

Synonymous codon preference and correlation analysis of *VvSTARD, AtSTARD* and *OsSTARD* genes. **A** The number of synonymous codon preference in grapes, *Arabidopsis* and rice. B The relative synonymous codon usage in grapes, *Arabidopsis* and rice. C Blue represents positive correlation; red represents negative correlation, and white represents no correlation. The larger circle with darker color is stronger correlation, and vice versa.

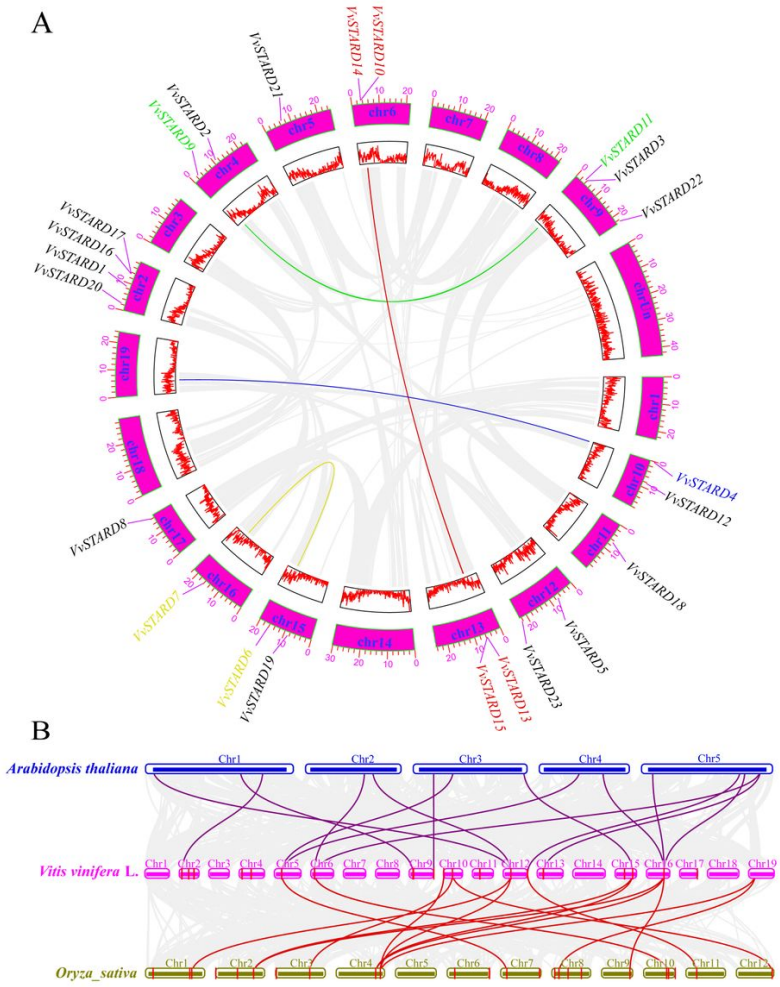

Figure 3

Inter chromosomal relationships of grape and synteny analysis of *STARD* genes between grapes and two representative plant species. **A** Chromosomal distribution and inter chromosomal relationships of *VvSTARD* genes. Gray lines indicate all synteny blocks in the grape genome, and the red, green, blue, and yellow lines indicate duplicated *STARD* gene pairs. The chromosome number is indicated at the bottom of each chromosome. **B** Synteny analysis of *STARD* genes among *Arabidopsis*, grapes, and rice. Gray lines in the background indicate the collinear blocks within *Arabidopsis*, grapes, and rice genomes, whereas the purple line highlights the syntenic *STARD* gene pairs in grapes and *Arabidopsis*,, and the red line highlights the syntenic *STARD* gene pairs in grapes and rice.

