Sari Agricultural & natural Resources University

JPMB

GABIT
Genetics and Agricultural Biotechnology Institute of Tabarestan

JOURNAL OF Plant Molecular Breeding

June 2023, Volume 11, Issue 1

ISSN: 2322-3332

Printed in February 2024

JPMB

Contact

Sari Agricultural Sciences and Natural Resources University, Genetics and Agricultural Biotechnology Institute of Tabarestan, Sari, P.O. Box: 578, Iran.
Tel: +98-11-33687577
Email: jpmbjournal@sanru.ac.ir
www.jpmb-gabit.ir

Message from the Editor-in-Chief

JPMB is an open-access journal that serves as an advanced forum for research findings in plant molecular genetics and breeding along with other related areas such as plant biology, physiology, taxonomy, stresses, and interactions with other organisms. The journal publishes original research articles, reviews, reports, conference proceedings (peer-reviewed full articles), and short communications. In original research papers, full experimental details must be provided. We also encourage the submission of lab protocols, data management protocols, and analytical procedures in sufficient detail on topics of interest to the plant research community.

Editorial Board

Aims and Scope

Journal of Plant Molecular Breeding | JPMB is an international, open-access, peer-reviewed, biannual scholarly publication that aims to offer comprehensive coverage of progress in the field of plant molecular breeding. It seeks to present findings to researchers, academics, and students, addressing the growing demand for applied plant improvement technologies, tools, and methodologies.

- Genetic engineering (transformation)
- Micropropagation and plant breeding
- Molecular markers and plant genetic diversity
- Plant functional genomics
- Germplasm assessments for genetic diversity
- Plant breeding technologies
- Plant mutation breeding
- Metabolomics and metabolites engineering in plant breeding
- Plant breeding for abiotic and biotic stresses
- Plant Molecular biology
- Plant Omics and Bioinformatics

Visit our main website https://www.jpmb-gabit.ir/ for more information including contact details.

Contents

2023 |
Volume 11 |
Issue 1

JPMB | Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

Molecular cloning and *in silico* analysis of a GTP cyclohydrolase I gene from grape

Nadia Eslami-Bojnourdi, Raheem Haddad*, Ghasem-Ali Garoosi, Reza Heidari-Japelaghi

Department of Biotechnology Engineering, Faculty of Agricultural and Natural Resources, Imam Khomeini International University, Qazvin, Iran

Original paper

OPEN ACCESS

Edited by

Prof. Ahmad Arzani,
Isfahan University of Technology, Iran

Date

Received 25 February 2023
Accepted: 8 October 2023
Published: 25 January 2024

Correspondence

Dr. Raheem Haddad
r.haddad@eng.ikiu.ac.ir

Citation

Eslami Bojnourdi, N., Haddad, R., & Heidari-Japelaghi, R. (2023). Molecular cloning and *in silico* analysis of a GTP cyclohydrolase I gene from grape.*J. Plant Mol. Breed 11* (1): 1-16. 10.22058/JPMB.2023.1990428.1271.

Abstract: An entire open reading frame (ORF) encoding for a polypeptide of GTP cyclohydrolase I (GTPCH I) was isolated and cloned from Askari cultivar of grape (*Vitis vinifera* L.) berries. The 1,338-nucleotide ORF yields a 445-residue amino acid sequence with a calculated molecular mass of 48.65 kDa and a predicted isoelectric point of 6.43. The *Vvgtpch* I genomic sequence with a length of 4,964 bp contains two exons (169 and 1,169 bp) and an intron (2,676 bp). The *gtpch* I sequence of grape displayed a strong similarity with *gtpch* I sequence found in other plants, including peach (72%), cocoa (72%), strawberry (70%), and poplar (69%). Analysis of mRNA secondary structure revealed that the start codon of *Vvgtpch* I is completely exposed, suggesting a robust binding of the ribosome and efficient translation. Similar to *gtpch*s I from diverse sources, molecular modeling uncovered that the monomer of VvGTPCH I adopts an $\alpha\beta$ structure, which includes 10 α-helices and 8 β-sheets. Moreover, *in silico* analysis of the *Vvgtpch* I gene promoter identified potential *cis*-acting elements responsive to environmental signals. This suggests that the *Vvgtpch* I gene has the capacity to be responsive to various environmental cues, such as heat, heavy metals, light, and plant hormones.

Keywords: folate, *in silico* analysis, regulatory elements, cloning, promoter region

Introduction

GTP cyclohydrolase I (GTPCH I, EC 3.5.4.16) plays a role in the conversion of GTP to dihydroneopterin triphosphate and formic acid. This reaction represents the initial and essential stage in the biosynthesis of tetrahydrofolate (FH_4) in plants and certain microorganisms, as well as tetrahydrobiopterin (BH_4) in mammals (Blau and van Spronsen, 2013) (Figure 1). Folates are part of a wide-ranging family of polyglutamates derived from pteroic acid and its related analogs. They play a vital role as cofactors in the synthesis of purines, pyrimidines, pantothenate, thymidylate, and in the metabolism of various amino acids, such as methionine, serine, and glycine (Blancquaert et al., 2014). Folates are comprised of a p-aminobenzoate (PABA) unit that is fused with a pterin ring derived from GTP, along with a varying number of glutamate moieties. PABA is synthesized from chorismate in the plastid, whereas the pterin unit is produced in the cytosol through the catalytic activity of *gtpch* I (Basset et al., 2004). The last stages of folate synthesis involve the conjugation of PABA and pterin, as well as the incorporation of glutamate moiety. These steps are catalyzed by a series of five enzymes in the mitochondria (Basset et al., 2002).

Insufficient intake of folate in the diet can lead to a decrease in the ability to produce DNA and sustain normal cell division rates. As a direct consequence, folate deficiency primarily leads to the development of megaloblastic anemia, characterized by a decrease in the production of cells in the bone marrow due to impaired biosynthesis. Insufficient levels of this nutrient have been linked to the occurrence of neural tube defects in infants, including conditions like spina bifida and anencephaly (Shlobin et al., 2020). Additionally, it has been found to elevate the risk of vascular disease, certain types of cancer, and cerebral folate deficiency syndromes during childhood (Shlobin et al., 2020; Rossignol and Frye, 2021; Gofir et al., 2022). Research studies with controlled conditions have demonstrated that the addition of folic acid to grain products has resulted in a notable decrease in the occurrence of neural tube defects, specific childhood cancers, and stroke (Gorelova et al., 2017; Strobbe and Van Der Straeten, 2017). Therefore, consuming folates in their natural form, which can be found in abundant quantities in various plant-based foods such as grapes, is an effective strategy to prevent folate deficiency.

Increasing folate biosynthesis through metabolic engineering was the first proposed strategy to biofortify plants (Agyenim-Boateng et al., 2023). The sole introduction of *gtpch* I has led to an enhancement of folate levels in lettuce, rice, Mexican common bean, maize, and wheat (Nunes et al., 2009; Dong et al., 2014; Ramírez Rivera et al., 2016; Liang et al., 2019). The overexpression of *gtpch* I increased folate levels in rice by 3.3 to 6.1-fold (Dong et al., 2014) and enhanced lettuce folates by 2.1 to 8.5-fold (Nunes et al., 2009). It has been reported that folates in stored rice grains are unstable, which reduces the potential benefits of folate biofortification. Blancquaert et al. (2015) obtained folate concentrations that are up to 150-fold higher than those of wild-type rice by complexing folate to folate-binding proteins to improve folate stability, thereby enabling long-term storage of biofortified high-folate rice grains. The Mexican common bean (*Phaseolus vulgaris* L.) was metabolically engineered by overexpressing *gtpch* I, which increased folate and pteridine levels in the seeds by 3-fold and 150-fold, respectively (Ramírez Rivera et al., 2016). Liang et al. (2019) cloned and co-overexpressed two key soybean folate biosynthesis genes, *Gm8gGCHI* (GTP cyclohydrolase I) and *GmADCS* (aminodeoxychorismate synthase) in maize and wheat. A 4.2-fold and 2.3-fold increase in folate levels were observed in transgenic maize and wheat grains, respectively. Folates serve as suppliers of methyl groups, playing a crucial role in methylation reactions. These reactions are not only vital for controlling gene expression but also play a significant role in the synthesis of proteins, lipids, chlorophyll, and lignin in the plant kingdom (Gorelova et al., 2017; Strobbe and Van Der Straeten, 2017). To the best of our knowledge, there has been limited research conducted on the characterization of the *gtpch* gene in grapes (*Vvgtpch*). Our objective was to isolate and characterize a *Vvgtpch* I gene from Askari cultivar of grape (*Vitis vinifera*) berries. Through *in silico* analysis, we found that the promoter region of *Vvgtpch* I gene is rich in potential regulatory elements, suggesting its potential response to various environmental signals.

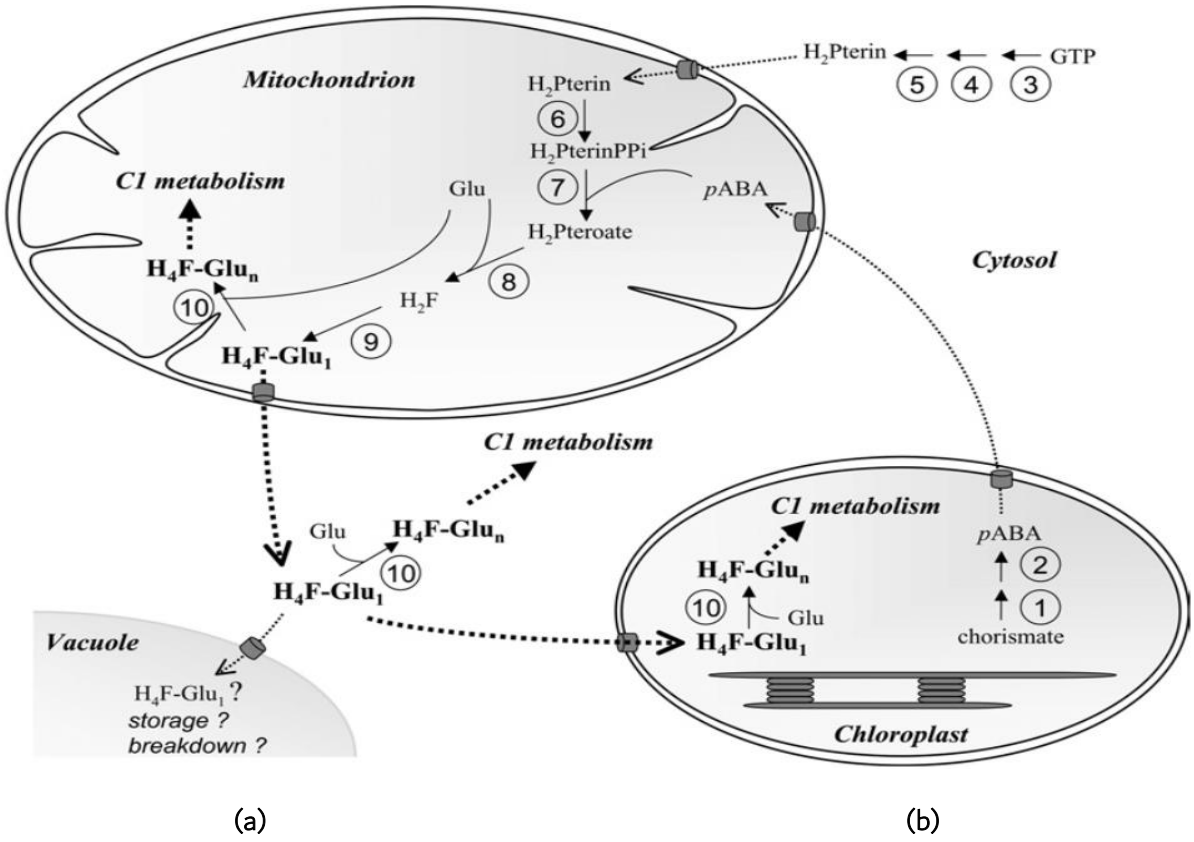


Figure 1. The FH_4 biosynthetic pathway in plants. The enzymes involved in the synthesis of FH_4-Glunare as follows: 1, Aminodeoxychorismate synthase; 2, Aminodeoxychorismate lyase; 3, GTP cyclohydrolase I; 4, Nudixhydrolase; 5, Dihydroneopterin aldolase; 6, Hydroxymethyl dihydropterin pyrophosphokinase (HPPK); 7, Dihydropteroate synthase (DHPS); 8, dihydrofolate synthetase (DHFS); 9, dihydrofolate reductase (DHFR); and 10, folylpolyglutamate synthetase (FPGS) (Sahr et al., 2005).

Materials and Methods

Plant materials and extraction of total RNA

Fresh berries from the cultivar Askari were obtained from grape plants cultivated in the grape collection of the Grape Research Station, Takistan-Qazvin, Iran, in the 2018. Upon collection, all samples were promptly frozen in liquid nitrogen and subsequently stored at -80 ºC. At the veraison stage, seeds were carefully removed from the berries by gently breaking them open in liquid nitrogen. Total RNA was also isolated from fresh grape berries using Cetyltrimethylamonium bromide (CTAB) procedure suggested by Japelaghi et al. (2011).

RT-PCR and molecular cloning

For first strand cDNA synthesis, 5 µg of total RNA treated with *DNase* I (Thermo Fisher Scientific) was used as a template using Oligo $(dT)_{18}$ primer (1 µg/µl, Qiagen) for 5 min at 70 ºC. The reaction mixture was incubated with RevertAid™ M-MuLV Reverse Transcriptase (200 u/µl, Thermo Fisher Scientific) for 60 min at 42 ºC. The mixture was brought to a halt by subjecting it to heat at 70 ºC for 10 min. The degenerate primers (Dgtpch IF: 5′-ATGGGNGCNCTNGAYGARGGN-3′; Dgtpch IR: 5′- NGANGTNGCNGTRTTNGGDAT-3′) used in this study were designed based on the available expressed sequence tags (ESTs) from *Prunus persica*, *Theobroma cacao*, and *Populus trichocarpa*, identified with the BLASTn program (http://www.ncbi.nlm.nih.gov) for the amplification of the *Vvgtpch* I gene by the reverse transcription-PCR (RT-PCR). 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 ºC for 30 s, annealing at 45 ºC for

min, and extension at 72 ºC for 1 min. The final extension was performed at 72 ºC for 5 min. The PCR products were purified using the AccuPrep Gel Purification kit (Bioneer, South Korea), following the manufacturer's instructions. Subsequently, the purified products were subcloned into the pTG19-T vector (Vivantis, Malaysia) as per the manufacturer's guidelines. The dideoxynucleotide sequencing (Bioneer, South Korea) was used to determine the nucleotide sequence of the inserts in both directions.

Sequence analysis

The estimated properties of the deduced amino acid sequence were obtained using ProtParam program (http://www.expasy.ch/tools/protparam.html), while its subcellular localization prediction was conducted by combining three different programs, TargetP (http://www.cbs.dtu.dk/services/TargetP/), iPSORT (http://ipsort.hgc.jp/) and YLOC (http://www-bs.informatik.uni-tuebingen.de/Services/YLoc/). The prediction of functional domains was carried out using the MotifScan program (https://myhits.isb-sib.ch/cgi-bin/motif_scan), while the identification of potential sites for post-translation modifications (PTM) was performed using the ScanProsite program (https://prosite.expasy.org/scanprosite/). The mRNA sequence was submitted to the RNAfold WebServer (Vienna RNAWebsuite) (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) to predict its secondary structure and calculated the minimum free energy (MFE) based on base-pair probabilities. PSIpred (http://bioinf.cs.ucl.ac.uk/psipred/) was used to predict the secondary structure of VvGTPCH I, while the I-TASSER program (http://zhang.bioinformatics.ku.edu/I-TASSER) was employed to predict its three-dimensional (3D) structure.

Analysis of promoter region

The promoter region of the grape *Vvgtpch* I gene was sourced from the Phytozome website (http://www.phytozome.net) and analyzed using PlantCare software (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) to identify regulatory elements associated with various types of plant stress responses.

Phylogenetic analysis

Using the *Vvgtpch* I sequence as a query, protein sequences from different organisms or microorganisms were selected by retrieving data from the GenBank via the BLASTp algorithm at the National Center for Biotechnology Information (NCBI). Sequences were aligned using the ClustalX software and a phylogenetic tree was also constructed with MEGA 4.0.2 software via Neighbor Joining method (Tamura et al., 2007).

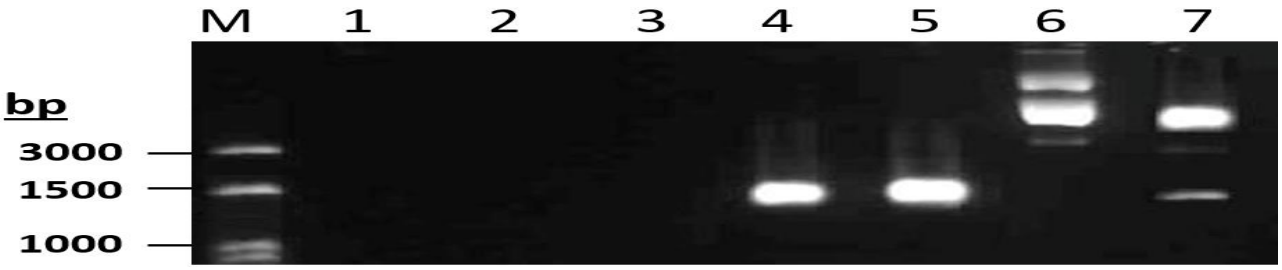


Figure 2. Confirmation of the recombinant pTG19-gtpch I plasmid by PCR, enzyme digestion and 1.2% agarose gel electrophoresis. M) 100 bp DNA ladder, 1) PCR negative control without primers, 2) PCR negative control without template, 3) PCR negative control with bacterial DNA as a template, 4) the *Vvgtpch* I gene amplified in PCR via screening of recombinant clones, 5) the *Vvgtpch* I gene amplified in PCR containing the recombinant pTG19-gtpch I plasmid as a template, 6) The recombinant pTG19-gtpch I plasmid, 7) The recombinant pTG19-gtpch I plasmid digested with *BamH* I.

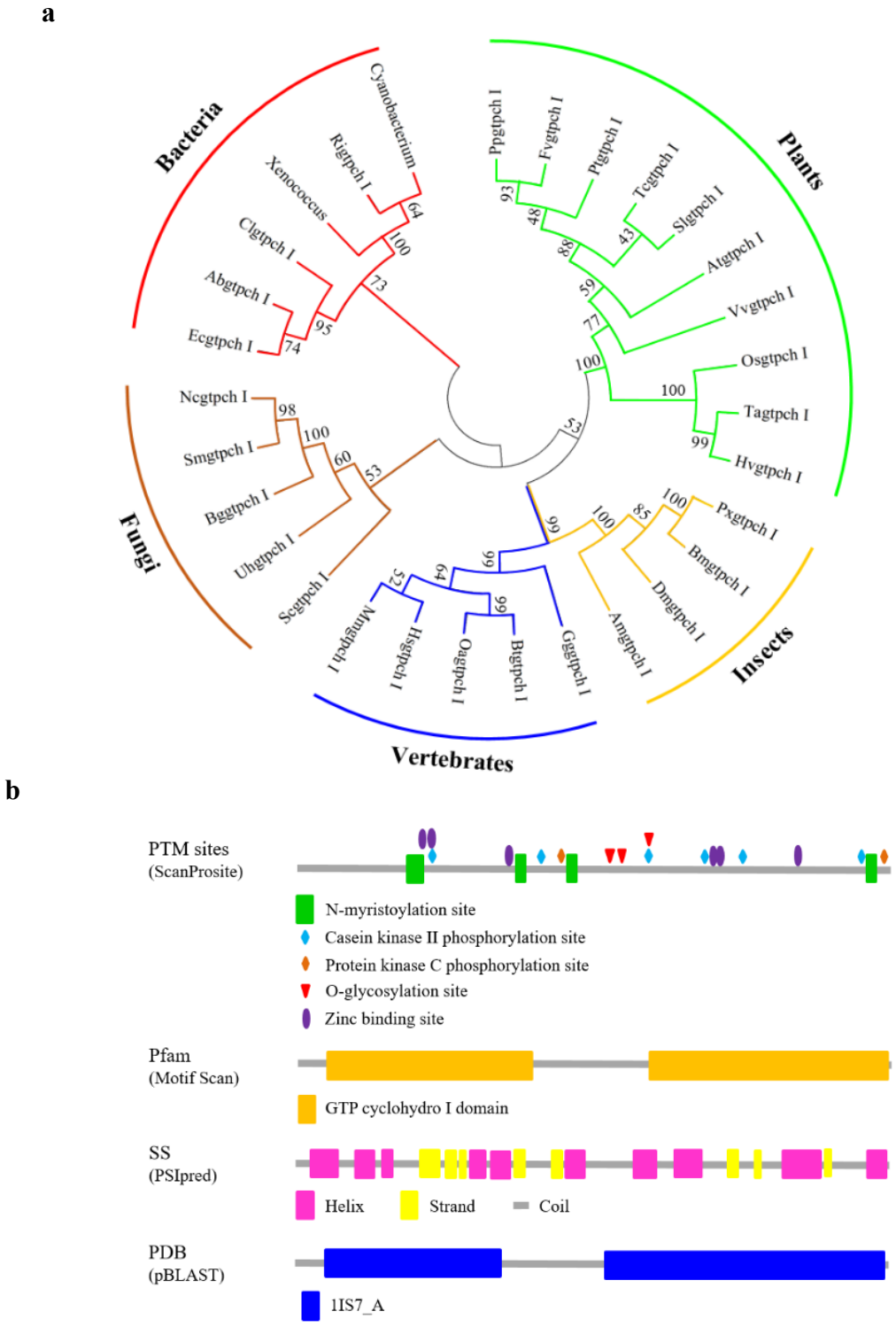
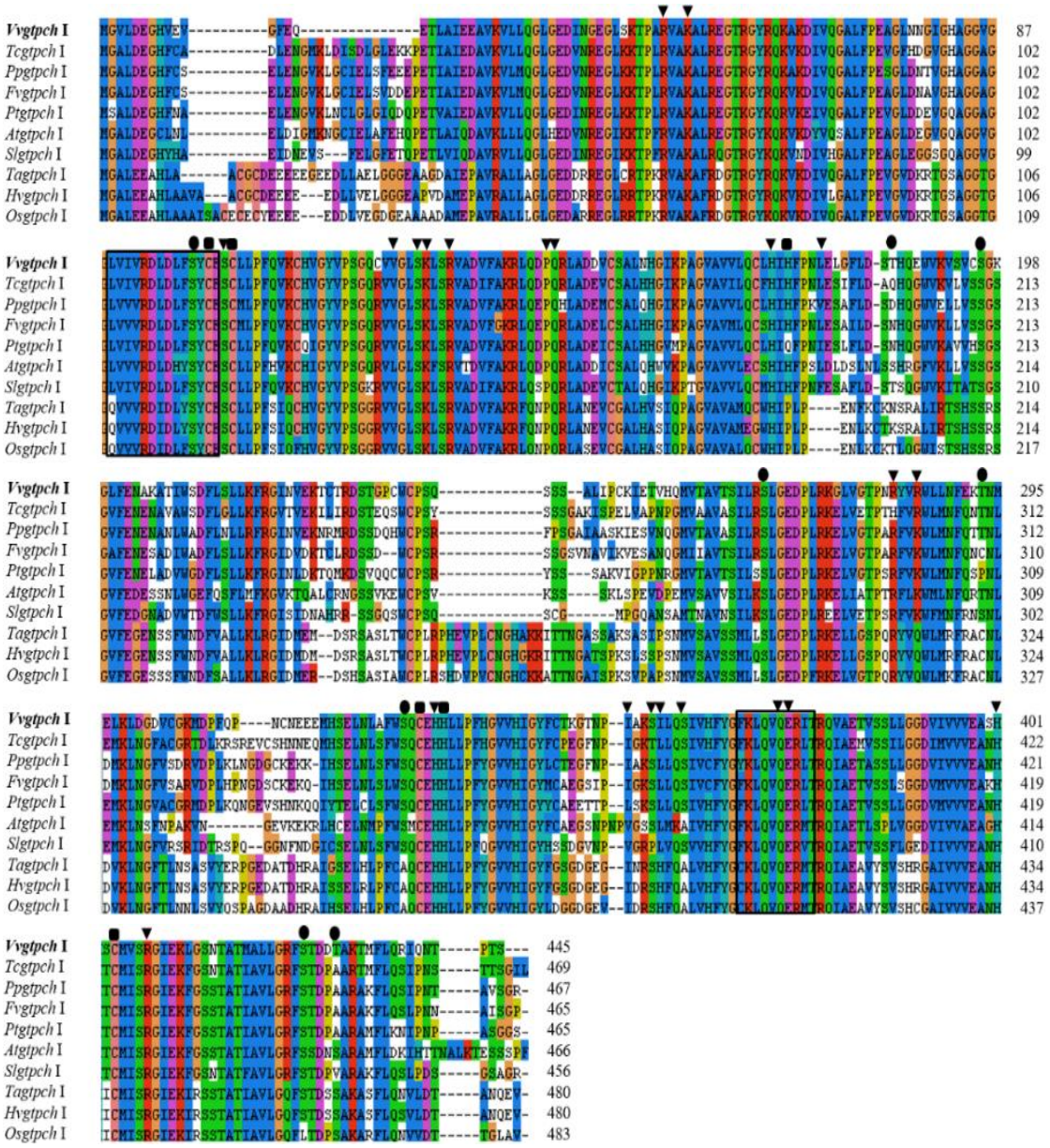


Figure 3. Phylogenetic analysis and structural characterization of *Vvgtpch* I. (**a**) Phylogenetic tree of *Vvgtpch* I and *gtpch*s I from different sources using MEGA 4.0.2 software. (**b**) Multiple sequence alignments of *Vvgtpch* I and *gtpch* I sequences from plants. The protein sequence deduced from the *Vvgtpch* I gene was aligned with its homologs from different plants using ClustalX. The *Vvgtpch* I signature sequences are indicated in black boxes. (Continue in next page)

c



The conserved active site residues, residues involvement in zinc binding, and the possible phosphorylation sites for protein kinase C and casein kinase II are shown by black triangles, squares, and circulars, respectively. The NCBI and EMBL accession numbers for the sequences described and mentioned in this study are as follows: *Apis mellifera* (*Amgtpch* I, XP_624456.2), *A. thaliana* (*Atgtpch* I, NP_187383), *Arcobacter butzleri* (*Abgtpch* I, YP_008331555), *Blumeria graminis* (*Bggtpch* I, EPQ61888), *Bombyx mori* (*Bmgtpch* I, NP_001166803), *Bos Taurus* (*Btgtpch* I, XP_002690995), *Cellulophaga lytica* (*Clgtpch* I, YP_004263693), *Chlamydomonas reinhardtii* (*Crgtpch* I, NW_001843677), *Cyanobacterium* spp. (WP_017322947), *E. coli* (*Ecgtpch* I, EU747840), *Fragaria vesca* (*Fvgtpch* I, XP_004290171), *G. gallus* (*Gggtpch* I, NP_990554), *Homo sapiens* (*Hsgtpch* I, AAN17459), *Hordeum vulgare* (*Hvgtpch* I, BAJ95470), *Neurospora crassa* (*Ncgtpch* I, XP_958695), *O. sativa* (*Osgtpch* I, EAZ32303), *Ovis aries* (*Oagtpch* I, XP_004011081), *Papilio xuthus* (*Pxgtpch* I, BAE66650), *P. trichocarpa* (*Ptgtpch* I, XP_002320619), *P. persica* (*Ppgtpch* I, EMJ02205), *Richelia intracellularis* (*Rigtpch* I, WP_008234507), *Saccharomyces cerevisiae* (*Scgtpch* I, CAA87397), *Solanum lycopersicum* (*Slgtpch* I, NP_001234141), *Sordaria macrospora* (*Smgtpch* I, XP_003345165), *Triticum aestivum* (*Tagtpch* I, ABM54074), *T. cacao* (*Tcgtpch* I, EOX93821), *Ustilago hordei* (*Uhgtpch* I, CCF48468), *Xenococcus* sp. (WP_006509939). (**c**) The potential sites for PTMs and the functional domains were predicted using ScanProsite and MotifScan. The secondary structure of *Vvgtpch* I was predicted by PSIpred program. The three-dimensional structure of *gtpch* I was retrieved from the PDB through the BLASTp algorithm at NCBI using the amino acid sequence of *Vvgtpch* I a query.

Results

Cloning and sequence analysis of Vvgtpch I gene

Using the degenerate primers, the *Vvgtpch* I gene was isolated and cloned into pTG19-T plasmid vector to generate the recombinant pTG19-gtpch I plasmid (Figure 2). The open reading frame (ORF) of *Vvgtpch* I (submitted at NCBI GenBank under accession number KF891965) was 1,338 nucleotides long and coding for a polypeptide of 445 amino acid residues. The calculated molecular mass and the predicated isoelectric point of the deduced polypeptide sequence are 48.65 kDa and 6.43, respectively. Using the ProtParam program, the Aliphatic, the Hydropathicity, and the Instability indexes were evaluated about 93.69, -0.016, and 38.32, repectively. The lack of an N-terminal extension indicates that the *Vvgtpch* I protein is likely localized in the cytosol. This proposal is further supported by the analysis conducted using three programs: TargetP, iPSORT, and YLOC.

Phylogenetic analysis

The *Vvgtpch* I and *gtpch* I sequences from various organisms or microorganisms were utilized to construct a phylogenetic tree (Figure 3a). The tree exhibits two prominent clusters. One cluster contains plants and the other includes non-plants. The first cluster consists of bacteria, fungi, insects, and vertebrates, while the second cluster has two main branches, including monocotyledons and dicotyledons. The grape *gtpch* I displays a strict identity with *gtpch* I sequences from other plants, such as peach (*P. persica*; *Ppgtpch* I, 72%), Cocoa (*T. cacao*; *Tcgtpch* I, 72%), Poplar. (*P. tricocarpa*; *Ptgtpch* I, 69%), and strawberry (*F. vesca*; *Fvgtpch* I, 70%) (Figure 3b). In contrast, the grape *gtpch* I shares lower degree of identity to *gtpch* I from human, *S. cerevisiae*, and *Cyanobacterium* with 40, 44, and 45%, respectively. The position of conserved amino acids in active site and other the conserved regions between *Vvgtpch* I and *gtpch*s I from other plants have also been represented in Figure 3b. The prediction of potential sites for PTM and the functional domains were also carried out using ScanProsite and MotifScan, respectively (Figure 3c). Analysis of secondary structure by PSIpred program revealed that the *Vvgtpch* I contains of ten α-helixes and eight β-sheets. By the amino acid sequence of *Vvgtpch* I as a query, the 3D structure of the *gtpch* I from *Rattus norvegicus* (PDB ID code 1IS7_A) with the highest score (E-value = 9e-23) was retrieved from the Protein DataBase (PDB) through the BLASTp algorithm at the NCBI (Figure 3c).

Prediction of mRNA secondary structure

The RNA secondary structure was analyzed to assess the mRNA stability of *Vvgtpch* I gene and determine situation of the start codon using RNAfold WebServer (Figure 4).

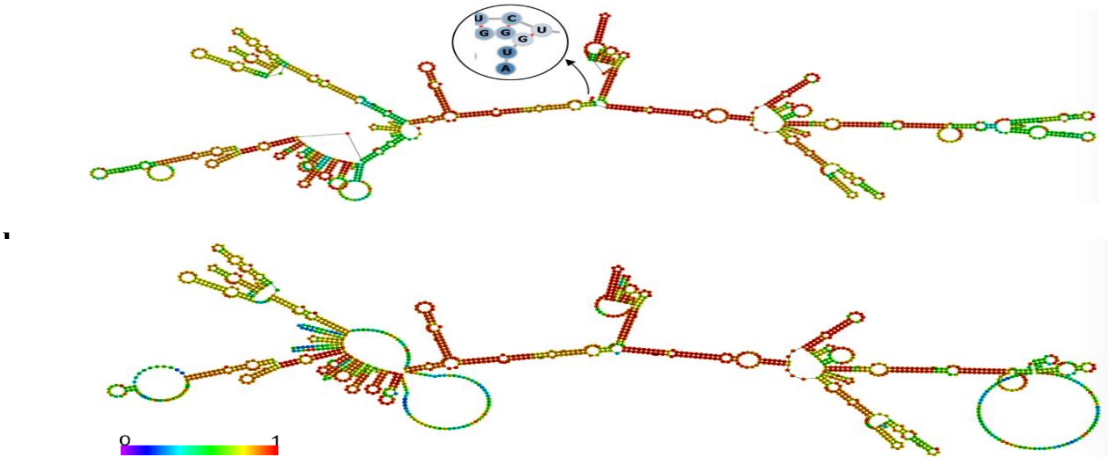


Figure 4. Prediction of mRNA secondary structure of *Vvgtpch* I gene based on the MFE. (**a**) The optimal secondary structure and (**b**) the centroid secondary structure.

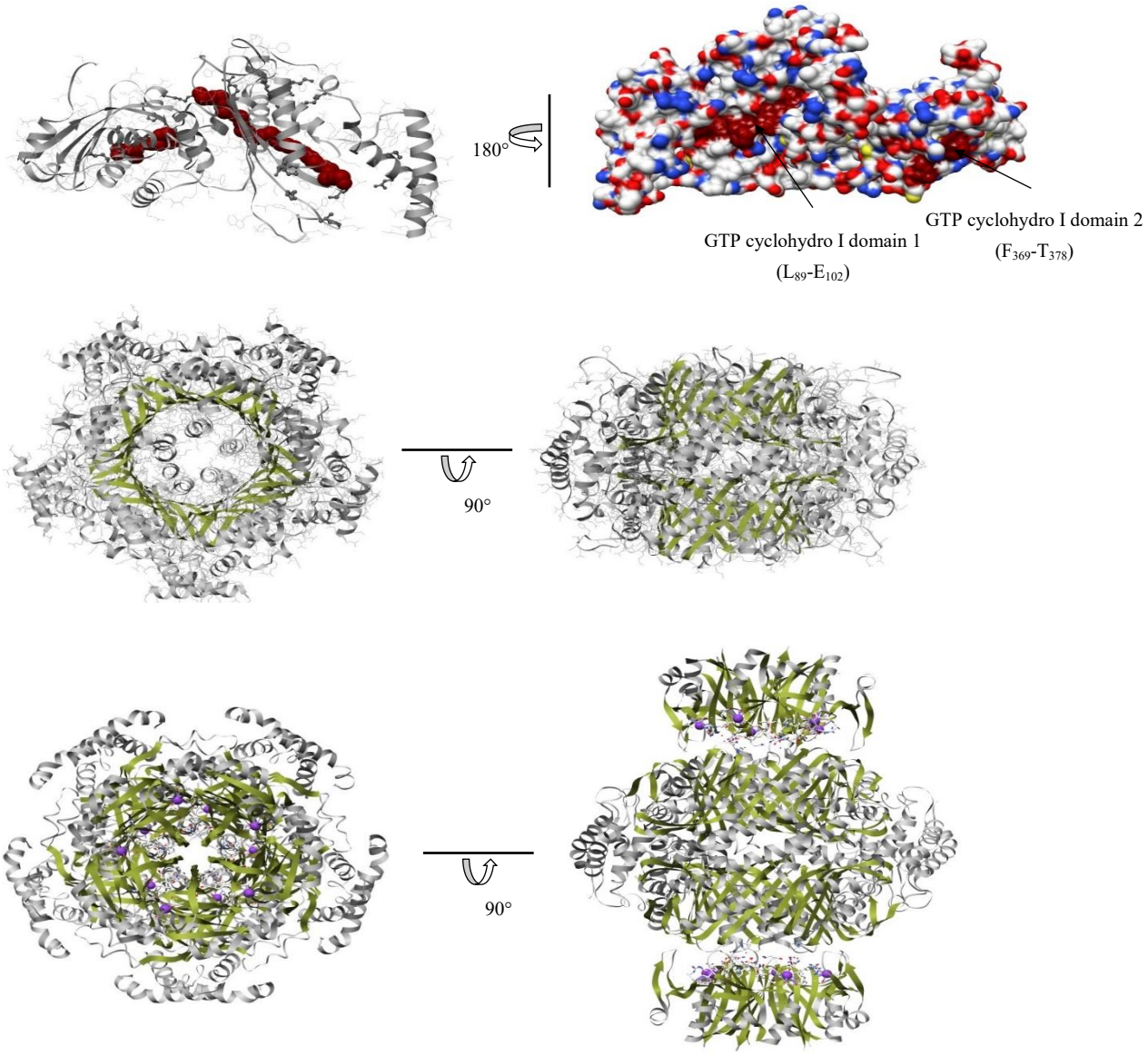


Figure 5. Prediction of three-dimensional model of *Vvgtpch* I and d_5-symmetric homodecamer model of for *gtpch* I. (**a**) Ribbon representation of the monomer structure of *Vvgtpch* I. The signature sequences and the conserved active site residues are indicated by spheres and ball and sticks, respectively. (**b**) View along the molecular 5-fold symmetry axis. The core of the active enzyme complex is formed by a 20-stranded antiparallel β-barrel surrounding five α-helices. (**c**) Complex of GTP cyclohydrolase I/GFRP, viewed along a molecular *c2* axis. The phenylalanine residues are indicated as sphere models.

The mRNA secondary structures emerged to be different in bi-dimensional models and displayed various predicted levels of the MFE. The MFE of the optimal secondary structure and the centroid secondary structure for *Vvgtpch* I mRNA were -405.50, and -344.95 kcal mol^{-1}, respectively. In addition, the start codon was fully exposed in the optimal secondary structure.

Molecular modeling analysis

A predicted 3D structure was determined for *Vvgtpch* I by applying I-TASSER simulation. Similar to all *gtpch*s I from different sources, the monomer of *Vvgtpch* I folds into a αβ structure with 10 α-helices and 8 β-sheets (Figure 5a). In eubacteria and animals, the GTP cyclohydrolases I are structurally

similar and they oligomerize to a toroid-shaped, d_5-symmetric homodecamer (Figure 5b). Figure 5c represents one homodecameric GTP cyclohydrolase I molecule in interaction with two GFRP molecules.

Regulatory elements in the promoter region of Vvgtpch I gene

In order to understand how oxidative stresses regulate the grape *gtpch* I gene, we retrieved its promoter region from the Phytozome website and used the PlantCare software to identify relevant *cis*-acting elements. Analyzing the grape *gtpch* I promoter region revealed the presence of several potential *cis*-acting elements that are known to respond to environmental signals (Figure 6). There are three heat stress responsive elements (HSE) in +223, +631, and -839. A putative basic motif (GGTCCAT) involved in auxin responsiveness was located at position +1045. A sequence similar to the conserved core sequence of metal responsive element (MRE) (TGCAGAC) is found that locates at position +1474. A potential TCA-element (-117) that is believed to be associated with the response to SA (Salicylic Acid) was also discovered. Lastly, in the promoter region of grape *gtpch* I, TC-rich repeats were identified at position +480, implying their potential involvement in stress response. The sequence also includes a proposed TATA box at +1460 relative to the transcriptional start site and a potential CAAT box at +1297. Table 1 show cases additional *cis*-acting elements in the *Vvgtpch* I promoter region, which are responsive to environmental signals.

Figure 6. The nucleotide sequence of promoter region of *Vvgtpch* I gene. The potential cis-acting regulatory elements identified using PlantCare software are shown by color boxes. The putative TATA box at +1460 and the putative CAAT box at +1297 are also underlined.

Table 1. The potential cis-acting regulatory elements identified in the promoter region of *Vvgtpch* I gene using PlantCare software.

Cis-acting elements	Function	Sequence	Numbers	Positions
ARE	Anaerobic induction	TGGTTT	1	-1286
AuxRR-Core	Auxin	GGTCCAT	1	+1045
Box 4	Light	ATTAAT	3	+542, +721, +1064
Box I	Light	TTTCAAA	1	-182
Box II	Light	GTGAGGTAATAT	1	-352
GATA-motif	Light	AAGGATAAGG	1	-693
HSE	Heat stress	AAAAAATTTC	3	+223, +631, -839
MNF1	Light	GTGCCC(A/T)	1	+149
MRE	Heavy metals	TGCAGAC	1	+1474
SP1	Light	CC(G/A)CCC	1	+1395
TC-rich repeats	Defense and stress	ATTTTCTCCA	1	+480
TCA-element	Salicylic acid	CAGAAAAGGA	1	-117

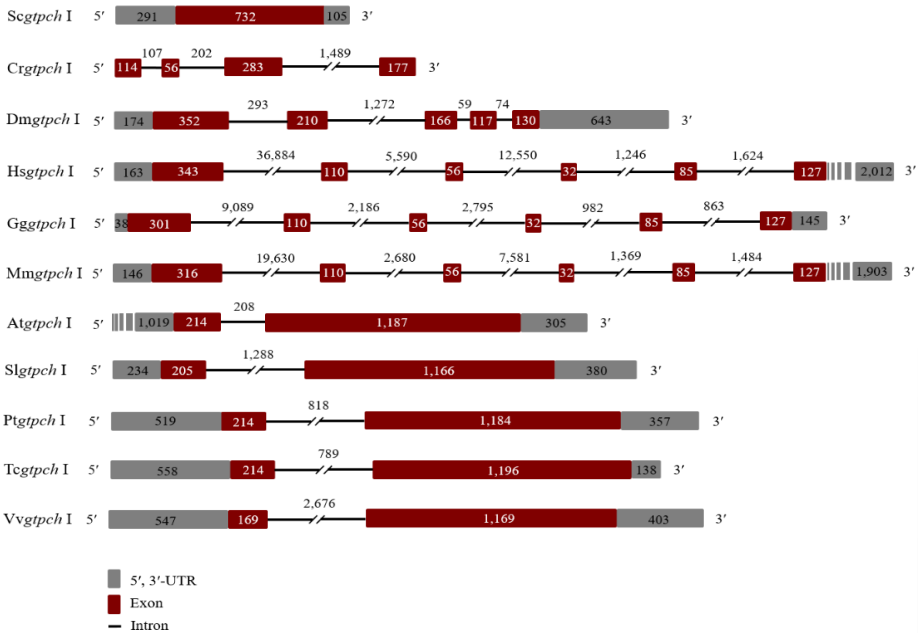
Table 2. Characteristics of nucleotide and amino acid sequences of *Vvgtpch* I and *gtpchs* I from different sources.

