Cloning and comparative modeling identifies a highly stress tolerant Cu/Zn cytosolic Super Oxide Dismutase 2 from a drought tolerant maize inbred line

Anuradha Gautam
ICAR- Indian Institute of Maize Research
Fatima Nazish Khan
Indian Agricultural Research Institute
Surabhi Priya
ICAR- Indian Institute of Maize Research
Krishan Kumar
ICAR- Indian Institute of Maize Research
Shivani Sharda
Amity University
Tanushri Kaul
International Centre for Genetic Engineering and Biotechnology
Ishwar Singh
ICAR- Indian Institute of Maize Research
Sapna Langyan
ICAR- National Bureau of Plant Genetic Resources
Pranjal Yadava (pranjal.yadava@icar.gov.in)
Indian Agricultural Research Institute

Article

Keywords: Zea mays L., superoxide dismutase, SOD2, drought stress, bioinformatic analysis

Posted Date: May 27th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1675452/v1

Abstract

Plants have a complex system of stress response that deals with different types of stresses. Maize (*Zea mays* L.), one of the most important crops grown throughout the world, across a range of agro- ecological environments, employs complex mechanisms of gene regulation in response to drought stress. HKI 335 is a tropical maize inbred line showing remarkable adaptation to drought stress. Abiotic stresses, like drought, trigger the production of reactive oxygen species (ROS) due to the incomplete reduction or excitation of the molecular oxygen, eventually leading to cell damage. Superoxide dismutase (SOD, EC 1.15.1.1) is a metalloenzyme that acts as the first line of defense against ROS. We cloned the *Sod2* gene from HKI 335 inbred line and analyzed its protein through detailed *in silico* characterization. Our comparative modeling revealed that at the level of tertiary structure, the HKI 335 SOD2 protein is highly similar to *Potentilla atrosanguinea* SOD2, which had been previously identified as highly thermostable SOD that can tolerate autoclaving as well as sub-zero temperatures. We performed phylogenetic analysis, estimated physico-chemical properties, post- translational modifications, protein- protein interactions, and domain composition of this SOD2.

Introduction

To meet the increasing demand due to rising population and economic development, the production of crops should maintain a sustainable growth [1, 2]. However, environmental stresses such as drought, have become a major limitation in crop production, under unpredictable weather and changing climate patterns [3]. Maize (*Zea mays* L.), is the most produced and consumed grain of the world, with diverse end uses as food, feed, fuel and industrial products [4]. It is grown worldwide across a range of agro-ecological environments, and is therefore prone to abiotic stresses, like drought. Although, the productivity of maize worldwide is high, yet the crop is highly sensitive to drought [5, 6]. Drought has been an ecological crisis throughout the world and is one of the major stresses that limits crop productivity [7]. It is also considered as the major serious environmental factor that limits the productivity of maize, especially under rain-fed conditions [8]. The maize crop is susceptible to drought at different growth stages such as grain-filling, pre-flowering and seedling [9]. Therefore, it is essential to understand the mechanism of drought-tolerance and breed drought-tolerant varieties [10].

Abiotic stress such as drought, extreme temperatures, high salinity, excessive light, pollutants such as ozone and herbicides, high concentrations of heavy metals, excessive UV radiation, and others [5] trigger the production of reactive oxygen species (ROS) due to incomplete reduction (hydrogen peroxide $-H_2O_2$; superoxide radical $-O^{\bullet -2}$; hydroxyl radical $-HO^{\bullet}$, etc.) or excitation (singlet oxygen $-{}^1O_2$) of molecular oxygen, eventually leading to cell damage [6]. Superoxide dismutase (SOD, EC 1.15.1.1) is a metalloenzyme that acts as the first line of defense against ROS. SODs play a major role in detoxification of ROS. SODs catalyze the dismutation of the superoxide radical ($O2^-$) into oxygen and hydrogen peroxide (H_2O_2). All aerobic organisms that are prone to oxidative stress require SODs to dismutase O_2 to yield H_2O_2. SODs are categorized in accordance with their metal cofactors and subcellular distribution: mitochondrial manganese SOD (MnSOD), cytosolic copper/zinc SOD (Cu/ZnSOD), and iron SOD (FeSOD) [5]. Maize contains 9 SOD isozymes viz. four Cu/Zn cytosolic isozymes (SOD2, SOD4, SOD4A and SOD5) four mitochondrial associated MnSODs (SOD3.1, SOD3.2, SOD3.3 and SOD3.4) and one Cu/Zn chloroplast-associated isozyme (SOD-1). There are many reports stating that SOD2 overexpressing transgenic plants show enhanced tolerance to oxidative stresses such as drought.