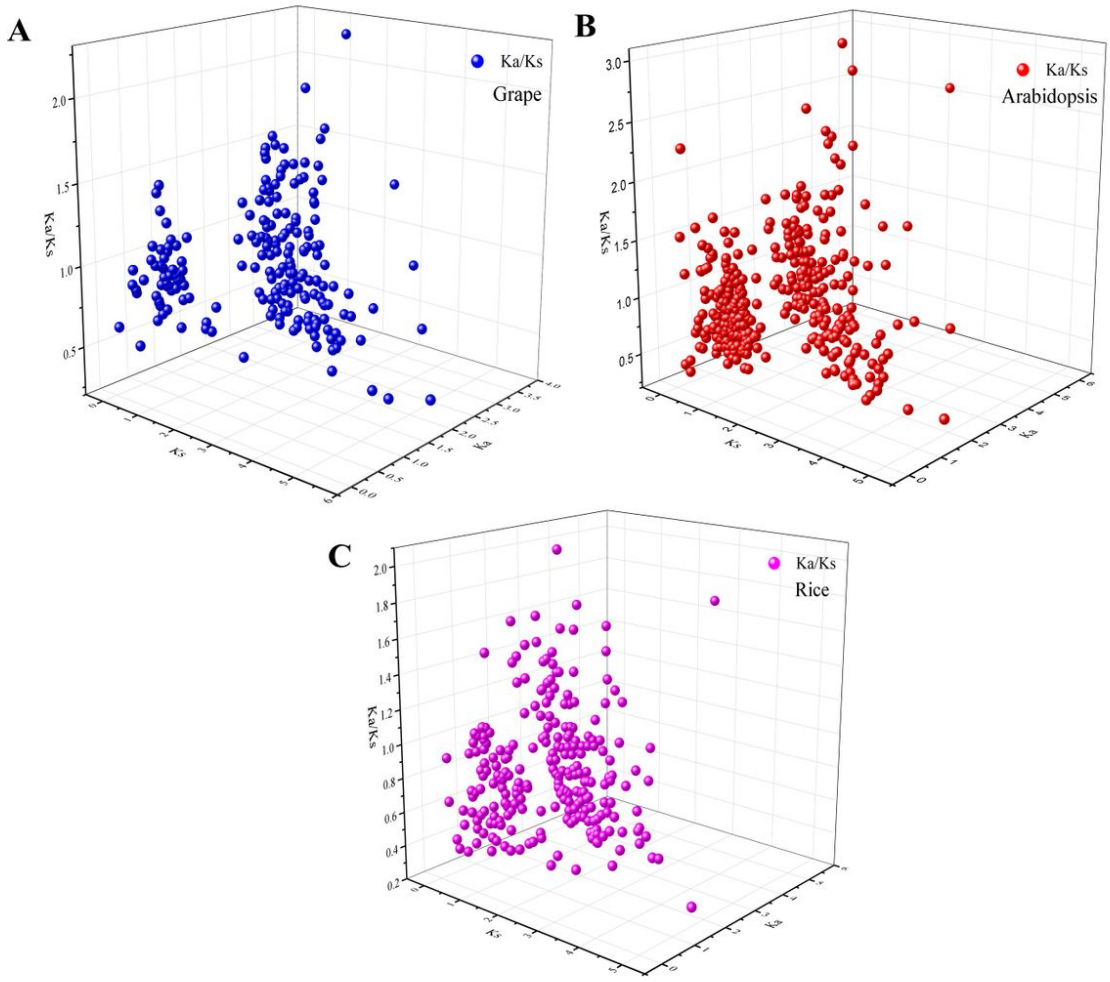

Figure 4

Ka/Ks analysis of *STARD* genes. **A** Ka/Ks analysis of grapes. **B** Ka/Ks analysis of *Arabidopsis*. **C** Ka/Ks analysis of rice.