Organism	Gene	RefSeq accession no.	Chromosomal position	Strand	Length (bp)				Molecular weight (Da)	Exon: Intron no.
					Gene	mRNA	ORF	Protein (aa)		
S. cerevisiae	*Scgtpch* I	NC_001139	VII	-	1,143	1,143	732	243	27,769	1:0
C. reinhardtii	*Crgtpch* I	NW_001843677	Unknown	+	2,428	630	630	209	22,928	4:3
D. melanogaster	*Dmgtpch* I	NT_033778	57C7-57C8	-	7,294	5,596	975	324	35,541	5:4
H. sapiens	*Hsgtpch* I	NC_000014	14q22.2	-	60,822	2,928	753	250	27,903	6:5
G. gallus	*Gggtpch* I	NC_006092	5	+	16,809	894	711	236	26,115	6:5
M. musculus	*Mmgtpch* I	NC_000080	14C1	-	35,519	2,775	726	241	27,014	6:5
A. thaliana	*Atgtpch* I	NC_003074	3	+	3,367	2,725	1,401	466	51,380	2:1
S. lycopersicum	*Slgtpch* I	NC_015443	6	-	3,273	1,985	1,371	456	49,951	2:1
P. trichocarpa	*Ptgtpch* I	XP_002320619	2	+	3,092	2,274	1,398	465	50,923	2:1
T. cacao	*Tcgtpch* I	CM001879	1	-	2,895	2,106	1,410	469	51,503	2:1
V. vinifera	*Vvgtpch* I	XM_002269229	1	+	4,964	2,288	1,338	445	48,651	2:1

Position of introns in Vvgtpch I and other gtpchs I

The genomic sequence of *Vvgtpch* I spans a length of 4,964 bp. Within this sequence, there are two exons, measuring 169 and 1,169 nucleotides in length, respectively. Additionally, there is an intron present, spanning a length of 2,676 bp. The exon-intron junctions follow the GT-AG rule, similar to how it occurs in higher plants. Additionally, in line with plant genes, the intron in this grape gene contains a relatively higher AT content when compared to the coding regions (Gallie, 1993). The characteristics of nucleotide and amino acid sequences of *gtpch* I genes from different sources have been demonstrated in Table 2. The plant *gtpch* I genes contain a single intron with the diverse size at the conserved position. The size of *Arabidopsis gtpch* I intron is far shorter than that of cacao (789 bp), populus (818 bp), tomato (1,288 bp), and grape (2,676 bp) (Figure 7a). Similar to other plant *gtpch* I genes, the *Vvgtpch* I contains an intron at the same position (50Gly) and the splicing site is between 1–2 nucleotides [1 (g/gt)] (Figure 7b). In plant *gtpch* I genes, the Gly residues place at different positions, nevertheless, their splicing sites are similar with each other. The size and position of introns were also investigated in yeast, *chlamydomonas*, fruit fly, and vertebrates *gtpch* I genes. In yeast, the *Scgtpch* I do not harbor any introns, whereas the *gtpch* I genes from *chlamydomonas*, fruit flies, and humans contain three, four, and five introns, respectively (Figure 7a).

a



b

```
                                                    10        20        30
Crgtpch I   ----------------------MAPCRDKDKVARMEQAVRVLIHGLGEDLEREGLRDTPKR  39
Dmgtpch I   PGHEKCTFHHDLELDH------KPPTREALLPDMARSYRLLLGGLGENPDRQGLIKTPER 170
Ggtpch I    SGGSGDGWR--GERPR------SEEDNELSLPSLAAAYTTILRALGEDPERQGLLKTPWR  83
Hsgtpch I   SAQPADGWK--GERPR------SEEDNELNLPNLAAAYSSILSSLGENPQRQGLLKTPWR  97
Mmgtpch I   GAQPADAWK--AGRHR------SEEENQVNLPKLAAAYSSILLSLGEDPQRQGLLKTPWR  88
Atgtpch I   GALDEGCLNLELDIGMKNGCIELAFEHQPETLAIQDAVKLLLQGLHEDVNREGIKKTPFR  61
Slgtpch I   GALDEGHYHAEIDNEVSF---ELGFETQPETLVIQDAVRVLLQGLGEDINREGIKKTPFR  58
Vvgtpch I   GVLDEGHVEVG----------------FEQETLAIEEAVKVLLQGLGEDINGEGLSKTPAR  46
Tcgtpch I   GALDEGHFCADLENGMKLDISDLGLEKKPETIAIEDAVKVLLQGLGEDVNREGLKKTPLR  61
Ptgtpch I   SALDEGHFNAELENGVKLNCLGLGIQDQPETVAIEDAVKVLLQGLGEDINREGLKKTPLR  61

                    40        50        60        70                  80        90
Crgtpch I   VAKALLDCTQGYHQDTSSTLGTALFHEPIVHD------GDEGVVLVRDIDFASTSEETLL  93
Dmgtpch I   AAKAMLYFTKGYDQSLEDVLNGAVFDED----------HDEMVVVKDIEMFSMCEHHLV 219
Ggtpch I    AATAMQFFTKGYQETIADVLNDAIFDED----------HDEMVIVKNIDMFSLCEHHLV 132
Hsgtpch I   AASAMQFFTKGYQETISDVLNDAIFDED----------HDEMVIVKDIDMFSMCEHHLV 146
Mmgtpch I   AATAMQYFTKGYQETISDVLNDAIFDED----------HDEMVIVKDIDMFSMCEHHLV 137
Atgtpch I   VAKALREGTRGYKQKVKDYVQSALFPEAGLDEGVGQAGGVGGLVVVRDLDHYSYCESCLL 121
Slgtpch I   VAKALRQGTRGYKQKVNDIVHGALFPEAGLEGGSGQAGGVGGLVIVRDLDLFSYCESCLL 118
Vvgtpch I   VAKALREGTRGYRQKAKDIVQGALFPEAGLNNGIGHAGGVGGLVIVRDLDLFSYCESCLL 106
Tcgtpch I   VAKALREGTRGYKQKVKDIVQGALFPEVGFHDGVGHAGGAGGLVIVRDLDLFSYCESCLL 121
Ptgtpch I   VAKALREGTKGYKQRVKEIVQGALFPEVGLDDEVGQAGGAGGLVIVRDLDLFSYCESCLL 121

                   100       110       120       130       140       150
Crgtpch I   PFHGRCHVAYRPKNGVVLGLSKLARLTKQYAKRLQTQERLAAEVAHALQQSLECYGIADA 153
Dmgtpch I   PFYGKVSIGYLPCN-KILGLSKLARIVEIFSRRLQVQERLTKQIAVAVTQAVQPAGVAV- 277
Ggtpch I    PFVGKVHIGYLPNK-QVLGLSKLARIVEIYSRRLQVQERLTKQIAIAITEALQPAGVGV- 190
Hsgtpch I   PFVGKVHIGYLPNK-QVLGLSKLARIVEIYSRRLQVQERLTKQIAVAITEALRPAGVGV- 204
Mmgtpch I   PFVGRVHIGYLPNK-QVLGLSKLARIVEIYSRRLQVQERLTKQIAVAITEALQPAGVGV- 195
Atgtpch I   PFHVKCHIGYVPSGQRVLGLSKFSRVTDVFAKRLQDPQRLADDICSALQHWVKPAGVAV- 180
Slgtpch I   PFQVKCHVGYVPSGKRVVGLSKLSRVADIFAKRLQSPQRLADEVCTALQHGIKPTGVAV- 177
Vvgtpch I   PFQVKCHVGYVPSGQCVVGLSKLSRVADVFAKRLQDPQRLADDVCSALNHGIKPAGVAV- 165
Tcgtpch I   PFQVKCHVGYVPSGQRVVGLSKLSRVADIFAKRLQDPQRLADEVCSALHHGIKPAGVAV- 180
Ptgtpch I   PFQVKCQIGYVPSGQRVVGLSKLSRVADVFAKRLQDPQRLADEICSALHHGVMPAGVAV- 180

                   160       170       180                           190
Crgtpch I   GCAGGVMVVCDAAHMCMVARGV---ENHSGSTTSF-------------AVRGAFASRPDM 197
Dmgtpch I   ------VVE--GVHMCMVMRGVQKINS-----KTVTSTMLGVFRDDP-KTREEFLNLVNS 323
Ggtpch I    ------VIE--ATHMCMVMRGVQKMNS-----KTATSTMLGVFREDP-KTREEFLTLIRS 236
Hsgtpch I   ------VVE--ATHMCMVMRGVQKMNS-----KTVTSTMLGVFREDP-KTREEFLTLIRS 250
Mmgtpch I   ------VIE--ATHMCMVMRGVQKMNS-----KTVTSTMLGVFREDP-KTREEFLTLIRS 241
Atgtpch I   ------VLECSHIHFPSLDLDSLNLSSHRGFVKLLVSSGSGVFEDESSNLWGEFQSFLMF 234
Slgtpch I   ------VLQCMHIHFPNFESAFLD-STSQGWVKITATSGSGVFEDGNADVWTDFWSLLKF 230
Vvgtpch I   ------VLQCLHIHFPNLELGFLD-STHQEWVKVSVCSGKGLFENAKATIWSDFLSLLKF 218
Tcgtpch I   ------ILQCFHIHFPNLESIFLD-AQHQGWVKVLVSSGSGVFENENAVAWSDFLGLLKF 233
Ptgtpch I   ------VLQCLHIQFPNIESLFLD-SNHQGWVKAVVHSGSGVFENELADVWGDFLSLLKF 233
```

Figure 7. The position and size of introns in the *Vvgtpch* I and *gtpch* I genes from different sources. (**a**) Structure of *Vvgtpch* I gene and genes encoding *gtpch* I from different organisms or micro-organisms. The UTRs, coding regions and non-coding regions are indicated in gray boxes, red boxes and black lines, respectively. (**b**) The position of introns in *Vvgtpch* I and other *gtpch*s I from different organisms or micro-organisms. The deduced amino acid sequences of *gtpch*s I were aligned using ClustalW2 and the position of the introns are represented by an arrow. The *Chlamydomonas Crgtpch* I is revealed for numeration and the starting residue of introns is distinguished by a circular and triangles for plants and other organisms or micro-organisms, respectively. Accession numbers are given in the Materials and methods section.

In *Crgtpch* I, the second intron [57S (2 ag/c)] is according to the second intron [57D (1 g/at)] and first intron [57D (1 g/at)] of *D. melanogaster* and human, respectively. Interestingly, the second, third [118R (2 cg/c)], and fourth [166V (2 gt/c)] introns of *D. melanogaster gtpch* I gene are located at the same positions as the first [57D (1 g/at)], third [118R (2 ag/g)], and fifth [166T (2 ac/a)] introns of *Hsgtpch* I, respectively (Figure 7b).

Discussion

The sequence analysis of *Vvgtpch* I gene indicated that it is a thermostable and hydrophilic protein with long half-life (Claverie and Notredame, 2011). Moreover, the phylogenetic analysis of the *Vvgtpch* I revealed that it is clustered with other plant homologs, and thus represents the *Vvgtpch* I orthologs in grape. It has been reported that the *gtpch* I sequences harbor the conserved amino acid signatures which describe essential residues in the two *gtpch* I catalytic regions (Sigrist et al., 2002). The two *gtpch* I signature sequences identified as the GTP cyclohydro I domains are: signature (1), [DENGQST]-[LIVMPF]-[LIVM]-x(1,2)-[KRNQELD]-[DENKGS]-[LIVM]-x(3)-[STG]-x-C-[EP]-H(2) and signature (2), [SA]-x-[RK]-x-Q-[LIVMT]-Q-E-[RNAK]-[LIM]-[TSNV] and consist of preserved residues that are crucial for both GTP binding and the formation of the GTP binding pocket (Nar et al., 1995).

The *Vvgtpch* I signature sequences identified using of the multiple sequence alignments are signature (1) L_{89}VIVRDLDLFSYCE$_{102}$ and signature (2) F_{369}KLQVQERIT$_{378}$. The grape *gtpch* I protein contains the possible phosphorylation sites by protein kinase C and casein kinase II, which are necessary for the correct enzyme activity and the protein folding (Maier et al., 1995). The grape *gtpch* I, like to other plant *gtpch*s I, has retained the amino acids essential for zinc binding including C_{101}, C_{104}, H_{173}, C_{331}, H_{334}, and C_{403}. In *E. coli*, the replacement of Cys_{110}, Cys_{181}, His_{112} or His_{113} by serine affords catalytically inactive mutant proteins with reduced capacity to bind zinc (Rebelo et al., 2003). Furthermore, it seems likely the grape GTPCH I has misplaced the EF-hand-like motif necessary for calcium interaction. The residues recognized as participants in coordinating Ca^{2+} within the conserved EF-hand-like motif are conserved in animals and *E. coli*, but they are not found in plants (Steinmetz et al., 1998). This suggests that calcium binding is not required for activity, although no data exists on its effects on catalytic action.

In an mRNA molecule, the secondary structure is made through the formation of hydrogen bonds between complementary pairs of nucleotides. The mRNA secondary structure is able to affect gene expression via adjustment of transcription, splicing RNA, transcript degradation, and translation. Thus, study of the mRNA secondary structure is crucial to recognize functional activity of a transcript (Proctor and Meyer, 2013). According to reports, the accessibility of the ribosome binding site and the start codon of the mRNA are crucial factors that significantly impact efficient translation (Morowvat et al., 2014). Furthermore, there have been demonstrations showing a direct dependence between the translation levels and the MFE of the ribosome binding site and the start codon. A reduction of MFE by 1.4 kcal mol^{-1} would decrease the gene expression by 10-fold (de Smit and van Duin, 1990). Furthermore, the manipulation of mRNA secondary structure via codon optimization increased the MFE level and improved translation efficiency (Prabhu et al., 2016). In prediction of mRNA secondary structure, the start codon was fully exposed in the optimal secondary structure, showing an efficient binding of the ribosome and active translation.

The determination of 3D structure of the GTP cyclohydrolase I without recognition of the metal requirement of the enzyme was first reported in *E. coli* (Nar et al., 1995). However, crystallographic study of the human GTP cyclohydrolase I illustrated the existence of crucial zinc ions at the active sites (Auerbach et al., 2000). The structure of d_5-symmetric homodecamer consists of a β-barrel that is formed by 20 strands arranged in an antiparallel manner, surrounding a core consisting of five α-helical segments (Steinmetz et al., 1998) and its molecular symmetry signifies one 5-fold and five 2-fold symmetry axes that are perpendicular to the molecular 5-fold axis (Gräwert et al., 2013). In contrast to eubacteria, the GTP cyclohydrolase of animals contains an extended N-terminal region that plays a critical role in the interaction of this enzyme with the GTP cyclohydrolase feedback regulatory protein (GFRP). The GFRP is a c_5-

symmetric, ring-shaped homopentamer and the entire enzyme/inhibitor complex follows d_5 symmetry (Gräwert et al., 2013). Phenylalanine has been revealed to bind at the enzyme/GFRP junctions with a total of 10 topologically equivalent binding sites (Maita et al., 2004).

It has been exhibited that the grape *gtpch* I promoter region harbors the several potential cis-acting elements responding to environmental signals including HSE, MRE, TCA-element, and TC-rich repeats (Amin et al., 1988; Sakai et al., 1996; Stuart et al., 1985; Goldsbrough et al., 1993; Diaz-de-Leon et al., 1993). It has been reported that a functional HSE consists of at least three basic repeats organized in alternating orientations (Amin et al., 1988). Thus, the HSE motif in the *Vvgtpch* I promoter appears to be functional. In addition, the *Vvgtpch* I gene is probably able to response to heavy metals because MRE plays a role in the induction of metallothionein gene expression in animals in response to heavy metal exposure (Stuart et al., 1985). Moreover, in the promoter region of grape *gtpch* I, a potential TCA-element and TC-rich repeats were identified, implying its potential involvement in response to SA and stress (Goldsbrough et al., 1993). The presence of these potential regulatory elements indicates that the promoter region of grape *gtpch* I has the capability to respond to multiple environmental cues such as heat, heavy metals, light, and plant growth regulators (PGR).

Analysis of size and position of introns in *Vvgtpch* I and other *gtpch*s I suggested that the size of the introns is related to the species (Sahrawy et al., 1996). Also, it confirms that although plant *gtpch* I genes have diverged in their amino acid sequences, they may share a common ancestor with having a single intron at position 50 (Meyer et al., 2002). Furthermore, in contrast to plant *gtpch*s I, the *gtpch* I genes from *chlamydomonas*, fruit flies, and humans contain three, four, and five introns, respectively, indicating a concordant reduction or extension of the various introns of a sequence during evolution (Sahrawy et al., 1996). However, no accommodations observed between the position of these introns and intron position of *Vvgtpch* I or higher plants *gtpch* I genes.

Conclusion

The *Vvgtpch* I gene, also known as grape *gtpch* I, was extracted from the grape berry organ of a specific Iranian cultivar named Askari. By constructing a phylogenetic tree, it was determined that the grape *gtpch* I gene grouped together with similar genes found in other plants, indicating a significant degree of similarity and shared identity with *gtpch* I genes from various plant species. The *in silico* analysis demonstrated that the promoter region of *Vvgtpch* I gene contains a number of potential *cis*-acting elements responding to environmental signals. The presence or absence of a *cis*-acting element in a gene promoter could suggest a specific mode of gene regulation. Collectively, possible strategies to target sensitive transcription factors of *gtpch* I gene or its cofactors are suggested based on the updated view of GTP-dependent gene regulation in grape.

Supplementary Materials

No supplementary material is available for this article.

Author Contributions

Methodology, investigation, N.E.B.; conceptualization, supervision, project administration, R.H.; validation, formal analysis, data curation, G.A.G. and R.H.J.; software, writing—original draft preparation, writing—review and editing, R.H.J. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Acknowledgments

No acknowledgments.

Conflicts of Interest

The authors declare no conflict of interest.

References

Agyenim-Boateng, K.G., Zhang, S., Shohag, M.J.I., Shaibu, A.S., Li, J., Li, B., and Sun, J. (2023). Folate biofortification in soybean: challenges and prospects. *Agronomy* 13(1): 241.

Amin, J., Ananthan, J., and Voellmy, R. (1988). Key features of heat shock regulatory elements. *Mol Cell Biol* 8(9): 3761-3769.

Auerbach, G., Herrmann, A., Bracher, A., Bader, G., Gütlich, M., Fischer, M., Neukamm, M., Garrido-Franco, M., Richardson, J., and Nar, H. (2000). Zinc plays a key role in human and bacterial GTP cyclohydrolase I. *Proc Natl Acad Sci* 97(25): 13567-13572.

Basset, G., Quinlivan, E.P., Ziemak, M.J., Díaz de la Garza, R., Fischer, M., Schiffmann, S., Bacher, A., Gregory III, J.F., and Hanson, A.D. (2002). Folate synthesis in plants: the first step of the pterin branch is mediated by a unique bimodular GTP cyclohydrolase I. *Proc Natl Acad Sci* 99(19): 12489-12494.

Basset, G.J., Quinlivan, E.P., Ravanel, S., Rébeillé, F., Nichols, B.P., Shinozaki, K., Seki, M., Adams-Phillips, L.C., Giovannoni, J.J., and Gregory III, J.F. (2004). Folate synthesis in plants: the p-aminobenzoate branch is initiated by a bifunctional PabA-PabB protein that is targeted to plastids. *Proc Natl Acad Sci* 101(6): 1496-1501.

Blancquaert, D., De Steur, H., Gellynck, X., and Van Der Straeten, D. (2014). Present and future of folate biofortification of crop plants. *J Exp Bot* 65(4): 895-906. doi: 10.1093/jxb/ert483.

Blancquaert, D., Van Daele, J., Strobbe, S., Kiekens, F., Storozhenko, S., De Steur, H., Gellynck, X., Lambert, W., Stove, C., and Van Der Straeten, D. (2015). Improving folate (vitamin B9) stability in biofortified rice through metabolic engineering. *Nat Biotechnol* 33(10): 1076-1078.

Blau, N., and van Spronsen, F.J. (2013). "Disorders of phenylalanine and tetrahydrobiopterin metabolism," in *Physician's guide to the diagnosis, treatment, and follow-up of inherited metabolic diseases*. Springer), 3-21.

Claverie, J.-M., and Notredame, C. (2011). *Bioinformatics for dummies.* John Wiley & Sons.

de Smit, M.H., and van Duin, J. (1990). Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proc Natl Acad Sci* 87(19): 7668-7672.

Dong, W., Cheng, Z.-j., Lei, C.-l., Wang, X.-l., Wang, J.-l., Wang, J., Wu, F.-q., Zhang, X., Guo, X.-p., and Zhai, H.-q. (2014). Overexpression of folate biosynthesis genes in rice (*Oryza sativa* L.) and evaluation of their impact on seed folate content. *Plant Foods Hum Nutr* 69: 379-385.

Gallie, D.R. (1993). Posttranscriptional regulation of gene expression in plants. *Annu Rev Plant Biol* 44(1): 77-105.

Gofir, A., Wibowo, S., Hakimi, M., Putera, D.D., and Satriotomo, I. (2022). Folic acid treatment for patients with vascular cognitive impairment: a systematic review and meta-analysis. *Int J Neuropsychopharmacol* 25(2): 136-143.

Goldsbrough, A.P., Albrecht, H., and Stratford, R. (1993). Salicylic acid‐inducible binding of a tobacco nuclear protein to a 10 bp sequence which is highly conserved amongst stress‐inducible genes. *Plant J* 3(4): 563-571.

Gorelova, V., Ambach, L., Rébeillé, F., Stove, C., and Van Der Straeten, D. (2017). Folates in plants: research advances and progress in crop biofortification. *Front Chem* 5: 21.

Gräwert, T., Fischer, M., and Bacher, A. (2013). Structures and reaction mechanisms of GTP cyclohydrolases. *IUBMB life* 65(4): 310-322.

Japelaghi, R.H., Haddad, R., and Garoosi, G.-A. (2011). Rapid and efficient isolation of high quality nucleic acids from plant tissues rich in polyphenols and polysaccharides. *Mol Biotechnol* 49: 129-137.

Liang, Q., Wang, K., Liu, X., Riaz, B., Jiang, L., Wan, X., Ye, X., and Zhang, C. (2019). Improved folate accumulation in genetically modified maize and wheat. *J Exp Bot* 70(5): 1539-1551.

Maier, J., Witter, K., Gutlich, M., Ziegler, I., Werner, T., and Ninnemann, H. (1995). Homology cloning of GTP-cyclohydrolase-I from various unrelated eukaryotes by reverse-transcription polymerase chain-reaction using a general set of degenerate primers. *Biochem Biophys Res Commun* 212(2): 705-711.

Maita, N., Hatakeyama, K., Okada, K., and Hakoshima, T. (2004). Structural basis of biopterin-induced inhibition of GTP cyclohydrolase I by GFRP, its feedback regulatory protein. *J Biol Chem* 279(49): 51534-51540.

Meyer, Y., Vignols, F., and Reichheld, J.P. (2002). Classification of plant thioredoxins by sequence similarity and intron position. *Meth Enzymol* 347: 394-402.

Morowvat, M.H., Babaeipour, V., Rajabi-Memari, H., Vahidi, H., and Maghsoudi, N. (2014). Overexpression of recombinant human beta interferon (rhINF-β) in periplasmic space of *Escherichia coli*. *Iran J Pharm Res* 13(Suppl): 151.

Nar, H., Huber, R., Auerbach, G., Fischer, M., Hösl, C., Ritz, H., Bracher, A., Meining, W., Eberhardt, S., and Bacher, A. (1995). Active site topology and reaction mechanism of GTP cyclohydrolase I. *Proc Natl Acad Sci* 92(26): 12120-12125.

Nunes, A.C., Kalkmann, D.C., and Aragao, F.J. (2009). Folate biofortification of lettuce by expression of a codon optimized chicken GTP cyclohydrolase I gene. *Transgenic Res* 18(5): 661-667.

Prabhu, A.A., Veeranki, V.D., and Dsilva, S.J. (2016). Improving the production of human interferon gamma (hIFN-γ) in Pichia pastoris cell factory: An approach of cell level. *Process Biochem* 51(6): 709-718.

Proctor, J.R., and Meyer, I.M. (2013). C o F old: an RNA secondary structure prediction method that takes co-transcriptional folding into account. *Nucleic Acids Res* 41(9): e102-e102.

Ramírez Rivera, N.G., García‐Salinas, C., Aragão, F.J., and Díaz de la Garza, R.I. (2016). Metabolic engineering of folate and its precursors in Mexican common bean (*Phaseolus vulgaris* L.). *Plant Biotechnol J* 14(10): 2021-2032.

Rebelo, J., Auerbach, G., Bader, G., Bracher, A., Nar, H., Hösl, C., Schramek, N., Kaiser, J., Bacher, A., and Huber, R. (2003). Biosynthesis of pteridines. Reaction mechanism of GTP cyclohydrolase I. *J Mol Biol* 326(2): 503-516.

Rossignol, D.A., and Frye, R.E. (2021). Cerebral folate deficiency, folate receptor alpha autoantibodies and leucovorin (folinic acid) treatment in autism spectrum disorders: a systematic review and meta-analysis. *J Pers Med* 11(11): 1141.

Sahr, T., Ravanel, S., and Rébeillé, F. (2005). Tetrahydrofolate biosynthesis and distribution in higher plants. *Biochem* 33(4): 758-762.

Sahrawy, M., Hecht, V., Lopez-Jaramillo, J., Chueca, A., Chartier, Y., and Meyer, Y. (1996). Intron position as an evolutionary marker of thioredoxins and thioredoxin domains. *J Mol Evol* 42: 422-431.

Shlobin, N.A., LoPresti, M.A., Du, R.Y., and Lam, S. (2020). Folate fortification and supplementation in prevention of folate-sensitive neural tube defects: a systematic review of policy. *J Neurosurg Pediatr* 27(3): 294-310.

Sigrist, C.J., Cerutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M., Bairoch, A., and Bucher, P. (2002). PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief Bioinformatics* 3(3): 265-274.

Steinmetz, M.O., Plüss, C., Christen, U., Wolpensinger, B., Lustig, A., Werner, E.R., Wachter, H., Engel, A., Aebi, U., and Pfeilschifter, J. (1998). Rat GTP cyclohydrolase I is a homodecameric protein complex containing high-affinity calcium-binding sites. *J Mol Biol* 279(1): 189-199.

Strobbe, S., and Van Der Straeten, D. (2017). Folate biofortification in food crops. *Curr Opin Biotechnol* 44: 202-211.

Stuart, G.W., Searle, P.F., and Palmiter, R.D. (1985). Identification of multiple metal regulatory elements in mouse metallothionein-I promoter by assaying synthetic sequences. *Nature* 317(6040): 828-831.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24(8): 1596-1599.

JPMB
Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

مقاله پژوهشی

همسانه‌سازی مولکولی و تجزیه و تحلیل *in silico* ژن *Vvgtpch I* از انگور

نادیا اسلامی بجنوردی، رحیم حداد *، قاسمعلی گروسی*، رضا حیدری--جاپلقی

گروه مهندسی بیوتکنولوژی، دانشکده کشاورزی و منابع طبیعی، دانشگاه بین‌المللی امام خمینی (ره)، قزوین، ایران

چکیده: توالی چارچوب خوانش باز (ORF) رمزگردان پلی‌پپتید GTP سیکلوهیدرولاز I (GTPCH I) از بافت حبه انگور (*Vitis vinifera* L. cv. Askari) جداسازی و همسانه‌سازی شد. توالی ORF به طول ۱۳۳۸ نوکلئوتید رمزگردان یک پلی‌پپتید به طول ۴۴۵ اسیدآمینه است. وزن مولکولی و نقطه ایزوالکتریک توالی پلی‌پپتیدی به ترتیب برابر ۴۸/۶۵ kDa و ۶/۴۳ محاسبه شد. توالی ژن *Vvgtpch* I به طول ۴۹۶۴ نوکلئوتید حاوی دو اگزون (۱۶۹ bp و ۱۱۶۹) و یک اینترون (۲۶۷۶ bp) است. بررسی فیلوژنتیکی نشان داد که ژن *Vvgtpch* I شباهت بالایی با توالی‌های *gtpch* I از سایر گیاهان مانند هلو (۷۲٪)، کاکائو (۷۲٪)، ، توت‌فرنگی (۷۰٪) و صنوبر (۶۹٪) دارد. بررسی ساختار دوم توالی mRNA حاصل از ژن *Vvgtpch* I نشان داد که رمز آغازین ATG به شکل آزادانه در ساختار دوم mRNA قرار گرفته و نشان‌دهنده اتصال و ترجمه کارآمد رونوشت توسط ریبوزوم است. مشابه با تمام *gtpch* Iها از منابع مختلف، بررسی مدل‌سازی مولکولی این ژن نشان داد که ساختار منومر توالی پلی‌پپتیدی VvGTPCH I حاوی ۱۰ مارپیچ α و ۸ صفحه β است. علاوه بر این، تجزیه و تحلیل *in silico* ناحیه راه‌انداز ژن *Vvgtpch* I وجود تعدادی عناصر *Cis-acting* پاسخ‌دهنده به پیام‌های محیطی را اثبات نمود. وجود تعدادی عناصر تنظیمی در ناحیه راه‌انداز نشان می‌دهد که ژن *Vvgtpch* I می‌تواند به انواع مختلفی از پیام‌های بیرونی و درونی از قبیل گرما، فلزات سنگین، نور و هورمون‌های گیاهی پاسخ دهد.

کلمات کلیدی: فولات، تجزیه و تحلیل *in silico*، عناصر تنظیمی، همسانه‌سازی، ناحیه راه‌انداز.

ویراستار علمی
دکتر احمد ارزانی،
دانشگاه صنعتی اصفهان، ایران

تاریخ
دریافت: ۶ اسفند ۱۴۰۱
پذیرش: ۱۶ مهر ۱۴۰۲
چاپ: ۵ بهمن ۱۴۰۲

نویسنده مسئول
دکتر رحیم حداد
r.haddad@eng.ikiu.ac.ir

ارجاع به این مقاله
Eslami Bojnourdi, N., Haddad, R., & Heidari-Japelaghi, R. (2023). Molecular cloning and *in silico* analysis of a GTP cyclohydrolase I gene from grape. *J. Plant Mol. Breed 11* (1): 1-16. 10.22058/JPMB.2023.1990428.1271.

JPMB | Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

Genetic diversity of Robusta coffee (*Coffea canephora* Pierre Ex. A. Froehner) using EST-SSR markers

Mohammed Baba Nitsa*[1], Alexander Chukwunweike Odiyi [2], Benjamin Oluwole Akinyele [2], Olaiya Pater Aiyelari [2], lawrence Stephen Fayeun [2]

[1.] Crop Improvement Division, Cocoa Research Institute of Nigeria, Ibadan, Nigeria
[2.] Department of Crop, Soil and Pest Management, Federal University of Technology, Akure, Ondo State, Nigeria

Original paper

OPEN ACCESS

Edited by

Dr. Ehsan Shokri,
Agricultural Biotechnology Research Institute of Iran (ABRII), Iran

Date

Received: 14 May 2023
Accepted: 19 December 2023
Published: 28 January 2024

Correspondence

Dr. Mohammed Baba Nitsa
bmohammed241@yahoo.com

Citation

Baba Nitsa, M., Odiyi, A., Akinyele, B. O., Aiyelari, O. P., Fayeun, L. S. (2023). Genetic diversity of Robusta coffee (*Coffea canephora* Pierre Ex. A. Froehner) using EST-SSR markers. *J. Plant Mol. Breed 11* (1): 17-27. 10.22058/JPMB.2023.2002371.1273

Abstract: Research on genetic diversity demonstrated insights into variations that can be valuable for enhancing food security. Expressed sequence tag (EST)-simple sequence repeat (SSR) markers were employed to assess variations among thirty-nine robusta coffee accessions. In this study, the EST-SSR markers utilized identified a total of 15 alleles, averaging 3.0 alleles per locus. Primer CESR02 exhibited the highest polymorphic information content value at 0.59, the greatest genetic diversity value at 0.66, and the highest number of alleles (4). Primers CESR04 and CESR05 revealed the highest percentage polymorphism reaching 42.86. The phylogenetic dendrogram clustered the accessions into three main groups and six subgroups. Accessions 3, 26, and 30 were identified as the most genetically distinct. The most genetically related accessions represented 43.59%, while the most distinct accession recorded 2.56%. In the principal coordinate analysis (PCoA), the least genetically similar accessions constituted 20.51%, whereas 28.21% of the accessions demonstrated genetic similarity, representing the highest grouping. The findings of this study highlight the utility of EST-SSR markers in identifying and categorizing coffee accessions. This approach offers valuable insights into the genetic diversity among robusta coffee accessions, facilitating further efforts in their improvement and preservation.

Keywords: robusta coffee, germplasm, accessions, EST-SSR markers, genetic diversity.

Introduction

Coffee is an important horticultural crop, beloved non-alcoholic beverage consumed worldwide (Samoggia and Riedel, 2019). Owing to its substantial economic impact on developing nations in the past, coffee held the position of being the second most valuable commodity in the global market following oil (Cheserek et al., 2020). A proper understanding of the crop species, including the cultivated robusta coffee species, necessitates essential genomic knowledge (Xu et al., 2017). The presence of natural genetic variation in robusta coffee is advantageous for improvement initiatives just because it provides the foundation for genetic diversity, enabling the emergence of new species (Kumar et al., 2016). Evolution is the process through which new species develop from older ones. The study of genetic diversity has a significant impact on how a species evolves, enabling it to adapt to a new environment. Low levels of adaptability of crop species to new environments can result in a loss of genetic diversity in those species, rendering them more vulnerable to various challenges including extinction and pests and diseases attack (Arzani and Ashraf, 2016). The success of breeding initiatives relies on the genetic diversity present in *Coffea* spp. populations (Machado et al., 2017).

The understanding of genetic diversity in crop species has tremendously benefited from the numerous efforts invested in utilizing molecular markers to identify accessions that are genetically diverse (Rauf et al., 2010). Due to their independence from environmental factors, molecular markers have enabled the detection of genetic variation at the DNA level (Ebrahimi et al., 2017). Numerous molecular markers have been developed and employed over an extended period as a method for assessing the genetic diversity of crop species. These markers encompass random amplified polymorphic (RAPD), single-primer amplification reaction (SRAP), inter simple sequence repeats (ISSR), amplified fragment length polymorphism reaction (AFLP), and microsatellite (SSR) markers (Aboukhalid et al., 2017). Due to its broad content distribution, high polymorphism, neutrality, transferability, and repeatability, expressed sequence tag SSRs (EST-SSRs) markers are considered as one of the most attractive markers (Liu et al., 2019). EST–SSR markers, situated in the coding region of the genome, exhibit high transferability between species, codominant inheritance, and cost-effective (Mishra et al., 2011; Xu et al., 2013). In various crops, including coffee plants, gSSR and EST-SSR markers have been utilized for exploring genetic diversity, genotyping, and cultivar fingerprinting, among other applications (Zhu et al., 2013; Sousa et al., 2022). Under the Brazilian project, the development of an EST-SSR marker date gene bank for robusta coffee was made possible (Vieira et al., 2006). The majority of robusta coffee grown in Nigeria, particularly landraces, lacked sufficient information about their genetic diversity. This contributes to the genetic erosion of the majority of landraces/ cultivars that possess desirable traits and could be valuable in breeding programs. Therefore, the objective of this study was to employ EST-SSR markers to assess the extent of genetic variation cultivated among robusta coffee accessions.

Materials and Methods

Plant materials

The study encompassed thirty-nine accessions of robusta coffee which comprises of accessions obtained from coffee germplasm of Cocoa Research Institute of Nigeria and accessions (landraces) sourced from farmers. A list of the coffee accessions, their corresponding codes, locations and coordinates of the collection areas are provided in Table 1. The collected accessions were planted and maintained in the field at Cocoa Research Institute of Nigeria from 2017 to 2018. Cocoa Research Institute of Nigeria is located at latitude 7° 12' 55" N and 3° 51' 45" E, at an altitude of 122 m above sea level, with an average annual rainfall ranging between 1200-1500 mm. During the dry season, the plants were irrigated using watering cans. Five EST-SSR markers were used for this study. About 500 g of fresh, young, disease-free leaf samples were collected from the coffee plants in the field. These samples were immediately placed into the cooler box containing ice blocks before undergoing lyophilization. Genomic DNA was extracted from young lyophilized leaf samples of the thirty nine robusta coffee accessions. The extraction process

involved grinding the samples, following the Sodium Dodecyl Sulphate (SDS) method (Dellaporta et al., 1983). The quantity and quality of the extracted DNA were checked by running of the individual accession in a mixture of 5 µL of DNA and 0.5 µg/mLof ethidium bromide in 1 % agarose gel electrophoresis. The electrophoresis was conducted for 30 min at 100 voltage. The quantity and purity of DNA were confirmed using a NanoDrop spectrophotometer.

Table 1. List of Robusta coffee: accessions number, code, sources and coordinate of collection.

Accession no.	Code	Source	GPS DD coordinate		
			Latitude (0)	Longitude (0)	Altitude (m)
T1	M36	Germplasm	7.20354	3.86183	422.25
T2	SG2	Germplasm	7.21880	3.86473	464.83
T3	IYA2	Landrace	7.79267	5.80625	1595.50
T4	ORA1	Landrace	7.86340	5.74112	1746.89
T5	EJU2	Landrace	8.05854	5.75192	1418.88
T6	DAC	Landrace	7.86157	6.07288	1407.06
T7	OMU1	Landrace	7.77586	5.77326	1643.60
T8	C105	Germplasm	7.20353	3.86472	422.25
T9	KB3	Landrace	7.82385	6.07772	1385.77
T10	ZN5	Germplasm	7.20660	3.86193	463.25
T11	OLU1	Landrace	7.49692	5.63078	1143.70
T12	OMU2	Landrace	7.77584	5.77324	1643.60
T13	OMA	Landrace	6.92361	3.44661	268.50
T14	OLU2	Landrace	7.49692	5.63078	1143.70
T15	ATK	Landrace	6.92245	3.44086	163.63
T16	T24	Germplasm	7.20353	3.86182	422.25
T17	ORA2	Landrace	7.86340	5.74112	1746.89
T18	AJA 1	Ajassor	5.87446	8.80908	452.21
T19	AJA 2	Ajassor	5.87444	8.81817	452.21
T20	OKU	Okundi	5.96118	8.77028	348.13
T21	E130	Germplasm	7.20353	3.86182	422.25
T22	C111	Germplasm	7.20354	3.86181	422.25
T23	EJU1	Landrace	8.05854	5.75192	1418.88
T24	IFE1	Landrace	7.89475	5.77761	1623.89
T25	C36	Germplasm	7.20357	3.86762	422.25
T26	D57	Germplasm	7.20359	3.86683	422.25
T27	OWE2	Landrace	7.19928	5.02489	1097.97
T28	IYA3	Iyamoye	7.79308	5.80627	1590.45
T29	OWE1	Landrace	7.19847	5.02451	848.81
T30	IYA1	Iyamoye	7.79306	5.80631	1595.50
T31	IBE	Landrace	5.56055	7.62828	165.56
T32	A81	Germplasm	7.20376	3.86572	422.25
T33	IFE2	Landrace	7.77611	5.77307	1641.23
T34	SEK	Landrace	7.48864	5.63687	1076.68
T35	IYA4	Iyamoye	7.79305	5.80631	1698.00
T36	C96	Germplasm	7.20445	3.87281	433.25
T37	SG1	Germplasm	7.21891	3.87571	465.81
T38	F63	Germplasm	7.21992	3.87632	432.35
T39	EJU3	Landrace	8.05854	5.75192	1418.21

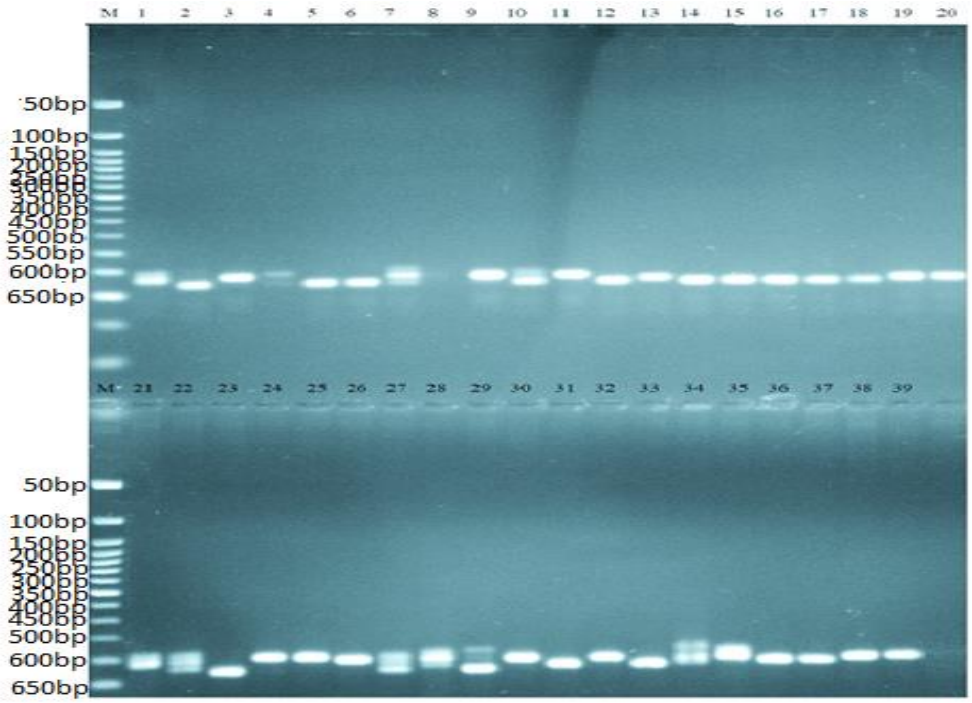


Figure 1. Image of DNA bands of the thirty-nine coffee accessions on CESR02 EST-SSR markers.

EST-SSR markers and PCR amplification

The DNA extracted from different samples were appropriately coded. PCR conditions for all the EST-SSR primers pairs were optimized using template DNA. The PCR reactions were carried out in a total volume of 20 µL, comprising 1.0 µL template DNA, 0.2 µL of each primer, 10 µL Mix (*Taq* polymerase, dNTPs, and PCR buffer), and 8.8 µL of sterile distilled water. PCR amplification was carried out with an initial denaturation step at temperature of 94°C for 5 min, followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 57°C for 30 s, extension at 72°C for 20 s, and a final extension step for 10 min at 72°C. The PCR products were separated on a 2.0% agarose gel, stained with ethidium bromide and the gels were visualized under ultraviolet (UV) light. A Polaroid camera was used to photograph the gels through a UV gel documentation system. Primers were obtained from the published article by Hendre and Aggarwal (2014). The list and details of EST-SSR primers are provided in Table 2.

Data Analysis

DNA bands corresponding to different alleles were identified through gel electrophoresis These observed DNA bands were then converted into a binary matrix where the presence of an allelic variation was scored as 1 and the absence of an allelic variation was scored as 0 (Hildebrand et al., 1992; He et al., 2021). Polymorphism among thirty-nine accessions of robusta coffee was evaluated using EST-SSR markers. To estimate genetic diversity the following parameters were determined; number of alleles, gene diversity, polymorphic information content (PIC) and percentage polymorphism at each EST-SSR locus. Power marker version 3.25 was employed for these analyses. (Liu and Muse, 2005). The resulting similarity matrix was used for plotting a phylogenetic dendrogram to illustrate the relationship among the accessions in cluster form using Power marker version 3.25 (Liu and Muse, 2005).

Principal coordinate analysis (PCoA) was constructed with the use of PAST Software version 3.26b (Hammer, 2001). The allelic data was used to calculate the polymorphic information content (PIC) value based on the formula described by Anderson et al. (1993)

$$\text{PIC} = 1 - \sum_{i=1}^{k} (P_i)^2$$

Where k is the total number of alleles detected for an EST-SSR marker, and Pi represents the frequency of the i^{th} allele.

The percentage of polymorphism loci (P) was calculated by dividing the number of polymorphic bands by the total number of amplified bands for each primer.

$P = \frac{\text{Np}}{\text{Nt}}$ x 100

Where Np is the number of polymorphic alleles and Nt is the total number of alleles (polymorphic and monomorphic alleles).

Table 2. EST-SSR Identification no, primers sequences, temperature and product sizes.

S/N	Markers ID	Primer sequence	Tm (C)	Product sizes (bp)
1	CCESSRO1	TGGTAGCACTGTCGGAAGCATAT GACCCATCTAACTTGCTGCATTTT	58	450
2	CCESSRO2	AGGGGCTGGTTATTTTTTGGG GGGGGTAAATACGGGAAAGCAGA	58	650
3	CCESSRO3	GCAGCAACAATCACTTCCACAGC TGCTGTTGTAACTGCGGGATTTG	60	650
4	CCESSRO4	TTCTGGCCGATTGATTGTGAT GCGACAAGGCTGACAAACTACTAC	58	600
5	CCESSRO5	GGCGCTAGAGTTGGTTGTTTGC CAGGCATTGGAACCAGCGAAC	60	550

Table 3. EST-SSR primers and allelic information of 39 Robusta coffee accessions.