In our earlier screens, HKI 335 maize inbred line (derived from Pool 10, All India Coordinated Research Project on Maize, Karnal centre, India) was found to be a highly drought tolerent line [11–13]. Previously, we reported that enhanced SOD activity in maize could be a key strategy employed by the plant to mitigate abiotic stress, as evident by transcriptional modulatiion of antioxidant genes during methyl viologen induced oxidative stress [14–20]. This prompted us to clone and analyze Cu/Zn cytosolic *Sod2* from the highly drought tolerant line HKI 335.

Results

Cloning of *Sod2* from HKI 335 maize inbred line

The *ZmSod2* gene was amplified from cDNA using gene specific primers and cloned in standard plasmid vector, followed by sequencing. The *ZmSod2* HKI 335 sequence was deposited in the GenBank database of National Centre for Biotechnology Information (NCBI), USA with accession number ALF00121.1. The cloned sequence (151 amino acid) contained all the characteristic conserved domains of Cu-Zn Superoxide Dismutase superfamily (PLN02386 family protein) (Figure S1).

Phylogenetic Analysis

The multiple sequence alignment shows aligning of different plant species for stress tolerance (Figure 1). Orthologous sequences of SOD2 from different species were aligned and further used for phylogenetic analysis using the neighbour joining (NJ) method. The results show that SOD2 in different plant species demarcated into two prominent clusters i.e., monocots (6) and dicots (5) (Figure 2).

Secondary Structure and Comparative Modeling

Using an online bioinformatics tool known as SABLE, we predicted the secondary structure of the target protein sequence (Figure 3). The secondary structure of the protein sequence viewed by polyview-2D show the absence of alpha-helices. At the same time, there are a large number of coils and beta-sheets with fully exposed relative solvent accessibility (RSA). The result indicate that it is a soluble protein and no transmembrane domains have been found (Figure S2). Most of the beta-sheets have RSA value greater than 2, which shows that the amino acids in the protein secondary structure are completely buried. The high value of the RSA of the target protein suggests that the protein is stable for a number of biological functions within the cell. The theoretical structure of SOD2 was generated using MOD web by comparative modeling of protein structure prediction through which identical or non- identical information about the target sequences is analyzed. We developed 14 models calculated for SOD2 out of which two models were selected. The model with LONGEST_DOPE had a score of -1.834 by using Modweb which was dynamically refined and validated. Comparative modeling revealed that the cloned SOD2 exhibits 83% identity with *Potentilla atrosanguinea* SOD2 (PDB ID: 2Q2L_A) with an e-value of 4e-86 (Figure 4).

Potentilla atrosanguinea, commonly known as Himalayan cinquefoil or ruby cinquefoil, is a vigorous herbaceous perennial of the rose family native to mountain slopes at lower elevations in the Himalayas. Surprisingly, SOD2 from *Potentilla atrosanguinea* have been previously reported as a highly thermostable SOD that can tolerate autoclaving as well as sub-zero temperatures [US Patent US20070269811A1; 19]. Further, a single C95A amino acid substitution in the wild type *Potentilla atrosanguinea* SOD2 is known to enhance thermal properties [20]. These findings prompted us to further investigate the SOD2 from HKI 335 maize inbred.