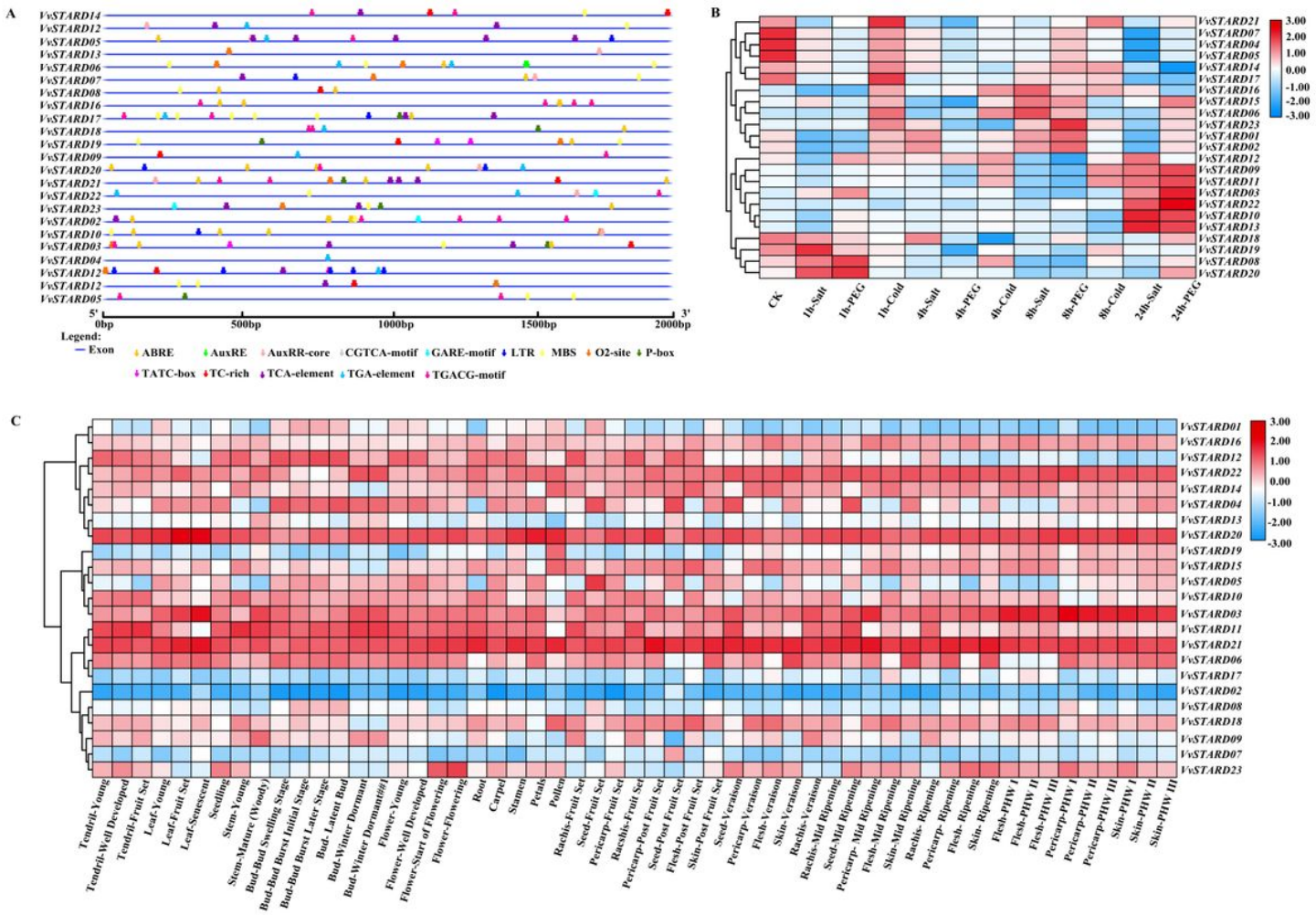

Figure 5

cis-Element of *STARD* genes and hierarchical clustering the expression profiles in grapes. **A** Major stress- and hormone-related *cis*-elements in the promoters of *VvSTARD* genes. TC-rich repeats (defense and stress), MBS (drought), and LTR (low-temperature responsive) elements, TGA-element/AuxRE/AuxRR-core (auxin), O2-site (zein metabolism), TCA-element (salicylic acid), ABRE (abscisic acid), GARE-motif /P-box /TATC-box (gibberellin), and CGTCA-motif /TGACG-motif (MeJA responsive element) are represented by different colors, as indicated in figure legend at the bottom. **B** Hierarchical clustering of the expression profiles of 23 *VvSTARD* genes at different abiotic stress experiments in grape. Abiotic stress experiments: salt, PEG, and cold. Heatmap experiments were performed with Gene Chip microarrays, which were from Affymetrix Gene Chip 16K with short-term abiotic stress 'Cabernet Sauvignon.' Red or blue shading represented the upregulated or downregulated expression level, respectively. The scale denoted the relative expression level. **C** Hierarchical clustering of the expression profiles of 23 *VvSTARD* genes at different organizations experiments in grape. Heatmap experiments were performed with Gene Chip microarrays, which were from Grape eFP Browser in grape. Red or blue shading represented the upregulated or downregulated expression level, respectively. The scale denoted the relative expression level.