Primer ID	Number of allele	Sample size	Polymorphic bands	Total bands scored	Allelic diversity	Gene diversity	PIC	Percentage polymorphic
CESRO1	2.00	39	2	6	0.05	0.50	0.37	33.33
CESRO2	4.00	39	4	11	0.10	0.66	0.59	36.36
CESRO3	3.00	39	1	7	0.02	0.56	0.49	14.28
CESRO4	3.00	39	3	7	0.07	0.56	0.47	42.86
CESRO5	3.00	39	6	14	0.15	0.53	0.43	42.86
Mean	3.0	39	3.2	9	0.08	0.56	0.47	33.94

Results

EST-SSR analysis

A cumulative of 15 alleles was identified across the 39 robusta coffee accessions at the five EST-SSR loci, as illustrated in Table 3. The allele count per primer varied between 2 to 4, with an average of 3.0 alleles at each locus. The polymorphism, as indicated by EST-SSR loci, demonstrated the Polymorphic Information Content (PIC) ranging from 0.37 (CESR01) to 0.59 (CESR02), averaging at 0.47. The lowest PIC value of 0.37 was observed in CESRO1 primer with 2 alleles and the highest PIC value of 0.59 was detected in CESRO2 with 4 alleles. The allelic diversity values ranged from 0.02 to 0.15 with an average of 0.08 diverse alleles per locus. Also, the highest gene diversity was revealed in loci CESR02 (0.66), and the lowest gene diversity was recorded in loci CESR01 (0.50) with an average value of 0.56 per locus. The five EST-SSR primers amplified a total of 45 scorable bands. The primer CESR05 had the highest scorable bands of 14, while primer CESR01 had the lowest scorable bands of 6. The total number of polymorphic bands was 16 with primer CESR05 indicating the highest number of bands (6), while primer CESR03 recorded the least number of bands (1). The highest percentage polymorphism of 42.86% was dictated in primer CESR04 and CESR05 and the lowest percentage polymorphism of 14.28% was revealed by primer CESR03. The percentage polymorphism ranged from 14.28% to 42.86% with average of 33.94%. Figure 1 shows the image of the amplified DNA of the thirty-nine robusta coffee accessions on the locus.

Genetic relationship clustering

The Jaccard's (Jaccard, 1908) similarity coefficients were used based on the genetic matrix data to construct the phylogenetic dendrogram which grouped the accessions based on similarity and dissimilarity into three main clusters and six sub-clusters with genetic distance values ranging from 0.33 to 1.00 respectively. The phylogenetic dendrogram which shows the genetic relationship is represented in (Figure 2). Accessions 3, 26, and 30 were the most genetically distinct among all the accessions. 43.59% of the total accessions were genetically related, while only 2.56% of the accessions exhibited high genetic distance.

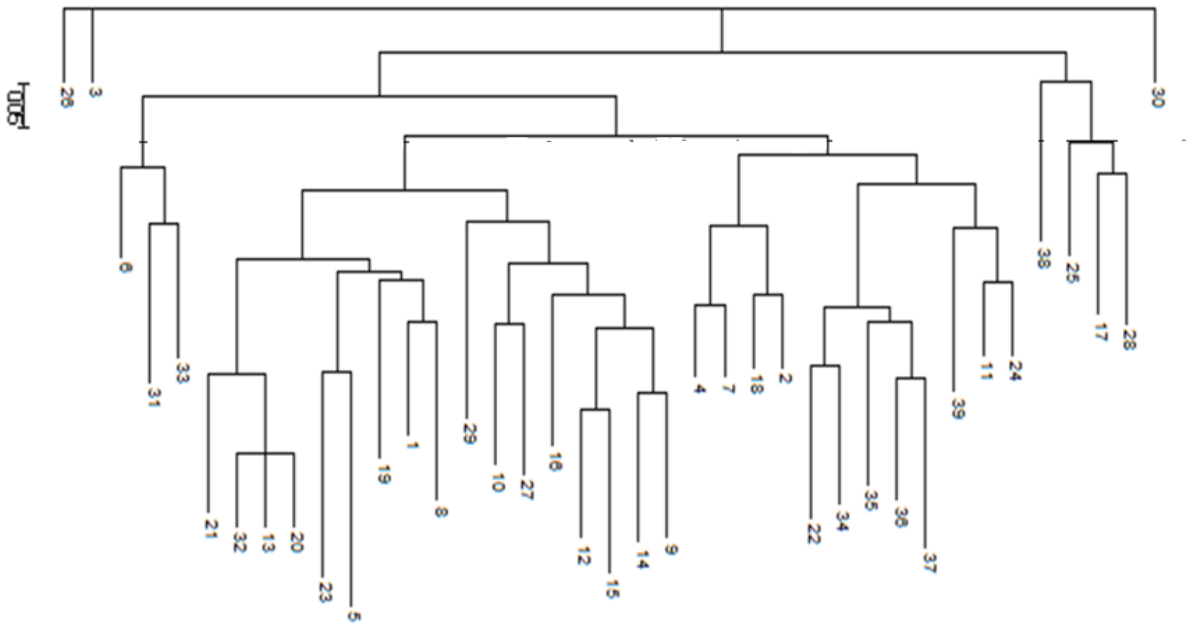


Figure 2. Phylogenetic dendrogram illustrating genetic relationship among the thirty-nine accessions of robusta coffee.

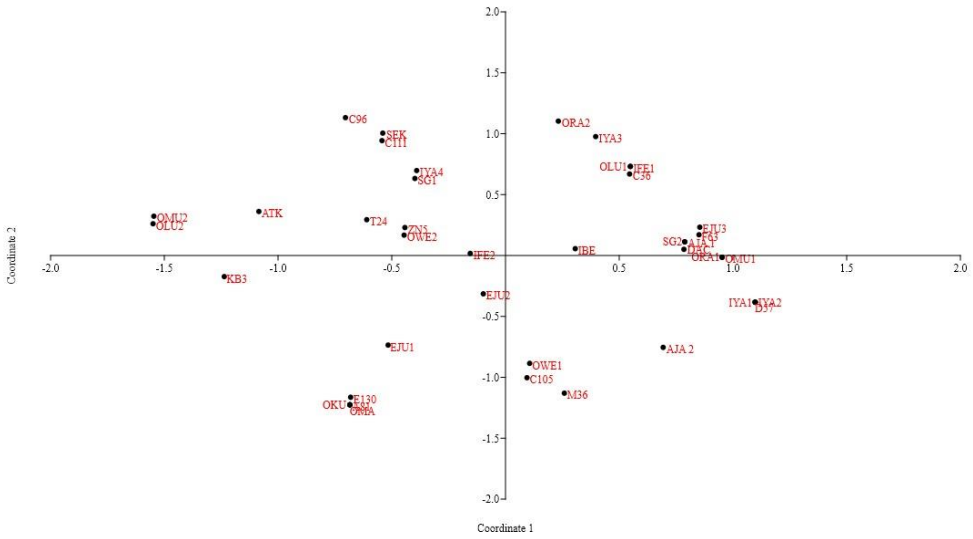


Figure 3. Principal coordinates analysis (PCoA) revealing the relationship among thirty nine Robusta coffee accessions based on EST-SSR markers

Principal coordinate analysis

Principal coordinate analysis (PCoA) is a factorial analysis used to identify the distance and relatedness among the coffee accessions. To understand the genetic diversity among accessions of robusta coffee, the principal coordinate analysis (PCoA) (Figure 3) was conducted. Based on their genetic similarities, the 39 accessions were divided into four main groups by the PCoA. Two of the groups had 11 accessions clustered together, representing 28.21% of the total accessions. The remaining two groups had 8 and 9 accessions grouped, which accounted for 20.51% and 23.08% respectively. This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn. All figures and tables should be cited in the main text (such as Figure 1 and Table 1) and numbered in the order they appear in the text.

Discussion

The EST-SSR markers provide useful information on the genetic diversity of coffee with the use of EST-SSR markers, as depicted in Figure 1. The primers CESR02 indicated a high polymorphic information content value of 0.59 as well as a high genetic diversity of 0.66 (Table 2). findings of the current study align withwith those of Missio et al. (2011) who similarly reported high polymorphism in robusta coffee. Prakash et al. (2005) noted elevated genetic diversity in the India robusta coffee gene pool when compared with the coffee originated from Africa. This observation was made using SSR and AFLP markers. The notable level of polymorphism observed may be attributed to the allogamous nature induced by self-incompatibility of the robusta coffee species (Depecker et al., 2023). These findings corroborate the wide genetic diversity observed by Sousa et al. (2022) who worked on *C. canaphora* using SSR markers. They justified the observed genetic diversity by attributing it to the receptivity of pollen grains from other plants, contributing to heterozygosity.

Among the primers, primer CESR02 stands out as more informative, as evidenced by its the highest PIC value and a greater number of alleles which are indicators used to measure the polymorphism of a marker (Guo and Elston, 1999). This implies that primer CESR02 demonstrated a heightened capacity to discern the presence of genetic diversity

among the studied coffee accessions. The mean value of allele number in this study was 3.0 different from the mean allele numbers of 10.5 reported by Poncet et al. (2006), with the use of 25 EST-SSR markers. As per Poncet et al. (2006) varied mean numbers of alleles have been documented in *Coffea* spp. According to Tshilenge et al. (2009), a high level of polymorphism was observed among robusta coffee accessions using RAPD markers. These results indicated that the EST-SSR markers used in this study revealed some levels of genomic DNA diversity in robusta coffee. Zhai et al. (2013) reported a low level of polymorphism in radish (*Rephanu sativus* L.) species with the use of EST-SSR markers. These findings demonstrate that the PIC value in a population depends highly on the genetic constitution assessed and the genotypes used (Tasma and Arumsari, 2013). Thus the observed PIC value becomes crucial (Missio et al., 2011). The genetic relationship among the 39 robusta coffee accessions was assessed through the phylogenetic dendrogram and the principal coordinate analysis (PCoA) which clustered the accessions into different groups. The clustering of landraces alongside with germplasm accessions suggests their genetic similarity, providing guidance to avoid inadvertent duplication during selection. Moreover, a significant proportion of the accessions clustered in the phylogenetic dendrogram exhibited similar groupings in the principal coordinate analysis (PCoA). Nonetheless, within the PCoA clustering, there were instances of overlapping accessions, suggesting the potential duplication of these entries. This overlap offers valuable insights into identifying duplicate accessions and serves as a guide to prevent the inadvertent of the same accession for the breeding program.

Conclusion

The presence of genetic variability in crops is crucial, as it creates a chance for the breeders to develop new varieties. Among the EST-SSR markers, primer CESR02 proved to be the most informative regarding the genetic diversity of robusta coffee studied. This is attributed to its capability to produce a substantial number of alleles, coupled with the highest PIC value. The identification of genetic variation in this study, suggests that the coffee accessions have the potential to serve as valuable genetic material for breeding programs. The data acquired through the application of EST-SSR markers, would aid in evaluating the genetic variation among coffee accessions. The congruence observed in clustering of the accessions through both phylogenetic dendrogram and PCoA suggests that these two methods can be utilized to assess the genetic relatedness of the robusta coffee accessions. This is because they complement each other effectively. The outcomes of this research will simplify the management, conservation, and identification of landrace accessions that exhibit genetic relatedness to those in the germplasm. It is therefore recommended that to incorporate genetically distinct accessions into coffee germplasm to increase the variation with coffee populations.

Supplementary Materials

No supplementary material is available for this article.

Author contributions

Conceptualization, M.B.N. and AC.O.: methodology, M.B.N.; A.C.O; BOA; OPA; L.S.F.: software, M.B.N.: validation, M.B.N.; A.C.O.; and L.S.F,: formal analysis, M.B.N.: investigation, M.B.N.: resources, M.B.N.; and CRIN.: data curation, M.B.N.; A.C.O.; and L.S.F.: writing—original draft preparation, M.B.N.: writing— review and editing, M.B.N.; A.C.O.; and L.S.F.: visualization, M.B.N.: supervision, A.C.O.; B.O.A.; O.P.A.; and CRIN.: project administration, M.B.N.; A.C.O.; and CRIN.: funding acquisition, M.B.N.: All authors have read and agreed to the published.

Funding

This research received no external funding.

Acknowledgments

The authors wish to acknowledge the management of the Cocoa Research Institute of Nigeria, Ibadan, Nigeria for the moral support and permission granted to carry out this research work.

Conflict of interest statement

The authors declare no conflict of interest.

References

Aboukhalid, K., Machon, N., Lambourdière, J., Abdelkrim, J., Bakha, M., Douaik, A., Korbecka-Glinka, G., Gaboun, F., Tomi, F., and Lamiri, A. (2017). Analysis of genetic diversity and population structure of the endangered *Origanum compactum* from Morocco, using SSR markers: Implication for conservation. *Biol Conserv* 212: 172-182.

Anderson, J.A., Churchill, G., Autrique, J., Tanksley, S., and Sorrells, M. (1993). Optimizing parental selection for genetic linkage maps. *Genome* 36(1): 181-186.

Arzani, A., and Ashraf, M. (2016). Smart engineering of genetic resources for enhanced salinity tolerance in crop plants. *Crit Rev Plant Sci* 35(3): 146-189.

Cheserek, J.J., Ngugi, K., Muthomi, J.W., and Omondi, C.O. (2020). Assessment of Arabusta coffee hybrids [*Coffea arabica* L. X *Tetraploid Robusta* (*Coffea canephora*)] for green bean physical properties and cup quality. *Afr J Food Sci* 14(5): 119-127.

Dellaporta, S.L., Wood, J., and Hicks, J.B. (1983). A plant DNA minipreparation: version II. *Plant Mol Biol Report* 1: 19-21.

Depecker, J., Verleysen, L., Asimonyio, J.A., Hatangi, Y., Kambale, J.-L., Mwanga Mwanga, I., Ebele, T., Dhed'a, B., Bawin, Y., and Staelens, A. (2023). Genetic diversity and structure in wild Robusta coffee (*Coffea canephora* A. Froehner) populations in Yangambi (DR Congo) and their relation to forest disturbance. *Heredity* 130(3): 145-153.

Ebrahimi, F., Majidi, M.M., Arzani, A., and Mohammadi-Nejad, G. (2017). Association analysis of molecular markers with traits under drought stress in safflower. *Crop Pasture Sci* 68(2): 167-175.

Guo, X., and Elston, R. (1999). Linkage information content of polymorphic genetic markers. *Hum Hered* 49(2): 112-118.

Hammer, O. (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontol* 4: 9.

He, Z., Wang, M., Xu, X., You, Y., Chen, Y., and Zhang, L. (2021). The Development of EST-SSR markers from transcriptome sequencing and genetic diversity of twenty genotypes with high yields of cornus wilsoniana, an important wood oil plant. *Molecul Plant Breed* 12.

Hendre, P.S., and Aggarwal, R.K. (2014). Development of genic and genomic SSR markers of robusta coffee (*Coffea canephora* Pierre Ex A. Froehner). *PLoS One* 9(12): e113661.

Hildebrand, C.E., Torney, D.C., and Wagner, R.P. (1992). Informativeness of polymorphic DNA markers. *Los Alamos Sci* 20(20): 100-102.

Jaccard, P. (1908). Nouvelles recherches sur la distribution florale. *Bull Soc Vaud Sci Nat* 44: 223-270.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33(7): 1870-1874.

Liu, K., and Muse, S.V. (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21(9): 2128-2129.

Liu, Y., Geng, Y., Song, M., Zhang, P., Hou, J., and Wang, W. (2019). Genetic structure and diversity of *Glycyrrhiza* populations based on transcriptome SSR markers. *Plant Mol Biol Rep* 37: 401-412.

Machado, C., Pimentel, N., Golynsk, A., Ferreira, A., Vieira, H., and Partelli, F. (2017). Genetic diversity among 16 genotypes of *Coffea arabica* in the Brazilian cerrado. *Genet Mol Res* 16(3).

Mishra, R.K., Gangadhar, B.H., Yu, J.W., Kim, D.H., and Park, S.W. (2011). Development and characterization of EST based SSR markers in Madagascar periwinkle (*Catharanthus roseus*) and their transferability in other medicinal plants. *Plant Omics* 4(3): 154-161.

Missio, R., Caixeta, E., Zambolim, E., Pena, G., Zambolim, L., Dias, L., and Sakiyama, N. (2011). Genetic characterization of an elite coffee germplasm assessed by gSSR and EST-SSR markers. *Genet Mol Res* 10(4): 2366-2381.

Poncet, V., Rondeau, M., Tranchant, C., Cayrel, A., Hamon, S., De Kochko, A., and Hamon, P. (2006). SSR mining in coffee tree EST databases: potential use of EST–SSRs as markers for the *Coffea* genus. *Mol Genet Genom* 276: 436-449.

Prakash, N.S., Combes, M.-C., Dussert, S., Naveen, S., and Lashermes, P. (2005). Analysis of genetic diversity in Indian robusta coffee genepool (*Coffea canephora*) in comparison with a representative core collection using SSRs and AFLPs. *Genetics* 52: 333-343.

Rauf, S., da Silva, J.T., Khan, A.A., and Naveed, A. (2010). Consequences of plant breeding on genetic diversity. *Int J Plant Breed* 4(1): 1-21.

Samoggia, A., and Riedel, B. (2019). Consumers' perceptions of coffee health benefits and motives for coffee consumption and purchasing. *Nutrients* 11(3): 653.

Sousa, P., Vieira, H., Santos, E., Viana, A., Boaechat, M., and Partelli, F. (2022). *Coffea canephora*: Heterotic Crosses Indicated by Molecular Approach. *Plants* 11(22): 3023.

Tasma, I.M., and Arumsari, S. (2013). Analisis diversitas genetik aksesi kelapa sawit Kamerun berdasarkan marka SSR. *Ind Crop Prod* 19(4): 194-202.

Tshilenge, P., Nkongolo, K., Mehes, M., and Kalonji, A. (2009). Genetic variation in *Coffea canephora* L.(Var. Robusta) accessions from the founder gene pool evaluated with ISSR and RAPD. *Afr J Biotechnol* 8(3).

Vieira, L.G.E., Andrade, A.C., Colombo, C.A., Moraes, A.H.d.A., Metha, Â., Oliveira, A.C.d., Labate, C.A., Marino, C.L., Monteiro-Vitorello, C.d.B., and Monte, D.d.C. (2006). Brazilian coffee genome project: an EST-based genomic resource. *Braz J Plant Physiol* 18: 95-108.

Xu, G., Liu, C., Huang, L., Wang, X., Zhang, Y., Liu, S., Liao, C., Yuan, Q., and Zhou, X. (2013). Development of new EST-derived SSRs in *Salvia miltiorrhiza* (Labiatae) in China and preliminary analysis of genetic diversity and population structure. *Biochem Syst Ecol* 51: 308-313.

Xu, Q., Zhu, J., Zhao, S., Hou, Y., Li, F., Tai, Y., Wan, X., and Wei, C. (2017). Transcriptome profiling using single-molecule direct RNA sequencing approach for in-depth understanding of genes in secondary metabolism pathways of *Camellia sinensis*. *Front Plant Sci* 8: 1205.

Zhai, L., Liu, L., Zhu, X., Xu, L., Jiang, L., and Gong, Y. (2013). Development, characterization and application of novel expressed sequence tag-simple sequence repeat (EST-SSR) markers in radish (*Raphanus sativus* L.). *Afr J Biotechnol* 12(9).

Zhu, X.C., Raman, H., Wu, H., Lemerle, D., Burrows, G.E., and Stanton, R. (2013). Development of SSR markers for genetic analysis of silverleaf nightshade (*Solanum elaeagnifolium*) and related species. *Plant Mol Biol Rep* 31: 248-254.

JPMB
Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

مقاله پژوهشی

تنوع ژنتیکی قهوه روبوستا (*Coffea canephora Pierre* Ex. A. Froehner) با استفاده از نشانگرهای EST-SSR

محمد بابانیتسا*[1]، الکس اودیبی [2]، بنجامین اولووله‌آکینیل [2]، اولایا پی‌آیلاری [2]، لارنس فایون [2]

[1] بخش بهبود محصول، موسسه تحقیقات کاکائو نیجریه، ایبادان، نیجریه
[2] گروه مدیریت زراعی، خاک و آفات، دانشگاه فناوری فدرال، آکوره، ایالت اوندو، نیجریه

چکیده: تحقیقات در مورد تنوع ژنتیکی، بینش‌هایی را در مورد تغییراتی که می‌تواند برای افزایش امنیت غذایی ارزشمند باشد نشان می‌دهد. نشانگرهای مبتنی بر توالی بیان‌شده (EST) و توالی تکرار ساده (SSR) برای ارزیابی تغییرات در بین ۳۹ نمونه قهوه روبوستا استفاده شد. در این مطالعه، نشانگرهای EST-SSR مورد استفاده، در مجموع ۱۵ آلل را شناسایی کردند که به طور متوسط ۳ آلل در هر مکان است. آغازگر CESR02 بالاترین مقدار محتوای اطلاعات چندشکلی را با ۰/۵۹، بیشترین مقدار تنوع ژنتیکی را با ۰/۶۶ و بیشترین تعداد آلل (۴) را نشان داد. آغازگرهای CESR04 و CESR05 بیشترین درصد چندشکلی را نشان دادند که به۴۲/۸۶ رسید. دندروگرام فیلوژنتیک، نمونه‌ها را به سه گروه اصلی و شش زیر گروه تقسیم نمود. ژنوتیپ‌های ۳، ۲۶ و ۳۰ به عنوان متمایزترین از نظر ژنتیکی شناسایی شدند. بیشترین قرابت بین ژنوتیپ‌های مورد بررسی ۴۳/۵۹ درصد بوده، در حالی که کمترین قرابت ۲/۵۶ درصد مشاهده شد. در تجزیه به مؤلفه اصلی (PCoA)، ژنوتیپ‌های با قرابت ژنتیکی کم و زیاد، به ترتیب ۲۰/۵۱ درصد و ۲۸/۲۱ درصد از نمونه‌ها را تشکیل داده که مبین قدرت بالای تفکیک آنالیز در گروه‌بندی است. در این تحقیق، کاربرد نشانگرهای EST-SSR در شناسایی و طبقه‌بندی ترکیبات قهوه مورد بررسی قرار گرفت. یافته‌های این تحقیق بینش ارزشمندی در مورد تنوع ژنتیکی میان قهوه‌های روبوستا ارائه می‌دهد و تلاش‌های بیشتر در بهبود و حفظ آنها را تسهیل می‌کند.

کلمات کلیدی: قهوه روبوستا، ژرم‌پلاسم، اکسیشن، نشانگرهای EST-SSR، تنوع ژنتیکی.

ویراستار علمی
دکتر احسان شکری،
پژوهشگاه بیوتکنولوژی کشاورزی ایران، سازمان تحقیقات، آموزش و ترویج کشاورزی، ایران

تاریخ
دریافت: ۲۴ اردیبهشت ۱۴۰۲
پذیرش: ۲۸ آذر ۱۴۰۲
چاپ: ۸ آذر ۱۴۰۲

نویسنده مسئول
دکتر محمد بابانیتسا
bmohammed241@yahoo.com

ارجاع به این مقاله
B Baba Nitsa, M., Odiyi, A., Akinyele, B. O., Aiyelari, O. P., Fayeun, L. S. (2023). Genetic diversity of Robusta coffee (Coffea canephora Pierre Ex. A. Froehner) using EST-SSR markers. J. Plant Mol. Breed 11 (1): 17-27. 10.22058/JPMB.2023.2002371.1273

JPMB | Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

Comparison of predicted protein sequences of the omega-3 fatty acid desaturase gene family in some of the oil seed crops

Original paper

Ammar Afkhami Ghadi *[1], Seyyed Jaber Hosseini [2] and Ali Ghanbari [3]

[1.] Department of Genetic Engineering and Plant Breeding, Imam Khomeini International University, Qazvin, Iran
[2.] Department of Agronomy, Tarbiat Modares University, Tehran, Iran
[3.] Department of Plant Production Engineering and Genetics, Zanjan University, Zanjan, Iran

Abstract: This study aimed to conduct a comparative analysis of protein-predicted sequences within the omega-3 fatty acid desaturase (FAD) gene family across various oilseed crops, such as cotton, soybean, rapeseed, and corn. Twenty-five omega-3 FAD genes were distinguished by the removal of similar sequences. Phylogenetic analysis of the omega-3 FAD gene family was detected. Omega-3 FAD gene information, and physical parameters including the number of amino acids, chromosome locations, exon count and other features were obtained. The results showed that the average length of the proteins encoded by the omega-3 fatty acid desaturase proteins was 388.58 amino acid. The molecular weights of these proteins ranged from 22.2 to 51.3 kDa. The phylogenetic tree divided the omega-3 FAD proteins into five clades. Clade 2 comprises the smallest number of omega-3 fatty acid desaturase gene members, whereas clade 1 encompasses the highest number of members. We identified five pairs of orthologous genes among the omega-3 FAD genes and identified a total of twenty distinct conserved motifs. Four of these motifs were associated with encoding the FAD domain, while an additional four were related to the DUF3474 domain. Undoubtedly, characterizing these FADs could offer valuable candidate genes for enhancing new oilseed varieties.

Keywords: conserved motifs, gene family, oil seed plants, omega-3 fatty acid

OPEN ACCESS

Edited by
Dr. Parviz Heidari,
Shahrood University of Technology,
Iran

Date
Received: 1 August 2023
Accepted: 19 October 2023
Published: 28 January 2024

Correspondence
Dr. Ammar Afkhami Ghadi
a.afkhami@sanru.ac.ir

Citation
Afkhami Ghadi, A., Hosseini, S. J., and Ghanbari, A. (2023). Comparison of protein predicted sequences of the omega-3 fatty acid desaturase gene family in some of the oil seed crops. *J. Plant Mol. Breed 11* (1): 28-40. doi: 10.22058/JPMB.2023.2008246.1281.

Introduction

A gene family consists of a group of genes that share significant characteristics. In numerous instances, genes within a family share identical sequences and may possess indistinguishable functions. These genes offer instructions for generating products, such as proteins, that share common components or biochemical functions. Alternatively, in some cases, heterologous genes are categorized into a family because the proteins they produce function together as a cohesive unit or participate in analogous biological processes. Researchers can employ gene families to forecast the functions of newly discovered genes by assessing their resemblance to established genes. Additionally, the similarities among genes within a family can aid to predict the locations and temporal patterns of expression of a particular gene.

Oil seed plants are important crops worldwide, and provide oilseeds for food, feed, and biofuel. Recently, FAD genes have also been identified in several oilseed crops, including olives (Hernández et al., 2016) cotton (Chi et al., 2011; Yurchenko et al., 2014; Liu et al., 2015) and soybean (Chi et al., 2011; Román et al., 2012). Fatty acid desaturase in oil seed plants is an enzyme that removes two hydrogen atoms from a fatty acid, creating a carbon/carbon double bond (Shanklin and Cahoon, 1998). These desaturases are classified as delta indicating that a double bond is created between the third and fourth carbon atoms of the methyl end of the fatty acid. Methyl end desaturases are responsible for produce FAs with two and three double bonds, establishing the n-6 and n-3 series of PUFAs (Alloatti and Uttaro, 2011) In contrast a front-end desaturase inserts a double bond between the pre-existing double bond and the carboxyl end of a fatty acid (Meesapyodsuk and Qiu, 2012).

Information about the fatty acid desaturase genes encoding these enzymes has permitted more detailed analyses of the role of these genes, and their fatty acid products, in plant lipid metabolism and abiotic stress responses. For example, omega-3 fatty acids are known to increase in plants in response to both drought (Martin et al., 1986) and cold temperature (Williams et al., 1988; Guschina and Harwood, 2006) and overexpression of omega-3 desaturases in several transgenic plants has been shown to increase both drought and chilling tolerance (Zhang et al., 2005; Domínguez et al., 2010). The endoplasmic reticulum (ER-localized) desaturases FAD2 and FAD3 are also involved in the production of PUFA components of seed oils (Ohlrogge and Browse, 1995). Given the value of these fatty acids to human nutrition, and to determine the stability of oils during cooking or other food applications, molecular markers for these genes have been developed to evaluate germplasm and identifying oilseed varieties with improved oil compositions (Bocianowski et al., 2012; Yang et al., 2012). a genome wide analysis of Gossypium's omega-3 fatty acid desaturase gene family indicated that 11 omega-3 FAD genes were identified (Ohlrogge and Browse, 1995; Yurchenko et al., 2014). They found that the omega-3 FAD family of cotton includes five distinct genes, two of which encode endoplasmic reticulum-type enzymes and three encode chloroplast-type enzymes. In soybean, the expression of FAD3 and FAD7 is tightly regulated in response to cold temperature (Román et al., 2012). This study aimed to compare the predicted protein sequences of the omega-3 fatty acid desaturase (FAD) gene family in oil seed crops.

Materials and Methods

Database search for properties of omega-3 fatty acid desaturase genes

Six species' omega-3 fatty acid desaturase genes (*Gossypium raimondii, Gossypium arboretum, Glycine soja, Brassica napus, Brassica juncea,* and *Zea mays*) were obtained from the National Centre for Biotechnology Information (NCBI) databases (**http://www.ncbi.nlm.nih.gov/**). Twenty-five Omega-3 fatty acid desaturase genes (5 *Gossypium raimondii*, 2 *Gossypium arboreum*, 9 *Glycine soja*, 3 *Zea mays*, 5 *Brassica napus*, and 1 *Brassica juncea*) were identified. The differentiation was achieved by removing similar sequences. The accession numbers of published omega-3 fatty acid desaturase genes from oil seed plants in this study are listed in Table 1. Phylogenetic analysis of comparative and functional genomics of the Omega-3 fatty acid desaturase gene family was detected of GreenPhylDB. This database contains a catalog of gene families based on gene predictions

of genomes, covering a broad taxonomy of green plants. Omega-3 fatty acid desaturase gene information, including the number of amino acids, chromosome locations, exon count and location coordinates (5′-3′) was obtained from the NCBI database. Physical parameters of the omega-3 fatty acid desaturase proteins, including molecular mass (kDa), and isoelectric point (pI) were calculated using compute pI/Mw tool in ExPASy (http://www.expasy.org/ tools/), with parameters (resolution) set to 'average' (Gasteiger et al., 2005). Several cysteine residues, aliphatic index, instability index, and grand average of hydropathicity were predicted using the ProtParam tool (Gasteiger et al., 2005). The aliphatic index of a protein is defined as the relative volume occupied by the aliphatic side chains of alanine, valine, isoleucine, and leucine residues. Proteins with a calculated instability index smaller than 40 were predicted as stable, while those with a value above 40 were considered as unstable. ProtParam calculates the GAH value as the sum of the amino acids' hydropathy value divided by the sequence's number of residues (Kyte and Doolittle, 1982).

Table 1. List of Omega-3 fatty acid desaturase genes identified in *Gossypium raimondii, Glycine soja, Brassica napus, Brassica juncea, Zea mays, Gossypium arboretum,* and their predicted physicochemical properties.

Accession no.	Species	Length (aa)	MW (Da)	PI	Instability index	Alphatic index
gi|728424424	*Gossypium*	435	50497.15	7.81	43.05(unstable)	87.33
gi|728424414	*Gossypium*	450	51171.86	8.91	36.18(unstable)	85.58
gi|728424404	*Gossypium*	446	50716.97	8.52	35.56(stable)	83.05
gi|728424394	*Gossypium*	376	43588.11	8.68	38.55(stable)	86.06
gi|728424384	*Gossypium*	388	45029.85	9.05	39.84(stable)	79.36
gi|408794	*Glycine soja*	380	44184.89	8.49	38.18(stable)	93.37
gi|734428355	*Glycine soja*	453	51215.88	8.17	33.80(stable)	90.97
gi|7734408411	*Glycine soja*	452	51376.92	7.40	38.56(stable)	87.94
gi|7734373505	*Glycine soja*	453	51247.81	8.45	32.78(stable)	88.01
gi|734344277	*Glycine soja*	453	51301.66	7.39	38.04(stable)	85.85
gi|734407330	*Glycine soja*	381	44147.63	8.72	33.04 (stable)	86.72
gi|734419357	*Glycine soja*	312	36753.14	7.48	36.50 (stable)	89.62
gi|734329362	*Glycine soja*	370	42763.07	8.87	37.41 (stable)	86.41
gi|734427032	*Glycine soja*	195	22236.48	8.93	46.83(unstable)	82.97
gi|728838741	*Gossypium*	311	35841.13	8.33	35.53 (stable)	92.70
gi|728837144	*Gossypium*	330	38132.82	8.79	33.02 (stable)	91.82
gi|195627062	*Zea mays*	443	49385.72	9.16	46.37 (unstable)	85.51
gi|195635609	*Zea mays*	384	44072.42	8.69	38.79(stable)	85.52
gi|195612756	*Zea mays*	408	45905.81	8.87	44.43 (unstable)	90.78
gi|408492	*Brassica napus*	377	43258.58	7.85	29.55 (stable)	90.74
gi|408490	*Brassica napus*	329	38089.77	6.59	25.72 (stable)	92.13
gi|49355342	*Brassica napus*	378	43377.84	8.52	27.62 (stable)	92.04
gi|47028567	*Brassica napus*	383	43850.36	7.86	33.08 (stable)	91.59
gi|46849975	*Brassica napus*	439	50273.62	8.21	34.81 (stable)	86.38
gi|7378667	*Brassica juncea*	429	49174.49	8.42	31.54 (stable)	83.85

Solubility upon overexpression, disulfide bond formation and potential glycosylation sites

Solubility upon overexpression in *Escherichia coli* and disulfide bond formation were predicted using SOLpro (Magnan et al., 2009) and DIpro (Scratch protein predictor), respectively (Cheng et al., 2006). Potential N-glycosylation sites were determined using the NetNGlyc 1.0 (Gupta and Brunak, 2001).

Phylogenetic analysis of members of omega-3 fatty acid desaturase family

Multiple sequence alignments of the full-length protein sequences from cotton, soybean, rapeseed, and corn were performed using MEGA 7.0 (Kumar et al., 2016) with default parameters. A phylogenetic tree based on the alignment was constructed using MEGA 7.0 and the Neighbor-Joining (NJ) method (Saitou and Nei, 1987). The Maximum Parsimony (MP) method (Tamura et al., 2011) was also used to create a phylogenetic tree and to validate the results from the N-J method. Bootstrap analysis was performed using 1000 replicates in pairwise gap deletion mode, which allows divergent domains to contribute to the topology of the NJ tree (Hu et al., 2010).

Conserved motif analysis

Conserved protein motifs were analyzed Online MEME (Multiple Expectation Maximization for Motif Elicitation) (Bailey and Elkan, 1995). The parameters were as follows: The number of repetitions for any, with maximum number of motifs = 20, and the optimum motif width was constrained to between 6 and 200 residues. In addition, structural motif annotation was performed using Pfam (http://pfam.sang er.ac.uk /search) and SMART (http://smart.embl-heidelb erg.de/) tools. Putative conserved domains within the omega-3 fatty acid desaturase were identified using the NCBI conserved domain database (CDD) (Marchler-Bauer et al., 2005; Wang et al., 2023).

Results

Physical properties and chromosomal distribution of omega-3 fatty acid desaturase gene family in five species genomes

Twenty-five genes were identified as members of the omega-3 fatty acid desaturase gene family, seven genes in cotton, nine genes in soybean, three genes in corn and six genes in rapeseed. These findings determined the physical location of individual omega-3 fatty acid desaturase genes on chromosomes. The chromosome location of the omega-3 fatty acid desaturase genes indicated that this gene in *Gossypium raimondii* were distributed on five out of 26 chromosomes (chr 1, 2, 6, 11 and 12). The chromosomal locations of some genes were unknown. Among the 25 genes, gi|734428355, gi|7734373505 and gi|734344277 encodes the longest protein (453 amino acids (aa)), while gi|734427032 encoded the shortest (195 aa) belongs to *Glycine soja*. The average length of the proteins encoded by omega-3 fatty acid desaturase proteins was 388.58 amino acid. The molecular weights of these proteins ranged from 22.2 kDa to 51.3 kDa, with an average of 44.7 kDa. Furthermore, the theoretical pI values of all proteins except gi|408490 were above 7, indicating that they were alkaline. The 25 omega-3 fatty acid desaturase genes contained different numbers of exons, ranging from 5 to 9. Furthermore, eight omega-3 fatty acid desaturase genes were found to possess eight exons. Detailed parameters are listed in Table 1. According to Protparam as shown in Table 1, the most unstable Omega-3 fatty acid desaturase gene is probably obtained from gi|734427032 and the most stable form is gi|408490. Stable enzymes can be used for a longer time to biocatalize a reaction. The most alphatic index omega-3 fatty acid desaturase gene is obtained from gi|408794. It may be regarded as a positive factor for the increase of thermostability of globular proteins. The chloroplast transit peptide (CTP) of these proteins ranged from 0.002 to 0.904, with an average of 0.342 (Table 2). CTPs can be used to transport non-chloroplastic proteins into the chloroplasts.

Solubility upon overexpression, disulfide bond formation and potential glycosylation sites

The omega-3 fatty acid desaturase protein from gi|734419357 and gi|728424394 were predicted to have the highest probabilities of solubility and insolubility, respectively (Table 3). Of 25 FAD, only eight sequences (three *G. raimondii,* two G. *soja,* one in each of *G. arboretum, Z. mays* and *B. juncea*) were predicted to be soluble while the other 17 FAD were predicted to be insoluble upon overexpression in E. coli. Table 3 shows the prediction by DIpro of

disulfide bond formation for different FAD. Most of FAD were predicted to have the lowest number (only one) of disulfide bonds (Table 3). Enzymes from G. *soja* (gi|734344277 and gi|7734408411) contained the highest amount of cysteine residues, 7 and 6, respectively. This means that according to our prediction, recombinant production of these FAD might be performed using expression systems that provide a suitable redox medium for disulfide bond formation. FAD from *Z. mays* (gi|195635609), has no potential Nglycosylation site. The highest numbers of N-glycosylation site were found for FAD from gi|46849975 and gi|7378667 (9 potential sites), and gi|728424384 (8 potential sites).

Table 2. Predicts the subcellular location of Omega-3 fatty acid desaturase proteins sequences.

Accession no.	Len	cTP	mTP	SP	other	Loc	RC
gi|728424424	435	0.193	0.370	0.043	0.233	M	5
gi|728424414	450	0.844	0.052	0.016	0.056	C	2
gi|728424404	446	0.831	0.037	0.011	0.159	C	2
gi|728424394	376	0.154	0.134	0.020	0.878	-	2
gi|728424384	388	0.021	0.091	0.155	0.976	-	1
gi|408794	380	0.051	0.194	0.017	0.846	-	2
gi|734428355	453	0.734	0.085	0.030	0.097	C	2
gi|7734408411	452	0.788	0.181	0.014	0.091	C	2
gi|7734373505	453	0.705	0.110	0.027	0.087	C	3
gi|734344277	453	0.923	0.188	0.008	0.047	C	1
gi|734407330	381	0.037	0.163	0.032	0.936	-	2
gi|734419357	312	0.004	0.066	0.938	0.057	S	1
gi|734329362	370	0.096	0.112	0.022	0.894	-	2
gi|734427032	195	0.098	0.338	0.228	0.278	M	5
gi|728838741	311	0.002	0.133	0.968	0.091	S	1
gi|728837144	330	0.002	0.133	0.968	0.091	S	1
gi|195627062	443	0.566	0.765	0.002	0.007	M	5
gi|195635609	384	0.131	0.137	0.029	0.871	-	2
gi|195612756	408	0.533	0.426	0.237	0.030	C	5
gi|408492	377	0.078	0.171	0.018	0.856	-	2
gi|408490	329	0.012	0.180	0.935	0.042	S	2
gi|49355342	378	0.078	0.171	0.018	0.856	-	2
gi|47028567	383	0.049	0.161	0.024	0.936	-	2
gi|46849975	439	0.904	0.162	0.001	0.021	C	2
gi|7378667	429	0.718	0.260	0.005	0.097	C	3

The location assignment is based on the predicted presence of any of the N-terminal presequences: Sequence length (**Len**), chloroplast transit peptide (**cTP**), mitochondrial targeting peptide (**mTP**) or secretory pathway signal peptide (**SP**), Prediction of localization (**Loc**), Reliability class (**RC**).

Table 3. Prediction of solubility upon overexpression, disulfide bond formation and N-glycosylation of linoleic acid isomerases (LAI).

Accession no.	Length (aa)	Solubility on Overexpression [a]	Cys	Disulfide bonds [b]	N glycosylation [c]
gi|728424424	435	Insoluble (0.712733)	4	1 (42-103)	6 (29, 71, 91, 171, 222, 257)
gi|728424414	450	Soluble (0.507310)	5	2 (9-42,115-171)	3 (30, 64, 183)
gi|728424404	446	Soluble (0.500000)	3	1 (113-169)	6 (37, 47, 57, 81, 254, 258)
gi|728424394	376	Insoluble (0.720550)	4	1 (67-94)	5 (106, 137, 150, 190, 194)
gi|728424384	388	Soluble (0.709559)	2	1 (39-155)	8 (43, 117, 122, 152, 168, 205, 209, 355)
gi|408794	380	Insoluble (0.597998)	4	1 (85,102)	3 (14, 162, 198)
gi|734428355	453	Insoluble (0.600799)	5	2 (9-52,118-174)	4 (39, 53, 106, 186)
gi|773440841	452	Insoluble (0.572344)	6	2 (9-27,117-173)	5 (39, 46, 56, 147, 236)
gi|773437350	453	Insoluble (0.616692)	5	2 (9-52,117-173)	6 (39, 53, 83, 105, 185, 219)
gi|734344277	453	Soluble (0.569171)	7	2 (9-27,118-174)	5 (39, 46, 84, 88, 237)
gi|734407330	381	Insoluble (0.698000)	3	1 (44-100)	7 (12, 22, 74, 160, 196, 200, 381)
gi|734419357	312	Soluble (0.713550)	2	1 (34-150)	2 (94, 130)
gi|734329362	370	Insoluble (0.663430)	3	1 (33-89)	5 (2, 149, 185, 189, 370)
gi|734427032	195	Insoluble (0.535586)	4	1 (186-193)	5 (14, 24, 72, 142, 158)
gi|728838741	311	Soluble (0.597674)	3	1 (21-48)	6 (60, 88, 91, 104, 144, 148)
gi|728837144	330	Insoluble (0.524630)	3	1 (21-48)	6 (60, 88, 91, 104, 144, 148)
gi|195627062	443	Insoluble (0.650700)	4	2 (9-10,107-163)	2 (175, 259)
gi|195635609	384	Soluble (0.572108)	5	2 (38-94,210-212)	0
gi|195612756	408	Insoluble (0.762427)	5	2 (35-73,129-245)	4 (141, 157, 214, 357)
gi|408492	377	Insoluble (0.617692)	3	1 (38-94)	3 (12, 190, 377)
gi|408490	329	Insoluble (0.612866)	3	1 (164-174)	4 (22, 23, 60, 156)
gi|49355342	378	Insoluble (0.645880)	3	1 (38-94)	3 (12, 191, 378)
gi|47028567	383	Insoluble (0.698698)	3	1 (44-100)	5 (10, 12, 160, 196, 383)
gi|46849975	439	Insoluble (0.614346)	4	1 (102-158)	9 (3, 29, 46, 50, 57, 90, 132, 170, 266)
gi|7378667	429	Soluble (0.514721)	5	2 (95-151,267-277)	9 (3, 28, 42, 52, 83, 125, 163, 247, 259)

a) Probability values are presented within parenthesis.
b) The position of disulfide bonds is written within the parenthesis.
c) The position of N -glycosylation sites is written within the parenthesis.

Phylogenetic analysis of omega-3 fatty acid desaturase family in cotton, soybean, rapeseed and corn

To estimate the similarity and evolutionary ancestry of omega-3 fatty acid desaturase genes in cotton, soybean, rapeseed and corn, we created an unrooted phylogenetic tree of the 25 omega-3 fatty acid desaturase protein sequences. The phylogenetic tree divided the omega-3 fatty acid desaturase proteins into five clades (Figure 1). Clade 2 has the fewest omega-3 fatty acid desaturase gene members (2), in contrast clade 1 contains the most members (12), followed by clades 3 and 4 (3) and clade 5 (5). The four species (except *G. arboreum*) contributed at least one omega-3 fatty acid desaturase gene to clade 1, while the members of clades 2, 3, 4 and clade 5 included one two or three species, for example, clade 5 consisted of *G. raimondii, Z. mays and B. napus*, which may correspond to some special in the evolutionary process.