Model Validation

With the PDBsum, it has been shown that generated model of SOD2 protein revealed (~98.0% expected):193(84.6%) residues falling in most favoured region, (~2.0% expected) : 34 (14.9%) residues in additionally allowed region, and 0.4% residues in generously allowed region with no residues in the disallowed region of the Ramachandran plot (Figure 5). Z-score of PROSA energy indicating overall model quality was used to check 3D models of protein structures for potential errors. In these plots displaying Z-scores, value (-6.51) of the target model was determined by X ray crystallography (represented in light blue) and nuclear magnetic resonance (represented in dark blue). This value was extremely close to the value of template 2Q2L (-6.44) (Figure 6). Root Mean Squared Deviation (RMSD) value indicates the degree to which the two three dimensional structures are similar. RSMD analysis of the SOD2 model was measured from its template (2Q2l_A). The Cα RMSD and backbone RSMD deviation for both the target and the template was 0.27 Å (Figure 7).

Physiochemical Properties

The target protein consists of very high percentages of glycine (18.5%), alanine (7.9%) and valine (9.9%) in comparison to other amino acids (Table 1). The high percentage of glycine (18.5%) shows that triple helical structure of the protein is likely to be more stable. Like that of glycine residues in protein, the proline also plays a major role in stability of helix of protein's secondary structure. Here, the residues of proline show less percentage (5.3%). On the other hand, the percentage of glycine, valine and leucine among the hydrophobic groups are 32.56% and 17.44% respectively, while serine and threonine (both hydrophilic group) are 20.51% and 28.21% respectively. The isoelectric point (pI) value was 5.43, while the instability index was 20.1 (Table 2). The aliphatic index was 80.66, which is extremely high. The value of grand average of hydropathy (GRAVY) was -0.132 that lies between -2 and +2, suggesting that the protein is hydrophobic in nature and rated positively.The atomic composition of the protein is shown in Figure 8.

Table 1

Amino acid composition of SOD2, *Zea mays,* HKI 335

Amino Acid (AA)	AA	AA (Number)	AA (%)	Hydrophobic Group (%)	Hydrophilic Group (%)
Ala	A	12	7.9	13.95	
Arg	R	3	2		
Asn	N	6	4		15.38
Asp	D	11	7.3		
Cys	C	2	1.3		5.13
Gln	Q	3	2		7.69
Glu	E	6	4		
Gly	G	28	18.5	32.56	
His	H	9	6		23.08
Ile	I	8	5.3	9.3	
Leu	L	9	6	10.47	
Lys	K	6	4		
Met	M	2	1.3	2.33	
Phe	F	4	2.6	4.65	
Pro	P	8	5.3	9.3	
Ser	S	8	5.3		20.51
Thr	T	11	7.3		28.21
Trp	W	0	0		0
Tyr	Y	0	0		0
Val	V	15	9.9	17.44	
Pyl	O	0	0		
Sec	U	0	0		
	B	0	0		
	Z	0	0		
	X	0	0		

Table 2

Physio-chemical properties of SOD2

Number of AA	Molecular weight	Theoretical pI	Formula	Negatively charged residues (Asp + Glu)	Positively charged residues (Arg + Lys)	Extinction coefficients	Instability index	Aliphatic index	GRAVY
151	15103.76	5.43	$C_{650}H_{1033}N_{193}O_{214}S_4$	17	9	125	20.1	80.66	-0.132

Post Translational Modification (PTM) Site Prediction

The post translational modification (PTM) is described as amino acid modification on the basis of protein sequence and also considered as vital issues for regulating the physiological and biological functions inside the cell. We could identify only one PTM site (Phosphorylation), located at position 20 having residue S (sequence: GTDVKGTIFFSQEGDGPTTVT) with score 0.692.