Note: Stamen, pool of stamens from undisclosed flowers at 10% and 50% open flowers; Berry Pericarp-FS, berry pericarp fruit set; Berry Pericarp-PFS, berry pericarp post-fruit set; Berry Pericarp-V, berry pericarp véraison; Berry Pericarp-MR, berry pericarp mid-ripening; Berry Pericarp-R, berry pericarp ripening; Bud-S, bud swell; Bud-B, bud burst (green tip); Bud-AB, bud after-burst (rosette of leaf tips visible); Bud-L, latent bud; Bud-W, winter bud; Berry Flesh-PFS, berry flesh post fruit set; Berry Flesh-V, berry flesh véraison; Berry Flesh-MR, berry flesh mid-ripening; Berry Flesh-R, berry flesh ripening; Berry Flesh-PHWI, berry flesh post-harvest withering I (1st month); Berry Flesh-PHWII, berry flesh post-harvest withering II (2nd month); Berry Flesh-PHWIII, berry flesh post-harvest withering III (3rd month); Inflorescence-Y, young inflorescence (single flower in compact groups); Inflorescence-WD, well developed inflorescence (single flower separated); Flower-FB, flowering begins (10% caps o); Flower-F, flowering (50% caps o); Root, root in vitro cultivation; Leaf-Y, young leaf (pool of leaves from shoot of 5 leaves); Leaf-FS, mature leaf (pool of leaves from shoot at fruit set); Leaf-S, senescence leaf (pool of leaves at the beginning of leaf fall); Carpel, pool of carpels from undisclosed flowers at 10% and 50% open flowers; Petal, pool of petals from undisclosed flowers at 10% and 50% open flowers; Berry Pericarp-PHWI, berry pericarp post-harvest withering I (1st month); Berry Pericarp-PHWII, berry pericarp post-harvest withering II (2nd month); Berry Pericarp-PHWIII, berry pericarp post-harvest withering III (3rd month); Pollen, pollen from disclosed flowers at more than 50% open flowers; Rachis-FS, rachis fruit set; Rachis-PFS, rachis post-fruit set; Rachis-V, rachis véraison; Rachis-MR, rachis mid-ripening; Rachis-R, rachis ripening; Seed-V, seed véraison; Seed-MR, seed mid-ripening; Seed-FS, seed fruit set; Seed-PFS, seed post-fruit set; Seedling, seedling pool of 3 developmental stages; Berry Skin-PFS, berry skin post-fruit set; Berry Skin-V, berry skin véraison; Berry Skin-MR, berry skin mid-ripening; Berry Skin-R, berry skin ripening; Berry Skin-PHWI, berry skin post-harvest withering I (1st

month); Berry Skin-PHWII, berry skin post-harvest withering II (2nd month); Berry Skin-PHWIII, berry skin post-harvest withering III (3rd month); Stem-G, green stem; Stem-W, woody stem; Tendril-Y, young tendril (pool of tendrils from shoot of 7 leaves); Tendril-WD, well developed tendril (pool of tendrils from shoot of 12 leaves); Tendril-FS, mature tendril (pool of tendrils at fruit set).