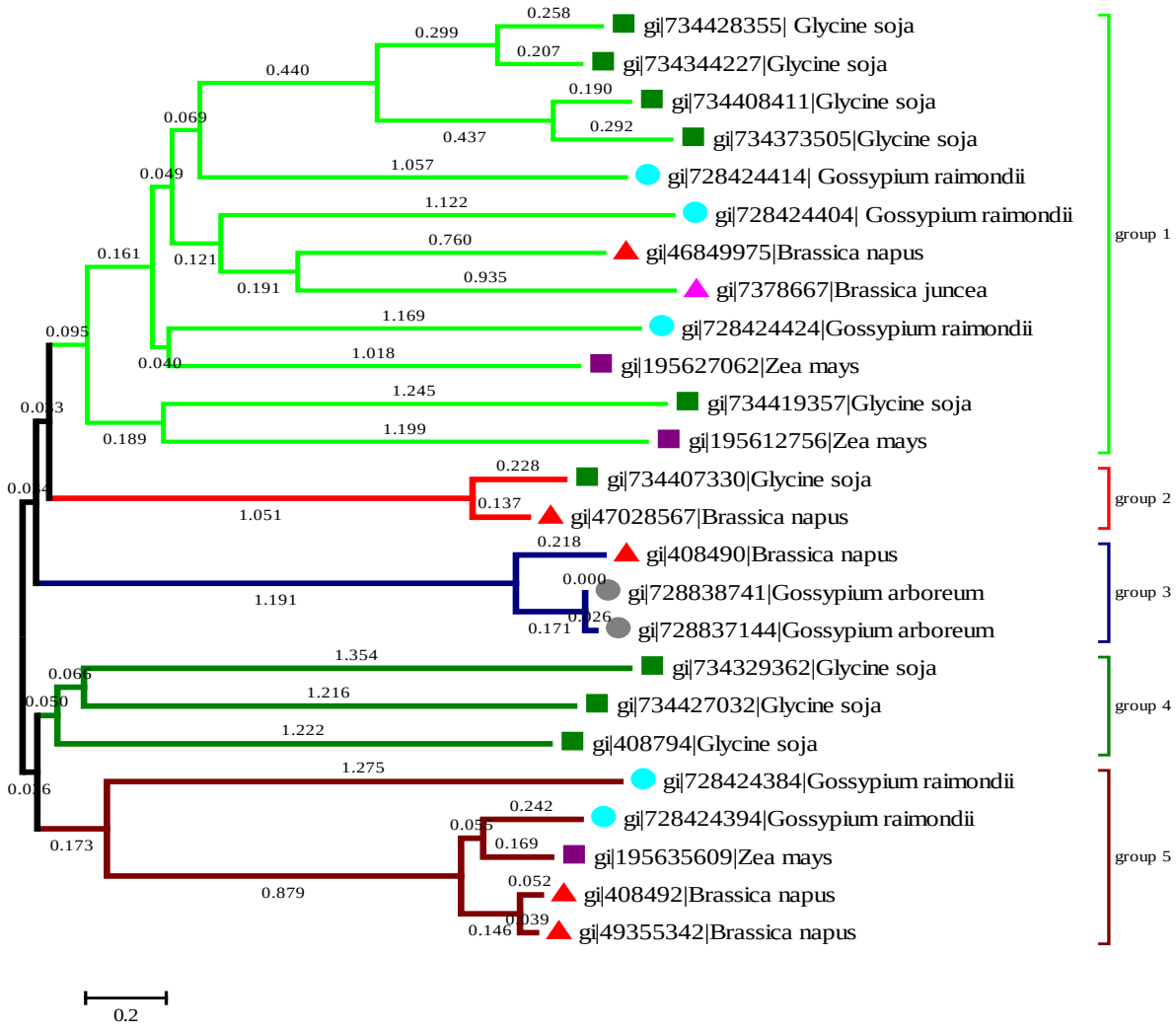


Figure 1. Phylogenetic tree of full-length Omega-3 fatty acid desaturase proteins from *G.soja* ( ), *B. napus* ( ), *B. juncea* ( ), *Z. mays* ( ), *G. arboretum* ( ) and *G. raimondii* ( ).

Based on phylogenetic analysis, five pairs of orthologous genes were identified among the omega-3 fatty acid desaturase genes: gi/195635609 and gi/728424394, gi/47028567 and gi/734407330, gi/195612756 and gi/734419357, gi/195627062 and gi/728424424, and gi/7378667 and gi/46849975. The remaining genes of the omega-3 fatty acid desaturase gene family were represented by paralogous pairs.

Conserved motifs of omega-3 fatty acid desaturase genes

Twenty distinct conserved motifs were identified, and the conserved amino acid sequences and length of each motif are shown in Figure 2. Each putative motif obtained from MEME was annotated by searching Pfam and SMART. While four motifs were found to encode the fatty acid desaturases domain, four different domains were found related to the DUF3474 domain, while the other motifs did not function in the annotation. Fatty acid desaturases are enzymes that catalyze the insertion of a double bond at the delta position of fatty acids. Additionally, the DUF3474 domain remains functionally uncharacterized (Figure 3). This domain is presented in both bacteria and eukaryotes. This domain is typically 126 - 140 amino

acids in length. Motifs 3 and 4 were the most common motifs, found in all twenty-five omega-3 fatty acid desaturase genes. Motif 15 was unique to the proteins in clade 4 and unique motifs 18 were found in clade 3; These motifs might be important to the functions of unique omega-3 fatty acid desaturase protein.

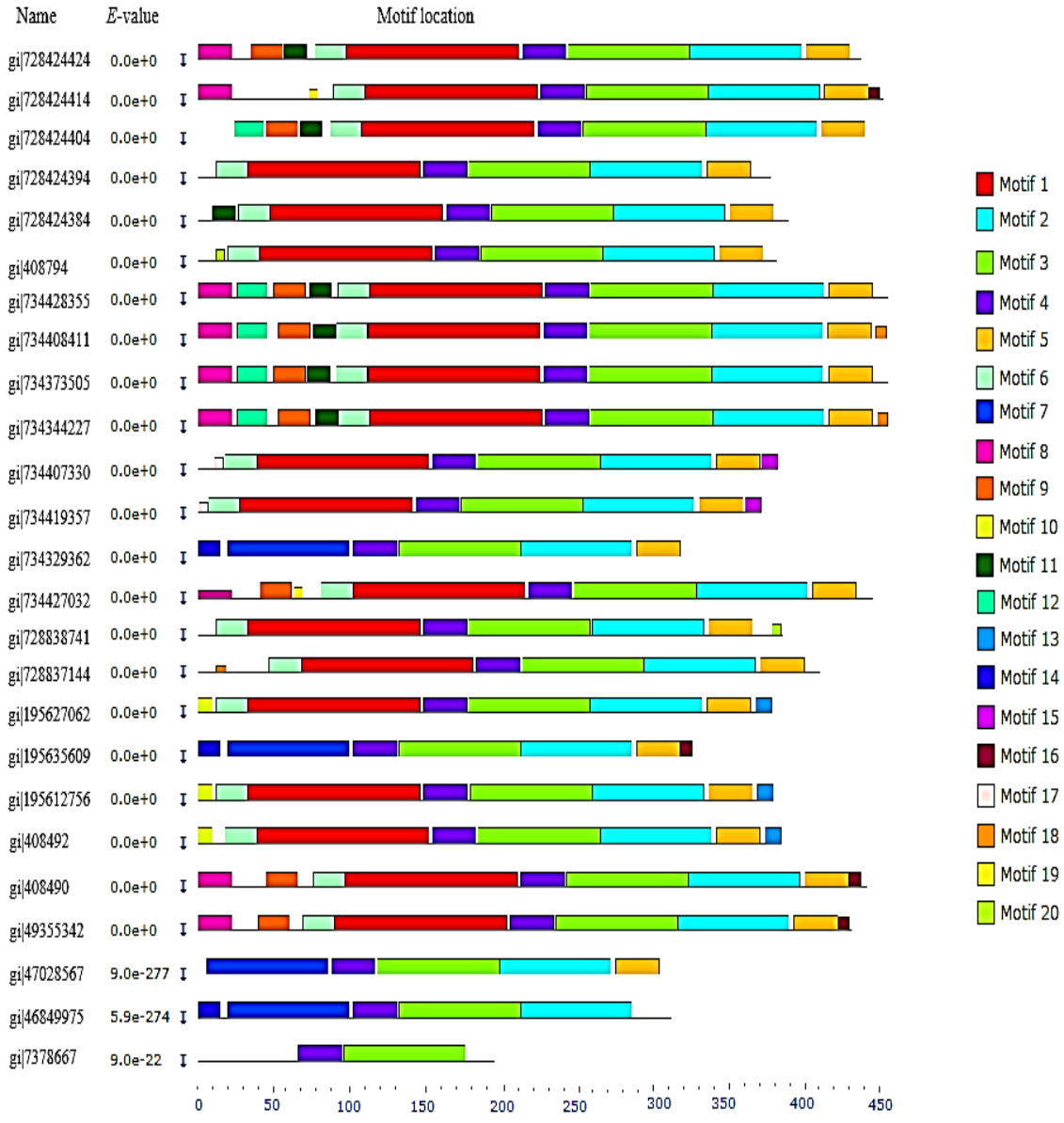


Figure 2. Distribution of conserved motifs in the omega-3 fatty acid desaturase family members. All motifs were identified by MEME using the complete amino acid sequences. Names of all members among the defined genes are shown on the left side of the figure, and motif lengths are indicated at the bottom of the figure. The positions of Zn-finger domains predicted by the SMART tool. Database are indicated by vertical tick marks below each protein model. The different-colored boxes represent different motifs and their position in each omega-3 fatty acid desaturase sequence.

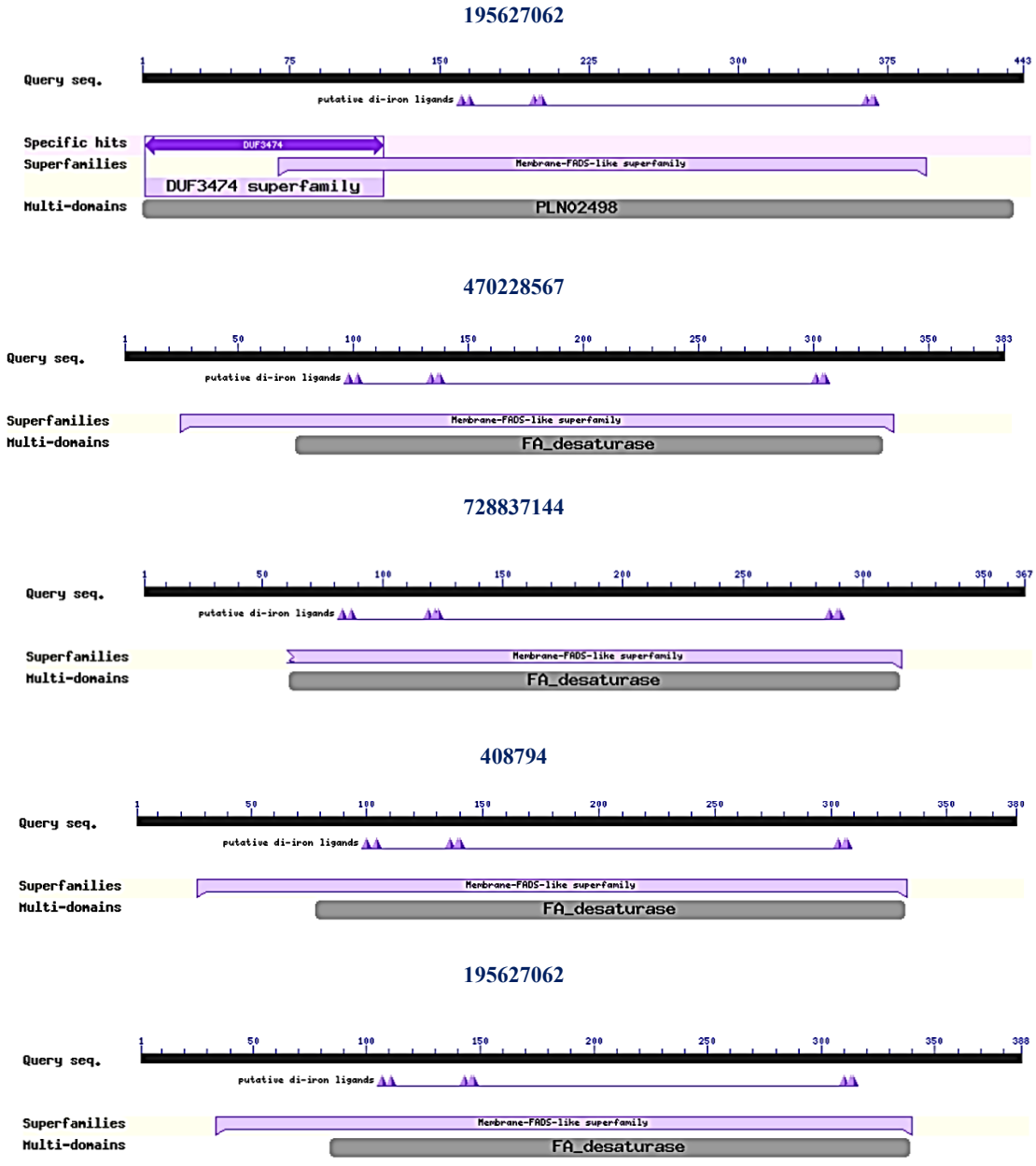


Figure 3. Conserved domain search on Omega-3 fatty acid desaturase gene sequences.

1

Discussion

Omega-3 fatty acid desaturase gene family is involved in the regulation of a variety of processes. In the present study, the complex features and functions of this group of proteins have been studied in the plant cotton, in soybean, in rapeseed and in corn. FAD genes encode enzymes that catalyze the desaturation of fatty acids, which affect the nutritional value and oxidative stability of seed storage oils (Clemente and Cahoon, 2009). Although the distribution of these omega-3 fatty acid desaturase genes is diverse, their genetic features and physicochemical properties tend to be identified. There are anatomical and physiological differences between the four species, which might be reflected in the diversity of omega-3 fatty acid

desaturase genes structure and conserved motifs. We identified that the 25 omega-3 fatty acid desaturase genes contain different numbers of amino acids and exons, indicating that there is some diversity in these four species. Protein length is shown to correlate with the probability that the protein is encoded by an essential gene. Increased protein length appears to be a notable mechanism by which the increasing complexity of protein–protein interaction networks is accommodated within the natural evolution of species (Tan et al., 2005). gi|728424384 and gi|408794 had the lowest and highest aliphatic indices among the protein sequences, respectively. This may be considered as positive factor for the increase in thermo stability of proteins (Ikai, 1980). MEME server identified the different conserved protein motifs that are present in each of the FAD proteins. Conserved motifs could be important for the diverse functions of omega-3 fatty acid desaturas proteins from cotton, soybean, rapeseed, and corn (Lynch and Conery, 2000). Some closely related members shared similar structures, implying functional similarities for these FAD proteins. The specific sequence motifs present in each clade may impart specific functions to the FAD proteins. The similarities in gene structure and motif composition of most FAD proteins consistented with phylogenetic analysis of the omega-3 fatty acid desaturase gene family. The differences in these characteristics among the different clades suggested that the FAD members were functionally diversified. Certainly, the characterization of FADs would provide candidate genes to improve new oilseed crops.

To explore how the omega-3 fatty acid desaturase gene family evolved, we performed a genome-wide comparison of plant FAD members from four species of various oilseed crops. Considerable phylogenetic analysis of FAD proteins has been conducted in cotton, soybean, rapeseed and corn, respectively. To obtain an overall picture of the 25 FAD proteins and their relationships with each other, a phylogenetic tree of FAD proteins was constructed, which divided the 25 FAD members into five clades. The plant FAD members from soybean appear to be more closely related to each other than to FAD genes of the other oilseed crops. Yurchenko et al in a genome wide analysis of the omega 3 tree fatty acid desaturase gene family in Gossypium indicated that species separated into three well-defined monophyletic groups in the phylogenetic analysis (Yurchenko et al., 2014). The result of connecting the thiol groups of two cysteine amino acids is formation of disulfides bond on the polypeptide chains. These bonds are responsible for stabilizing the globular structure and correct conformation of the protein (Bulleid and Ellgaard, 2011). In addition, extracellular proteins often have several disulfide bonds, whereas intracellular proteins usually lack them. Disulfide bonds' prediction can be beneficial in selecting recombinant expression system (Bulleid and Ellgaard, 2011). Glycosylation is one of the most current and structurally diverse forms of post-translational alteration of proteins. The bacterial host lacks post-translational modification systems such as glycosylation, but species from plants, fungi and animals have the ability to glycosylate proteins. Glycosylation may increase the stability of the protein, and in some cases, decrease the biological activity because of masking of the active site (Easton and Leader, 2011).

Conclusion

In the current study, 25 members of the FAD3 were analyzed, a comprehensive analysis including their chromosomal location, physical parameters, aliphatic index, instability index phylogeny, and conversed motifs. Phylogenetic analysis revealed that the genes could be grouped into five major clades. In each clade, the characteristics of exon/intron structure and motif compositions were relatively conserved. Five pairs of orthologous genes among the omega-3 fatty acid desaturase genes and a total of twenty distinct conserved motifs were identified. Four of these motifs were associated with encoding the fatty acid desaturase domain, while an additional four were related to the DUF3474 domain. The remaining motifs did not exhibit any functional annotation. Certainly, the identification of these FADs can help improve new varieties of oilseeds by introducing valuable candidate genes.

Supplementary Materials

There is no supplementary material available for this article.

Author contributions

Conceptualization, A.A.G.; software and formal analysis, A.A.G.; investigation, A.A.G and S.J.H.; data curation, A.A.G.; writing—original draft preparation, A.A.G. and S.J.H.; writing—review and editing, A.G.; All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Acknowledgments

The authors would like to express their gratitude to the Dr. Ebrahim Sharafi and Dr. Jafar Ahmadi for their valuable technical assistance.

Conflict of interest statement

The authors declare no conflict of interest.

References

Alloatti, A., and Uttaro, A.D. (2011). Highly specific methyl-end fatty-acid desaturases of trypanosomatids. *Mol Biochem Parasitol* 175(2): 126-132.

Bailey, T.L., and Elkan, C. (Year). "The value of prior knowledge in discovering motifs with MEME", in: *Proceedings of the Third International Conference on Intelligent Systems for Molecular Biology July, 1995 (Ismb)*), 21-29.

Bocianowski, J., Mikołajczyk, K., and Bartkowiak-Broda, I. (2012). Determination of fatty acid composition in seed oil of rapeseed (*Brassica napus* L.) by mutated alleles of the *FAD3* desaturase genes. *J Appl Genet* 53: 27-30.

Bulleid, N.J., and Ellgaard, L. (2011). Multiple ways to make disulfides. *Trends Biochem Sci* 36(9): 485-492.

Cheng, J., Saigo, H., and Baldi, P. (2006). Large‐scale prediction of disulphide bridges using kernel methods, two‐dimensional recursive neural networks, and weighted graph matching. *Proteins: Struct* 62(3): 617-629.

Chi, X., Yang, Q., Lu, Y., Wang, J., Zhang, Q., Pan, L., Chen, M., He, Y., and Yu, S. (2011). Genome-wide analysis of fatty acid desaturases in soybean (*Glycine max*). *Plant Mol Biol Rep* 29: 769-783.

Clemente, T.E., and Cahoon, E.B. (2009). Soybean oil: genetic approaches for modification of functionality and total content. *Plant Physiol* 151(3): 1030-1040.

Domínguez, T., Hernández, M.L., Pennycooke, J.C., Jiménez, P., Martínez-Rivas, J.M., Sanz, C., Stockinger, E.J., Sánchez-Serrano, J.J., and Sanmartín, M. (2010). Increasing *ω-3* desaturase expression in tomato results in altered aroma profile and enhanced resistance to cold stress. *Plant Physiol* 153(2): 655-665.

Easton, R., and Leader, T. (2011). Glycosylation of proteins–structure, function and analysis. *Life Sci Tech Bull* 60: 1-5.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S.e., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005). *Protein identification and analysis tools on the ExPASy server.* Humana press. Springer.

Gupta, R., and Brunak, S. (2001). "Prediction of glycosylation across the human proteome and the correlation to protein function," in *PubMed. Pacific Symp. Biocomput.*), 310-322.

Guschina, I.A., and Harwood, J.L. (2006). Mechanisms of temperature adaptation in poikilotherms. *FEBS Lett* 580(23): 5477-5483.

Hernández, M.L., Sicardo, M.D., and Martínez-Rivas, J.M. (2016). Differential contribution of endoplasmic reticulum and chloroplast *ω-3* fatty acid desaturase genes to the linolenic acid content of olive (*Olea europaea*) fruit. *Plant Cell Physiol* 57(1): 138-151.

Hu, R., Qi, G., Kong, Y., Kong, D., Gao, Q., and Zhou, G. (2010). Comprehensive analysis of *NAC* domain transcription factor gene family in *Populus trichocarpa*. *BMC Plant Biol* 10(1): 145.

Ikai, A. (1980). Thermostability and aliphatic index of globular proteins. *J Biol Chem* 88(6): 1895-1898.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33(7): 1870-1874.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157(1): 105-132.

Liu, W., Li, W., He, Q., Daud, M.K., Chen, J., and Zhu, S. (2015). Characterization of 19 genes encoding membrane-bound fatty acid desaturases and their expression profiles in *Gossypium raimondii* under low temperature. *PloS One* 10(4): e0123281.

Lynch, M., and Conery, J.S. (2000). The evolutionary fate and consequences of duplicate genes. *Sci* 290(5494): 1151-1155.

Magnan, C.N., Randall, A., and Baldi, P. (2009). SOLpro: accurate sequence-based prediction of protein solubility. *Bioinformatics* 25(17): 2200-2207.

Marchler-Bauer, A., Anderson, J.B., Cherukuri, P.F., DeWeese-Scott, C., Geer, L.Y., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., and Ke, Z. (2005). CDD: a Conserved Domain Database for protein classification. *Nucleic Acids Res* 33(suppl_1): D192-D196.

Martin, B.A., Schoper, J.B., and Rinne, R.W. (1986). Changes in soybean (*Glycine max* [L.] Merr.) glycerolipids in response to water stress. *Plant Physiol* 81(3): 798-801.

Meesapyodsuk, D., and Qiu, X. (2012). The front-end desaturase: structure, function, evolution and biotechnological use. *Lipids* 47: 227-237.

Ohlrogge, J., and Browse, J. (1995). Lipid biosynthesis. *Plant Cell* 7(7): 957.

Román, Á., Andreu, V., Hernández, M.L., Lagunas, B., Picorel, R., Martínez-Rivas, J.M., and Alfonso, M. (2012). Contribution of the different omega-3 fatty acid desaturase genes to the cold response in soybean. *J Exp Bot* 63(13): 4973-4982.

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

Shanklin, J., and Cahoon, E.B. (1998). Desaturation and related modifications of fatty acids. *Annu Rev Plant Biol* 49(1): 611-641.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28(10): 2731-2739.

Tan, T., Frenkel, D., Gupta, V., and Deem, M.W. (2005). Length, protein–protein interactions, and complexity. *Physica A* 350(1): 52-62.

Wang, J., Chitsaz, F., Derbyshire, M.K., Gonzales, N.R., Gwadz, M., Lu, S., Marchler, G.H., Song, J.S., Thanki, N., and Yamashita, R.A. (2023). The conserved domain database in 2023. *Nucleic Acids Res* 51(D1): D384-D388.

Williams, J.P., Khan, M.U., Mitchell, K., and Johnson, G. (1988). The effect of temperature on the level and biosynthesis of unsaturated fatty acids in diacylglycerols of *Brassica napus* leaves. *Plant Physiol* 87(4): 904-910.

Yang, Q., Fan, C., Guo, Z., Qin, J., Wu, J., Li, Q., Fu, T., and Zhou, Y. (2012). Identification of *FAD2* and *FAD3* genes in *Brassica napus* genome and development of allele-specific markers for high oleic and low linolenic acid contents. *Theor Appl Genet* 125: 715-729.

Yurchenko, O.P., Park, S., Ilut, D.C., Inmon, J.J., Millhollon, J.C., Liechty, Z., Page, J.T., Jenks, M.A., Chapman, K.D., and Udall, J.A. (2014). Genome-wide analysis of the omega-3 fatty acid desaturase gene family in *Gossypium*. *BMC Plant Biol* 14(1): 1-15.

Zhang, M., Barg, R., Yin, M., Gueta‐Dahan, Y., Leikin‐Frenkel, A., Salts, Y., Shabtai, S., and Ben‐Hayyim, G. (2005). Modulated fatty acid desaturation via overexpression of two distinct *ω‐3* desaturases differentially alters tolerance to various abiotic stresses in transgenic tobacco cells and plants. *Plant J* 44(3): 361-371.

JPMB
Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

مقاله پژوهشی

مقایسه توالی‌های پروتئینی پیش‌بینی شده از خانواده ژنی اسیدهای چرب غیراشباع امگا۳ در برخی از گیاهان دانه روغنی

عمار افخمی قادی*[۱]، سید جابر حسینی[۲]، علی قنبری[۲]

[۱] گروه مهندسی ژنتیک و اصلاح نباتات، دانشکده کشاورزی، دانشگاه امام خمینی (ره)، قزوین، ایران
[۲] گروه زراعت، دانشگاه تربیت مدرس، تهران، ایران
[۳] گروه ژنتیک و مهندسی تولیدات گیاهی، دانشگاه زنجان، زنجان، ایران

چکیده: هدف از این مطالعه، تجزیه مقایسه‌ای بین توالی‌های پروتئینی پیش‌بینی‌شده در خانواده ژنی اسید چرب غیراشباع امگا۳ (FAD) در گیاهان مختلف دانه روغنی، همچون پنبه، سویا، کلزا و ذرت بود. بیست و پنج ژن اسید چرب غیراشباع امگا۳ با حذف توالی‌های مشابه، متمایز شدند. با تجزیه و تحلیل فیلوژنتیک، خانواده ژنی اسید چرب غیراشباع امگا۳ شناسایی شد. اطلاعات ژنی اسید چرب غیراشباع امگا۳ با استخراج ویژگی‌های فیزیکی شامل تعداد اسیدهای آمینه، مکان کروموزومی، تعداد اگزون و غیره به دست آمد. نتایج نشان داد میانگین طول توالی‌های پروتئینی کدگذاری شده اسید چرب غیراشباع امگا۳، ۳۸۸/۵۸ اسید آمینه بود. وزن مولکولی پروتئین‌های بررسی‌شده از ۲۲/۲ تا ۵۱/۳ کیلو دالتون متغیر بود. درخت فیلوژنتیکی، پروتئین‌های اسید چرب امگا۳ را در پنج گروه قرار داد. گروه دوم، کمترین تعداد اعضای ژن اسید چرب غیراشباع امگا۳ را شامل شده، در حالی که گروه اول بیشترین تعداد را در بر می‌گیرد. بنابراین پنج جفت ژن ارتولوگ در میان ژن‌های اسید چرب غیراشباع امگا۳ و در مجموع بیست موتیف حفاظت‌شده مجزا شناسایی شد. چهار مورد از این موتیف‌ها با رمزگذاری دمین اسید چرب غیراشباع مرتبط بوده، در حالی که چهار مورد دیگر مربوط به دمین DUF3474 بودند. بدون شک، مشخص کردن خصوصیات این اسیدهای چرب غیراشباع، می‌تواند ژن‌های گزینشی ارزشمندی را برای بهبود واریته‌های جدید دانه روغنی آشکار سازد.

کلمات کلیدی: اسید چرب امگا۳، خانواده ژنی، گیاهان دانه روغنی، موتیف‌های حفاظت‌شده.

ویراستار علمی
دکتر پرویز حیدری،
دانشگاه صنعتی شاهرود، ایران

تاریخ
دریافت: ۱۰ مرداد ۱۴۰۲
پذیرش: ۲۷ مهر ۱۴۰۲
چاپ: ۸ بهمن ۱۴۰۲

نویسنده مسئول
دکتر عمار افخمی قادی
a.afkhami@sanru.ac.ir

ارجاع به این مقاله
Afkhami Ghadi, A., Hosseini, S. J., and Ghanbari, A. (2023). Comparison of protein predicted sequences of the omega-3 fatty acid desaturase gene family in some of the oil seed crops. *J. Plant Mol. Breed 11* (1): 28-40. doi: 10.22058/JPMB.2023.2008246.1281.

JPMB | Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

Alterations in antioxidant enzyme activities in rice plants treated with various abiotic inducers against the bacterial blight agent *Xanthomonas oryzae* pv. *oryzae*

Original paper

Leila Esfahani [1], Valiollah Babaeizad [*,1], Heshmatollah Rahimian and [1], Ali Dehestani [2,]

[1.] Sari Agricultural Sciences and Natural Resources University, Sari, Iran
[2.] Genetics and Agricultural Biotechnology Institute of Tabarestan, Sari Agricultural Sciences and Natural Resources University, Sari, Iran.

Abstract: Rice is the most important staple food in the world. Bacterial leaf blight of rice, caused by *Xanthamonos oryzae* pv. *oryzae* (Xoo), is a highly destructive and widespread disease. Chemical management approach to control this disease appears ineffective. In this experiment, the effects of three treatments of salicylic acid, potassium phosphite, and chitosan on susceptible rice plants inoculated with the bacteria were investigated to assess the induction of resistance and activity of antioxidant enzymes including catalase (CAT), guaiacol peroxidase (GPX) and superoxide dismutase (SOD) during four days. The results showed that the highest and lowest activity of CAT was recorded in the chitosan and salicylic acid treatment, respectively. The maximum amount of catalase activity was 72 hours after inoculation. Comparison of GPX and SOD enzyme activities at different sampling times revealed that these enzymes reached their highest level at 48 and 72 hours after inoculation across all treatments, respectively. However, among different treatments, the highest activity of these enzymes was observed in plants infected with bacteria under potassium phosphite treatment. The findings show that potassium phosphite increases the activity of plant defense enzymes against the pathogen, ultimately reducing the symptoms of the disease.

Keywords: chitosan, induced resistance, *Oryza sativa*, potassium phosphite, salicylic acid.

OPEN ACCESS

Edited by
Dr. Shideh Mojerloum,
Aarhus University, Denmark
Shahrood university of technology, Iran

Date
Received: 5 September 2023
Accepted: 18 January 2024
Published: 1 February 2024

Correspondence
Dr. Valiollah Babaeizad
babaeizad@yahoo.com

Citation
Esfahani, L., Babaeizad, V., Rahimian, H. and Dehestani A. (2023). Investigation of changes in antioxidant enzyme activities in rice plants treated with various abiotic inducers against the bacterial blight agent *Xanthomonas oryzae* pv. *oryzae* . *J Plant Mol Breed* 11 (1): 41-53. 10.22058/JPMB.2024.2010427.1284 .

Introduction

Rice (*Oryza sativa* L.) is one of the most important human food sources, which is widely cultivated all over the world (Ainsworth, 2008). Since the introduction of high-yielding semi-dwarf varieties, diseases and insect pests have increasingly caused severe yield losses year after year (Siddiq and Vemireddy, 2021). Bacteria, viruses, and fungi can cause rice to be susceptible to many diseases (Dai et al., 2007). Bacterial blight of rice *Xanthamonos oryzae* pv. *oryzae* (*Xoo*) is one of the most destructive bacterial diseases of rice in some rice-growing areas in the world, especially tropical Asian areas. This disease can cause as high as 50-70% reduction in rice yield in severe epidemics and it is the second most important rice disease in the world, after Rice Blast (Mew et al., 1992). In response to the invasion of microorganisms, plants use different defense mechanisms to deal with the pathogen. The mechanism of enzymatic antioxidant activities includes the regulation of enzymes like Superoxide dismutase, Catalase, Peroxidase, Glutathione reductase, Glutathione S threonate and Guaiacol peroxidase (Hussain et al., 2016).

Catalase is an important antioxidant inhibitory enzyme which plays a role in the removal of H_2O_2. Catalase plays the role of a specific peroxidase (POX) that protects cells against the toxic effects of H_2O_2. Catalase has a great affinity for H_2O_2, which in turn enables it to form H_2O and O_2 (Mittler, 2002; Mastouri et al., 2012). Under stress conditions, guaiacol peroxidase (GPX) reduces the level of H_2O_2 in cells (Asada, 1999; Basu et al., 2010). The plant has an immunity system that can be induced by abiotic or abiotic inducers (Buensanteai et al., 2009). Previous studies have shown that priming plants with chemical inducers, such as SA (Ganesan and Thomas, 2001; Li and Zhang, 2012; Saba Anwar et al., 2013), potassium phosphite (Huang et al., 2020) and chitosan (Orzali et al., 2014), is capable of inducing resistance. In rice, the application of SA can induce a SA-dependent signaling pathway that leads to disease resistance. The SA-dependent pathway is associated with systemic acquired resistance (SAR) activated by rice plant pathogens (Sticher et al., 1997). One of the important features of induced resistance is the priming phenomenon, in which rice plants show faster and higher defense responses against bacterial contamination compared to untreated rice plants (Conrath et al., 2002). Chitosan is a linear polysaccharide that may be obtained by deacetylation of chitin, long-chain N-Acetyl- Gluconicamine polymers which can easily be extracted from fungal cell walls and crustacean shells (Badawy and Rabea, 2011). This natural compound plays a role in controlling plant diseases. The mode of action includes a direct antimicrobial activity and an indirect induction of resistance that induces several defense responses in plants (Falcón-Rodríguez et al., 2012). Various studies showed that chitosan prevented the growth of a number of pathogens, including *Xanthomonas* sp. (Li et al., 2008), *Pseudomonas syringae* (Mansilla et al., 2013), *Agrobacterium tumefaciens* and *Erwinia carotovora* (Badawy et al., 2014).

Potassium phosphite is used against soil and airborne fungal and bacterial pathogens. This compound moves in a symplastic way in the plant and has the properties of prevention, immunization, and treatment (Thao and Yamakawa, 2009). Phosphites can have a direct or indirect effect on the pathogen (Deliopoulos et al., 2010). The direct effect includes preventing the growth of fungi and reducing or changing the metabolism of pathogens, and the indirect effect includes stimulating the plant's defense system, such as increasing the production of phytoalexins and reactive oxygen species (ROS), inducing Pathogenesis-Related Proteins (PRs), and strengthening the cell wall (Lobato et al., 2008).

There is no effective chemical control of bacterial blight in rice, and cultivars with good resistance to this disease agent have not been offered (Derakhshan et al., 2020). One of the low cost, healthy ways to fight a number of diseases in plants is through induction of resistance genes using different living and non-living factors. In this study, the possibility of inducing resistance to bacterial leaf blight of rice by using non-living stimuli (salicylic acid, chitosan and potassium phosphite) was evaluated.

Materials and Methods

Building upon the findings of (Derakhshan et al., 2020), who evaluated the resistance of 24 commercial cultivars of Iranian rice against blight bacteria, the local Tarom cultivar was identified as a

sensitive variety. Consequently, this cultivar was selected for the intended research.

In this study, the standard isolate of *Xanthomonas oryzae* pv. *oryzae* (442) was selected, cultured and propagated in Nutrient Agar Sucrose (NAS) at 29 °C (Derakhshan et al., 2020). To perform the pathogenicity test, a fresh culture of bacteria was used and a suspension with different concentration (OD=0.1, 0.3, 0.5) was prepared. Pathogenicity was confirmed by different methods (suspension injection, sterile scissors, and spray) applied to the local Tarom cultivar (Backer, 2002). In this research, salicylic acid (Merck company) with a concentration of 3 mM, chitosan with a concentration of 400 ppm and potassium phosphite with a concentration of 4 g/liter were used (Katoch et al., 2005; Valadi et al., 2013; Heidarzade et al., 2017). The seeds of local Tarom cultivars were obtained from the Seed Breeding Research Department of the Rice Research Institute of Iran and were disinfected with Tiram's Carboxin 2% solution before cultivation.

Seedlings of local Tarom cv. were inoculated with pathogenic bacteria at the 5-6 leaf stage with a concentration of 10^7 (OD = 0.1) (17). In the case of salicylic acid, chitosan and potassium phosphite treatments, foliar spray was used 48 hours before pathogen inoculation (Sodhi et al., 2003; Mkhoshkdaman and Pedramfar, 2009). Seedlings were kept in the greenhouse after inoculation with non-living elicitors and pathogenic bacteria. Sampling of leaf tissue of treated and control seedlings (inoculated with bacteria) were done at 0, 48, 72 and 96 hours after inoculation (hai). To investigate the mechanism of resistance, the activity of defense-related enzymes such as catalase (CAT), guaiacol peroxidase (GPX), and superoxide dismutase (SOD) was studied in treatments. This study carried out as factorial experiment in a randomized complete design. All data was analyzed by SAS Software version 9.00. Comparison of the means was performed by Tukey test ($P \leqslant 0.01$). To extract the enzyme solutions, 0.5 g of the leaf sample were ground with a mortar in liquid nitrogen and homogenised with icecold 5ml phosphate buffer (pH 7.5) containing 0.5 mM EDTA. The homogenate was centrifuged at 20,000 RPM for 15 min at 4 °C and supernatant used for assays of the activities of CAT and GPX (Reuveni, 1995) .

Catalase enzyme activity was measured based on Aebi's method (Aebi, 1984). The reaction mixture consisted of 1.5 ml of 100 mM phosphate buffer (pH = 7), 0.5 ml of 7.5 mM hydrogen peroxide and 50 µl of enzyme extract. The final volume was adjusted to 3 ml by adding distilled water. The absorbance changes of the reaction solution were measured by a spectrophotometer (Camspec M501, UV-Vis single beam scanning spectrophotometer, UK) at a wavelength of 240 nm for 2 minutes. To determine the activity of GPX (Chance and Maehly, 1955), the reaction mixture (2 ml) included 1 ml of 100 mM phosphate buffer (pH=7), 250 µl of 0.1 mM EDTA, 1 ml of 5 mM guaiacol, one ml of 15 mM peroxide, and 50 µl of the enzyme solution. The reaction was started by adding the enzyme solution and the increase in absorbance was recorded by a spectrophotometer at a wavelength of 470 nm for 1 minute. Enzyme activity was determined based on the amount of tetragaiacol and obtained using the extinction coefficient of 1.33 µmol/cm. The activity of SOD enzyme was performed using the method of Beyer and Fridovich (Beyer Jr and Fridovich, 1987). 1000 µl of 50 mM phosphate buffer containing 1.5 mM EDTA, 10 mM methionine, and 75 µM nitroblue tetrazolium chloride were mixed with 100 µL of 1 µM riboflavin and 100 µl of enzyme extract. Then the lamp was turned off and the absorbance changes of the reaction mixture were read by a spectrophotometer at a wavelength of 560 nm.

Results

Morphological and pathological features

In this experiment, the local Tarom cultivar was used as a very susceptible variety against *Xoo* bacterial pathogen (Derakhshan et al., 2020). The pathogenicity of the desired bacterium was proven with a concentration of 10^7 cfu/ml at the 5-6 leaf stage (Backer, 2002) (Figure 1). Two standard methods of sterile scissors and fogging method with OD = 0.1 equivalent to a suspension with a concentration of 10^7 cfu/ml were chosen to conduct the experiment. Morphological and pathological characteristics based on Koch's principles were consistent with the results of other researchers (Goto, 1964; Schaad et al., 1996; Backer, 2002). This bacterium causes scorched spots on rice leaves and yellow spots and necrosis appear after 2 weeks.

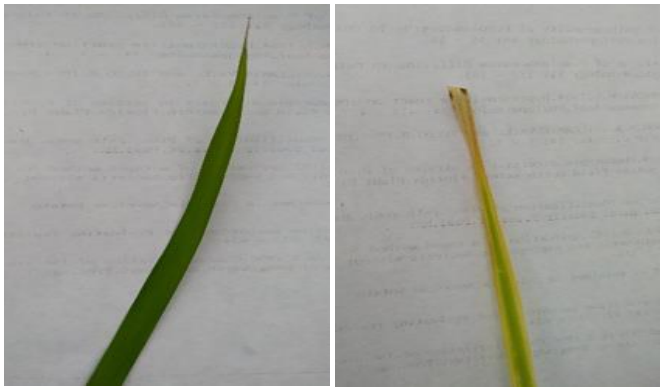


Figure 1. Pathogenicity test (Right figure: mock plant; Left: rice plant inoculated with *Xanthomonas oryzae* pv. *oryzae*).

Then the yellow spots spread towards the base of the leaf, and finally more parts of the leaf surface showed yellow symptoms. This bacterium causes scorched spots on rice leaves and within two weeks, yellow spots and necrosis appear. Then, the yellow spots progress in the leaf and spread towards the base, and finally, a larger part of the leaf surface showed yellow symptoms. While in the leaves of the control plants (treated with distilled water), no signs of water-soaked and yellowness were observed.

Disease progression

The results of the analysis of variance of the data showed that there were a significant difference between the treatments. The average comparison between the treatments showed the highest disease progress in the control treatment inoculated with bacteria and the lowest for potassium phosphite treatment (Figure 5). The results of analysis of variance of the data showed that the effect of non-living stimuli treatment (salicylic acid, chitosan, and potassium phosphite), sampling times and the interaction effect of the treatment with time on the activity of CAT, GPX and SOD enzymes in the leaves of local Tarom rice inoculated with *Xoo* at a probability level of one percentage was significant (Table 1). The amount of enzyme activity compared to the control was significant in different treatments and at different times, and it showed that these three treatments increased the activity of defense enzymes (Table 2).

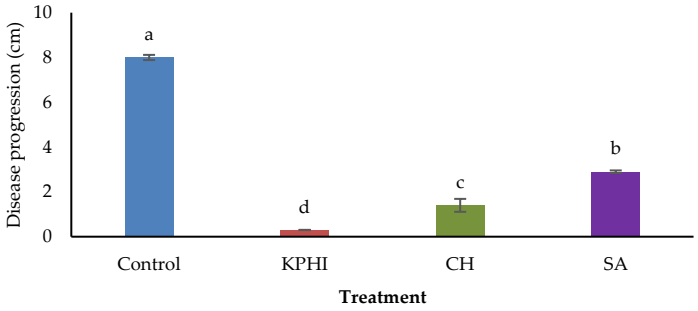


Figure 5. Disease progression in the leaves of rice plants treated with potassium phosphite, chitosan and salicylic acid, 11 days after inoculation with Xanthomonas oryzae pv. oryzae. (Control: rice plant inoculated with bacteria, KPHI: rice plant treated with potassium phosphite, CH: rice plant treated with chitosan, SA: rice plant treated with salicylic acid).

Table 1. Analysis of variance for activity of defense enzymes of rice plants treated with potassium phosphite, chitosan and salicylic acid against *Xanthamonos oryzae* pv. *oryzae.*

Source of variations	df	MS		
		CAT	GPX	SOD
Treatment	3	11.44**	37.35**	54.46**
time	3	0.508**	22.26**	36.96**
Treatment*time	9	2.43**	1.80**	4.26**
Error	32	0.005	0.19	0.009
CV		3.93	7.64	1.77

* * Significant at P≤0.01.

Table 2. Mean comparison between inducers for CAT, GPX and SOD enzymes CAT, GPX and SOD enzymes.

Enzymes	Treatments			
	Potassium phosphite	Chitosan	Salicylic acid	Control (rice plant inoculated with bacteria)
CAT	2.22^{b}	2.62^{a}	2.13^{b}	0.42^{c}
GPX	7.48^{a}	6.74^{b}	5.47^{c}	3.45^{d}
SOD	7.69^{a}	6.48^{b}	4.82^{c}	2.78^{d}

Comparison of the means was performed by Tukey test (P ≤ 0.01).
Means followed by different letters in each row show significant difference (P≤0.01).