Protein-Protein Interaction (PPI) Analysis

It has been observed that the interacting partner proteins with query protein are *IDP712* (uncharacterized LOC100283786 in *Zea mays*), *sod3*, *pco095461* (Superoxide dismutase in *Zea mays*), *GRMZM2G139680_P01* (2-cys peroxiredoxin BAS1 in *Zea mays*), *SODA.3*, *541646*, *GRMZM2G157018_P01* (ATP synthase subunit d, mitochondrial in *Zea mays*), *GRMZM2G125151_P01* (uncharacterized LOC100283203 in *Zea mays*) and *PER1* (period circadian regulator 1 in humans) with similarity scores 0.979, 0.883, 0.852, 0.847, 0.847, 0.837, 0.837, 0.835, 0.789 and 0.761 respectively (Figure 9, Table 3). The process and function of SOD2 has also been compared with other proteins (Table 4). Other analysis such as protein interaction descriptions and other parameters are furnished in supplementary files (Table S1 and Table S2)

Table 3

Interacting proteins and their similarity scores with our query protein SOD2

Interacting Partner	Score
IDP712	0.979
sod3	0.883
pco095461	0.852
GRMZM2G139680_P01	0.847
1E+08	0.847
SODA.3	0.837
541646	0.837
GRMZM2G157018_P01	0.835
GRMZM2G125151_P01	0.789
PER1	0.761

Table 4

Process and function of SOD2 compared with other proteins from protein-protein interaction

Process					
#term ID	term description	observed gene count	background gene count	strength	false discovery rate
GO:0098869	cellular oxidant detoxification	4	14	3.01	6.57E-10
GO:0019430	removal of superoxide radicals	3	5	3.33	6.60E-09
GO:0006950	response to stress	4	68	2.32	1.29E-08
GO:0055114	oxidation-reduction process	4	116	2.09	6.96E-08
Functions					
GO:0016209	antioxidant activity	4	14	3.01	1.29E-10
GO:0004784	superoxide dismutase activity	3	5	3.33	5.85E-09
GO:0016491	oxidoreductase activity	4	107	2.13	6.25E-08
GO:0046872	metal ion binding	3	136	1.9	1.50E-05

Domain Composition Prediction

Domains play a significant role in the functional activities of proteins in the cell, and are either active regularly or during the process of evolution. The domain present in SOD2 protein sequence is Sod_Cu, which is metalloprotein that helps in the prevention of damage by free radicals with the catalysation of SOD into O^2 and H_2O_2. Two outliers homologous and homologous of known structures are also found (Table

S3). The fragment GFHVHALGDTT was identified as the plant Dismutase signature motif (42-52 residues). This was a consensus pattern in both target sequences of *Zea mays* (HKI 335) and template *Potentilla atrosanguinea*. GNAGGRUACGII signature motif was identified from heme ligand site which was conserved in both SOD2 of target and template sequences**.**

Discussion

Plants are very sensitive to number of environmental stresses such as drought, salinity and cold that can cause extreme and rapid ROS accumulation [21, 22]. Plants also have important enzymatic and non-enzymatic mechanism that regulate oxidation process and help cells to protect from oxidative damage by ROS scavenging [23]. The cloning and sequencing of SOD2 from drought tolerant maize inbred line HKI 335 shows that the length of sequence is 151 amino acids with a theoretical isoelectric point of 5.43 and a predicted molecular weight of 15103.76 kD. The utilization of BLASTp tool of NCBI we have done the homologous search of target protein with the template/reference sequences. From this analysis, best 10 homologous (reference) protein sequences on the basis of numbers of criteria like percent identity (> 80%) were selected and utilized for further analysis of the protein. The analysis of amino acid sequence demonstrated that SOD_Cu were high conserved domains in the SOD2 sequence. The phylogenetic analysis of SOD2 of maize showed that orthologous sequences of SOD from different species were clustered into two clusters (i.e., cluster A and cluster B). Previous study showed that there were two Cu/ZnSODs in eukaryotes, including intracellular and extracellular Cu/ZnSODs, encoded by two different genes [24].