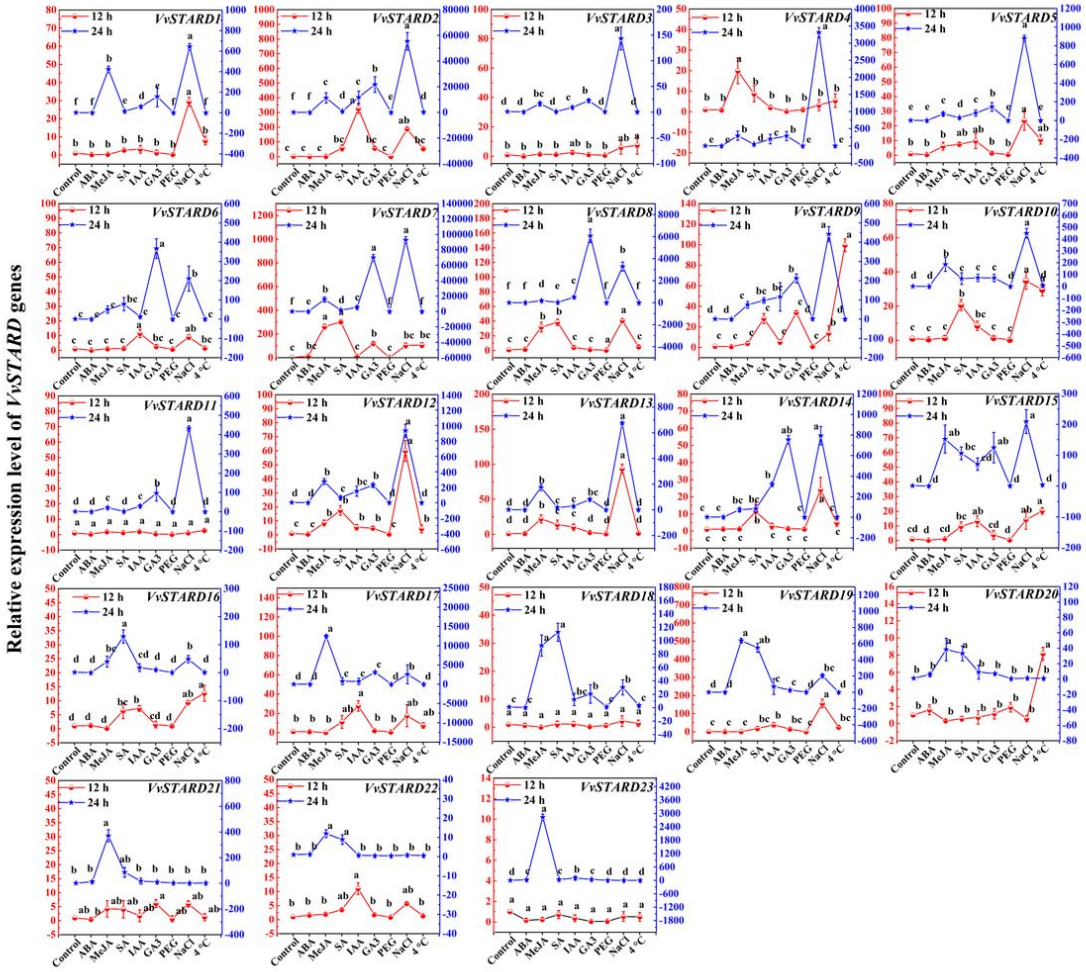

Figure 6

Expression levels of *VvSTARD* in grape leaves after 12 h and 24 h under different treatments. The treatment includes 0.2 mmol·l^{-1} ABA, 150 μmol·l^{-1} MeJA, SA, 50 mg·l^{-1} SA, 100 μmol·l^{-1} IAA, 50 mg·l^{-1} gibberellin 3 (GA3), 10% PEG6000, 400 mmol l^{-1} NaCl, 4 °C low temperature, and control. The red axis on the left represents 12 h treatment, and the blue axis on the right represents 24 h treatment. Values represent the average of three independent experiments ± SD. Standard errors are shown as bars above the columns. a, b, c, d, e, and f denote a significant difference at the level of $p < 0.05$.