CAT activity

The results of the interaction effect of comparing the average data show that the highest activity of this enzyme was observed in the chitosan treatment (2.62 U mg-1 protein) and the lowest level was observed in the control treatment inoculated with bacteria (0.42 U mg-1 protein), and they were significantly different from each other (Table 1). The highest level of enzyme activity was observed at 72 hours after inoculation (2.08 U mg-1 protein) and the lowest at 0 hours (1.60 U mg-1 protein) (Table 2). Also, the highest effect in increasing the amount of CAT enzyme was observed in the chitosan treatment, which was recorded 48 hours after inoculation (3.48 U mg-1 protein), which was increased 6 times compared to the control treatment infected with bacteria (0.43 U mg-1 protein), and the lowest increase was related to the salicylic acid treatment (1.23 U mg-1 protein) (Figure 2). Salicylic acid treatment recorded the highest amount of CAT enzyme activity at 0 (2.69 U mg-1 protein) and 96 (2.93 U mg-1 protein) hours. Potassium phosphite treatment recorded the maximum activity of CAT enzyme at 72 hours (3.82 U mg-1 protein) and reached the highest value only at this time. In measuring the amount of CAT enzyme among the four sampling times, the maximum average activity of this enzyme was recorded at 72 hours (2.08 U mg-1 protein) after inoculation, after which the enzyme activity decreased at 48 (1.94 U mg-1 protein), 96 (1.78 U mg-1 protein), and 0 (1.60 U mg-1 protein) hours, respectively. Sampling times were significantly different at the one percent probability level (Figure 2).

GPX activity

According to the obtained results of the interaction effect of comparing the average data (Figure 3), the activity level of this enzyme in 48 hours (7.50 U mg-1 protein) after inoculation is at the highest level in all treatments, which has a significant difference with other sampling times at the level of 1% probability (Table 1). The results of the activity of GPX showed that the most effective treatment in increasing the average activity of this enzyme was the potassium phosphite treatment (7.48 U mg-1 protein), which reached its highest level at 48 hours, which was almost three times the control treatment infected with bacteria, which had a significant difference with other treatments at the probability level of 1% (Table 2).

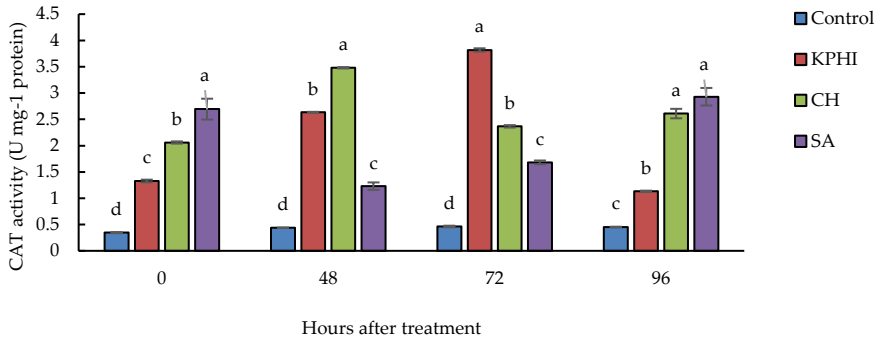


Figure 2. Change in catalase activity (CAT) in leaves of rice plants treated with potassium phosphite, chitosan and salicylic acid at hours 0, 48, 72 and 96 after inoculation with *Xanthomonas oryzae* pv. *oryzae*. (Control: rice plant inoculated with bacteria, KPHI: rice plant treated with potassium phosphite, CH: rice plant treated with chitosan, SA: rice plant treated with salicylic acid).

Chitosan (6.74 U mg-1 protein) and salicylic acid (5.47 U mg-1 protein) treatments were ranked second and third, respectively, which had significant differences from each other (Figure 3). The chitosan treatment peaked at 48 hours (9.07 U mg-^{1} protein), which was almost 2.5 times the control treatment infected with bacteria. Salicylic acid treatment recorded the highest activity of GPX enzyme at 48 hours (7.14 U mg-^{1} protein), which was almost twice the control treatment infected with bacteria. The results obtained from the average activity of GPX enzyme showed that among the four sampling times, the activity of the enzyme reached its maximum at 48 hours (7.50 U mg-^{1} protein) and then the enzyme activity decreased at 72 (6.17 U mg-^{1} protein), 96 (5.08 U mg-^{1} protein) and 0 (4.38 U mg-^{1}.protein) hours, respectively. Statistically significant difference at the level of 1% was observed between different times (Figure 3).

The results of the interaction effect of comparing the average data showed that the highest activity level of this enzyme (7.69 U mg-^{1} protein) was observed in the potassium phosphite treatment and the lowest level was observed in the control treatment inoculated with bacteria (2.78 U mg-^{1} protein) and they had significant differences with each other (Table 1). The highest activity of this enzyme was observed at 72 hours after inoculation (7.89 U mg-^{1} protein) and the lowest at 0 hours (3.71 U mg-^{1} protein) (Table 3).

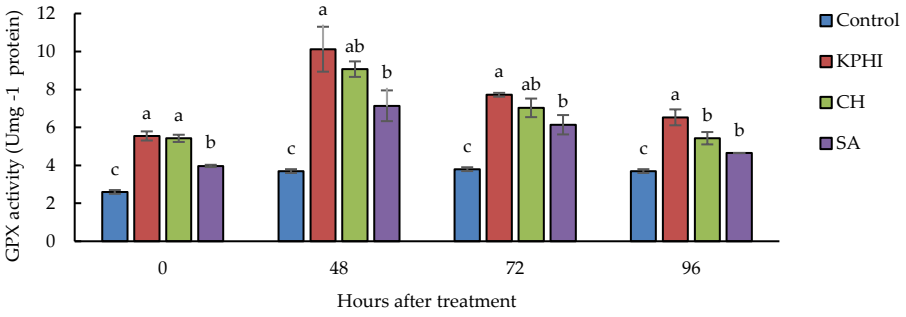


Figure 3. Changes in the activity of guaiacol peroxidase (GPX) in the leaves of rice plants treated with potassium phosphite, chitosan and salicylic acid at hours 0, 48, 72 and 96 after inoculation with *Xanthomonas oryzae* pv. *oryzae*. (Control: rice plant inoculated with bacteria, KPHI: rice plant treated with potassium phosphite, CH: rice plant treated with chitosan, SA: rice plant treated with salicylic acid).

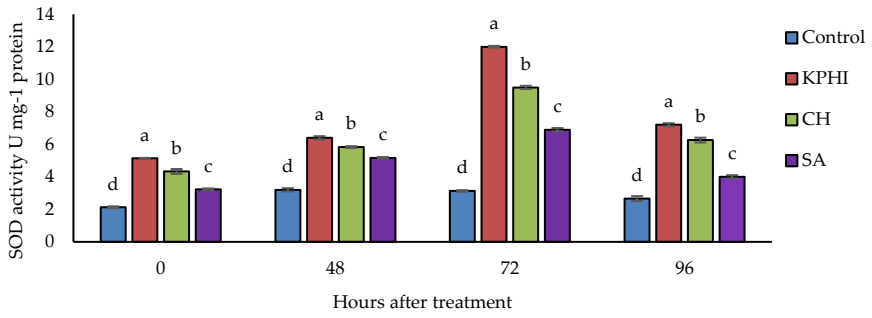


Figure 4. Changes in the activity superoxide dismutase (SOD) in the leaves of rice plants treated with potassium phosphite, chitosan and salicylic acid at hours 0, 48, 72 and 96 after inoculation with *Xanthomonas oryzae* pv. *oryzae*. (Control: rice plant inoculated with bacteria, KPHI: rice plant treated with potassium phosphite, CH: rice plant treated with chitosan, SA: rice plant treated with salicylic acid).

Table 3. Mean comparison between times after treatment for CAT, GPX and SOD enzymes.

Enzymes	Hours after treatment			
	0	48	72	96
CAT	1.60^{d}	1.94^{b}	2.08^{a}	1.78^{c}
GPX	4.38^{d}	7.50^{a}	6.17^{b}	5.08^{c}
SOD	3.71^{c}	5.15^{b}	7.89^{a}	5.03^{b}

Comparison of the means was performed by Tukey test ($P \leqslant 0.01$).
Means followed by different letters in each row show significant difference ($P \leqslant 0.01$).

SOD activity

Also, the highest effect in increasing the SOD enzyme activity was recorded in the potassium phosphite treatment at 72 hours after inoculation (12 U mg-^{1} protein), which was 4 times higher than the control treatment infected with bacteria, and the lowest increase was related to the salicylic acid treatment (6.9 U mg-^{1} protein) (Figure 4). Comparing the average data among the four sampling times, the highest average activity of this enzyme was recorded at 72 hours after inoculation, after which the enzyme activity decreased at 48, 96 and 0 hours, respectively. Statistically significant difference at the level of 1% was observed between different times (Figure 4).

Discussion

All plants, both resistant and susceptible, respond to pathogen attack by inducing an appropriate signaling system, which leads to the accumulation of different gene products. The response to pathogen attack is effective at different levels: initially, pathogen recognition leads to a rapid local cell death in the plant, also known as the hypersensitive response (HR), which causes necrosis at the site of infection (local response). Then, even in uninfected plant parts, a wide range of systemic expression and long-term resistance to further infection of the pathogen is established. This leads to the production of reactive oxygen species (ROS), the activation of defense-related genes, as well as the increased expression of genes related to the production of molecules, such as phytoalexins, terpenes, pathogenicity-related proteins (PR) and many enzymes involved in defense mechanisms (phenyl Alanine ammonia lyase (PAL), polyphenol oxidases (PPOs) and peroxidases such as guaiacol peroxidase (G-POD) and ascorbate peroxidase (APX) (Heil and Bostock, 2002; Iriti and Faoro, 2009; Pieterse et al., 2009) .

In this experiment, for the CAT enzyme activity, the best treatment in increasing the amount of this enzyme was the chitosan treatment, which recorded the highest amount in 48 hours after inoculation. In plants infected with bacteria and treated with chitosan, the amount of enzyme increased by 87.41% compared to the control plants inoculated with *Xoo* (Figure 2). In the investigation of GPX enzyme activity of plants infected with *Xoo* in chitosan treated plants, it has increased by 59.2% compared to the control plants inoculated with *Xoo* (Figure 3).

Chitosan, with its low molecular weight, acts as a strong biostimulant, capable to inducing plant defense responses and activating various pathways that increase crop resistance to diseases. The most response of the plant to chitosan treatment is the formation of chemical and mechanical barriers and the synthesis of new molecules and enzymes involved in the defense response (Iriti and Faoro, 2009; Siddaiah et al., 2018). In some cases, chitosan induces a hypersensitive reaction, mainly around the site of contamination, leading to programmed cell death. This hypersensitive reaction can be followed by the systemic response of the plant's defense mechanisms. It mainly includes the synthesis and accumulation of secondary metabolites with an active role in defense: phenolic compounds such as lignin, callose, phytoalexins, PR proteins (proteins related to pathogenicity) and modulating the activity of key enzymes of metabolic pathways involved in the defense response, such as PAL, peroxidases and chitinase (Li and Zhu, 2013; Orzali et al., 2014). The results of investigating the effect of chitosan and salicylic acid on the response of rice plants against *Fusarium fujikuroi*, the cause of root and crown rot disease, showed that the two compounds of salicylic acid and chitosan can play an effective role in inducing resistance and reducing the severity of the disease caused by this pathogen (Ebrahimi et al., 2020). The use of chitosan on rice seedlings causes the production of H_2O_2, the accumulation of glucanases and chitinase, phenylalanine aminolyase, and also increases the expression of genes involved in pathogenesis (Amborabé et al., 2008; Khatami et al., 2018). The results obtained from this experiment showed that chitosan, in addition to increasing the enzymes mentioned above research, also increases the enzyme CAT and GPX in sensitive rice cultivar (local Tarom) against bacterial leaf blight. These results were consistent with other studies mentioned above (Amborabé et al., 2008; Li and Zhu, 2013; Orzali et al., 2014; Ebrahimi et al., 2020).

In this study, the potassium phosphite treatment was the best treatment in increasing the activity of the GPX enzyme, the maximum of which was observed at 48 hours, which caused a 43.63% increase of this enzyme compared to the control treatment (Figure 3). One of the most effective intracellular antioxidants is SOD enzyme. This enzyme can keep many plants safe from the attack of oxygen free radicals and make the plant stable against many environmental stresses. In a complex reaction, superoxide dismutase combines two superoxide molecules with the help of two hydrogen molecules and turns them into two hydrogen peroxide molecules which in the next step is removed by other antioxidants (Lin et al., 2011). In this research, potassium phosphite increases SOD enzyme activity 72 hours after infection, which caused a 73.98% increase of this enzyme compared to the control treatment. This enzyme can limit the pathogenic mechanism of bacteria by creating unfavorable conditions. If phosphite (PHI) is applied at appropriate concentrations, it can stimulate defense responses in plants. Some chemicals such as PHI at lower concentrations have been widely used in the management of *Phytophthora* spp. (Guest and Grant, 1991) . The protective pattern of PHI application mainly involves the direct inhibition of pathogen growth, the accumulation of stress-related metabolism, and the expression of defense genes (Daniel and Guest, 2005; Habibi Daronkolaei et al., 2023). The protection of PHI application on pathogens is dose-dependent. A low concentration of PHI (10 µg/ml) causes the synthesis of defense genes, phytoalexins and phenolic compounds. While under high concentrations (100 µg/ml), PHI can directly inhibit the growth of pathogens (Daniel and Guest, 2005; Dalio et al., 2014). PHI induces resistance to pathogen attack through priming. Priming is a process in which plants gain more resistance to pathogens by the action of some plant hormones and other chemical compounds (Conrath et al., 2002). ROS production increases and callose and pathogenicity-related genes accumulate in priming

responses. Under low concentrations of PHI and SA, pathogenicity-associated protein 1 accumulated and increased resistance to future pathogen attack (Van Loon et al., 1998; Conrath et al., 2002). Another report (Machinandiarena et al., 2012) showed that the application of Phi reduces the accumulation of H_2O_2 in fresh plant tissues.

Potassium phosphite (KH_2PO_3 or KPhi) has a high potential to control *Phytophthora infestans* due to its direct and indirect effects, which causes indirect effects by stimulating plant defense mechanisms, inducing hypersensitive reactions (HRs) and accumulation of phenylpropanoid biosynthetic enzymes to ultimately prevent the development of late blight (Eshraghi et al., 2011). The results of investigating the effect of five levels of potassium phosphite treatment at five different times in cucumber plants inoculated with *Fusarium oxysporum* f. sp. *cucumerinum-radicis* showed that compared to the control plants, the activity of all defense enzymes and metabolites investigated including GPX, CAT, SOD and also the amount of metabolites such as malondialdehyde (MDA) and hydrogen peroxide (H_2O_2) increased significantly as a result of different potassium phosphite and *Fusarium oxysporum* f. sp. *cucumerinum-radicis* (Heidarzade et al., 2017). The results obtained from our research were consistent with other studies mentioned above.

In this experiment, the activity of CAT enzyme in plants infected with *Xoo* under salicylic acid treatment increased by 87.15% at zero hour and by 84.53% at 96 hours after teatment compared to the control (inoculated with bacteria) (Figure 2). Also, the activity of GPX enzyme in plants infected with bacteria and treated with salicylic acid increased by 48.17% compared to control (inoculated with bacteria) at 48 hours (Figure 3). Foliar spraying with SA causes the production of proteins with different functions, including signal transduction, antioxidant and defense activity. Glucosyltransferase enzyme production was maximally induced by 1 mM SA, with a 7-fold increase in enzyme activity at 6 h in both roots and shoots after SA treatment (Silverman et al., 1995). However, increased peroxidase activity was recorded after foliar application of 8 mM SA in rice to induce resistance to rice blast (Daw et al., 2008). The results obtained from our experiment were consistent with the studies mentioned above.

Conclusion

In this research, the applied treatments have led to an increase in the activity of defense enzymes, effectively impeding the advancement of *Xanthomonas oryzae* pv. *oryzae* bacteria. The results of this research indicate that pretreating rice plants with chitosan, potassium phosphite and salicylic acid enhances resistance against the bacterium *Xanthomonas oryzae* pv. *oryzae*, the causative agent of bacterial leaf blight of rice, through increasing the activity of defense enzymes. The obtained results showed that chitosan was the most effective in increasing CAT enzyme activity, while potassium phosphite demonstrated superior efficacy in promoting GPX and SOD enzyme activities. The highest activity of antioxidant enzymes CAT and GPX occurred 48 hours after inoculation, while for the SOD enzyme, the peak activity was observed 72 hours after inoculation of rice plants with *Xanthomonas oryzae* pv. *oryzae* bacteria. Increasing the amount of antioxidant enzymes is one of the mechanisms of plant resistance against pathogens. In the treated plants, the enzyme activity increased compared to the infected control plants. This suggests that the applied treatments stimulate greater enzyme activity, consequently enhancing the plant's resistance to the pathogen.

Supplementary Materials

No supplementary material is available for this article.

Author contributions

Conceptualization, L.E. and V.B.; methodology, L.E.; software, L.E.; validation, L.E., V.B. and A.D.; formal analysis, L.E. and V.B.; investigation, L.E.; resources, V.B.; data curation, , L.E. and V.B.; writing—original draft preparation, , L.E. and V.B.; writing—review and editing, , L.E. and V.B.; visualization, , L.E. and V.B.; supervision, L.E., V.B., H.R. and A.D.; project administration, L.E., V.B., H.R. and A.D.; funding acquisition, L.E., V.B. and A.D.

Funding

This research received no external funding.

Acknowledgments

Conflict of interest statement

The authors declare no conflict of interest.

References

Aebi, H. (1984). "[13] Catalase in vitro," in *Methods in enzymology*. (Amsterdam: Elsevier), 121-126.

Ainsworth, E.A. (2008). Rice production in a changing climate: a meta‐analysis of responses to elevated carbon dioxide and elevated ozone concentration. *Glob Chang Biol* 14(7): 1642-1650.

Amborabé, B.-E., Bonmort, J., Fleurat-Lessard, P., and Roblin, G. (2008). Early events induced by chitosan on plant cells. *J Exp Bot* 59(9): 2317-2324.

Asada, K. (1999). The water-water cycle in chloroplasts: Scavenging of active oxygens and dissipation of excess photons. *Annu Rev Plant Biol* 50(1): 601-639.

Backer, D. (2002). Method for inoculating rice with *Xanthomonas*. *Plant Pathol*.

Badawy, M.E., and Rabea, E.I. (2011). A biopolymer chitosan and its derivatives as promising antimicrobial agents against plant pathogens and their applications in crop protection. *Int J Carbohydr Chem* 2011: 1-29.

Badawy, M.E., Rabea, E.I., and Taktak, N.E. (2014). Antimicrobial and inhibitory enzyme activity of N-(benzyl) and quaternary N-(benzyl) chitosan derivatives on plant pathogens. *Carbohydr Polym* 111: 670-682.

Basu, S., Roychoudhury, A., Saha, P.P., and Sengupta, D.N. (2010). Differential antioxidative responses of indica rice cultivars to drought stress. *Plant Growth Regul* 60: 51-59.

Beyer Jr, W.F., and Fridovich, I. (1987). Assaying for superoxide dismutase activity: some large consequences of minor changes in conditions. *Analytical biochemistry* 161(2): 559-566.

Buensanteai, N., Yuen, G.Y., and Prathuangwong, S. (2009). Priming, signaling, and protein production associated with induced resistance by *Bacillus amyloliquefaciens* KPS46. *World J Microbiol Biotechnol* 25: 1275-1286.

Chance, B., and Maehly, A. (1955). "Assay of catalases and peroxidases". Meth Enzymol).

Conrath, U., Pieterse, C.M., and Mauch-Mani, B. (2002). Priming in plant–pathogen interactions. *Trends Plant Sci* 7(5): 210-216.

Dai, L.Y., Liu, X.L., Xiao, Y.H., and Wang, G.L. (2007). Recent advances in cloning and characterization of disease resistance genes in rice. *J Integr Plant Biol* 49(1): 112-119.

Dalio, R.J., Fleischmann, F., Humez, M., and Osswald, W. (2014). Phosphite protects *Fagus sylvatica* seedlings towards *Phytophthora plurivora* via local toxicity, priming and facilitation of pathogen recognition. *PLOS One* 9(1): e87860.

Daniel, R., and Guest, D. (2005). Defence responses induced by potassium phosphonate in *Phytophthora palmivora*-challenged *Arabidopsis thaliana*. *Physiol Mol Plant Pathol* 67(3-5): 194-201.

Daw, B., Zhang, L., and Wang, Z. (2008). Salicylic acid enhances antifungal resistance to *Magnaporthe grisea* in rice plants. *Australas Plant Pathol* 37: 637-644.

Deliopoulos, T., Kettlewell, P.S., and Hare, M.C. (2010). Fungal disease suppression by inorganic salts: a review. *Crop Prot* 29(10): 1059-1075.

Derakhshan, A., Babaeizad, V., Panjekeh, N., and Taheri, A. (2020). Study of biochemical and molecular changes of iranian rice cultivars in interaction with bacterial pathogen *Xanthomonas oryzae* pv. oryzae causes leaf blight disease. *J Crop Breed* 12(36): 77-89. (In Persian).

Ebrahimi, A., Taliei, F., and Zolfaghari, A. (2020). Effect of salicylic acid and chitosan on response of rice against *Fusarium fujikuroi* the causal agentof rice root and crown rot. *Appl entomol phytopathol* 88(1): 23-37.

Eshraghi, L., erson, Anderson, J., Aryamanesh, N., Shearer, B., McComb, J., Hardy, G.S., and O'Brien, P. (2011). Phosphite primed defence responses and enhanced expression of defence genes in *Arabidopsis thaliana* infected with *Phytophthora cinnamomi*. *Plant Pathol* 60(6): 1086-1095.

Falcón-Rodríguez, A.B., Wégria, G., and Cabrera, J.-C. (2012). "Exploiting plant innate immunity to protect crops against biotic stress: chitosaccharides as natural and suitable candidates for this purpose," in *New perspectives in plant protection*.), 139-166.

Ganesan, V., and Thomas, G. (2001). Salicylic acid response in rice: influence of salicylic acid on H2O2 accumulation and oxidative stress. *Plant Sci* 160(6): 1095-1106.

Goto, M. (1964). Kresek and pele yellow leaf systemic symptoms of bacterial leaf blight of rice caused by *Xanthomonas oryzae* (Uyeda et Ishiyama) Dawson. *PI Dis Rep* 48: 858-861.

Guest, D., and Grant, B. (1991). The complex action of phosphonates as antifungal agents. *Biol Rev* 66(2): 159-187.

Habibi Daronkolaei, M., Rahimian, H., and Dehestani, A. (2023). The effect of potassium phosphite in the induction of some rice resistance genes in following the sheet blight disease agent *Rhizoctinia solani*. *J Crop Breed* 15(46): 62-72. (In Persian).

Heidarzade, S., Gharanjik, S., Dehestani, A., and Shahriari, D. (2017). *Fusarium oxysporum* f. sp. radicis-cucumerinum. *Iranian J Hortic Sci* 48(3): 601-611. (In Persian).

Heil, M., and Bostock, R.M. (2002). Induced systemic resistance (ISR) against pathogens in the context of induced plant defences. *Ann Bot* 89(5): 503-512.

Huang, Y., Cai, S., Zhang, G., and Ruan, S. (2020). Transcriptome-based analysis of phosphite-induced resistance against pathogens in rice. *Plants* 9(10): 1334.

Hussain, S., Khan, F., Cao, W., Wu, L., and Geng, M. (2016). Seed priming alters the production and detoxification of reactive oxygen intermediates in rice seedlings grown under sub-optimal temperature and nutrient supply. *Front Plant Sci* 7: 439.

Iriti, M., and Faoro, F. (2009). Chitosan as a MAMP, searching for a PRR. *Plant Signal Behav* 4(1): 66-68.

Katoch, R., Mann, A., and Sohal, B. (2005). Enhanced enzyme activities and induction of acquired resistance in pea with elicitors. *Int J Veg Sci* 11(1): 67-83.

Khatami, M., Ahangar, L., Taliei Tabari, F., Sabouri, H., and Babaeizad, V. (2018). Assay of NPR1, MLO and BI-1 genes expression in susceptible wheat to powdery mildew after treatment with chitosan. *Cell Mol Res* 31(4): 471-483. (In Persian).

Li, B., Wang, X., Chen, R., Huangfu, W., and Xie, G. (2008). Antibacterial activity of chitosan solution against *Xanthomonas* pathogenic bacteria isolated from *Euphorbia pulcherrima*. *Carbohydr Polym* 72(2): 287-292.

Li, S.j., and Zhu, T.h. (2013). Biochemical response and induced resistance against anthracnose (*Colletotrichum camelliae*) of camellia (*Camellia pitardii*) by chitosan oligosaccharide application. *For Pathol* 43(1): 67-76.

Li, X., and Zhang, L. (2012). "SA and PEG-induced priming for water stress tolerance in rice seedling," in *Information Technology and Agricultural Engineering. Advances in Intelligent and Soft Computing*. (Berlin: Springer), 881-887.

Lobato, M., Olivieri, F., Altamiranda, E.G., Wolski, E., Daleo, G., Caldiz, D., and Andreu, A. (2008). Phosphite compounds reduce disease severity in potato seed tubers and foliage. *Eur J Plant Pathol* 122: 349-358.

Machinandiarena, M.F., Lobato, M.C., Feldman, M.L., Daleo, G.R., and Andreu, A.B. (2012). Potassium phosphite primes defense responses in potato against *Phytophthora infestans*. *J Plant Physiol* 169(14): 1417-1424.

Mansilla, A.Y., Albertengo, L., Rodríguez, M.S., Debbaudt, A., Zúñiga, A., and Casalongué, C. (2013). Evidence on antimicrobial properties and mode of action of a chitosan obtained from crustacean exoskeletons on *Pseudomonas syringae* pv. tomato DC3000. *Appl Microbiol Biotechnol* 97: 6957-6966.

Mastouri, F., Björkman, T., and Harman, G.E. (2012). *Trichoderma harzianum* enhances antioxidant defense of tomato seedlings and resistance to water deficit. *Mol Plant Microbe Interact* 25(9): 1264-1271.

Mew, T., Vera Cruz, C., and Medalla, E. (1992). Changes in race frequency of *Xanthomonas oryzae* pv. oryzae in response to rice cultivars planted in the Philippines. *Plant Dis* 76(10): 1029-1032.

Mittler, R. (2002). Oxidative stress, antioxidants and stress tolerance. *Trends Plant Sci* 7(9): 405-410.

Mkhoshkdaman, M., and Pedramfar, H. (2009). Identification of causal agent of bacterial bight of rice in the fields of Guilan province. *J Agric Mach* 23(1).

Orzali, L., Forni, C., and Riccioni, L. (2014). Effect of chitosan seed treatment as elicitor of resistance to *Fusarium graminearum* in wheat. *Seed Sci Technol* 42(2): 132-149.

Pieterse, C.M., Leon-Reyes, A., Van der Ent, S., and Van Wees, S.C. (2009). Networking by small-molecule hormones in plant immunity. *Nat Chem Biol* 5(5): 308-316.

Reuveni, R. (1995). Biochemical marker of disease resistance. *Molecular methods in plant pathology*: 99-114.

Saba Anwar, M.I., Raza, S.H., and Iqbal, N. (2013). Efficacy of seed preconditioning with salicylic and ascorbic acid in increasing vigor of rice (*Oryza sativa* L.) seedling. *Pak J Bot* 45(1): 157-162.

Schaad, N., Wang, Z., Di, M., McBeath, J., Peterson, G., and Bonde, M. (1996). An improved infiltration technique to test the pathogenicity of *Xanthomonas oryzae* pv. oryzae in rice seedlings. *Seed Sci Technol* 24: 449-456.

Siddaiah, C.N., Prasanth, K.V.H., Satyanarayana, N.R., Mudili, V., Gupta, V.K., Kalagatur, N.K., Satyavati, T., Dai, X.-F., Chen, J.-Y., and Mocan, A. (2018). Chitosan nanoparticles having higher degree of acetylation induce resistance against pearl millet downy mildew through nitric oxide generation. *Sci Rep* 8(1): 2485.

Siddiq, E., and Vemireddy, L.R. (2021). *Advances in genetics and breeding of rice: an overview.* Springer Nature Switzerland AG.

Silverman, P., Seskar, M., Kanter, D., Schweizer, P., Métraux, J.-P., and Raskin, I. (1995). Salicylic acid in rice (biosynthesis, conjugation, and possible role). *Plant Physiol* 108(2): 633-639.

Sodhi, M., Vikal, Y., George, M.L.C., Bala, G., Mangat, G., Garg, M., Sidhu, J., and Dhaliwal, H. (2003). DNA fingerprinting and virulence analysis of *Xanthomonas oryzae* pv. oryzae isolates from Punjab, northern India. *Euphytica* 130: 107-115.

Sticher, L., Mauch-Mani, B., Métraux, and JP (1997). Systemic acquired resistance. *Annu Rev Phytopathol* 35(1): 235-270.

Thao, H.T.B., and Yamakawa, T. (2009). Phosphite (phosphorous acid): fungicide, fertilizer or bio-stimulator? *J Soil Sci Plant Nutr* 55(2): 228-234.

Valadi, S., Soleimani, M.J., Khoda Karamian, G., and Ghiasvand, T. (2013). Effect of salicylic acid & chitosan on induction of resistance in chickpea against fusarial wilt & root rot. *Iran J Plant Pathol* 49(2): 181-199. (In Persian).

Van Loon, L., Bakker, P., and Pieterse, C. (1998). Systemic resistance induced by rhizosphere bacteria. *Annu Rev Phytopathol* 36(1): 453-483.

JPMB
Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

مقاله پژوهشی

تغییرات فعالیت آنزیم‌های آنتی‌اکسیدانی در گیاهان برنج تیمارشده با القاکننده‌های غیرزنده مختلف در برابر عامل سوختگی باکتریایی *Xanthomonas oryzae* pv. *oryzae*

لیلا اصفهانی [۱]، ولی‌اله بابایی‌زاد *[۱]، حشمت‌اله رحیمیان [۱] و علی دهستانی [۲]

[۱]. گروه گیاه‌پزشکی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران

[2]. پژوهشکده ژنتیک و زیست‌فناوری کشاورزی طبرستان، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران

چکیده: برنج (*Oryzae sativa* L.) مهم‌ترین غذای اصلی در جهان است. سوختگی برگ باکتریایی (BLB) برنج ناشی از *X. oryzae* pv. *oryzae* یک بیماری بسیار مخرب و گسترده است. مدیریت شیمیایی برای کنترل این بیماری ناکارآمد به نظر می‌رسد. در این آزمایش تاثیر سه تیمار سالیسیلیک‌اسید، فسفیت‌پتاسیم و کیتوزان در گیاه برنج حساس تلقیح شده با باکتری *Xanthamonos oryzae* pv. *Oryzae* بر القای مقاومت و فعالیت آنزیم‌هایی مانند کاتالاز (CAT)، گایاکول‌پراکسیداز (GPX) و سوپراکسیددیسموتاز (SOD) در طی چهار روز بررسی شد. نتایج نشان داد که بیشترین و کمترین فعالیت CAT به ترتیب در تیمار کیتوزان و اسیدسالیسیلیک ثبت شد. حداکثر میزان فعالیت کاتالاز ۷۲ ساعت پس از تلقیح بود. مقایسه فعالیت آنزیم‌های GPX و SOD در زمان‌های مختلف نمونه‌برداری نشان داد که این آنزیم‌ها به ترتیب در ۴۸ ساعت و ۷۲ ساعت ساعت پس از تلقیح در تمامی تیمارها در بالاترین سطح قرار دارند. اما در بین تیمارهای مختلف، بیشترین فعالیت این آنزیم‌ها در گیاهان آلوده به باکتری تحت تیمار فسفیت‌پتاسیم مشاهده شد. یافته‌ها نشان می‌دهد که از بین سه تیمار سالیسیلیک‌اسید، فسفیت‌پتاسیم و کیتوزان، تیمار فسفیت‌پتاسیم باعث افزایش فعالیت آنزیم‌های دفاعی گیاه در برابر عامل بیماری‌زا می‌شود و سپس منجر به محدود کردن عامل بیماری‌زا و کاهش علائم بیماری می‌شود.

کلمات کلیدی: *Oryzae sativa*، اسیدسالیسیلیک، فسفیت‌پتاسیم، کیتوزان، مقاومت القایی.

ویراستار علمی
دکتر شیده موجرلو،
دانشگاه آرهوس، دانمارک
دانشگاه صنعتی شاهرود، ایران

تاریخ
دریافت: ۱۴ شهریور ۱۴۰۲
پذیرش: ۲۸ دی ۱۴۰۲
چاپ: ۱۱ بهمن ۱۴۰۲

نویسنده مسئول
دکتر ولی‌اله بابایی‌زاد

babaeizad@yahoo.com

ارجاع به این مقاله

Esfahani, L., Babaeizad, V., Rahimian, H. and Dehestani A. (2023). Investigation of changes in antioxidant enzyme activities in rice plants treated with various abiotic inducers against the bacterial blight agent *Xanthomonas oryzae* pv. *oryzae* . *J Plant Mol Breed* 11 (1): 41-53. 10.22058/JPMB.2024.2010427.1284 .

JPMB | Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

Enhancing haploid wheat induction efficiency through the wheat × maize cross utilizing silver nitrate and calcium phosphate

Hamed Modirrousta [1,*] Raheleh Khademian [1] and Reza Bozorgipour [2]

[1.] Department of Genetic and plant breeding faculty of Agriculture and Natural Resources, Imam Khomeini International University, Qazvin, Iran
[2.] Seed and Plant Improvement Institute, Agricultural Research, Education and Extension Organization Karaj, Iran

Original paper

Abstract: This study aimed to assess the efficacy of silver nitrate and calcium phosphate in inducing chromosome elimination and promoting haploid wheat generation in the context of wheat-maize crosses. The experiment explored the impact of various treatments, including non-use (control) and the utilization of silver nitrate (3 gr/l) combined with calcium phosphate (0.1 gr/l). Additionally, different durations of treatments post-pollination, spanning 48 and 72 hours, were considered. The study evaluated key characteristics in wheat x maize system, such as the number of florets, number as well as the quantity and percentage of successful seed formations. The findings underscored significant distinctions among the treatments, achieving a level of significance of 1%. The use of silver nitrate in conjunction with calcium phosphate for wheat tillers demonstrated an increased propensity for haploid embryo formation. Specifically, maintaining wheat tillers in a solution of silver nitrate and calcium phosphate for 72 and 48 hours, in comparison to the control group, resulted in haploid embryo formations with proportions of 12.53%, 5.60%, and 3.82%, respectively. The findings suggest that prolonged exposure to the silver nitrate-calcium phosphate treatment increases the development of haploid embryos.

Keywords: haploid, chromosome elimination, interspecific hybridization, embryo rescue, tissue culture.

OPEN ACCESS

Edited by

Dr. Esmaeil Bakhshandeh,
Genetics and Agricultural Biotechnology Institute of Tabarestan (GABIT), Sari Agricultural Sciences and Natural Resources University (SANRU), Iran

Date

Received: 15 July 2023
Accepted: 19 December 2023
Published: 11 February 2024

Correspondence

Mr. Hamed Modirrousta
hamed.modirrousta@ut.ac.ir

Citation

Modirrousta, H., Khademian, R., and Bozorgipoor, R. (2023). Enhancing haploid wheat induction efficiency through the wheat× maize crossbreeding system utilizing silver nitrate and calcium phosphate. *J. Plant Mol. Breed 11* (1): 54-62. doi:10.22058/jpmb.2023.2007085.1279.

Introduction

Haploid plant production serves several crucial purposes, including the acceleration of the breeding process, efficient selection of genotypes, and simplification of genetic analysis (Sharma et al., 2002). This method offers the advantage of achieving complete homozygosity in a single generation, as it results in double haploid plants characterized by pure genetic lines. However, it's important to note that this process can be time-consuming. The application of double haploids is pivotal in the cultivation of many crop varieties, relying on efficient protocols for haploid creation (Niu et al., 2014). Successful production of haploid wheat has been demonstrated through interspecific or intergeneric cross (Sourour et al., 2011). Wheat plants have been induced to become haploid via a chromosomal elimination method, involving crosses with various herbaceous plants such as *Teosinte, Sorghum, Tripsacumpearl, Millet, and Hordeum* Notably, crosses between wheat and maize have shown higher efficiency compared to other crosses, increasing the likelihood of obtaining more haploid plants through this approach (Khan and Ahmad, 2011). However, it's essential to acknowledge that this system comes with its challenges, including slight incompatibility issues, the need for specific conditions to maintain wheat and maize plants, limited embryo differentiation, and early maturity. These challenges must be addressed to optimize the effectiveness of this method.

The application of some growth regulators impresses on the frequency of embryo formation and embryo growth and development (Knox et al., 2000). The most common growth regulators used for this purpose are 2, 4-D, which play an essential role in wheat haploid embryo induction (Suenaga and Nakajima, 1989). Using auxins in the external environment causes to increase in ethylene production (Mensuali - Sodi et al., 1992). Silver nitrate ion controls the external ethylene in the plant and plant parts (Bidmeshkipour et al., 2007). The use of silver nitrate in the culture medium can increase embryogenesis (Sridevi et al., 2010). Calcium phosphate is a practical element in transferring foodstuffs to organs in the plant. Calcium plays a role in response to cell signals and responds to various stimuli; furthermore, it plays a role in metabolism and enzymatic absorption, and hormonal processes. Phosphorus, on the other hand, is an essential element for storing and transferring energy in plants (Upadhyaya et al., 2017). Although the chromosome elimination method in wheat × maize crosses has some disadvantages in this study, we tried to increase the chromosomal elimination method's efficiency to improve the production of wheat haploid embryos in crosses with maize by using silver nitrate with calcium phosphate.

Materials and Methods

Plant materials and Plant growth conditions

Plant material used in this study comprised six genotypes of the F_2 generation resulting from hybrid crosses between hexaploid wheat "Mortared" and "Man" as well as"Mortared" × "Pishgam". These genotypes served as the maternal parent, while three commercial maize cultivars, namely SC400, SC260 and SC301, were utilized as the pollinator parent. This research was conducted at the Seed and Plant Improvement Institute of Karaj, Iran. Maize seeds were cultivated 45 days earlier than wheat seeds to synchronize their flowering for pollination. The seeds were treated with carbon thiram 40% and then were sowed in pots with a diameter of 8 centimeters and a height of 10 centimeters with a bed of sand in a greenhouse (250c with 16h/8 light/dark).

The sowing of maize seeds was done continuously until the end of pollination time. After germination and growth sufficiently, the seedlings (2-3 leaf stage) were transferred in the pots with a diameter of 23 cm and a height of 18 cm with a bed of farm soil, sand, and peat moss, respectively, with ratios of 5: 1: 1. Because the wheat seeds were elite, twenty seeds of each genotype were cultivated in laboratory conditions in Petri dishes with 10 centimeters span diameter on filter paper after sterilization with 2% sodium hypochlorite for 5 minutes and 70% ethanol for 1 minute and three times washing with distilled water. The Petri dishes were then exposed to a temperature of 4 degrees Celsius; if they had seed dormancy, they would be broken, and the other growth condition was the same. After 24 hours, they were kept in the

germinator at 22C for 2-3 weeks. After that, they were transferred in pots with a diameter of 13 centimeters and a height of 13 centimeters supplemented with farm soil, sand, and peat moss with ratios of 5: 1: 1, respectively, at 200C and 16h/8h light/dark. Wheat and maize plants were fed with commercial fertilizers containing all the elements needed for plant growth.

Pollination, treatment, and maintenance tillers

The crosses were randomized regarding the availability of spikes from each genotype and fresh pollen from maize plants. When two-thirds of the spikes emerged from their boot leaf, and the middle-floral spikelet had two branches and a fluffy stigma, it was considered the best condition for accepting pollen. The appropriate spikes were harvested from the stem's last part and transferred to the laboratory. Then, the upper and lower florets on the spikes were removed, and two-thirds of the florets in each spike were cut with scissors (without emasculation and removing central florets). Immediately afterward, fresh pollen from each maize variety was collected in the greenhouse on an aluminum foil and brought to the laboratory. Pollination was conducted using a very soft brush. The pollinated spikes on the stems were then transferred to an MS medium supplemented with 0.1 g/l 2, 4-D, 0.1 g/l silver nitrate, three g/l calcium phosphate, 40 g/l of sucrose, and 8 ml sulfurous acid. After 48 and 72 hours, the ends of the stems were shortened to size 2 to 3 centimeters. This medium was replaced with another medium containing 40 g/l sucrose and 8 ml sulfurous acid until the harvesting of the seeds, cutting the end of stems was done every 2 to 3 days, and a fresh medium was replaced. The seeds were harvested after 20 to 22 days, and the self-pollinated seeds were removed. Embryo rescue was done for seeds in a laminar airflow after sterilization of seeds with 2% sodium hypochlorite for 5 minutes and 70% ethanol for 1 minute. The seeds were washed three times with distilled water to remove the effects of sodium hypochlorite and ethanol. Then, they were placed on binoculars so that the line on the surface of the seeds was downwards. Due to the absence of endosperm, the formed embryos were definitely haploid, so they were cultured in an MS medium.

Sampling and statistical analysis

The number of pollinated florets, the number of seeds, and the number of haploid embryos were recorded. The following formula calculated the percentage of seed and embryo formation.

$$\text{Frequency of seed formation} = \frac{\textit{Number of seed formed}}{\textit{Number of pollinated florets}} \times 100$$

$$\text{Frequency of embryoformation} = \frac{\textit{Number of embryo formed}}{\textit{Number of seed formed}} \times 100$$

The data were compared using the Chi-square x^2 test.

Results

Seed formation

According to Table 1 seed formation was significantly greater in the treated florets than in the control. Furthermore, the highest seed formation (71%) was observed in the silver nitrate treatment containing calcium phosphate for 48 hours (see Figure 1).

Embryo formation

Table 2 shows that treatments with silver nitrate containing calcium phosphate for 48 and 72 hours were effective significantly in enhancing wheat embryo formation in the wheat x maize hybridization. The formation of wheat embryos increased in alignment with longer with longer treatment time when a chemical combination of silver nitrate and calcium phosphate was applied. The highest frequency of embryos (12.53%) was obtained from 72 hours of treatment which showed a significant difference compared to both 48 hours (5.61%) and control (3.82 %) treatments (refer to Figure 2). Therefore, maintaining tillers for 72 hours immediately after pollination in the MS medium supplemented with 0.1 g/l silver nitrate, 3 g/l calcium phosphate was the best. Since the frequency of haploid embryos obtained from seeds reached 12.53%, which was 2.2 times more than the haploid production in the 48-hour treatment and 3.3 times more than in the control treatment. Figure 3 illustrates the various steps of haploid embryo induction in wheat, starting from the emasculation of the florets to seed formation.

Table 1. The effect of treatments in seed formation (caryopsis) in the wheat x maize crosses.

Treatments	Number of pollinated florets	Seed formation (no)	df	χ^2
Control (without $AgNO_3$+$Ca_3(PO_4)_2$)	22,561	16,143		17.95
48 hours use of $AgNO_3$+$Ca_3(PO_4)_2$	14,286	10,698		46.83
72 hours use of $AgNO_3$+$Ca_3(PO_4)_2$	8,533	5,960		175.80
Total	45,380	32,801	**10**	240.58**

**, ns: significant at 1% level and nonsignificant, respectively

Table 2. The effect of treatments in embryo formation in the wheat x maize crosses.