In secondary structure of SOD2, we found absence of alpha-helices and large number of coils and beta-sheets with fully exposed relative solvent accessibility, showing that the amino acids in the protein secondary structure are completely buried. The high value of the RSA of the target protein suggests that the protein is stable for a number of biological functions within the the cell. The Ramachandran plot from PDBsum has revealed that most of the residues falling in highly favoured region and some are in additional allowed region, very less in generously allowed region with no in the disallowed region. Independently, the high percentages of these amino acids in protein regulate the synthesis of protein as well as pathways involved in signalling. Alanine with highest percentage considered as a major component of cell wall biosynthesis. Valine is involved in increase the glutelin, globulin, and albumin fractions in maize kernels. The high percentage of glycine shows that triple helical structure of protein is more stable. Like that of glycine residues in protein, the proline also plays a major role in stability of helix of protein's secondary structure. In this study, the residues of proline show less percentage that is lower than 10, indicates less efficient for stability of protein structure. For perfect three-dimensional structure of proteins, hydrophobic as well as hydrophilic group play important roles. Among them, the percentage of glycine and valine and leucine (both hydrophobic group) are 32.56% and 17.44% respectively, while serine and threonine (both hydrophilic group) are 20.51% and 28.21% respectively, are majorly involved. The pI value was lesser than 7, indicating that the target protein is acidic in nature. It was found that the value of instability index was less than 40, which shows that the protein is stable. The value of aliphatic index was found to be extremely high, and therefore the target protein is considered as thermostable protein, which means that the target protein SOD2 is resistant to decay at high temperature. This is further affirmed by comparative modeling, which led us to highly thermo-tolerant SOD2 from *Potentilla atrosanguinea.* The value of GRAVY lies between − 2 and + 2, suggesting that the target is hydrophobic in nature and rated positively. Therefore, during the cellular processes, this protein decreases the region of linking between non-polar and water molecules, and also utilize the hydrogen bonding between water molecules within the cell. There are many reports establishing that tolerance to drought and salt or drought and cold stress can be improved by exotic gene transferring from plants [25–27].

Overall, the unique stress tolerant properties of HKI 335 SOD2, may be one of the reasons contributing to the highly drought tolerance trait exhibited by HKI 335 maize inbred line. Further work on HKI 335, especially in relation to SOD2, like *in vitro* characterization of the protein, overexpression, and allele mining, may lead to greater insights on the drought adapatation mechanism in maize and eventual deployment of the trait in maize hybrids.

Materials And Methods

RNA isolation

Leaf tissue was sampled from 21-day old maize seedlings (*Zea mays,* HKI 335 inbred line) and RNA was isolated using RNA isolation kit (Genetix Life Sciences). Quality and concentration of RNA was checked on Nanodrop Spectrophotometer. 1µg of RNA was used for cDNA synthesis using oligodT primers.

PCR amplification

Gene specific primers were designed and synthesized for maize (*Sod2*) and amplication was done using polymerase chain reaction (PCR). The amplified product was ligated into pJET vector and transformed into competent *E. coli* cells (DH5α strain). Colony PCR was done to identify transformed colonies and plasmid was isolated. This plasmid was sent for sequencing. BLAST analysis of sequencing results

confirmed our sequences as SOD2 mRNA which were then submitted in NCBI (ALF00121.1). The sequences were retrieved for further bioinformatic analysis.

Sequence Search

The target sequence was searched for similar sequence using the BLAST (Basic Local Alignment Search Tool) (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) against Protein Database (PDB) [28]. Only full-length protein sequences were considered for *in silico* analysis.

Multiple Sequence Alignment and Phylogenetic Analysis

Orthologous sequences of SOD2 were searched using BLAST and multiple sequence alignment was done using Clustal W (https://www.genome.jp/tools-bin/clustalw), in order to infer the phylogenetic relationship between members of the antioxidant gene family [29]. The final alignment was completed through manual editing. The FASTA multiple sequence alignment was used to infer the neighbor-joining (NJ) phylogenetic tree using the MEGA 6.0 software [30]. The bootstrap interior branch test, superior to bootstrap analysis in testing confidence in a given tree node, was applied to test the statistical significance of the SOD2 of NJ phylogenetic tree.