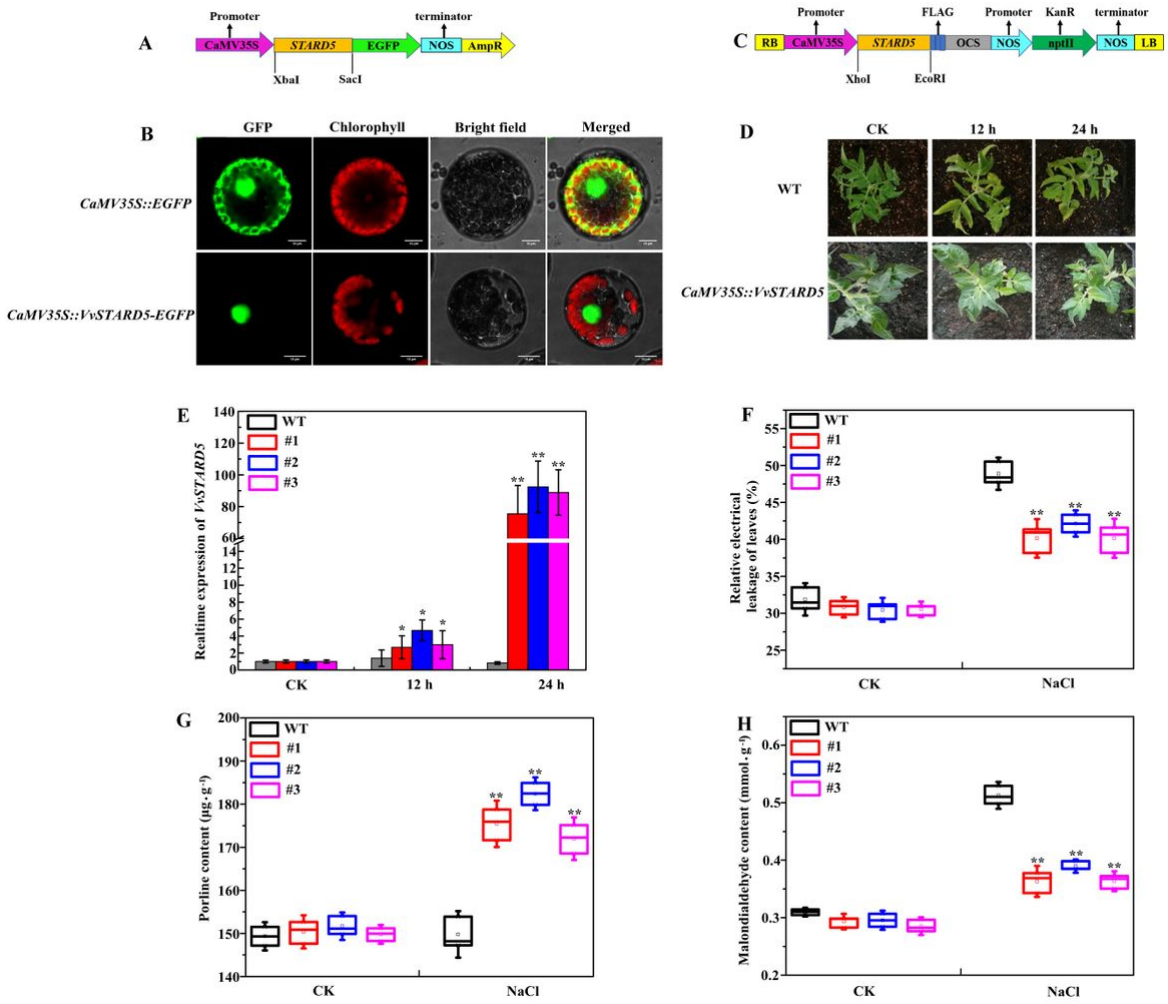

Figure 7

Subcellular localization and verification of salt stress resistance of transgenic tomato (*STARD5*). **A** Schematic of the 35S::*VvSTARD5*-EGFP construct. **B** Subcellular localization results of *STARD5* in *Arabidopsis* protoplast. **C** Schematic of the pCAMBIA1300-*VvSTARD5* construct. **D** The phenotypic difference of WT and transgenic tomato plants at different time periods under NaCl treatment. **E** The expression levels of *VvSTARD5* in WT and transgenic lines (#1, #2, and #4) under the 12 h or 24 h with 400 mmol·l^{-1} of NaCl. Asterisks (**) and (*) indicate significant differences compared with the CK (control) at P<0.01 and P<0.05 (Student's t-test), respectively. **F** Relative electrical conductivity (%) of tomato leaves of WT and transgenic lines (#1, #2, and #4) under natural conditions and after 24 h with 400 mmol·l^{-1} of NaCl. **G** The content of proline in tomato leaves of WT and transgenic lines (#1, #2, and #4) under natural conditions and after 24 h with 400 mmol·l^{-1} of NaCl. **H** The content of Malondialdehyde in tomato leaves of WT and transgenic lines (#1, #2, and #4) under natural conditions and after 24 h with 400 mmol·l^{-1} of NaCl**.** Values represent the means ± SD of three replicates. Asterisks (**) and (*) indicate significant differences of transgenic lines (#1, #2, and #4) compared with the WT at p <0.01 and p <0.05.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Fig.S1.tif
- Fig.S2.tif
- Fig.S3.tif
- Fig.S4.tif
- Fig.S5.tif
- Fig.S6.tif
- TableS1.docx
- TableS2.docx
- TableS3.docx

- TableS4.docx
- TableS5.xlsx
- TableS6.docx
- TableS7.docx
- TableS8.xlsx