Treatment	Number of seeds	Embryo formation (no)	df	χ^2
Control (without $AgNO_3$+$Ca_3(PO_4)_2$	16,143	617		26.18
48 hours use of $AgNO_3$+$Ca_3(PO_4)_2$	10,698	600		45.18
72 hours use of $AgNO_3$+$Ca_3(PO_4)_2$	5,960	747		41.15
Totals	32,801	1,964		10
			10	112.50**

**,ns: significant at 1% leveland nonsignificant, respectively

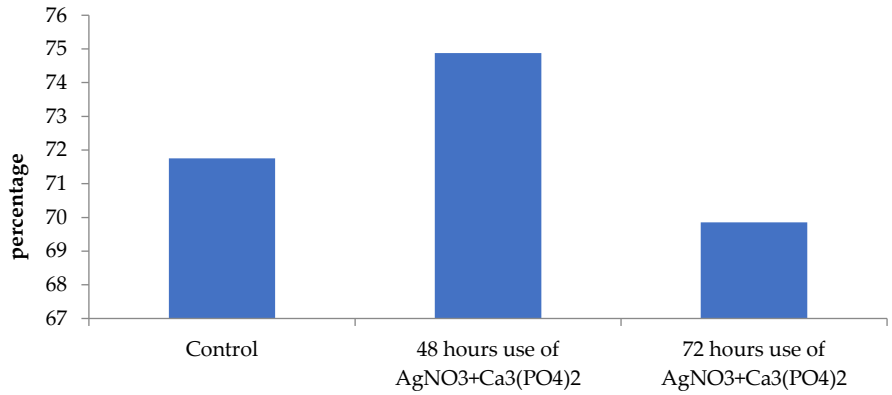


Figure 1. Percentage of seed formation with and without (control) exogenous application of $AgNO_3$+$Ca_3(PO_4)$.

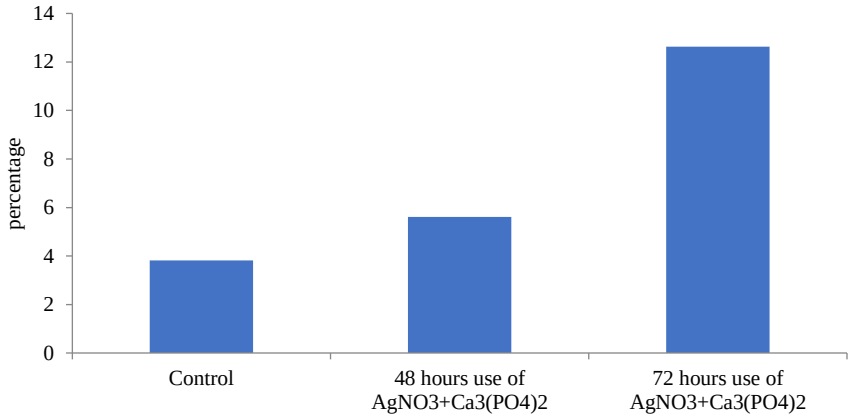


Figure 2. Percentage of embryo formation with and without (control) exogenous application of AgNO3+Ca3(PO4).

Figure 3. Different steps of haploid induction in wheat embryos, A: wheat spike in the optimal growth stage for pollination, B: emasculated spike, C: Spikes ready to receive pollen D: Collection of maize pollen E: Pollination of wheat spike with a soft brush F: pollinated spike G: Maintenance of pollinated spikes in a liquid medium in germinator H: seed formed after 20-22 days I: Haploid embryo.

Discussion

The highest seed formation was seen in silver nitrate treatment with calcium phosphate for 48 hours. Upadhyaya et al. (2017) reported that calcium phosphate is an influential factor in transferring foodstuffs to plant organs. Also,

calcium plays a role in response to cell signals and various stimuli in metabolic processes and absorbing enzymes and growth regulators (Upadhyaya et al., 2017). Nevertheless, phosphorus is an essential element that needs to store and transfer energy in plants. Therefore, using these two substances results in an improved percentage of seed formation. However, according to the present study's results, the treatment time influenced the frequency of seeds significantly. So, with increasing treatment time, the result was reversed, and the seed frequency decreased.

Although 72 hours treatment resulted in the lowest hybrid seed formation in our experiment, it is important to note that the primary goal of the present experiment was to improve production of wheat haploid embryos capable of generating haploid plants via *in vitro* culture. The production ability to form haploid embryos might depend on the application of silver nitrate and growth regulators treatment, alone or in combination (Slama-Ayed et al., 2019). It was reported that treatment with silver nitrate and 2,4-D is an efficient method in wheat and maize cross-system, and also, keeping crossed tillers in 2,4-D solution saves time and reduces problems associated with injection of this substance into the plant (Brazauskas et al., 2005). Bidmeshkipour et al. (2007) used silver nitrate along with 2,4-D in their research, and they expressed that using these substances in optimum concentrations can increase the production of the embryo and haploid plant (Bidmeshkipour et al., 2007). Applying silver nitrate in the culture medium can increase embryogenesis (Sridevi et al., 2010). Increased haploid embryo production through silver nitrate with 2,4-D treatment obtained in the study was conducted by Khan et al. (2012) They reported that treating wheat crossed tillers with 100 milligrams of silver nitrate in maintenance media resulted in a frequency of 52.53% and 28.95% seed and embryo formation, respectively. Also, increased silver nitrate in 150 mg causes more production of seed (54.5%) and embryo (26.96%) (Khan et al., 2012). Khan and Ahmad (2011) declared that the optimum concentration of silver nitrate for seed and embryo formation is 100 milligrams per liter (Khan and Ahmad, 2011). In another study, the optimal concentration of silver nitrate and 2,4-D for this aim was 100 mg/l 2,4-D combined with 75 mg/l silver nitrates (Sourour et al., 2011). In another study, Patial et al. reported in 2020, according to their research, that the response to silver nitrate as an ethylene inhibitor of silver nitrate in the medium does not affect the production of haploid embryos.

Producing haploid embryos is a significant factor in chromosomal elimination techniques. This method's frequency of haploid embryo production in bread wheat is shallow compared to the number of pollinated florets and seeds formed. While the efficiency of haploid plant production from these embryos is high, and the resulting plants are often routine and stable. In most reports, the frequency of haploid embryos obtained from seeds forming by crossing between maize and wheat is less than 5%, and it causes to increase in crosses and embryo rescue for seed production. This will laborious work whit improving the process for embryo formation in crossing between wheat and maize that some treatments for wheat-crossed tillers can reduce. In this study, the frequency of embryo formation in the control treatment was 3.82, but using silver nitrate plus calcium phosphate in tillers maintenance media during the initial 48 hours after pollination improved embryo formation by a frequency of 5.61%. The maintenance of these tillers over 48 hours in this media led to a significant increase in embryo production. So, by maintaining tillers for 72 hours immediately after pollination in this media, the frequency of haploid embryos obtained from seeds reached 12.53%, 2.2, and 3.3 times more than haploid production in 48 hours and control treatments, respectively.

Conclusion

The efficiency of haploid embryo induction is critical to implementing chromosomal elimination techniques aiming at haploid plant production in wheat. However, the frequency of haploid embryo production in bread wheat is low to considering the ratio of the number of pollinated florets and the number of seeds set. While the efficiency of haploid plant production from the induced embryos is high, and the resulting plants are often normal and stable. In most reports, the frequency of haploid embryos obtained from seeds forming by crossing between maize and wheat is less than 5%. It translates the requirement for a high in number of crosses and

embryo rescue for seed production. To reduce this laborious work we implement a treatment of wheat florets following their pollination with maize pollens . In this study, the frequency of embryo formation in comparison with the control was increased 1.5 fold when tillers were kept in medium containing silver nitrate plus calcium phosphate for 48 hours after pollination. Nonetheless, by maintaining tillers for 72 hours immediately after pollination in this treatment, the frequency of haploid embryos obtained from seeds reached 12.53%. This was 2.2, and 3.3 times more than haploid embryo production in 48 hours and control treatments, respectively.

Supplementary Materials

No supplementary material is available for this article.

Author Contributions

Conceptualization R.B, R.K; methodology, H.M, R.K, R.B; software, H.M; validation, H.M, R.K, R.B; formal analysis, R.B; investigation, H.M, R.K, R.B; data curation, H.M; writing—original draft preparation, H.M; writing—review and editing, H.M, R.K, R.B; supervision, R.B; project administration, H.M, R.B; funding acquisition, R.K, R.B; Authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by Seed Plant Improvement Institute, Agricultural Research, Education and Extension Organization and Khomeini International University.

Acknowledgments

We thank the following institutes which supported and facilitated our work: Research, Seed, Plant Improvement Institute, Agricultural Research, Education and Extension Organization Karaj, Iran, and Imam Khomeini International University, Qazvin, Iran.

Conflicts of Interest

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

Bidmeshkipour, A., Thengane, R., Bahagvat, M., Ghafari, S., and Rao, V. (2007). Production of haploid wheat via maize pollination. *J Sci I R Iran* 18(1): 5-11.

Brazauskas, G., Pašakinskienė, I., and Ruzgas, V. (2005). Improved approaches in wheat× maize crossing for wheat doubled haploid production. *Biologija* 51(4): 15-18.

Khan, M.A., and Ahmad, J. (2011). In vitro wheat haploid embryo production by wheat x maize cross system under different environmental conditions. *Pak J Agric Sci* 48(1): 49-53.

Khan, M.A., Shaukat, S., Ahmad, J., Kashif, M., Khan, A.S., and Iqbal, M.Z. (2012). Use of intergeneric cross for production of doubled haploid wheat (*Triticum aestivum* L.). *J Sci Technol* 31(4): 295-300.

Knox, R., Clarke, J., and DePauw, R. (2000). Dicamba and growth condition effects on doubled haploid production in durum wheat crossed with maize. *Plant Breed* 119(4): 289-298.

Mensuali‐Sodi, A., Panizza, M., and Tognoni, F. (1992). Quantification of ethylene losses in different container‐seal systems and comparison of biotic and abiotic contributions to ethylene accumulation in cultured tissues. *Physiol Plant* 84(3): 472-476.

Niu, Z., Jiang, A., Abu Hammad, W., Oladzadabbasabadi, A., Xu, S.S., Mergoum, M., and Elias, E.M. (2014). Review of doubled haploid production in durum and common wheat through wheat× maize hybridization. *Plant Breed* 133(3): 313-320.

Sharma, H., Yang, Y., and Ohm, H. (2002). An assessment of doubled haploid production in soft red winter wheat by wheat x corn wide crosses. *Cereal Res Commun* 30: 269-275.

Slama-Ayed, O., Bouhaouel, I., Ayed, S., De Buyser, J., Picard, E., and Amara, H.S. (2019). Efficiency of three haplomethods in durum wheat (*Triticum turgidum* subsp. durum Desf.): Isolated microspore culture, gynogenesis and wheat× maize crosses. *Czech J Genet Plant Breed* 55(3): 101.

Sourour, A., Olfa, S.-A., da Silva, J.A.T., and Hajer, S.-A. (2011). Effect of different factors on haploid production through embryo rescue in durum wheat× maize crosses. *Int J Plant Breed* 5(2): 118-121.

Sridevi, V., Giridhar, P., Simmi, P., and Ravishankar, G. (2010). Direct shoot organogenesis on hypocotyl explants with collar region from in vitro seedlings of *Coffea canephora* Pierre ex. Frohner cv. C× R and *Agrobacterium tumefaciens*-mediated transformation. *Plant Cell, Tissue Organ Cult* 101: 339-347.

Suenaga, K., and Nakajima, K. (1989). Efficient production of haploid wheat (*Triticum aestivum*) through crosses between Japanese wheat and maize (*Zea mays*). *Plant Cell Rep* 8: 263-266.

Upadhyaya, H., Begum, L., Dey, B., Nath, P., and Panda, S. (2017). Impact of calcium phosphate nanoparticles on rice plant. *J Plant Sci Phytopathol* 1(1): 001-010.

JPMB
Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

مقاله پژوهشی

بهبود کارایی القای گندم هاپلوئید از طریق روش تلاقی گندم × ذرت با استفاده از نیترات نقره و فسفات کلسیم

حامد مدیرروستا[1*]، راحله خادمیان [1]، رضا بزرگی‌پور[2]

[1]. گروه ژنتیک و به‌نژادی گیاهی، دانشکده کشاورزی و منابع طبیعی دانشگاه بین‌المللی امام خمینی (ره)، قزوین، ایران

[2]. موسسه اصلاح و تهیه نهال و بذر ، سازمان تحقیقات آموزش و ترویج کشاورزی، ایران

چکیده: این مطالعه با هدف بررسی اثر نیترات نقره و فسفات کلسیم در بهبود تولید گندم هاپلوئید با روش حذف کروموزوم در تلاقی گندم با ذرت انجام شد. در این مطالعه تأثیر تیمارهای مختلف، از جمله عدم استفاده (شاهد) و استفاده از نیترات نقره (۳ گرم در لیتر) همراه با فسفات کلسیم (۰.۱ گرم در لیتر) مورد بررسی قرار گرفتند، همچنین تیمارهای زمانی مختلف پس از گرده افشانی، شامل ۴۸ و ۷۲ ساعت، در نظر گرفته شد. در این مطالعه ویژگی‌های کلیدی مانند تعداد گلچه‌ها، همچنین تعداد و درصد تشکیل دانه مورد ارزیابی قرار گرفت. بین تیمارها تفاوت‌های قابل ملاحظه در سطح معنی‌دار ۱ درصد مشاهده شد. استفاده از نیترات نقره همراه با فسفات کلسیم در محیط نگهداری ساقه‌های گندم تولید بیشتر جنین هاپلوئید را به دنبال داشت. نگهداری ساقه‌های گندم در محلول نیترات نقره و فسفات کلسیم به مدت ۷۲ و ۴۸ ساعت تشکیل جنین هاپلوئید را با ۱۲.۵۳، ۵.۶۰ درصد نسبت به شاهد با ۳.۸۲ درصد بهبود بخشید.

کلمات کلیدی: هاپلوئید، حذف کروموزوم، هیبریداسیون بین گونه‌ای، نجات جنین، کشت بافت.

ویراستار علمی
دکتر اسماعیل بخشنده،
پژوهشکده ژنتیک و زیست‌فناوری کشاورزی طبرستان،
دانشگاه علوم کشاورزی و منابع طبیعی ساری، ایران

تاریخ
دریافت: ۲۴ تیر ۱۴۰۲
پذیرش: ۲۸ آذر ۱۴۰۲
چاپ: ۲۲ بهمن ۱۴۰۲

نویسنده مسئول
مهندس حامد مدیر روستا
hamed.modirrousta@ut.ac.ir

ارجاع به این مقاله
Modirrousta, H., Khademian, R., and Bozorgipoor, R. (2023). Enhancing haploid wheat induction efficiency through the wheat× maize crossbreeding system utilizing silver nitrate and calcium phosphate. *J. Plant Mol. Breed 11* (1): 54-62. doi:10.22058/jpmb.2023.2007085.1279.

JPMB | Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

Magnesium transporter family: sequence, evolution and expression analysis in soybean (*Glycine max* L.)

Parviz Heidari*, Bahar Saberi and Ariana Seifi

Faculty of Agriculture, Shahrood University of Technology, Shahrood 3619995161, Iran

Original paper

OPEN ACCESS

Edited by

Dr. S Hamidreza Hashemipetroudi, Genetics and Agricultural Biotechnology Institute of Tabarestan (GABIT), Sari Agricultural Sciences and Natural Resources University (SANRU), Iran

Date

Received: 10 January 2024
Accepted: 30 January 2024
Published: 3 February 2024

Correspondence

Dr. Parviz Heidari
heidarip@shahroodut.ac.ir

Citation

Heidari,P. Saberi, B. and Seifi A. (2023). Magnesium transporter family: sequence, evolution and expression analysis in soybean (*Glycine max* L.). *J. Plant Mol. Breed 11* (1): 63-74. doi:10.22058/JPMB.2024.2020070.1287.

Abstract: *Magnesium transporter (MGT)* genes play a critical role in plant growth, development, and stress responses. These genes involve Mg uptake, transport, and distribution within cells and organs. In this study, 48 MGT family members were screened from the soybean genome, and three subfamilies, including MRS2-like (18 members), CorA (3 members), and NIPA (22 members). The length of soybean MGTs ranged from 160 amino acids (aa) (GLYMA_06G214500) to 555 aa (GLYMA_15G125900). The results revealed that NIPA subfamily proteins have more genetic distance than MRS2-like and CorA proteins. In addition, the NIPA subfamily members showed a high variation in physiochemical properties such as theoretical isoelectric point (pI), grand average of hydropathicity (GRAVY), and instability index. All NIPAs were identified as hydrophobic proteins, while the MRS2-like and CorA members were predicted as hydrophilic. Moreover, the instability index revealed that NIPA members are more stable, while more MRS2-like and CorA proteins are predicted to be unstable. Additionally,, several duplication events were recognized among MGT family members, and all duplicated genes have been created through segmental duplication. The expression profile of soybean *MGT* genes showed significant differences in expression levels across various s. These results confirm that MGTs' widespread distribution across organs, regulating magnesium homeostasis.

Keywords: ion transporter; plant genome; gene family; physicochemical properties.

Introduction

Magnesium (Mg) is an essential cation that plays a crucial role in various physiological processes in plants, including photosynthesis, growth, and product quality (Chen et al., 2018). Mg deficiency can lead to leaf chlorosis and negatively impact plant growth and development (Lin and Nobel, 1971; Ko et al., 1999). Mg transporter (MGT) proteins facilitate the absorption, transport, and distribution of Mg in plants. MGT proteins are characterized by two conserved transmembrane domains in the C-terminal region and a Gly-Met-Asn (GMN) motif at the end of the first transmembrane domain (Lin and Nobel, 1971; Ko et al., 1999; Li et al., 2016; Chen et al., 2018; Heidari et al., 2021). The presence of the GMN motif is essential for the activity of MGT proteins (Palombo et al., 2013). Based on sequence structure, MGTs are classified into three subfamilies: CorA, MRS2-like, and NIPA (Heidari et al., 2022). While the CorA subfamily has been extensively studied, limited information is available on the functions of MGTs in the MRS2 and NIPA subfamilies. The CorA protein was firstly identified in *Salmonella typhimurium* (Silver, 1969), although its homologues have also been identified in plants, animals and yeast (Knoop et al., 2005). CorA proteins are involved in both efflux and influx of magnesium (Franken et al., 2022).

MGTs play diverse roles in plant growth and development. Some MGT proteins are involved in the distribution and accumulation of Mg within cells. For example, in *Arabidopsis thaliana*, AtMGT2 and AtMGT3 contribute to Mg accumulation in the vacuole (Lenz et al., 2013), while AtMGT10 regulates Mg homeostasis in chloroplasts (Drummond et al., 2006). It stated that AtMGT2 and AtMGT3 proteins increase Mg accumulation in the vacuole under high concentration of Mg and enable plant growth under low calcium concentration (Faraji et al., 2021).

MGT genes are expressed in different plant organs, such as the leaf, root, stem, and flower, to maintain optimal magnesium concentrations (Faraji et al., 2021). AtMGT5 and AtMGT9 are specifically involved in pollen growth and chloroplast development, respectively (Chen et al., 2009). AtMGT9 acts as a transporter that translocates Mg from roots to shoot parts in *Arabidopsis* (Gebert et al., 2009). In root tissues, MGTs like OsMGT1 in *Oryza sativa* and AtMGT6 in *Arabidopsis* are responsible for Mg uptake from the soil (Chen et al., 2012; Mao et al., 2014). Additionally, MGTs have been found to enhance plant tolerance to environmental stresses. For instance, OsMGT1 in rice is associated with salt stress tolerance (Chen et al., 2017), while the expression of MGT genes correlates with aluminum (Al) stress tolerance. Transgenic lines expressing AtMGT1 in *Nicotiana benthamiana* exhibited reduced aluminum toxicity (Deng et al., 2006).

MGT gene family has been identified and characterized in *O. sativa* (Mohamadi et al., 2023), *A. thaliana* (Schock et al., 2000), *Poncirus trifoliata* (Liu et al., 2019), *Zea mays* (Li et al., 2016), *Cucumis sativus* (Heidari et al., 2022), *Saccharum spontaneum* (Wang et al., 2019), and *Triticum turgidum* and *Camelina sativa* (Faraji et al., 2021), *Vitis vinifera* (Ge et al., 2022), *Gossypium hirsutum*, and *Theobroma cacao* (Heidari et al., 2021), and *Solanum lycopersicum* (Regon et al., 2019). Due to the importance of MGTs in regulating the growth and development of plants, the MGT gene family has not been identified and characterized in soybean (*Glycine max* L.). In the present study, MGT family members were analyzed based on sequence structure, evolution and expression in soybean.

Materials and Methods

Identification of MGT family members

The putative MGT proteins from the dicot model plant, *A. thaliana*, were used as a query in BLASTp tool of the Ensembl Plants database (https://plants.ensembl.org/index.html) against the soybean genome. All introduced proteins were checked to have the Mg transporter domain based on the output of the pfam tool in Ensembl Plants database. Finally, the peptide, CDS, and genomic sequence of confirmed genes were downloaded from the Ensembl Plants database.

Sequence analysis of MGTs

The Protparam tool (Gasteiger et al., 2005) was applied to predict the physiochemical properties of MGTs, including the grand average hydropathicity (GRAVY), instability index, theoretical isoelectric point (pI), and molecular weight (kDa).

Upstream analysis of MGT genes

To screen the putative cis-regulatory element related to response to stress, hormones and light, the upstream region, a 1500 bp fragment before start codon, of each MGT gene was analyzed using an online server, PlantCARE (Lescot et al., 2002). Finally, the identified elements were classified based on their function.

Phylogeny analysis

MGTs from soybean with along their orthologs in *Arabidopsis* and rice were aligned using an online multiple alignment tool, the Clustal Omega (Sievers and Higgins, 2014). Then, the phylogeny tree was constructed by Neighbor-Joining (NJ) method. Finally, tree was reconstructed using the iTOL tool (Letunic and Bork, 2021).

Chromosome location and duplication events in MGTs

The location of each *MGT* was pictured on soybean genome using the TBtools software v1.132 (Chen et al., 2020). To screen duplicated genes, the similarity of more than 70% in the coding sequence of MGT pairs was considered as a selection criterion. Then, the Ka (non-synonymous) and Ks (synonymous) indexes were calculated using TBtools between the duplicated MGT pairs.

Expression profile of MGTs

The expression profile of soybean MGTs in various soybean organs were extracted from the SoyBase database (https://soybase.org/soyseq/) based on the available RNA-seq information. The collected raw data were normalized and illustrated in the heatmaps.

Results

Physiochemical properties of MGTs

In this study, 48 MGT family members were screened from soybean genome (Table S1). The length of MGTs ranged from 160 aa (GLYMA_06G214500) to 555 aa (GLYMA_15G125900) in soybean, 241 aa to 540 aa in *Arabidopsis*, and 165 aa to 572 aa in rice (Table 1). The prediction of physicochemical characteristics of MGT proteins disclosed that isoelectric points (pI) of all soybean MGT proteins varied from 4.65 (GLYMA_10G126500) to 9.91 (GLYMA_06G053100), and ranged from 4.78 to 9.48 in *Arabidopsis*, and 4.61 to 11.43 in rice. Interestingly, instability index revealed that MGT proteins in the studied dicots, soybean and *Arabidopsis*, are more stable than the monocot model, rice (Table 1). In addition, the GRAVY (grand average of hydropathicity) index showed that MGTs are both hydrophobic and hydrophilic. Overall, these results suggested that MGT family members are diverse in term of physiochemical properties.

Phylogeny analysis

Phylogenetic analysis determined relationships between 43 soybean MGTs and orthologs in *A. thaliana* and *O. sativa*. MGT family members were separated into three subfamilies: NIPA, MRS2-like and CorA (Figure 1).

Accordingly, 22 NIPs, 18 MRS2-like proteins and 3 CorA proteins were identified in soybean. The results revealed that NIPA subfamily proteins showed more distance than MRS2-like and CorA proteins, indicating that NIPAs have had an evolutionary divergence and that CorAs and MRS2-like proteins probably had the same ancestral group.

Structure and physiochemical comparison between the subfamilies of MGTs

MGT subfamilies in soybean, including NIPA, CorA, and MRS2-like, were compared based on exon number, pI, GRAVY, and stability (Figure 2). The results showed significant differences between the members of these subfamilies.

Table 1. Comparison of the physicochemical characteristics of MGTs in soybean with monocot and dicot model plants. Full details are provided in the Supplementary Table 1.

Plant	Length (aa)	MW (KDa)	pI	GRAVY	Instability
Soybean	160-555	17.85-62.27	4.65-9.91	-0.464-1.057	49% stable
Arabidopsis	241-540	26.56-61.05	4.78-9.48	-0.369-0.869	53% stable
Rice	165-572	17.67-62.91	4.61-11.43	-0.451-0.789	27% stable

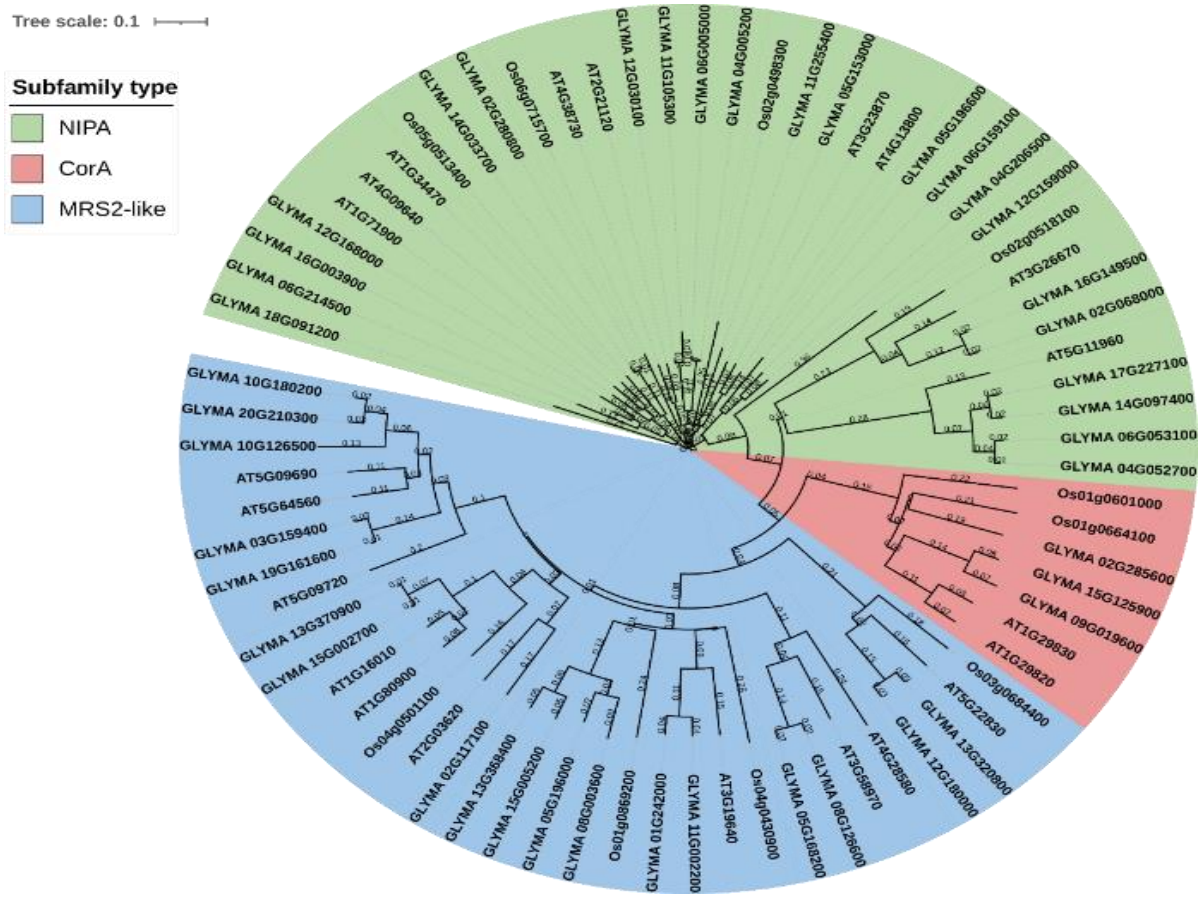


Figure 1. Evolutionary analysis of MGT family members.

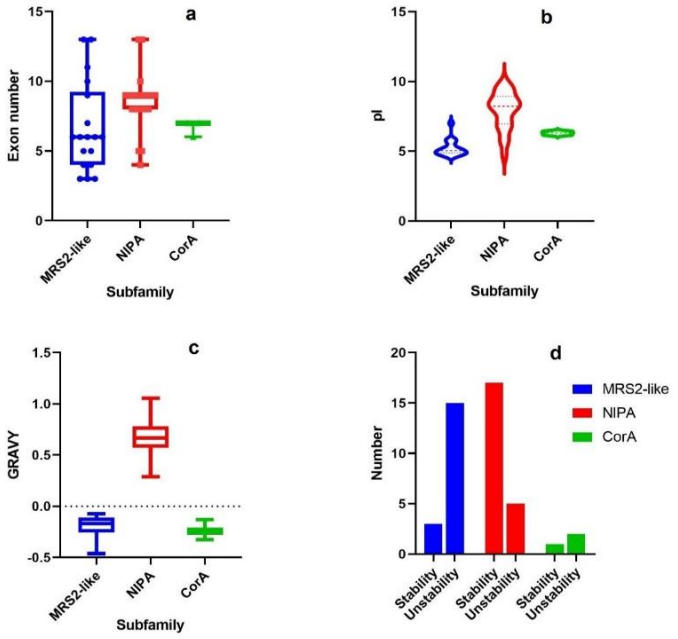


Figure 2. Comparison between the subfamilies of MGTs in terms of exon number (a), pI (b), GRAVY (c), and instability index (d).

The number of exons in the NIPA subfamily was higher than that of the other subfamilies, although the members of the MRS2-like subfamily had significant variation based on the number of exons together (Figure 2a). In addition, the NIPA subfamily members showed a high variation for pI, and more NIPAs were predicted activating under alkaline conditions (Figure 2b). In contrast, it was predicted that the CorA and MRS2-like subfamily members are more active in acidic environments. Based on GRAVY index, all NIPA proteins were predicted to be hydrophobic, but the MRS2-like and CorA members were identified as hydrophilic (Figure 2c). Besides, the instability index disclosed that more NIPA members are stable, while more MRS2-like and CorA proteins are predicted as unstable proteins (Figure 2c). These results indicated that NIPA subfamily members have noticeable structural differences from other MGT subfamilies and have probably undergone a different evolutionary process.

Location and duplication events in MGT family members

Chromosomal localization of genes can provide information related to genetic linkage, regulation and evolutionary history. In the present study, soybean *MGT* genes were located on 19 chromosomes, and an uneven distribution of *MGT* genes was observed (Figure 3). Only in soybean chromosome 7, there was no *MGT* gene, while four *MGT* genes were located in each of chromosomes 2, 5, 6, and 12. Duplication events of *MGTs* were predicted based on the similarity percentage between pair genes. Results revealed seven duplications between pairs of *MGT* genes (Table 2). Besides, all MGT genes were predicted to have been segmentally duplicated in soybean genome. Based on Ka/Ks ratio, six pairs of duplicated genes were under purifying selection, and a pair of duplicated *NIPA* genes (*GLYMA_12G168000-GLYMA_18G091200*) was under positive selection.

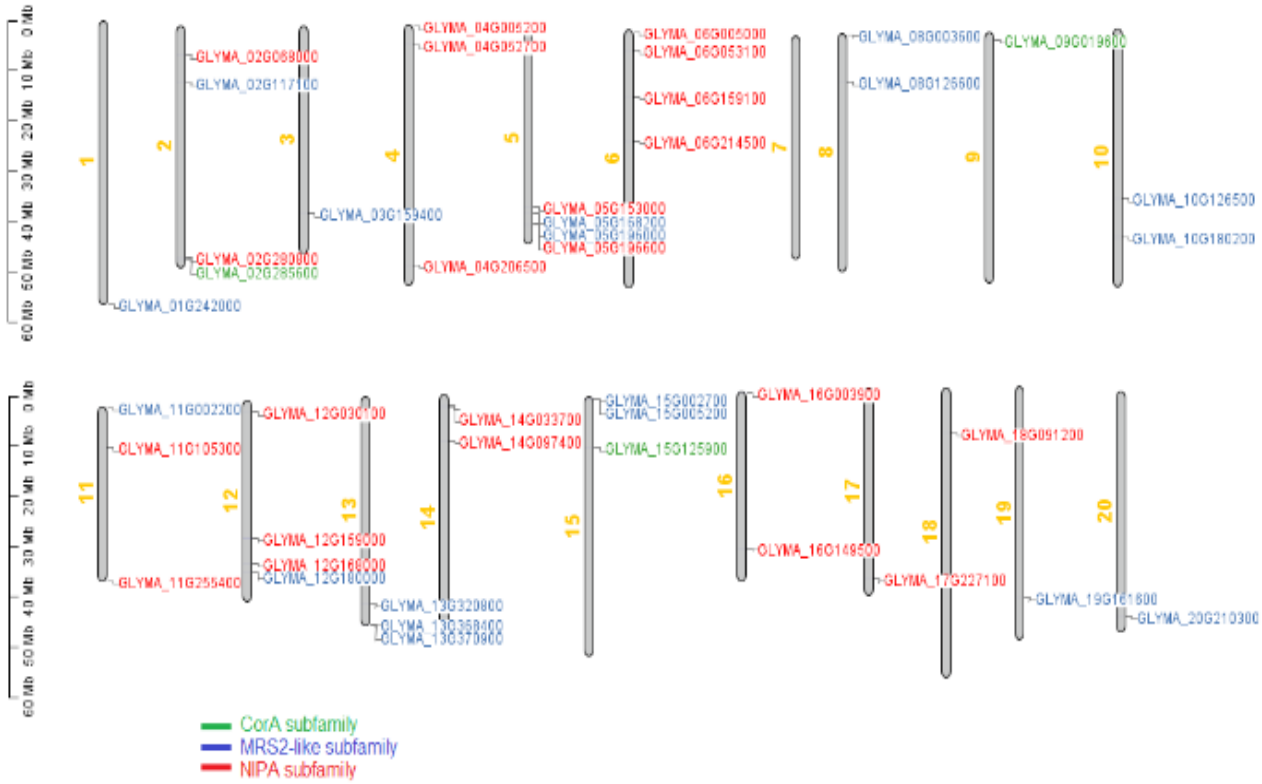


Figure 3. Distribution of *MGTs* on chromosomes of Soybean (*Glycine max*).

Table 2. List of duplicated *MGT* genes in soybean.

Gene 1	Gene 2	Subfamily	Ka	Ks	Ka/Ks	Duplication type
GLYMA_05G196000	GLYMA_08G003600	MRS2-MRS2	0.0817	0.1769	0.4618	Segmental
GLYMA_05G196000	GLYMA_15G005200	MRS2-MRS2	0.1197	0.2425	0.4936	Segmental
GLYMA_06G214500	GLYMA_16G003900	NIPA- NIPA	0.1253	0.1636	0.7659	Segmental
GLYMA_10G126500	GLYMA_10G180200	MRS2-MRS2	0.1089	0.2107	0.5168	Segmental
GLYMA_10G126500	GLYMA_20G210300	MRS2-MRS2	0.1010	0.1194	0.8459	Segmental
GLYMA_12G168000	GLYMA_16G003900	NIPA- NIPA	0.0102	0.0659	0.1548	Segmental
GLYMA_12G168000	GLYMA_18G091200	NIPA- NIPA	0.1654	0.1389	1.1908	Segmental

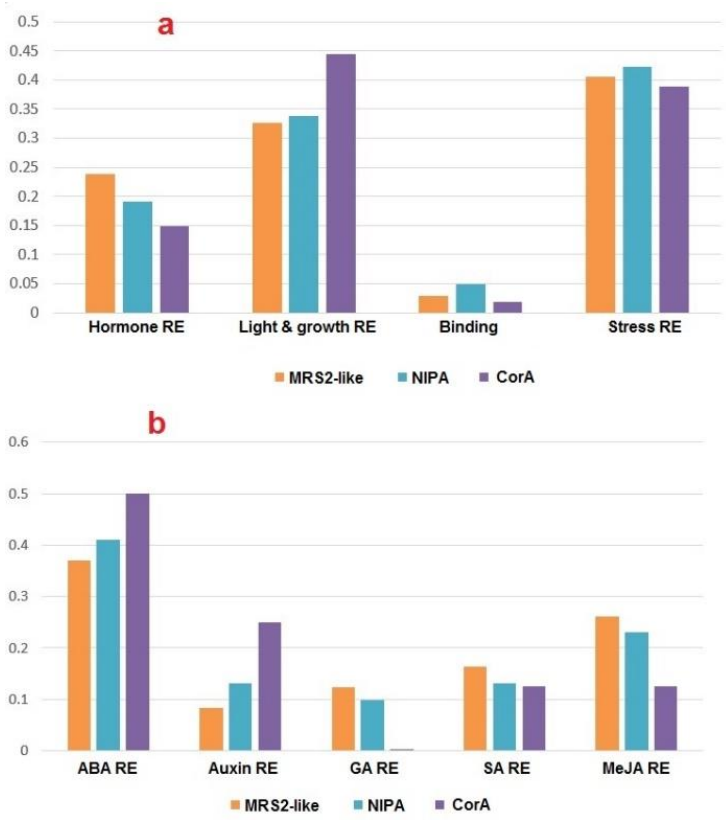


Figure 4. Frequency of cis-regulatory elements based on their function (a) in subfamilies of soybean *MGTs*. Frequency of cis-regulatory elements related to hormone responsiveness in subfamilies of soybean MGTs (b).

Overall, it seems that segmental duplication events have extended the *MGT* family members in soybean.

Promoter analysis of MGTs

In order to better understand the regulatory system in soybean *MGT* genes, the upstream region of these genes was screened. The identified *cis*-regulatory elements were classified based on their functions in four classes: hormone responsiveness elements (Res), light and growth Res, protein binding site, and stress Res (Figure 4a). The results disclosed that soybean *MGT* genes are involved in various processes and play roles in response to phytohormones and stress conditions as well as response to light and regulation of plant growth. *Cis*-regulatory elements-related stresses were

recognized more in the upstream site of *MGTs*. However, the distribution of cis-elements was different among the MGT subfamilies. For example, hormone-responsive elements were observed more in the promoter of *MRS2-like* genes than those of *NIPA a*nd *CorA* genes. In addition, *cis*-regulatory elements related to ABA responsiveness were more recognized in upstream regions of soybean *MGTs* (Figure 4b). Elements linked with GA, MeJA, and SA responsiveness were more detected in *MRS2-like* genes. These results disclosed that soybean *MGTs* are controlled by upstream signals related to hormones and stimuli.

Expression profile of MGTs

Expression profile of soybean *MGT* genes was illustrated based on the RNA-seq data in different tissues (Figure 5). *MGT* genes showed significant differences in expression levels in different soybean organs. In root tissues, one *CorA* gene (*GLYMA_09G019600*), two *MRS2-like* genes (*GLYMA_11G002200* and *GLYMA_13G370900*) and six *NIPA* genes (*GLYMA_04G052700, GLYMA_06G005000, GLYMA_06G214500, GLYMA_20G210300* and *GLYMA_11G105300*) were more expressed, indicating that these genes are involved in the absorption of magnesium.

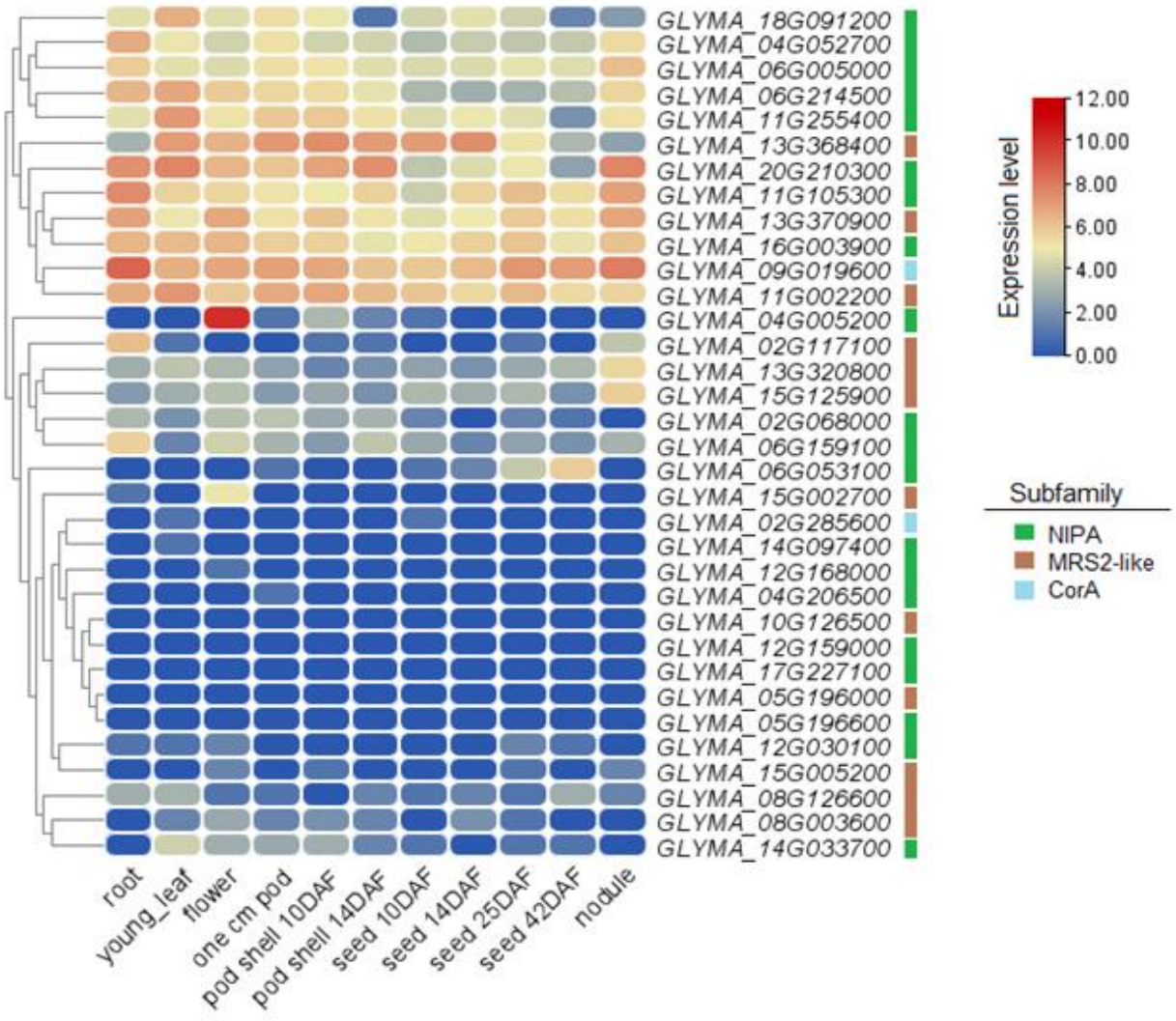


Figure 5. Expression profile of *MGT* genes in different organs of *G. max*. DAF: days after flowering.

In young leaf, four *NIPA* genes including *GLYMA_20G210300*, *GLYMA_06G214500*, *GLYMA_11G255400*, and *GLYMA_18G091200* and two *MRS2-like* genes, *GLYMA_13G368400* and *GLYMA_11G002200* showed high expression levels and *GLYMA_04G005200* as a *NIPA* gene sharply expressed in flower tissues. In pod and pod shell, two *MRS2-like* genes including *GLYMA_13G368400*, and *GLYMA_11G002200* and a *CorA* gene (*GLYMA_09G019600*) and a *NIPA* gene (*GLYMA_20G210300*) were more induced. *GLYMA_13G368400*, as a *MRS2-like* gene, and *GLYMA_09G019600*, as a *CorA* gene, were more expressed in young seed tissues (10 to 14 days after flowering (DAF)) and old seed tissues (25-42 DAF), respectively. In addition, *GLYMA_09G019600*, as a *CorA* gene, *GLYMA_20G210300* and *GLYMA_11G105300*, as *NIPA* genes, and *GLYMA_13G370900*, as a *MRS2-like* gene were sharply expressed in nodule tissues.