Secondary Structure Prediction

The target sequence in FASTA format have been used for prediction of secondary structure of protein. With the utilization of online bioinformatics tool known as SABLE (http://sable.cchmc.org/), one can identify the secondary structure of target protein sequence [31]. This tool generally predicts real valued relative Solvent AccessiBiLitiEs of amino acid residues in protein. It uses evolutionary profiles for improved prediction of secondary structure. The result of secondary structure was visualized using POLYVIEW-2D.

Modelling of Protein's Three-Dimensional (3D) Structure

The 3D structure of the protein was identified using online bioinformatics tool called MODWEB (https://modbase.compbio.ucsf.edu/modweb/) [32]. It modeled the structure using comparative protein structure modeling, which can calculate comparative models for a large number of protein sequences, using many different template structures and sequence-structure alignments.

Model Validation

The model was evaluated on the basis of geometrical and stereo-chemical constraints using PDBsum (http://www.ebi.ac.uk/pdbsum) [33] and ProSA-Web (https://prosa.services.came.sbg.ac.at/) [24]. PDBsum is used for validation of PDB structure through which stereochemical quality and accuracy of the predicted model is evaluated, which is based on Ramachandran plot calculation. Amino acid residues in proteins were examined by torsion angles Ø and Ψ and a percentage quality measurement of the protein structure was used, in which four sorts of occupancy were called core, allowed, generously allowed and disallowed regions.
SuperPose (http://superpose.wishartlab.com/) was used for protein superposition that calculates the superpositions using a modified quaternion approach [35]. We calculated Root Mean Squared Deviation (RMSD) to identify the distance between two objects by superposing the two structures and is performed between equivalent atom pairs

Physiochemical Properties

The physiochemical properties of the target protein such as molecular weight, number of amino acids, theoretical isoelectric point (pI), total number of negatively charged (Asp + Glu) and positively charged (Arg + Lys) residues, amino acid and atomic composition, extinction coefficient, estimated half-life, Grand average of hydropathicity (GRAVY), instability and aliphatic index are identified using online tool ExPASy's ProtParam (http://expasy.org/tools/protparam.html) [36].

PTM Site Prediction

For the prediction of PTM sites, an online bioinformatics tool PTM-ssMP (http://bioinformatics.ustc.edu.cn/PTM-ssMP/index/) was explored [37].

Protein-Protein Interaction Analysis

For predicting the function of query sequence with the reference sequence, protein-protein interaction (PPI) networks analysis has been done using online tool STRING (http://string-db.org/) [38].

Domain and Motif Identification

We identified the domain composition of target protein sequence by using the bioinformatics tool SMART (http://smart.embl-heidelberg.de/) [39]. It annotates and identifies the genetically mobile domains and analyses the domain structure. It can also detect extracellular, signalling and chromatin associated proteins bearing more than 500 domain families. ScanProsite server (https://prosite.expasy.org/scanprosite/) have been used for identification of signature motif in the query protein sequence [40].

Regulatory statement

All methods were performed in accordance with the relevant guidelines and regulations.

Declarations

Author contribution

AG cloned the reported gene; FNK, SP and KK conducted *in silico* studies; SS guided AG and contributed in manuscript drafting; IS phenotyped HKI 335 line used in this study; PY, TK and SL contributed in funding acquisition and manuscript preparation; PY and IS designed and supervised the project; AG, FNK and PY wrote the manuscript, which was read, improved and accepted by all the authors.

Funding

This work was supported by a research grant from National Agricultural Science Fund of Indian Council of Agricultural Research (NASF/GTR-5004/2015-16/204) to PY and TK. FNK is supported by Junior Research Fellowship from Core Research Grant project of Science and Engineering Research Board (CRG/2019/005451) to PY. AG was supported by a Senior Research Fellowship from Indian Council of Agricultural Research (NPTC 3015).