Discussion

In the current study, 48 MGT family members were identified from the soybean genome. The number of MGT family genes varies among plant species; for instance, 8 MGTs in *Poncirus trifoliata* (Liu et al., 2019), 12 MGTs in *Z. mays* (Li et al., 2016), 20 MGTs in *C. sativus* (Heidari et al., 2022), 10 MGTs in *S. spontaneum* (Wang et al., 2019), 41 MGTs in *T. turgidum* and 61 MGTs in *C. sativa* (Faraji et al., 2021), 9 MGTs in *V. vinifera* (Ge et al., 2022), and 18 MGTs in *T. cacao* (Heidari et al., 2021). The polyploidy level and duplication rate have more affected the number of members of plant gene family (Faraji et al., 2020). The identified soybean-MGTs showed diversity in terms of physicochemical properties. Differences in physicochemical properties can affect the activity conditions and performance of plant transporters (Heidari et al., 2022). According to the role and distribution of MGTs, it can be suggested that the difference in physicochemical properties affects the differentiation of their functions. Based on the phylogenetic results, three main subfamilies including NIPs, MRS2-like and CorA were recognized that NIPA subfamily proteins showed more distance than MRS2-like and CorA proteins, indicating that NIPAs have had an evolutionary divergence and that CorAs and MRS2-like proteins probably had the same ancestral group. Previous studies have also emphasized that NIPA subfamily has significant differences compared to MRS2-like and CorA subfamilies (Heidari et al., 2021). Based on phylogeny results, it seems that the expansion and diversity occurred more in the members of MGT gene family after the derivation of monocots and dicots. In addition MGT subfamilies showed diversity in term of number of exons. The soybean-MGTs /introns affects the duration of the editing process of primary mRNAs (Hashemipetroudi et al., 2023; Puresmaeli et al., 2023). Besides, the NIPA subfamily members showed a high variation for pI, GRAVY index, and instability index, than MRS2-like and CorA members. These results indicated that NIPA subfamily members have noticeable structural differences from other MGT subfamilies and have probably undergone a different evolutionary process. Several duplication events were predicted in soybean-MGTs. It seems that segmental duplication events have extended the MGT family members in soybean. The expression profile of soybean *MGT* genes disclosed that MGTs are expressed in all organs. These results confirm that MGTs have been distributed in all organs for magnesium distribution and homeostasis regulation of this key element. In addition, few differences were observed in the expression level of duplicated genes, possibly due to the mutations in the genes' regulatory or coding regions (Arab et al., 2023; Hashemipetroudi et al., 2023). Based on upstream analysis, it found that soybean MGTs are involved in various processes related to phytohormones, adverse conditions and growth. Magnesium is involved in various cellular processes. The distribution of these genes in different organs shows the importance of magnesium distribution in all tissues. On the other hand, the presence of *cis*-regulatory elements related to hormones shows that the activity of this transporter is controlled by hormone-dependent signaling pathways

Conclusion

Understanding the functions and regulation of *MGT* genes in plants provides valuable insights into plant growth, development, and stress responses. This study screened 48 MGT family members from the soybean genome, showing diverse

physiochemical properties, promoter structure, and expression levels. NIPA subfamily members were predicted as a different subfamily transporter of magnesium with unique characteristics. It was found that segmental duplication events have extended the MGT family members in soybean. In addition, we conclude that *MGTs* are induced by upstream signals linked to phytohormones and stresses to control the Mg concentration in different organs. Further research into the function of MGT genes and their role in oil production of soybean is suggested.

Supplementary Materials

The Supplementary Material for this article can be found online at: https://www.jpmb-gabit.ir/article_710571.html .

Supplementary Table 1. Physicochemical characteristics of MGTs in soybean, rice and arabidopsis.

Not true?! Lines 131 & 174 refer to Table S1.

Author Contributions

Conceptualization, P.H. and B.S.; methodology, B.S.; software, B.S. and A.S; validation, P.H.; formal analysis, B.S. and A.S; writing—original draft preparation, P.H.; writing—review and editing, P.H.; project administration, P.H. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Acknowledgments

Not applicable.

Conflicts of Interest

The authors declare no conflict of interest.

References

Arab, M., Najafi Zarrini, H., Nematzadeh, G., Heidari, P., Hashemipetroudi, S.H., and Kuhlmann, M. (2023). Comprehensive analysis of calcium sensor families, CBL and CIPK, in *Aeluropus littoralis* and their expression profile in response to salinity. *Genes* 14(3): 753.

Chen, C., Chen, H., Zhang, Y., Thomas, H.R., Frank, M.H., He, Y., and Xia, R. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant* 13(8): 1194-1202. doi: 10.1016/j.molp.2020.06.009.

Chen, J., Li, L.-g., Liu, Z.-h., Yuan, Y.-j., Guo, L.-l., Mao, D.-d., Tian, L.-f., Chen, L.-b., Luan, S., and Li, D.-p. (2009). Magnesium transporter AtMGT9 is essential for pollen development in *Arabidopsis*. *Cell Res* 19(7): 887-898.

Chen, Z.C., Peng, W.T., Li, J., and Liao, H. (2018). Functional dissection and transport mechanism of magnesium in plants. *Semin Cell Dev Biol* 74: 142-152.

Chen, Z.C., Yamaji, N., Horie, T., Che, J., Li, J., An, G., and Ma, J.F. (2017). A Magnesium transporter OsMGT1 plays a critical role in salt tolerance in rice. *Plant Physiol* 174(3): 1837-1849. doi: 10.1104/pp.17.00532.

Chen, Z.C., Yamaji, N., Motoyama, R., Nagamura, Y., and Ma, J.F. (2012). Up-regulation of a magnesium transporter gene OsMGT1 is required for conferring aluminum tolerance in rice. *Plant Physiol* 159(4): 1624-1633.

Deng, W., Luo, K., Li, D., Zheng, X., Wei, X., Smith, W., Thammina, C., Lu, L., Li, Y., and Pei, Y. (2006). Overexpression of an *Arabidopsis* magnesium transport gene, AtMGT1, in *Nicotiana benthamiana* confers Al tolerance. *J Exp Bot* 57(15): 4235-4243. doi: 10.1093/jxb/erl201.

Drummond, R., Tutone, A., Li, Y.-C., and Gardner, R. (2006). A putative magnesium transporter AtMRS2-11 is localized to the plant chloroplast envelope membrane system. *Plant Sci* 170(1): 78-89.

Faraji, S., Ahmadizadeh, M., and Heidari, P. (2021). Genome-wide comparative analysis of Mg transporter gene family between *Triticum turgidum* and *Camelina sativa*. *BioMetals* 34: 639-660.

Faraji, S., Filiz, E., Kazemitabar, S.K., Vannozzi, A., Palumbo, F., Barcaccia, G., and Heidari, P. (2020). The ap2/erf gene family in *Triticum durum*: genome-wide identification and expression analysis under drought and salinity stresses. *Genes (Basel)* 11(12): 1464. doi: 10.3390/genes11121464.

Franken, G.A.C., Huynen, M.A., Martinez-Cruz, L.A., Bindels, R.J.M., and de Baaij, J.H.F. (2022). Structural and functional comparison of magnesium transporters throughout evolution. *Cell Mol Life Sci* 79(8): 418. doi: 10.1007/s00018-022-04442-8.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S.e., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005). *Protein identification and analysis tools on the ExPASy server.* Springer.

Ge, M., Zhong, R., Sadeghnezhad, E., Hakeem, A., Xiao, X., Wang, P., and Fang, J. (2022). Genome-wide identification and expression analysis of magnesium transporter gene family in grape (*Vitis vinifera*). *BMC Plant Biol* 22(1): 217. doi: 10.1186/s12870-022-03599-5.

Gebert, M., Meschenmoser, K., Svidova, S., Weghuber, J., Schweyen, R., Eifler, K., Lenz, H., Weyand, K., and Knoop, V. (2009). A root-expressed magnesium transporter of the MRS2/MGT gene family in A*rabidopsis thaliana* allows for growth in low-Mg2+ environments. *Plant Cell* 21(12): 4018-4030. doi: 10.1105/tpc.109.070557.

Hashemipetroudi, S.H., Arab, M., Heidari, P., and Kuhlmann, M. (2023). Genome-wide analysis of the laccase (LAC) gene family in *Aeluropus littoralis*: A focus on identification, evolution and expression patterns in response to abiotic stresses and ABA treatment. *Front Plant Sci* 14: 1112354. doi: 10.3389/fpls.2023.1112354.

Heidari, P., Abdullah, Faraji, S., and Poczai, P. (2021). Magnesium transporter gene family: genome-wide identification and characterization in *Theobroma cacao*, *Corchorus capsularis*, and *Gossypium hirsutum* of family *Malvaceae*. *Agronomy* 11(8): 1651.

Heidari, P., Puresmaeli, F., and Mora-Poblete, F. (2022). Genome-wide identification and molecular evolution of the magnesium transporter (MGT) gene family in *Citrullus lanatus* and *Cucumis sativus*. *Agronomy* 12(10): 2253.

Knoop, V., Groth-Malonek, M., Gebert, M., Eifler, K., and Weyand, K. (2005). Transport of magnesium and other divalent cations: evolution of the 2-TM-GxN proteins in the MIT superfamily. *Mol Genet Genom* 274(3): 205-216. doi: 10.1007/s00438-005-0011-x.

Ko, Y.H., Hong, S., and Pedersen, P.L. (1999). Chemical mechanism of ATP synthase: magnesium plays a pivotal role in formation of the transition state where ATP is synthesized from ADP and inorganic phosphate. *J Biol Chem* 274(41): 28853-28856.

Lenz, H., Dombinov, V., Dreistein, J., Reinhard, M.R., Gebert, M., and Knoop, V. (2013). Magnesium deficiency phenotypes upon multiple knockout of *Arabidopsis thaliana* MRS2 clade B genes can be ameliorated by concomitantly reduced calcium supply. *Plant Cell Physiol* 54(7): 1118-1131.

Lescot, M., Dehais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouze, P., and Rombauts, S. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res* 30(1): 325-327. doi: 10.1093/nar/30.1.325.

Letunic, I., and Bork, P. (2021). Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 49(W1): W293-W296. doi: 10.1093/nar/gkab301.

Li, H., Du, H., Huang, K., Chen, X., Liu, T., Gao, S., Liu, H., Tang, Q., Rong, T., and Zhang, S. (2016). Identification, and functional and expression analyses of the CorA/MRS2/MGT-Type magnesium transporter family in Maize. *Plant Cell Physiol* 57(6): 1153-1168. doi: 10.1093/pcp/pcw064.

Lin, D.C., and Nobel, P.S. (1971). Control of photosynthesis by Mg2+. *Archives of biochemistry and biophysics* 145(2): 622-632.

Liu, X., Guo, L.X., Luo, L.J., Liu, Y.Z., and Peng, S.A. (2019). Identification of the magnesium transport (MGT) family in *Poncirus trifoliata* and functional characterization of PtrMGT5 in magnesium deficiency stress. *Plant Mol Biol* 101(6): 551-560. doi: 10.1007/s11103-019-00924-9.

Mao, D., Chen, J., Tian, L., Liu, Z., Yang, L., Tang, R., Li, J., Lu, C., Yang, Y., and Shi, J. (2014). Arabidopsis transporter MGT6 mediates magnesium uptake and is required for growth under magnesium limitation. *The Plant Cell* 26(5): 2234-2248.

Mohamadi, S.F., Babaeian Jelodar, N., Bagheri, N., Nematzadeh, G., and Hashemipetroudi, S.H. (2023). New insights into comprehensive analysis of magnesium transporter (MGT) gene family in rice (*Oryza sativa* L.). *3 Biotech* 13(10): 322.

Palombo, I., Daley, D.O., and Rapp, M. (2013). Why is the GMN motif conserved in the CorA/Mrs2/Alr1 superfamily of magnesium transport proteins? *Biochem* 52(28): 4842-4847.

Puresmaeli, F., Heidari, P., and Lawson, S. (2023). Insights into the sulfate transporter gene family and its expression patterns in durum wheat seedlings under salinity. *Genes* 14(2): 333.

Regon, P., Chowra, U., Awasthi, J.P., Borgohain, P., and Panda, S.K. (2019). Genome-wide analysis of magnesium transporter genes in *Solanum lycopersicum*. *Comput Biol Chem* 80: 498-511. doi: 10.1016/j.compbiolchem.2019.05.014.

Schock, I., Gregan, J., Steinhauser, S., Schweyen, R., Brennicke, A., and Knoop, V. (2000). A member of a novel *Arabidopsis thaliana* gene family of candidate Mg2+ ion transporters complements a yeast mitochondrial group II intron-splicing mutant. *Plant J* 24(4): 489-501. doi: 10.1046/j.1365-313x.2000.00895.x.

Sievers, F., and Higgins, D.G. (2014). Clustal omega. *Curr Protoc Bioinformatics* 48(1): 3.13. 11-13.13. 16.

Silver, S. (1969). Active transport of magnesium in escherichia coli. *Proc Natl Acad Sci U S A* 62(3): 764-771. doi: 10.1073/pnas.62.3.764.

Wang, Y., Hua, X., Xu, J., Chen, Z., Fan, T., Zeng, Z., Wang, H., Hour, A.L., Yu, Q., Ming, R., and Zhang, J. (2019). Comparative genomics revealed the gene evolution and functional divergence of magnesium transporter families in *Saccharum*. *BMC Genom* 20(1): 83. doi: 10.1186/s12864-019-5437-3.

JPMB
Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

مقاله پژوهشی

خانواده انتقال‌دهنده منیزیم: تجزیه و تحلیل توالی، تکامل و بیان در سویا (*Glycine max* L.)

پرویز حیدری*، بهار صابری، آریانا سیفی

دانشکده کشاورزی، دانشگاه صنعتی شاهرود، شاهرود، ایران

چکیده: ژن‌های انتقال دهنده منیزیم (*MGT*)، نقش مهمی در رشد، نمو و پاسخ به تنش گیاه دارند. این ژن‌ها در جذب، انتقال و توزیع منیزیم در سلول‌ها و اندام‌ها نقش دارند. در این مطالعه، ۴۸ عضو خانواده ژنی *MGT* از ژنوم سویا شناسایی شدند و به سه زیر خانواده شامل MRS2 (با ۱۸ عضو)، CorA (با ۳ عضو) و NIPA (با 22 عضو) تفکیک شدند. طول پروتئین‌های MGTs در سویا از ۱۶۰ (GLYMA_06G214500) تا ۵۵۵ اسیدآمینه (GLYMA_15G125900) متغیر بود. نتایج نشان داد که پروتئین‌های زیرخانواده NIPA فاصله ژنتیکی بیشتری نسبت به پروتئین‌های MRS2 و CorA دارند. علاوه بر این، اعضای زیرخانواده NIPA تنوع بالایی در خواص فیزیکوشیمیایی مانند pI ، GRAVY و شاخص ناپایداری نشان دادند. تمام NIPA ها به عنوان پروتئین‌های آبگریز شناسایی شدند، در حالی که اعضای MRS2 و CorA به عنوان آبدوست پیش‌بینی شدند. علاوه بر این، شاخص بی‌ثباتی نشان داد که اعضای NIPA پایدارتر هستند، در حالی که پروتئین‌های MRS2 و CorA بیشتر به عنوان پروتئین‌های ناپایدار پیش‌بینی شدند. علاوه بر این، چندین رویداد دوبرابر شدگی بین اعضای خانواده *MGT* شناسایی شد که احتمالا همه ژن‌های تکراری به‌صورت قطعه‌ای ایجاد شده‌اند. پروفایل بیانی ژن‌های MGT در سویا تفاوت‌های معنی‌داری را در اندام های مختلف نشان داد. این نتایج تایید کرد که MGT ها در تمام اندام‌ها برای کنترل هموستازی منیزیم توزیع شده‌اند.

کلمات کلیدی: انتقال دهنده‌های یونی، ژنوم گیاهی، خانواده ژنی، خواص فیزیکوشیمیایی.

ویراستار علمی
دکتر سیدحمیدرضا هاشمی‌پطرودی،
پژوهشکده ژنتیک و زیست‌فناوری کشاورزی طبرستان،
دانشگاه علوم کشاورزی و منابع طبیعی ساری

تاریخ
دریافت: ۲۰ دی ۱۴۰۲
پذیرش: ۱۰ بهمن ۱۴۰۲
چاپ: ۱۴ بهمن ۱۴۰۲

نویسنده مسئول
دکتر پرویز حیدری
heidarip@shahroodut.ac.ir

ارجاع به این مقاله
Heidari,P. Saberi, B. and Seifi A. (2023). Magnesium transporter family: sequence, evolution and expression analysis in soybean (*Glycine max* L.). *J. Plant Mol. Breed 11* (1): 63-74. doi:10.22058/JPMB.2024.2020070.1287.

JPMB | Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

Association mapping of morpho-physiological traits in barley (*Hordeum vulgare* L.) under salinity stress

Mahdiyeh Zare-Kohan *[1], Nadali Babaeian Jelodar [1], Reza Aghnoum [2], Seyed Ali Tabatabaee [3] and Mohammadreza Ghasemi-Nezhadraeini[4]

1. Department of Plant Breeding and Biotechnology, Sari Agricultural Sciences and Natural Resources University (SANRU), Sari, Iran
2. Seed and Plant Improvement Research Department, Khorasan Razavi Agricultural and Natural Resources Research and Education Center, AREEO, Mashhad, Iran
3. Seed and Plant Improvement Research Department, Yazd Agricultural and Natural Resources Research and Education Center, AREEO, Yazd, Iran
4. Department of Water Engineering, Islamic Azad University, Kerman Branch, Iran

Original paper

OPEN ACCESS

Edited by

Prof. Ahmad Arzani,
Isfahan University of Technology, Iran

Date

Received: 24 May 2023
Accepted: 02 February 2024
Published: 15 February 2024

Correspondence

Mahdiyeh Zare-Kohan
mahdiyehzare65@gmail.com

Citation

Zare-Kohan, M., Babaeian Jelodar, N., Aghnoum, R., Tabatabaee, S. A., and Ghasemi-Nezhadraeini, M. (2023). Association mapping of morpho-physiological traits in barley (*Hordeum vulgare* L.) under salinity stress. *J Plant Mol Breed 11* (1): 75-88. doi:10.22058/JPMB.2024.2003151.1274.

Abstract: In this study molecular markers associated with morpho-physiological traits were identified using 14 AFLP primer combinations and 32 SSRs primer pairs across a cohort of 148 barley cultivars employing the association mapping approach. Phenotypic analysis was carried out using an alpha-lattice design with five incomplete blocks replicated twice under normal and salinity stress conditions (EC = 12 dS m^{-1}) in two growing seasons. Population genetic structure was divided into two subpopulations (K = 2). In the present association panel, the mean of D´and r^2, indicators for linkage disequilibrium (LD) were estimated at 0.25 and 0.02, respectively. The mixed linear model identified 194 significant marker-trait associations for nine studied traits under normal and salinity stress conditions. Several quantitative trait loci (QTLs) were stable for plant height, number of grains per spike, grain weight per spike, and leaf proline content traits under each of the environmental conditions, and termed stable QTLs. In addition, some stable QTLs were common to several traits and thereby enable barley breeder to undertake a concurrent selection of multiple traits to develop high-yielding cultivars. The identified markers could be useful in the implementation of marker-assisted selection in barley to improve the efficiency of selecting genotypes for salinity tolerance.

Keywords: association mapping, barley, linkage disequilibrium, mixed linear model, salinity stress, stable QTLs.

Introduction

Barley (*Hordeum vulgare* L.) belongs to the cereal group of Gramineae family. It today ranks fourth in importance after wheat, rice, and corn. Barley serves as a model cereal for studying mechanisms of salinity tolerance due to its simpler genome than wheat and its notably higher salinity tolerance compared to wheat and rice (Gharaghanipor et al., 2022). Salinity stress poses a significant threat to agricultural production worldwide, exacerbated by climate change, salt intrusion into irrigation from surface and groundwater sources, and depletion of genetic resources (Arzani and Ashraf, 2016). Ellis et al. (2000) and Kilian et al. (2006) motioned that the new barley cultivars contain only 15-40% of the alleles in the barley gene pool. Thus, a part of barley's gene pool is tapped by breeders to improve salinity tolerance.

Salinity tolerance in crops is a quantitative trait with complex genetic and physiological architectures controlled by many gene loci (Flowers, 2004; Arzani, 2008; Omrani et al., 2022). With the advent of biotechnological tools such as molecular markers and transformation techniques, the science of plant breeding has evolved into a new realm (Arzani and Ashraf, 2016). The two most common methods for identifying and locating quantitative trait locus (QTL) are linkage mapping and association mapping (Flint - Garcia et al., 2005). In association mapping, QTL identification performs in a general population instead of a specific and segregating population (Zhu et al., 2008). It has advantages over linkage mapping, including examining more alleles and saving time and money because there is no need to create two-parent populations. Another advantage of association mapping is its high accuracy; due to many recombinations during the Ancestorl pedigree, genetic mapping has a high resolution and can easily use in marker-assisted selection (Flint-Garcia et al., 2003; Moose and Mumm, 2008). Therefore, this method avoids the disadvantages and limitations associated with linkage mapping. Association mapping does by the general linear model (GLM) and mixed linear model (MLM) methods. In the MLM method, population structure (Q-matrix) and kinship relationships between individuals (K-matrix) are predicted using several markers and used as a covariate in the model. Therefore, this method minimizes the results of false marker-trait associations. Fan et al. (2016) experiment showed that 206 barley genotypes with 408 markers were genotyped and tested for salinity stress tolerance. In their study, association analysis was performed by both GLM and MLM models based on population structure and kinship relationships. Finally, 24 markers that were highly associated with traits were identified.

Irrigated agricultural lands in arid and semi-arid regions contribute to the accumulation of soluble salts and exchangeable sodium in the soil where the roots grow (Arzani and Ashraf, 2016). Salinity imposes primary stresses such as osmotic stress and specific ion toxicity (predominantly from Na^+ and Cl^-); as well as secondary stresses like nutritional disorder and oxidative stress (Arzani, 2008). These stresses ultimately impair plant growth and development. Eleuch et al. (2008) experimented in two different environments (Egypt and India) to investigate the barley's genetic diversity and association analysis of salinity tolerance. Their study evaluated traits using 22 SSR markers and 48 barley genotypes. Their results showed that some QTLs were identified as responsible for salinity tolerance in each experimental environment, but only a small number of QTLs were identified in both environments. Also, Inostroza et al. (2009), El-Denary et al. (2012), Long et al. (2013), Sbei et al. (2014), Elakhdar et al. (2016a), and Elakhdar et al. (2016b) used association mapping under salinity stress in barley.

This study aimed to analyze the population structure of barley germplasm cultivars and investigate the relationship between AFLP and SSR markers and morpho-physiological traits of barley under salinity stress conditions. Breeding stress-tolerant barley cultivars is a complex and time-consuming activity. Therefore, introducing markers associated with these traits can facilitate marker-assisted selection in barley breeding programs.

Materials and Methods

Plant material (germplasm)

This study used 148 modern European two-row spring barley cultivars (Supplementary Table 1), representing commercial germplasm used all over

North and West Europe (Kraakman et al., 2006). The seeds of the association panel were received from the Khorasan Razavi Agricultural and Natural Resources Research and Education Center.

Phenotyping

An alpha-lattice design with five incomplete blocks replicated twice was used. Each block includes 30 plots in normal (water EC 2, soil EC 3.4 dS m^{-1}) and salinity stress (water EC 12, soil EC 14 dS m^{-1}) environments at the Agriculture and Natural Resources Research Station of Yazd (31° 55′ N, 54° 16′ E, 1213 m of sea level), Iran, for the two years. Salinity treatment was applied with water. The field soil in this experiment was naturally saline. Soil salinity was measured regularly during the growth period. The soil salinity was kept constant in each plot at the desired treatment level through the amount of water used and the need for soil leaching. The studied traits include Plant height (PH), Thousand-grain weight (TGW), Harvest index (HI), number of grains per spike (NGS), Grain weight per spike (GWS), number of total tillers (NTT), Relative water content (RWC), leaf proline content (LPC), and leaf chlorophyll content (LChC). The data normality test was first performed based on the Kolmogorov-Smirnov method using SPSS software. Then, a combined analysis of variance was performed with SAS 9.1 software.

Genotyping

In this study, a genetic map of molecular markers, including 407 AFLP and SSR markers, was prepared by Kraakman et al. (2006), and Aghnoum et al. (Unpublished data) were used.

(Kraakman et al., 2006) used 14 AFLP primers (E33M54, E35M48, E35M54, E35M55, E35M61, E37M33, E38M50, E38M54, E38M55, E39M61, E42M32, E42M48, E45M49, and E45M55) for genotyping and identified 286 polymorphic markers. Then, in 2006, 11 SSR primers (Bmac0018, Bmag0009, HVM14, HVM22, HVM65, HVM74, Bmag0223, Bmac0134, HVM54, Bmac0163, and Bmac0316) were added to the genotyping map (Kraakman et al., 2006). Also, Aghnoum et al. (Unpublished data) mapped 21 SSR molecular markers (EBmac0603, GBMS035, HVM36, scssr10559, Bmag0225, Bmag0841, Bmag0606, Bmag0013, HVM40, GBM1482, GBM1015, GBMS062, Bmac0399, EBmac0560, HvHVA1, Bmag0500, GBM1021, Bmag0173, scssr07106, Bmag0357, and Bmag0222) in this population. Finally, in total, and considering all the alleles of AFLP and SSR markers, 407 polymorphic markers were used in the present population. In this study, the sites of mapped QTLs were obtained from an integrated barley genetic map consisting of 6990 molecular markers (Aghnoum et al., 2010). This integrated genetic map included 7 linkage groups, and the molecular markers density was 0.125 markers per cM.

Population structure (Q-matrix) and kinship relationships (K-matrix)

In association analysis studies using natural populations, it is important to avoid population structure, as its presence can hinder the attainment of reliable results. Therefore, if the effect of population structure and kinship relationships is not considered to determine the trait-marker associations in association mapping, LD increases. As a result, false-positive results occur, leading to false marker-trait associations (Breseghello and Sorrells, 2006; Yu and Buckler, 2006; Zhang et al., 2012). Therefore, to determine the population structure (Q-matrix), the Bayesian method and Structure 2.3.4 software (Pritchard et al., 2000; Falush et al., 2003) were used on genotypic data. This analysis was performed on 148 barley genotypes in the Admixture model. The length of the Burnin period was 100,000, and the number of Markov Chain Monte Carlo (MCMC) replications was 100,000. Set K from 1 to 10, and the number of iterations 10 was considered. The optimal K was determined based on the delta K method. Finally, the Q-matrix was calculated with the same software by determining the optimal K, related to the highest value of delta K. Also, using genotypic data, the kinship relationships (K-matrix) were determined by TASSEL4.3.15 software.

LD and association analysis

Associations mapping was used to identify the markers related to the studied traits under normal and salinity stress conditions. For this purpose, LD for each pair of markers was estimated by the r2 statistic for each linkage group and D′statistic with LD plot by TASSEL 4.3.15 and TASSEL 2.1 software [5]. Marker-trait associations were performed using MLM with TASSEL 4.3.15 software. In the MLM

method, in addition to genotypic data, phenotypic data, and population structure (Q-matrix), kinship relationships (K-matrix) were also used as covariates in the model (Yu et al., 2006). In association analysis, just markers with a frequency of more than 10% were used, and the p-value with 1000 permutations was estimated. Also, the selection basis of the associated marker was the existence of the lowest P-value. The distribution of markers was examined based on the determination coefficient of marker (R2) in the regression model, that R2 is the ratio of calculated phenotypic variance for QTL in each location. Finally, MapChart software was used to show the mapped gene loci.

Results

Analysis of variance

A combined analysis of variance revealed high levels of genetic variability among genotypes across all traits except for harvest index and relative water content, indicating variations among genotypes in the environment. (Supplementary Table 2). The effect of the environment was significant on all of the studied traits. Also, the year effect was significant on all traits except the number of grains per spike and grain weight per spike. The environment × year, environment × genotype, year× genotype, and environment × year× genotype were significant for some traits. G × E interaction usually affects the efficiency of phenotypic selection in breeding programs (Sallam et al., 2019).

Population structure

This study determined the population's genetic structure by the Bayesian method. This method attributes each genotype to hypothetical subpopulations with a probability that in each subpopulation, the linkage disequilibrium is minimum and the gamete equilibrium is maximum. According to Supplementary Table 3 and Figure 1, the K = 2, which corresponds to the highest value of Delta K, was determined as the optimum K. Therefore, it is the most appropriate number to use for calculating the Q-matrix. Finally, the Q-matrix was obtained by placing K = 2 in the Structure 2.3.4 software.

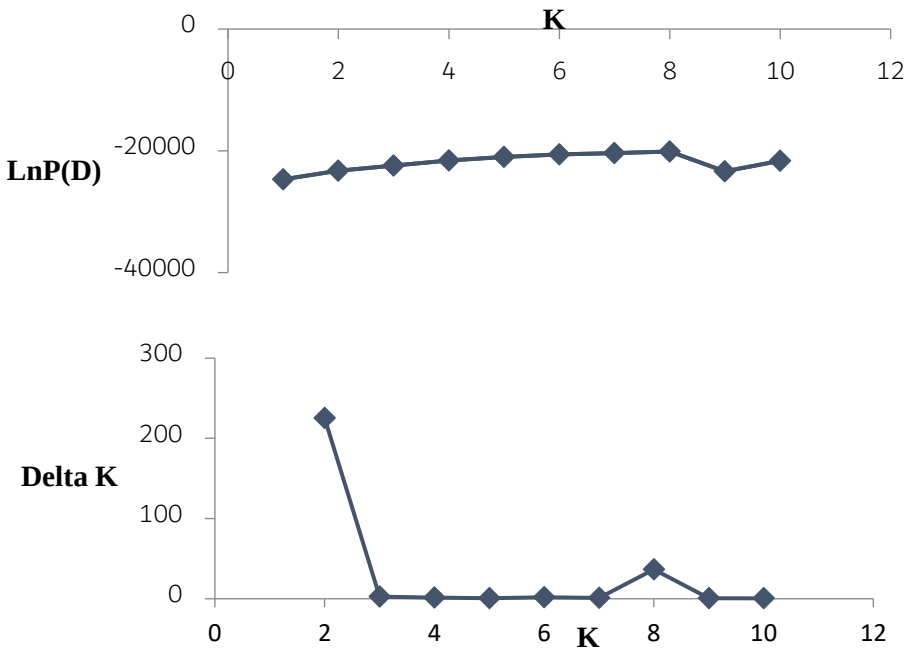


Figure 1. The two-way graphs to determine the optimum K value using 2.3.4 Structure software.

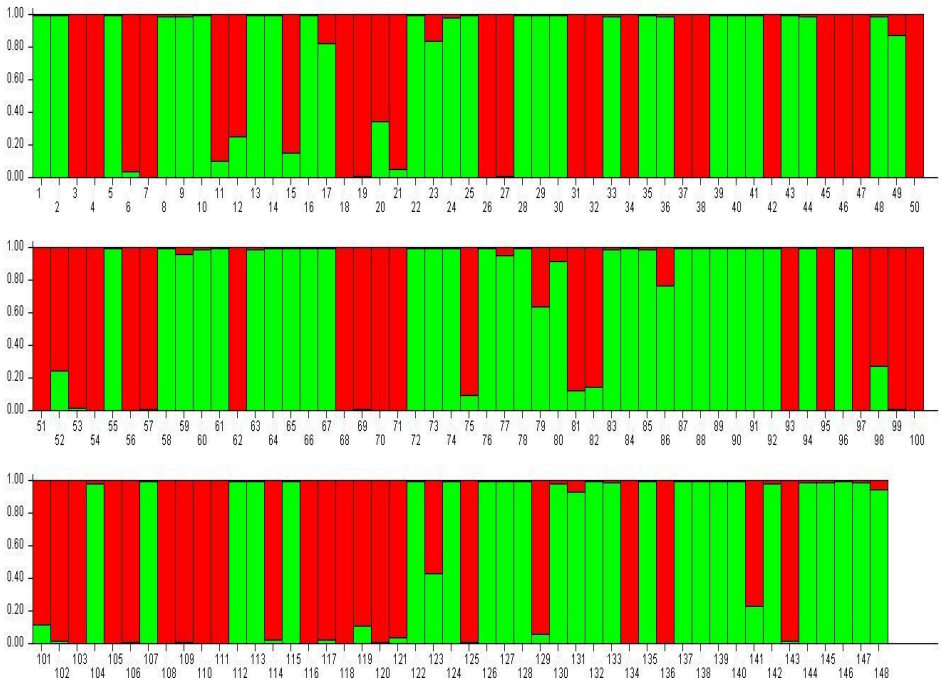


Figure 2. Bar plot generated using 407 AFLP and SSR markers by Structure 2.3.4 software. The horizontal axis represents genotypes, while the vertical axis shows the share of each genotype in each group.

The bar plot provided by Structure 2.3.4 software for 148 barley genotypes (Figure 2) also confirms the optimum K value. The horizontal axis is related to genotypes, and the vertical axis shows the share of each genotype in each group. In this bar plot, when the percentage of genotype membership in one cluster is more than or equal to 0.7, the genotype is assigned to that cluster. If the membership percentage is less than this value, it is considered a mixed genotype (Spataro et al., 2011). Here, each group is marked with a distinct colour that two separate colours for each genotype indicate that the genotype belongs to one of the two groups or both groups. Then, the number of clusters that better represent the population structure (kinship relationships defined by the K-matrix) was determined by TASSEL4.3.15 software for use in the MLM method.

LD and association mapping

LD associated with each pair of markers was estimated by D′statistic shown in the LD plot (Figure 3) and the r^2 statistic for each linkage group. The average D′ was 0.25 and the average r^2 was 0.02. The upper part of the diameter indicates the linkage disequilibrium using the D′ statistic, and the lower part of the diameter indicates the P-value for the pair of markers. The presence of red colour in the P-value study indicates the high statistical probability of LD, and green, blue and white are at lower levels of LD statistical probability, respectively. This study used MLM by association analysis to identify associated markers with the studied traits. The results showed that 194 significant marker-trait associations ($P<0.001$) were observed under normal and salinity stress conditions (Tables 1, 2, 3, 4).

Thirty-seven DNA markers were found to be significantly associated with PH, from which 33 markers were associated with the trait in normal conditions, and 4 markers were associated with the trait in salinity stress conditions (Table 1). Two DNA markers were identified for TGW, from which 1 marker was associated with the trait in normal conditions and 1 marker associated with the trait in salinity stress conditions (Table 1). Also, 2 DNA markers were found to be significantly associated with HI; 1 associated with the trait under normal conditions and 1 marker associated with the trait under salinity stress conditions (Table 1).

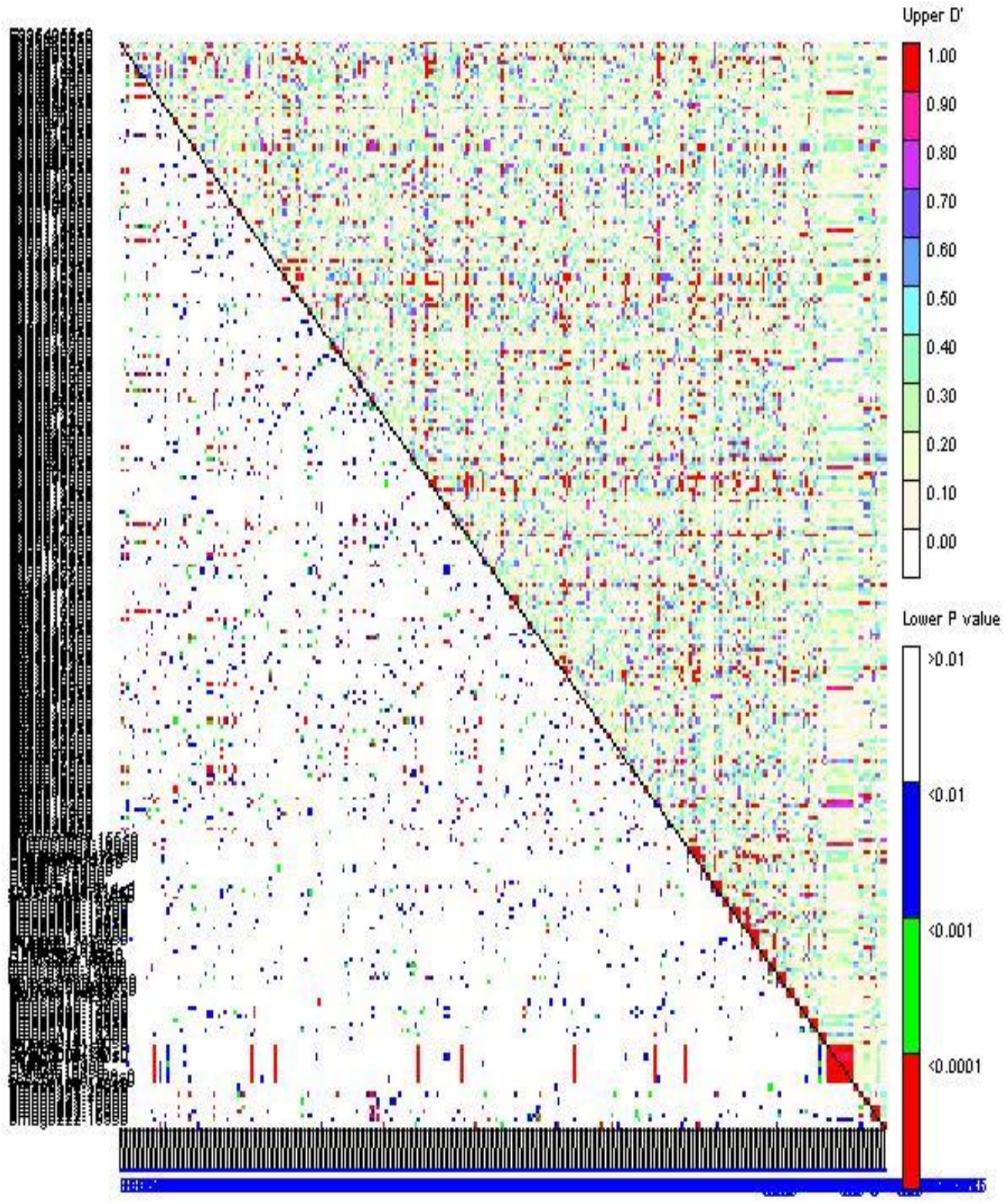


Figure 3. LD plot of barley genotypes generated using TASSEL 2.1 software. The upper part of the diameter represents the linkage disequilibrium using the D′ statistic, while the lower part of the diameter represents the corresponding *P*-value for each pair of markers.

Table 1. Markers associated with PH, TGW, HI, and NTT in barley genotypes under normal and salinity stress conditions, using the MLM model.

Trait	Conditions	Year	Marker	R^2	P-value	Position (cM)	chromosome
PH	Normal	1	EBmac0603-157	0.10	0.0007	38.3	7H
			EBmac0603-183	0.10	0.001	38.3	7H
			EBmac0603-178	0.11	0.0007	38.3	7H
			GBMS035-147	0.16	0.00002	49	7H
			GBMS035-137	0.16	0.00002	49	7H
			Bmag0606-151	0.12	0.0003	112.5	3H
			Bmag0606-138	0.11	0.0005	112.5	3H
			Bmag0606-126	0.11	0.00046	112.5	3H
			Bmag0606-147	0.14	0.00006	112.5	3H
			Bmag0606-118	0.11	0.0004	112.5	3H
			Bmag0606-122	0.11	0.00037	112.5	3H
			Bmag0606-269	0.11	0.0001	112.5	3H
			HVM40-144	0.20	0.000002	32.3	4H
			HVM40-147	0.20	0.000001	32.3	4H
			HVM40-152	0.20	0.000002	32.3	4H
			HVM40-162	0.21	0.000001	32.3	4H
			Bmag0500-110	0.10	0.0009	29.2	6H
			Bmag0500-146	0.11	0.0005	29.2	6H
			Bmag0500-166	0.10	0.001	29.2	6H
			Bmag0500-181	0.10	0.001	29.2	6H
			Bmag0500-192	0.10	0.0009	29.2	6H
			Bmag0500-194	0.10	0.00085	29.2	6H
			scssr07106-168	0.15	0.00004	23.9	5H
			scssr07106-172	0.15	0.00004	23.9	5H
			E42M48-087	0.11	0.0005	-	unmapped
		2	EBmac0603-183	0.11	0.0007	38.3	7H
			EBmac0603-143	0.10	0.001	38.3	7H
			GBMS035-147	0.11	0.0005	49	7H
			GBMS035-137	0.13	0.00018	49	7H
			HVM40-144	0.13	0.00016	32.3	4H
			HVM40-147	0.14	0.00006	32.3	4H
			HVM40-152	0.13	0.00013	32.3	4H
			HVM40-162	0.14	0.00006	32.3	4H
	Salinity	1	-	-	-	-	-
		2	HVM40-144	0.16	0.00002	32.3	4H
			HVM40-147	0.16	0.00002	32.3	4H
			HVM40-152	0.16	0.00002	32.3	4H
			HVM40-162	0.17	0.00001	32.3	4H
TGW	Normal	1	E33M54-230	0.10	0.001	131	2H
		2	-	-	-	-	-
	Salinity	1	-	-	-	-	-
		2	E45M55-103	0.12	0.00026	-	unmapped
HI	Normal	1					
		2	E33M54-214	0.54	0.000000	83.4	7H
	Salinity	1					
		2	Bmag0606-118	0.1	0.000879	112.5	3H
NTT	Normal	1	-	-	-	-	-
		2	E35M55-434	0.10	0.00085	-	unmapped
	Salinity	1	E35M54-265	0.11	0.0004	-	unmapped
			E35M61-162	0.11	0.0005	-	unmapped
			E45M55-262	0.11	0.0005	61.2	6H
		2	E35M55-434	0.11	0.0005	-	unmapped

PH: Plant height, TGW: Thousand-grain weight, HI: Harvest index, NTT: Number of total tillers, R2: Coefficient of determination, cM: Centimorgan.

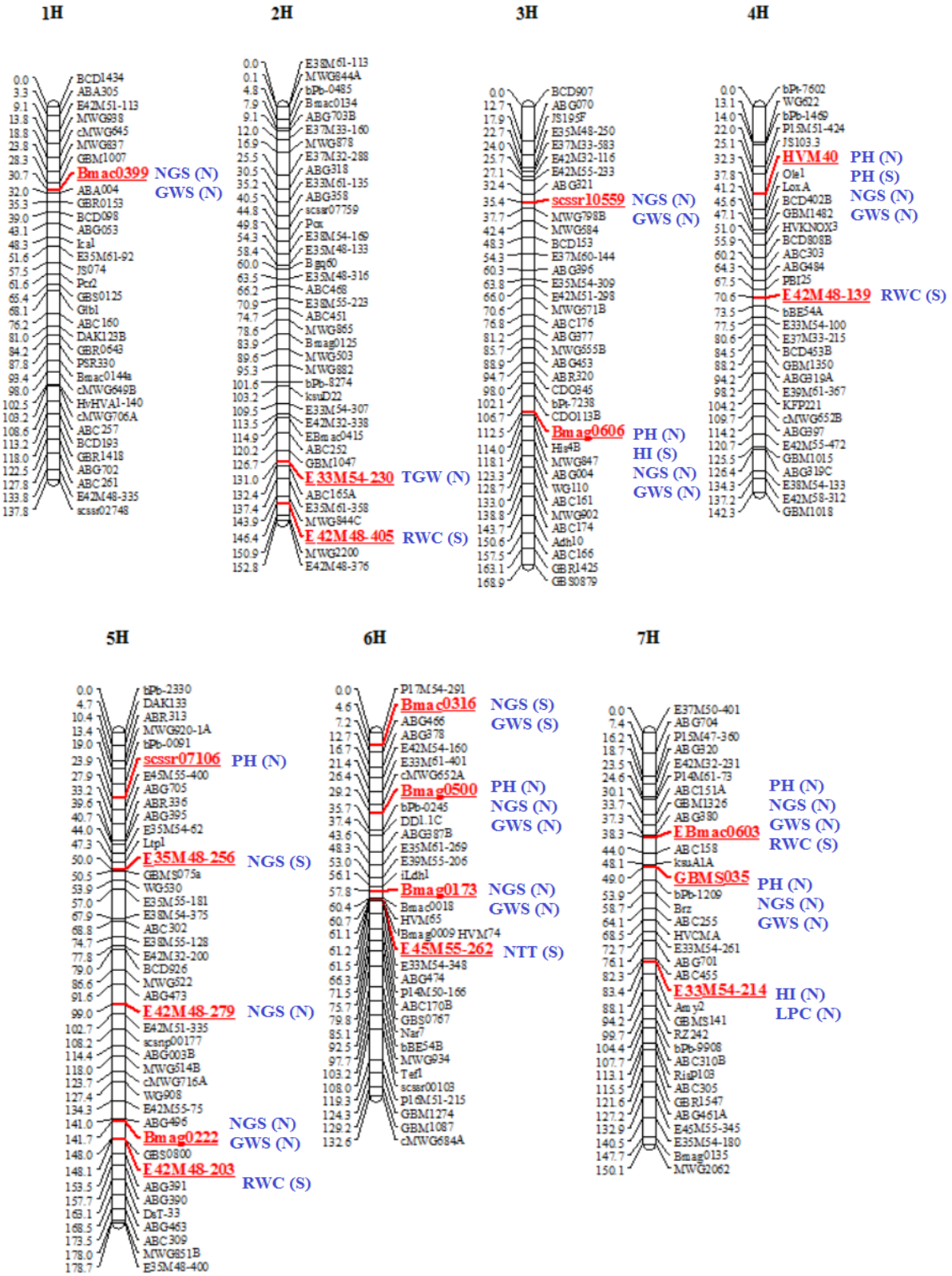


Figure 4. The genetic map of SSR and AFLP markers and genomic location of significant associated markers with studied traits in the barley (refer to Materials and Methods for the abbreviation of the traits used here; S: salinity stress, N: normal).

Table 2. Markers associated with NGS in barley genotypes using the MLM model under normal and salinity stress conditions.

Trait	Conditions	Year	Marker	R^2	P-value	Position (cM)	chromosome
NGS	Normal	1	EBmac0603-155	0.13	0.00018	38.3	7H
			EBmac0603-180	0.13	0.00018	38.3	7H
			EBmac0603-157	0.12	0.00019	38.3	7H
			EBmac0603-159	0.12	0.00019	38.3	7H
			EBmac0603-170	0.13	0.00018	38.3	7H
			EBmac0603-183	0.13	0.00014	38.3	7H
			EBmac0603-143	0.12	0.00019	38.3	7H
			EBmac0603-178	0.12	0.00019	38.3	7H
			EBmac0603-153	0.13	0.00011	38.3	7H
			GBMS035-147	0.13	0.00014	49	H
			GBMS035-137	0.13	0.00012	49	7H
			scssr10559-210	0.10	0.00071	35.4	3H
			Bmag0606-151	0.15	0.00005	112.5	3H
			Bmag0606-138	0.15	0.00005	112.5	3H
			Bmag0606-126	0.15	0.00005	112.5	3H
			Bmag0606-147	0.16	0.00002	112.5	3H
			Bmag0606-118	0.15	0.00005	112.5	3H
			Bmag0606-122	0.15	0.00004	112.5	3H
			Bmag0606-269	0.15	0.00001	112.5	3H
			HVM40-144	0.14	0.00005	32.3	4H
			HVM40-147	0.14	0.00006	32.3	4H
			HVM40-152	0.14	0.00006	32.3	4H
			HVM40-162	0.15	0.00005	32.3	4H
			Bmac0399-130	0.10	0.00077	30.7	1H
			Bmac0399-138	0.11	0.00051	30.7	1H
			Bmac0399-152	0.10	0.00099	30.7	1H
			Bmag0500-110	0.11	0.00058	29.2	6H
			Bmag0500-146	0.11	0.0006	29.2	6H
			Bmag0500-166	0.11	0.00052	29.2	6H
			Bmag0500-181	0.11	0.00047	29.2	6H
			Bmag0500-192	0.11	0.00048	29.2	6H
			Bmag0500-194	0.11	0.00059	29.2	6H
			Bmag0173-153	0.10	0.00071	57.79	6H
			Bmag0173-156	0.14	0.00009	57.79	6H
			Bmag0222-153	0.11	0.00054	141.7	5H
			Bmag0222-185	0.11	0.00054	141.7	5H
		2	E42M48-087	0.11	0.00063	-	unmapped
			E42M48-279	0.10	0.00096	99	5H
			EBmac0603-155	0.11	0.00053	38.3	7H
			EBmac0603-180	0.11	0.00052	38.3	7H
			EBmac0603-157	0.11	0.00044	38.3	7H
			EBmac0603-159	0.11	0.00055	38.3	7H
			EBmac0603-170	0.11	0.00055	38.3	7H
			EBmac0603-183	0.11	0.00045	38.3	7H
			EBmac0603-143	0.11	0.00046	38.3	7H
			EBmac0603-178	0.11	0.00054	38.3	7H
			EBmac0603-153	0.11	0.00053	38.3	7H
			GBMS035-147	0.13	0.00014	49	7H
			GBMS035-137	0.13	0.00015	49	7H
			Bmag0606-151	0.14	0.00007	112.5	3H
			Bmag0606-138	0.14	0.00007	112.5	3H
			Bmag0606-126	0.15	0.00004	112.5	3H
			Bmag0606-147	0.17	0.00001	112.5	3H
			Bmag0606-118	0.14	0.00007	112.5	3H
			Bmag0606-122	0.14	0.00007	112.5	3H
			Bmag0606-269	0.14	0.00001	112.5	3H
			HVM40-144	0.20	0.000002	32.3	4H
			HVM40-147	0.19	0.000002	32.3	4H
			HVM40-152	0.20	0.000002	32.3	4H
			HVM40-162	0.20	0.000002	32.3	4H
			Bmac0399-138	0.10	0.00094	30.7	1H
			Bmac0399-143	0.10	0.00071	30.7	1H
			Bmag0500-110	0.11	0.00045	29.2	6H
			Bmag0500-146	0.11	0.00055	29.2	6H
			Bmag0500-166	0.11	0.00058	29.2	6H
			Bmag0500-181	0.11	0.00045	29.2	6H
			Bmag0500-192	0.11	0.0006	29.2	6H
			Bmag0500-194	0.11	0.0006	29.2	6H
			Bmag0173-156	0.11	0.00052	57.79	6H
	Salinity	1	E35M48-256	0.12	0.001	50	5H
			E35M48-408	0.12	0.0009	-	unmapped
		2	Bmac0316-168	0.18	0.00001	4.6	6H

NGS: Number of grains per spike, R2: Coefficient of determination, cM: Centimorgan.

Table 3. Markers associated with GWS in barley genotypes using the MLM model under normal and salinity stress conditions.

Trait	Conditions	Year	Marker	R^2	P-value	Position (cM)	chromosome
GWS	N	1	E42M48-087	0.10	0.00081	-	unmapped
			EBmac0603-155	0.12	0.00021	38.3	7H
			EBmac0603-180	0.12	0.00021	38.3	7H
			EBmac0603-157	0.12	0.00021	38.3	7H
			EBmac0603-159	0.13	0.00017	38.3	7H
			EBmac0603-170	0.12	0.00021	38.3	7H
			EBmac0603-183	0.13	0.00017	38.3	7H
			EBmac0603-143	0.12	0.00022	38.3	7H
			EBmac0603-178	0.12	0.00021	38.3	7H
			EBmac0603-153	0.13	0.00016	38.3	7H
			GBMS035-147	0.14	0.0001	49	7H
			GBMS035-137	0.14	0.0001	49	7H
			scssr10559-214	0.10	0.00071	35.4	3H
			scssr10559-213	0.11	0.0006	35.4	3H
			scssr10559-216	0.12	0.00034	35.4	3H
			scssr10559-210	0.11	0.0006	35.4	3H
			Bmag0606-151	0.12	0.00019	112.5	3H
			Bmag0606-138	0.12	0.00021	112.5	3H
			Bmag0606-126	0.12	0.00025	112.5	3H
			Bmag0606-147	0.15	0.00004	112.5	3H
			Bmag0606-118	0.12	0.00021	112.5	3H
			Bmag0606-122	0.12	0.00025	112.5	3H
			Bmag0606-269	0.12	0.00005	112.5	3H
			HVM40-144	0.17	0.000013	32.3	4H
			HVM40-147	0.17	0.000013	32.3	4H
			HVM40-152	0.17	0.00001	32.3	4H
			HVM40-162	0.17	0.000013	32.3	4H
			Bmac0399-138	0.11	0.00056	30.7	1H
			Bmag0500-110	0.11	0.00047	29.2	6H
			Bmag0500-146	0.11	0.00052	29.2	6H
			Bmag0500-166	0.12	0.00034	29.2	6H
			Bmag0500-181	0.11	0.00048	29.2	6H
			Bmag0500-192	0.11	0.00052	29.2	6H
			Bmag0500-194	0.11	0.00051	29.2	6H
			Bmag0173-156	0.12	0.0002	57.79	6H
			Bmag0222-153	0.10	0.001	141.7	5H
			Bmag0222-185	0.10	0.001	141.7	5H
		2	E42M48-087	0.11	0.00037	-	unmapped
			GBMS035-147	0.12	0.00025	49	7H
			GBMS035-137	0.11	0.0005	49	7H
			scssr10559-213	0.10	0.00086	35.4	3H
			scssr10559-216	0.11	0.00039	35.4	3H
			Bmag0606-151	0.13	0.0001	112.5	3H
			Bmag0606-138	0.13	0.0001	112.5	3H
			Bmag0606-126	0.15	0.00003	112.5	3H
			Bmag0606-147	0.16	0.00002	112.5	3H
			Bmag0606-118	0.13	0.0001	112.5	3H
			Bmag0606-122	0.13	0.00009	112.5	3H
			Bmag0606-269	0.13	0.00002	112.5	3H
			HVM40-144	0.18	0.00001	32.3	4H
			HVM40-147	0.17	0.00001	32.3	4H
			HVM40-152	0.17	0.00001	32.3	4H
			HVM40-162	0.17	0.00001	32.3	4H
			Bmac0399-143	0.11	0.0006	30.7	1H
			Bmag0500-110	0.11	0.00047	29.2	6H
			Bmag0500-146	0.11	0.00061	29.2	6H
			Bmag0500-166	0.11	0.00057	29.2	6H
			Bmag0500-181	0.11	0.00061	29.2	6H
			Bmag0500-192	0.11	0.00051	29.2	6H
			Bmag0500-194	0.11	0.00057	29.2	6H
			Bmag0173-156	0.10	0.00074	57.79	6H
			Bmag0222-153	0.10	0.00079	141.7	5H
			Bmag0222-185	0.10	0.00079	141.7	5H
	S	1	-	-	-	-	-
		2	Bmac0316-168	0.20	0.000002	4.6	6H

GWS: Grain weight per spike, R2: Coefficient of determination, cM: Centimorgan, N: Normal, S: Salinity stress.

Table 4. Markers associated with RWC, LPC, and LChC in barley genotypes based on the MLM model under normal and salinity stress conditions.

Trait	Conditions	Year	Marker	R^2	P-value	Position (cM)	chromosome
RWC	Normal	1	-	-	-	-	-
		2	E38M54-091	0.08	0.0009	-	unmapped
			E45M49-339	0.11	0.00032	-	unmapped
	Salinity	1	-	-	-	-	-
		2	E35M48-111	0.13	0.00018	-	unmapped
			E42M48-139	0.15	0.00003	70.6	4H
			E42M48-195	0.1	0.00069	-	unmapped
			E42M48-196	0.1	0.00076	-	unmapped
			E42M48-203	0.1	0.00072	148.1	5H
			E42M48-405	0.15	0.00004	146.4	2H
			EBmac0603-178	0.11	0.00053	38.3	7H
LPC	Normal	1	E33M54-214	0.15	0.00003	83.4	7H
		2	E33M54-214	0.15	0.00003	83.4	7H
	Salinity	1	-	-	-	-	-
		2	-	-	-	-	-
LChC	Normal	1	E35M48-251	0.09	0.00037	-	unmapped
		2	-	-	-	-	-
	Salinity	1	-	-	-	-	-
		2	-	-	-	-	-

RWC: Relative water content, LPC: Leaf Proline content, LChC: Leaf chlorophyll content, R2: Coefficient of determination, cM: Centimorgan.

Five markers were found to be significantly associated with NTT, of which 1 marker was associated with the trait in normal conditions, while the remaining 4 markers were associated with the trait under salinity stress conditions (Table 1). Seventy-two markers were identified for NGS, from which 69 markers were linked with the trait under normal conditions, and 3 markers were associated with the trait under salinity stress conditions (Table 2). For the GWS trait, 64 DNA markers were identified, from which 63 markers were associated with the trait in normal conditions and 1 marker associated with the trait under salinity stress conditions (Table 3). Nine DNA markers were identified for RWC. Two markers were associated with this trait under normal conditions, and seven were associated under salinity stress conditions (Table 4). Totally 2 DNA markers were identified for LPC, associated with the trait in normal conditions (Table 4). For LChC, only one marker was detected in the normal conditions (Table 4). The genetic map of SSR and AFLP markers and genomic location of significant markers with studied traits showed in figure 4.

Discussion

ANOVA revealed significant genetic variation among genotypes across all traits except harvest index and relative water content, suggesting distinctions among genotypes within the environment. The environment × year, environment × genotype, year× genotype, and environment × year× genotype were significant for some traits. G × E interaction usually affects the efficiency of phenotypic selection in breeding programs [38].

In association mapping, false-positive results obtain if the population structure and kinship relatedness are not considered (Breseghello and Sorrells, 2006). Hence, estimating population structure as a prerequisite in association mapping can prevent false-positive associations between markers and traits (Pritchard and Donnelly, 2001). This study subdivided barley cultivars into two subpopulations. Some reports suggest that the

population structure of barley cultivars is related to spike morphology (two-rowed versus six-rowed cultivars) (Pasam et al., 2012). In the association mapping method, QTLs are located based on LD (Gupta et al., 2005). In the present association panel, the mean of D′and r^2, indicators for LD, were 0.25 and 0.02, respectively. According to the LD plot, LD had a significant difference between barley chromosomes, which indicates that this factor can affect the accuracy of association mapping of identified QTLs on different chromosomes.

Several studies have previously reported different rates of LD in different barley populations (Caldwell et al., 2006; Ramsay et al., 2011) and among different chromosomes Rostoks et al. (2006). Caldwell et al. (2006) reported rapid decay of LD in barley landraces compared to superior barley cultivars. Eleuch et al. (2008), Inostroza et al. (2009), El-Denary et al. (2012), Long et al. (2013), Sbei et al. (2014), Elakhdar et al. (2016a), Elakhdar et al. (2016b) and Fan et al. (2016) used association mapping under salinity stress in the barley. This study identified 194 significant marker-trait associations for nine studied morphophysiological traits under normal and salinity stress conditions.

This study detected 33 and 4 significant marker-trait associations for PH in normal and salinity stress conditions, respectively. Seven QTLs on chromosomes 3H (112.5 cM), 8 QTLs on 4H (32.3 cM), 2 QTLs on 5H (23.9 cM), 6 QTLs on 6H (29.2 cM), 5 QTLs on 7H (38.3 cM), 4 QTLs on 7H (49 cM), 1 QTL with unknown gene location in the normal experiment and 4 QTLs on chromosome 4H (32.3 cM) under salinity stress conditions were observed for PH. Elakhdar et al. (2016a), in a study on barley for mean normal and salinity stress conditions, showed that this trait had a significant association with marker EBmac0603 on chromosome 7H at 35.39 cM position, which is similar to our results. Sayed et al. (2021) reported PH on chromosome 7H, Long et al. (2013) on chromosomes 2H (59.2 cM), 6H (60.2 cM) and 7H (4.9 cM) and 7H (61.3 cM), Eleuch et al. (2008) on 1H (62 cM) and 6H (10 cM), Inostroza et al. (2009) on 2H (5, 50, and 44 cM), 4H (78 and 118 cM), 5H (66 and 126 cM), 6H (79), and 7H (80, 85 and 107 cM), El-Denary et al. (2012) on 2H, Xue et al. (2009) on 3H, Saade et al. (2020) on 6H (51.93), under salinity stress conditions in the barley. Xu et al. (2012) detected this trait on chromosome 7H under normal conditions in the barley, which is consistent with our results.This study found 1 QTL for TGW on chromosome 2H (131 cM) in normal conditions and 1 QTL with unknown gene location in salinity stress conditions. Elakhdar et al. (2016a) identified this trait on chromosomes 6H (75.42 cM), 6H (7.16 cM), and 1H (30.81 cM) for mean the normal and salinity conditions in the barley. Wang et al. (2016) observed TGW on 2H, 5H, and 7H in the barley under normal conditions.

In the present study, one QTL was identified for HI on chromosome 7H at 83.4 cM under normal conditions and 1 QTL on chromosome 3H at 112.5 cM under salinity stress conditions. The marker E33M54-214 on chromosome 7H (83.4 cM) has a high coefficient of determination (R^2 = 0.54) with QTL controlling the HI, indicating a strong association between the marker and the trait. Under salinity stress conditions in the barley, Elakhdar et al. (2016a) on 2H and 5H and Saade et al. (2020) on 7H at 28.46 cM reported this trait.

According to the results, 1 QTL on chromosome 6H (61.2 cM), 3 QTLs with unknown gene location in salinity stress conditions, and 1 QTL with unknown gene location in normal conditions were detected for NTT. Long et al. (2013), under salinity stress conditions in the barley, found this trait on chromosomes 4H, 6H, and 7H, which were located in the positions of 79.6, 60.2, and 54.4 cM, respectively. As can be seen, our results' position of 61.2 cM is almost close to 60.2 cM in Long et al. (2013), so NTT is probably located in this gene locus. Xue et al. (2009), in a study on barley under both normal and salinity stress conditions, identified NTT on chromosome 4H, which was located in the positions of 72 cM. Long et al. (2013), under salinity stress conditions in the barley, observed this trait on chromosomes 4H, 6H, and 7H, which were located in the positions of 79.6, 60.2, and 54.4 cM, respectively.

NGS, one of the important yield components, has a major effect on the final yield. This study detected 69 and 3, a significant marker-trait association for NGS under normal conditions and salinity stress. 5 QTLs on chromosomes 1H (30.7 cM), 14 QTLs on 3H (112.5 cM), 1 QTL on 3H (35.4 cM), 8 QTLs on 4H (32.3 cM), 2 QTLs on 5H (141.7 cM), 1 QTL on 5H (99 cM), 12 QTLs on 6H (29.2 cM), 3 QTLs on 6H (57.79 cM), 18 QTLs on 7H (38.3 cM), 4 QTLs on 7H

(49 cM), and 1 QTL with unknown gene location in normal conditions were identified for NGS. Under salinity stress conditions, 1 QTL on 5H (50 cM), 1 QTL on 6H (4.6 cM), and 1 QTL with unknown gene location were observed for this trait. Xue et al. (2009), in a study on barley under both normal and salinity stress conditions, reported NGS on chromosome 2H. Elakhdar et al. (2016a) on chromosomes 1H (64.84 cM), 2H (89.83 cM), 4H (96.17 cM), 6H (7.16 cM), 7H (81.78 cM), 7H (97 cM) and Saade et al. (2020) on chromosome 7H at 128.35 cM observed this trait under salinity stress conditions in the barley. Also, Sun et al. (2011) detected NGS on chromosomes 1H, 4H, and 5H under normal conditions in the barley.

This study found for GWS 2 QTLs on chromosomes 1H (30.7 cM), 6 QTLs on 3H (35.4 cM), 14 QTL on 3H (112.5 cM), 8 QTLs on 4H (32.3 cM), 4 QTLs on 5H (141.7 cM), 12 QTLs on 6H (29.2 cM), 2 QTLs on 6H (57.79 cM), 9 QTLs on 7H (38.3 cM), 4 QTLs on 7H (49 cM), and 2 QTL with unknown gene location in normal conditions. Under salinity conditions, 1 QTL on chromosome 6H (4.6 cM) was identified with the marker for the GWS.

In the present study, 1 QTL on chromosomes 2H at 146.4 cM, 1 QTL on 4H at 70.6 cM, 1 QTL on 5H at 148.1 cM, 1 QTL on 7H at 38.3 cM, and 3 QTLs with unknown gene location was detected for RWC in salinity stress conditions. Under normal conditions, 2 QTL with unknown gene location was observed with the marker for this trait. Liu et al. (2015) on chromosomes 6H (57.8 cM), 6H (53.8 cM), and 7H (62.3 cM) reported RWC under salinity stress conditions in the barley. Mohamed et al. (2015) identified QTLs for this trait in barley under the normal conditions on chromosomes 1H, 3H, and 6H and QTLs for the trait under salt stress conditions on chromosomes 2H, 3H, 5H, 7H, and 6H. Also, Jabbari et al. (2021) observed this trait on chromosomes 2H and 7H under normal conditions.

According to the results, 2 QTLs were identified for LPC on Chromosome 7H (83.4 cM) in normal conditions. Under salinity stress conditions, no significant association was observed with the marker for the LPC. Jabbari et al. (2021), under normal conditions in the barley, detected LPC on chromosomes 2H, 4H, 5H, 6H, and 7H.

Abundant nutrition production is essential to sustain crop growth, which depends on the LChC (Yap and Harvey, 1972; Liu et al., 2015). This study found one QTL with an unknown gene location for LChC in normal conditions. Under salinity stress conditions, no significant association was observed with the markers for this trait. Elakhdar et al. (2016a) on 1H (64.84 cM), 1H (54.6 cM), 4H (58.6 cM), and 4H (96.17 cM), Elakhdar et al. (2016b) on 1H, 4H, Long et al. (2013), on 1H (31.1 cM), 5H (6.4 cM), 6H (45.4 cM), 6H (60.2 cM), 7H (4.9 cM), Liu et al. (2015) on chromosomes 2H (75.9 cM), 7H (47.5 cM), and 7H (58.9 cM) identified this trait under salinity stress conditions in the barley. Barati et al. (2017) reported two and four QTLs for LChC in barley under normal and stress conditions on chromosomes 3H, 4H, 5H, and 6H. Jabbari et al. (2021) observed QTLs for this trait on chromosomes 1H, 2H, 3H, and 4H under normal conditions.

Some identified DNA markers were common among some studied traits in this study. In normal conditions, EBmac0603-157, EBmac0603-183, EBmac0603-178, GBMS035-147, GBMS035-137, Bmag0606-151, Bmag0606-126, Bmag0606-147, Bmag0606-118, Bmag0606-122, Bmag0606-269, Bmag0606-138, HVM40-144, HVM40-147, HVM40-152, HVM40-162, Bmag0500-110, Bmag0500-146, Bmag0500-166, Bmag0500-181, Bmag0500-192, Bmag0500-194, and E42M48-087 were common for PH, GWS and NGS traits, EBmac0603-155, EBmac0603-180, EBmac0603-159, EBmac0603-170, EBmac0603-143, EBmac0603-153, Bmag0173-156 , scssr10559-210 , Bmac0399-138, Bmag0222-153, Bmag0222-185, Bmac0399-143 and E33M54-214 were common for GWS and NGS traits. Under salinity stress conditions, Bmac0316-168 was common for GWS and NGS traits. Identifying common markers is very important in plant breeding because it allows the simultaneous selection of several traits (Tuberosa et al., 2002; Hittalmani et al., 2003). The common markers among traits are helpful because they increase the efficiency of marker-assisted selection. Common markers among traits can be due to pleiotropic effects or linkage between genomic regions involved in these traits (Jun et al., 2008). Of course, the presence of common markers is valuable when they are associated with large-effect QTLs, and secondly, they are stable and can be identified by repeated testing. However, in this experiment, the value of the coefficient of determination (R^2) was negligible in most traits. Although this

phenomenon was not unexpected because the nature of QTLs is such that several positions are involved in one trait, and a high R^2 for a marker is unexpected.

In the present study, some common markers were identified for a particular trait or several traits under normal and salinity stress conditions, called stable QTLs. Bmag0606-151, Bmag0606-126, Bmag0606-147, Bmag0606-118, Bmag0606-122, Bmag0606-269, Bmag0606-138, Bmag0500-110, Bmag0500-146, Bmag0500-166, Bmag0500-181, Bmag0500-192, Bmag0173-156 were common for GWS and NGS, EBmac0603-183 was common for PH and NGS, HVM40-144, HVM40-147, HVM40-152, HVM40-162 were common for PH in both normal and salinity stress conditions. GBMS035-147, GBMS035-137, HVM40-144, HVM40-147, HVM40-162 were common for PH, GWS and NGS, Bmag0606-151, Bmag0606-138, Bmag0606-126, Bmag0606-147, Bmag0606-118, Bmag0606-122, Bmag0606-269, Bmag0500-110, Bmag0500-146, Bmag0500-166, Bmag0500-181, Bmag0500-192, Bmag0500-194, GBMS035-147, GBMS035-137, HVM40-144, HVM40-147, HVM40-152, HVM40-162, Bmag0173-156 were common for GWS and NGS, EBmac0603-183, GBMS035-147, GBMS035-137 were common for PH, scssr10559-213, scssr10559-216, Bmag0222-153, Bmag0222-185, E42M48-087, were common for GWS, Bmac0399-138, EBmac0603-155, EBmac0603-180, EBmac0603-157, EBmac0603-159, EBmac0603-170, EBmac0603-183, EBmac0603-143, EBmac0603-178, EBmac0603-153 were common for NGS, E33M54-214 was common for LPC in normal conditions across two years, Which indicates the stability of these gene loci in normal environments and have no effect on the trait under stress conditions. Gene loci that act the same in different environments can be introduced as stable QTLs. Stable QTLs provide relative stability to genetic control and overcome the interaction between genotype and environment; therefore, their selection for a trait under normal conditions also improves the trait value under stress conditions. The stability of QTLs in different environments is due to the control of traits by a small number of large-effect gene loci, so marker-assisted selection efficiency is highly effective in this population. Common stable QTLs can be used in plant breeding to select several traits simultaneously

Conclusion

Salinity tolerance in crop plants is governed by a multifaceted interplay of genetic and physiological factors, with a quantitative and intricate nature influenced by numerous gene loci.. Results of the present study revealed that association mapping is a powerful tool for identifying DNA markers for morpho-physiological traits in barley. Specifically, 194 significant marker-trait associations were identified for the studied traits. Out of 194 QTLs 171 and 23 QTLs were observed for traits under normal and salinity stress conditions, respectively. Identified markers could be helpful in marker-assisted breeding programs for salinity stress tolerance in barley. It is suggested that markers with a higher determination coefficient (R^2) can use in the saturation of genetic maps. In this study, the marker E33M54-214 on chromosome 7H (83.4 cM) has a high coefficient of determination (R^2 = 0.54) with QTL controlling the HI, indicating a strong association between the marker and the trait. Several QTLs were stable for plant height, the number of grains per spike, grain weight per spike, and leaf proline content under different environmental conditions, introduced as stable QTLs. The results showed that some stable QTLs were common to several traits, providing an opportunity to improve several traits simultaneously and facilitate the development of high-yielding barley cultivars.

Supplementary Materials

The Supplementary Material for this article can be found online at: https://www.jpmb-gabit.ir/article_710678.html.

Supplementary Table 1. Genotypes used in the present study.

Supplementary Table 2. Combined analysis of variance of the studied traits in both non-stress and salinity stress conditions during two cropping years.

Supplementary Table 3. Statistics calculated for optimum K values using 2.3.4 Structure software.

Author contributions

M.Z: Conceptualization, methodology, software, formal analysis, investigation, data curation, writing - original draft preparation, Writing - review and editing; Nadali Babaeian-Jelodar: Supervision, conceptualization, Writing - review

and editing; R.A: Resources; S.A.T: Project administration; M.G.N: Data curation.

Funding

This research received no external funding.

Acknowledgments

The authors thank the Department of Plant Breeding and Biotechnology of Sari Agricultural Sciences and Natural Resources University and the Department of Seed and Plant Improvement Research of Yazd Agricultural and Natural Resources Research and Education Center for supporting this study.

Conflict of interest statement

The authors declare no conflict of interest.

Abbreviations

GLM: General linear model; GWS: Grain weight per spike; HI: Harvest index; LChC: Leaf chlorophyll content; LD: Linkage disequilibrium; LPC: Leaf Proline content; MLM: Mixed linear model; NGS: Number of grains per spike; NTT: Number of total tillers; PH: Plant height; QTLs: quantitative trait loci; RWC: Relative water content; TGW: Thousand-grain weight.

References

Aghnoum, R., Marcel, T.C., Johrde, A., Pecchioni, N., Schweizer, P., and Niks, R.E. (2010). Basal host resistance of barley to powdery mildew: connecting quantitative trait loci and candidate genes. *Mol Plant-Microbe Interact* 23(1): 91-102.

Arzani, A. (2008). Improving salinity tolerance in crop plants: a biotechnological view. *In Vitro Cell Dev Biol Plant* 44: 373-383.

Arzani, A., and Ashraf, M. (2016). Smart engineering of genetic resources for enhanced salinity tolerance in crop plants. *CRC Crit Rev Plant Sci* 35(3): 146-189.

Barati, A., Moghadam, M., Mohammadi, S., Ghazvini, H., and Sadeghzadeh, B. (2017). Identification of QTLs associated with agronomic and physiological traits under salinity stress in barley. *J Agric Sci Tech* 19: 185-200.

Breseghello, F., and Sorrells, M.E. (2006). Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172(2): 1165-1177.

Caldwell, K.S., Russell, J., Langridge, P., and Powell, W. (2006). Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species, *Hordeum vulgare*. *Genetics* 172(1): 557-567.

El-Denary, M., Noaman, M., Abdelkhalek, A., and Mariey, S. (2012). Marker traits association of some barley genotypes under soil salinity condition using SSR markers. *Egypt J Medical Hum Genet* 41(2).

Elakhdar, A., Abd EL-Sattar, M., Amer, K., Rady, A., and Kumamaru, T. (2016a). Population structure and marker–trait association of salt tolerance in barley (*Hordeum vulgare* L.). *C R Biol* 339(11-12): 454-461.

Elakhdar, A., El-Sattar, M.A., Amer, K., and Kumamaru, T. (2016b). Genetic diversity and association analysis among Egyptian barley ('*Hordeum vulgare*'L.) genotypes with different adaptations to saline conditions analyzed by SSR markers. *Aust J Crop Sci* 10(5).

Eleuch, L., Jilal, A., Grando, S., Ceccarelli, S., von Korff Schmising, M., Tsujimoto, H., Hajer, A., Daaloul, A., and Baum, M. (2008). Genetic diversity and association analysis for salinity tolerance, heading date and plant height of barley germplasm using simple sequence repeat markers. *J Integr Plant Biol* 50(8): 1004-1014.

Ellis, R.P., Forster, B.P., Robinson, D., Handley, L., Gordon, D.C., Russell, J.R., and Powell, W. (2000). Wild barley: a source of genes for crop improvement in the 21st century? *J Exp Bot* 51(342): 9-17.

Falush, D., Stephens, M., and Pritchard, J.K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164(4): 1567-1587.

Fan, Y., Zhou, G., Shabala, S., Chen, Z.-H., Cai, S., Li, C., and Zhou, M. (2016). Genome-wide association study reveals a new QTL for salinity tolerance in barley (*Hordeum vulgare* L.). *Front Plant Sci* 7: 946.

Flint-Garcia, S.A., Thornsberry, J.M., and Buckler IV, E.S. (2003). Structure of linkage disequilibrium in plants. *Annu Rev Plant Biol* 54(1): 357-374.

Flint‐Garcia, S.A., Thuillet, A.C., Yu, J., Pressoir, G., Romero, S.M., Mitchell, S.E., Doebley, J., Kresovich, S., Goodman, M.M., and Buckler, E.S. (2005). Maize association population: a high‐resolution platform for quantitative trait locus dissection. *Plant J* 44(6): 1054-1064.

Flowers, T. (2004). Improving crop salt tolerance. *J Exp Bot* 55(396): 307-319.

Gharaghanipor, N., Arzani, A., Rahimmalek, M., and Ravash, R. (2022). Physiological and transcriptome indicators of salt tolerance in wild and cultivated barley. *Front Plant Sci* 13: 819282.

Gupta, P.K., Rustgi, S., and Kulwal, P.L. (2005). Linkage disequilibrium and association studies in higher plants: present status and future prospects. *Plant Mol Biol* 57: 461-485.

Hittalmani, S., Huang, N., Courtois, B., Venuprasad, R., Shashidhar, H., Zhuang, J., Zheng, K., Liu, G., Wang, G., and Sidhu, J. (2003). Identification of QTL for growth-and grain yield-related traits in rice across nine locations of Asia. *Theor Appl Genet* 107: 679-690.

Inostroza, L., del Pozo, A., Matus, I., Castillo, D., Hayes, P., Machado, S., and Corey, A. (2009). Association mapping of plant height, yield, and yield stability in recombinant chromosome substitution lines (RCSLs) using *Hordeum vulgare* subsp. spontaneum as a source of donor alleles in a *Hordeum vulgare* subsp. vulgare background. *Mol Breed* 23: 365-376.

Jabbari, M., Fakheri, B.A., Aghnoum, R., Darvishzadeh, R., Mahdi Nezhad, N., Ataei, R., Koochakpour, Z., and Razi, M. (2021). Association analysis of physiological traits in spring barley (*Hordeum vulgare* L.) under water‐deficit conditions. *Food Sci Nutr* 9(3): 1761-1779.

Jun, T.-H., Van, K., Kim, M.Y., Lee, S.-H., and Walker, D.R. (2008). Association analysis using SSR markers to find QTL for seed protein content in soybean. *Euphytica* 162: 179-191.

Kilian, B., Özkan, H., Kohl, J., von Haeseler, A., Barale, F., Deusch, O., Brandolini, A., Yucel, C., Martin, W., and Salamini, F. (2006). Haplotype structure at seven barley genes: relevance to gene pool bottlenecks, phylogeny of ear type and site of barley domestication. *Mol Genet Genom* 276: 230-241.

Kraakman, A., Martinez, F., Mussiraliev, B., Van Eeuwijk, F., and Niks, R. (2006). Linkage disequilibrium mapping of morphological, resistance, and other agronomically relevant traits in modern spring barley cultivars. *Mol Breed* 17: 41-58.

Liu, L., Sun, G., Ren, X., Li, C., and Sun, D. (2015). Identification of QTL underlying physiological and morphological traits of flag leaf in barley. *BMC Genet* 16(1): 1-10.

Long, N.V., Dolstra, O., Malosetti, M., Kilian, B., Graner, A., Visser, R.G., and van der Linden, C.G. (2013). Association mapping of salt tolerance in barley (H*ordeum vulgare* L.). *Theor Appl Genet* 126: 2335-2351.

Mohamed, N., Said, A., Mustafa, A., and Léon, J. (2015). Association mapping for salinity tolerance related traits in a structured barley population. *Egypt J Agron* 37(1): 11-33.

Moose, S.P., and Mumm, R.H. (2008). Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiol* 147(3): 969-977.

Omrani, S., Arzani, A., Esmaeilzadeh Moghaddam, M., and Mahlooji, M. (2022). Genetic analysis of salinity tolerance in wheat (*Triticum aestivum* L.). *PloS One* 17(3): e0265520.

Pasam, R.K., Sharma, R., Malosetti, M., van Eeuwijk, F.A., Haseneyer, G., Kilian, B., and Graner, A. (2012). Genome-wide association studies for agronomical traits in a world wide spring barley collection. *BMC Plant Biol* 12(1): 1-22.

Pritchard, J.K., and Donnelly, P. (2001). Case–control studies of association in structured or admixed populations. *Theor Popul Biol* 60(3): 227-237.

Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155(2): 945-959.

Ramsay, L., Comadran, J., Druka, A., Marshall, D.F., Thomas, W.T., Macaulay, M., MacKenzie, K., Simpson, C., Fuller, J., and Bonar, N. (2011). *INTERMEDIUM-C*, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene *TEOSINTE BRANCHED 1*. *Nat Genet* 43(2): 169-172.

Rostoks, N., Ramsay, L., MacKenzie, K., Cardle, L., Bhat, P.R., Roose, M.L., Svensson, J.T., Stein, N., Varshney, R.K., and Marshall, D.F. (2006). Recent history of artificial outcrossing facilitates whole-genome association mapping in elite inbred crop varieties. *Proc Natl Acad Sci USA* 103(49): 18656-18661.

Saade, S., Brien, C., Pailles, Y., Berger, B., Shahid, M., Russell, J., Waugh, R., Negrão, S., and Tester, M. (2020). Dissecting new genetic components of salinity tolerance in two-row spring barley at the vegetative and reproductive stages. *PLoS One* 15(7): e0236037.

Sallam, A., Alqudah, A.M., Dawood, M.F., Baenziger, P.S., and Börner, A. (2019). Drought stress tolerance in wheat and barley: advances in physiology, breeding and genetics research. *Int J Mol Sci* 20(13): 3137.

Sayed, M.A., Nassar, S.M., Moustafa, E.S., Said, M.T., Börner, A., and Hamada, A. (2021). Genetic mapping reveals novel exotic and elite QTL alleles for salinity tolerance in barley. *Agronomy* 11(9): 1774.

Sbei, H., Sato, K., Shehzad, T., Harrabi, M., and Okuno, K. (2014). Detection of QTLs for salt tolerance in Asian barley (*Hordeum vulgare* L.) by association analysis with SNP markers. *Breed Sci* 64(4): 378-388.

Spataro, G., Tiranti, B., Arcaleni, P., Bellucci, E., Attene, G., Papa, R., Spagnoletti Zeuli, P., and Negri, V. (2011). Genetic diversity and structure of a worldwide collection of *Phaseolus coccineus* L. *Theor Appl Genet* 122: 1281-1291.

Sun, D., Ren, W., Sun, G., and Peng, J. (2011). Molecular diversity and association mapping of quantitative traits in Tibetan wild and worldwide originated barley (Hordeum vulgare L.) germplasm. *Euphytica* 178: 31-43.

Tuberosa, R., Salvi, S., Sanguineti, M.C., Landi, P., Maccaferri, M., and Conti, S. (2002). Mapping QTLs regulating morpho‐physiological traits and yield: Case studies, shortcomings and perspectives in drought‐stressed maize. *Ann Bot* 89(7): 941-963.

Wang, J., Sun, G., Ren, X., Li, C., Liu, L., Wang, Q., Du, B., and Sun, D. (2016). QTL underlying some agronomic traits in barley detected by SNP markers. *BMC Genet* 17(1): 1-13.

Xu, R., Wang, J., Li, C., Johnson, P., Lu, C., and Zhou, M. (2012). A single locus is responsible for salinity tolerance in a Chinese landrace barley (*Hordeum vulgare* L.). *PLoS One* 7(8): e43079.

Xue, D., Huang, Y., Zhang, X., Wei, K., Westcott, S., Li, C., Chen, M., Zhang, G., and Lance, R. (2009). Identification of QTLs associated with salinity tolerance at late growth stage in barley. *Euphytica* 169: 187-196.

Yap, T., and Harvey, B. (1972). Relations between grain yield and photosynthetic parts above the flag leaf node in barley. *Can J Plant Sci* 52(2): 241-246.

Yu, J., and Buckler, E.S. (2006). Genetic association mapping and genome organization of maize. *Curr Opin Biotechnol* 17(2): 155-160.

Yu, J., Pressoir, G., Briggs, W.H., Vroh Bi, I., Yamasaki, M., Doebley, J.F., McMullen, M.D., Gaut, B.S., Nielsen, D.M., and Holland, J.B. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* 38(2): 203-208.

Zhang, Q., Wu, C., Ren, F., Li, Y., and Zhang, C. (2012). Association analysis of important agronomical traits of maize inbred lines with SSRs. *Aust J Crop Sci* 6(6): 1131-1138.

Zhu, C., Gore, M., Buckler, E.S., and Yu, J. (2008). Status and prospects of association mapping in plants. *Plant Genome* 1(1).

JPMB
Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 1

مقاله پژوهشی

نقشه‌یابی ارتباطی صفات مورفوفیزیولوژیک در جو تحت تنش شوری

مهدیه زارع‌کهن*[۱]، نادعلی بابائیان‌جلودار [۱]، رضا اقنوم [۲]، سیدعلی طباطبایی [۳]، و محمدرضا قاسمی‌نژادرائینی [۴]

[۱] گروه به‌نژادی و بیوتکنولوژی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران
[۲] گروه تحقیقات اصلاح بذر و نهال، سازمان تحقیقات، آموزش و ترویج کشاورزی و منابع طبیعی خراسان رضوی، مشهد، ایران
[۳] گروه تحقیقات اصلاح بذر و نهال، سازمان تحقیقات، آموزش و ترویج کشاورزی و منابع طبیعی یزد، یزد، ایران
[۴] گروه مهندسی آب، دانشگاه آزاد اسلامی واحد کرمان، کرمان، ایران

چکیده: مطالعه حاضر برای شناسایی نشانگرهای مولکولی مرتبط با صفات مورفوفیزیولوژیک، از ۱۴ ترکیب آغازگر AFLP و ۳۲ جفت آغازگر SSR در ۱۴۸ ژنوتیپ جو از روش نقشه‌یابی ارتباطی استفاده کرد. این آزمایش در قالب طرح آلفا لاتیس با پنج بلوک ناقص در دو تکرار تحت شرایط معمول و تنش شوری (EC=12 dS m^{-1}) در مزرعه مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی یزد به مدت دو سال اجرا شد. ساختار ژنتیکی جمعیت به دو زیرجمعیت (K=۲) تقسیم شد. در جمعیت ارتباطی حاضر، میانگین D' و r^2 که شاخص های عدم تعادل پیوستگی هستند، به ترتیب ۰/۲۵ و ۰/۰۲ بودند. مدل خطی مخلوط ۱۹۴ ارتباط معنی‌دار نشانگر – صفت برای ۹ صفت مورد مطالعه در شرایط معمول و تنش شوری شناسایی کرد. تعدادی از جایگاه‌های ژنی کمی برای صفات ارتفاع بوته، تعداد دانه در سنبله، وزن دانه در سنبله و مقدار پرولین برگ در شرایط محیطی مختلف، پایدار بودند که به عنوان جایگاه های ژنی کمی پایدار معرفی شدند. نتایج نشان داد که تعدادی از جایگاه‌های ژنی کمی پایدار در چندین صفت مشترک بودند که می‌توان از آن‌ها در به‌نژادی گیاهان زراعی به‌منظور گزینش همزمان چند صفت استفاده کرد و بهبود ارقام جو پرمحصول را تسهیل کرد. نشانگرهای شناسایی شده را می‌توان در انتخاب به کمک نشانگر در برنامه‌های اصلاحی تحمل به تنش شوری جو استفاده نمود.

کلمات کلیدی: تنش شوری، جایگاه‌های ژنی کمی پایدار، جو، عدم تعادل پیوستگی، مدل خط مخلوط، نقشه‌یابی ارتباطی.

ویراستار علمی
دکتر احمد ارزانی،
دانشگاه صنعتی اصفهان، ایران

تاریخ
دریافت: ۳ خرداد ۱۴۰۲
پذیرش: ۱۳ بهمن ۱۴۰۲
چاپ: ۲۶ بهمن ۱۴۰۲

نویسنده مسئول
مهدیه زارع‌کهن
mahdiyehzare65@gmail.com

ارجاع به این مقاله
Zare-Kohan, M., Babaeian Jelodar, N., Aghnoum, R., Tabatabaee, S. A., and Ghasemi-Nezhadraeini, M. (2023). Association mapping of morpho-physiological traits in barley (*Hordeum vulgare* L.) under salinity stress. *J Plant Mol Breed 11* (1): 75-88. doi:10.22058/JPMB.2024.2003151.1274.

Contents:

Address:

Genetics and Agricultural Biotechnology Institute of Tabarestan (GABIT)
Sari Agricultural Sciences & Natural Resources University (SANRU)
Khazar Abad road, Sari, Mazandaran, Iran P.O.Box: 578

www.jpmb-gabit.ir Tel : +981133687744