Data availability

The cloned gene sequence has been deposited in NCBI (Accession # ALF00121.1). All other data reported here, has been furnished in tables and figures in the main text and supplementary materials.

Competing Interests

The authors declare that they have no conflict of interest.

References

1. Brown, M. E., & Funk, C. C. Food security under climate change (2008).
2. Liu, X., Zhang, X., Sun, B., Hao, L., Liu, C., Zhang, D., ... & Li, Y. Genome-wide identification and comparative analysis of drought-related microRNAs in two maize inbred lines with contrasting drought tolerance by deep sequencing. *PloS one.* **14**, e0219176 (2019).
3. Takeda, S., & Matsuoka, M. Genetic approaches to crop improvement: responding to environmental and population changes. *Nat Rev Genet.* **9**, 444-457 (2008).
4. Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., ... & Presting, G. G. The B73 maize genome: complexity, diversity, and dynamics. S*cience.* **326**, 1112-1115 (2009).
5. Lee, S. Y., Cheon, K. S., Kim, S. Y., Kim, J. H., Kwon, O. H., Lee, H. J., & Kim, W. H. Expression of SOD2 enhances tolerance to drought stress in roses. *Hortic Environ Biotechnol.* **61**, 569-576 (2020).
6. Rodrigues, V. A., Crusciol, C. A. C., Bossolani, J. W., Portugal, J. R., Moretti, L. G., Bernart, L., ... & Lollato, R. P. Foliar nitrogen as stimulant fertilization alters carbon metabolism, reactive oxygen species scavenging, and enhances grain yield in a soybean–maize rotation. *Crop Sci.* **61**, 3687-3701 (2021).
7. Bartels, D., & Sunkar, R. Drought and salt tolerance in plants. *Crit Rev Plant Sci.* **24**, 23-58 (2005).
8. Rao, G. J., Reddy, J. N., Variar, M., & Mahender, A. Molecular breeding to improve plant resistance to abiotic stresses. In Advances in plant breeding strategies: Agronomic, abiotic and biotic stress traits (pp. 283-326). Springer, Cham (2016).
9. Liang, C., Wang, W., Ma, J., Wang, J., Zhou, F., Li, W., ... & Huang, X. Identification of differentially expressed microRNAs of sunflower seedlings under drought stress. *Agron J.* **112**, 2472-2484 (2020).
10. Ferdous, J., Hussain, S. S., & Shi, B. J. Role of micro RNA s in plant drought tolerance. *Plant Biotechnol J.* **13**, 293-305 (2015).

11. Nepolean T, Singh I, Hossain F, Pandey N, Gupta HS. Molecular characterization and assessment of genetic diversity of inbred lines showing variability for drought tolerance in maize. *J Plant Biochem and Biotech.* **22**, 71-79 (2013).
12. Yadava P, Kaur P, Singh I. Exogenous application of ascorbic acid alleviates oxidative stress in maize. *Indian J Plant Physiol.* **18**, 339-343 (2013).
13. Yadava P. Salicylic acid alleviates methyl viologen induced oxidative stress through transcriptional modulation of antioxidant genes in Zea mays L. *Maydica.* **60,** 1-9 (2016).
14. Singroha, G., Sharma, P., & Sunkur, R. Current status of microRNA-mediated regulation of drought stress responses in cereals. *Physiol Plant.* (2021).
15. Lobell, D. B., Roberts, M. J., Schlenker, W., Braun, N., Little, B. B., Rejesus, R. M., & Hammer, G. L. Greater sensitivity to drought accompanies maize yield increase in the US Midwest. *Science.* **344**, 516-519 (2014).
16. Lee SY, Han BH, Noh EW, Kwak SS. Transfer of SOD2 or NDP kinase 2 genes into purebred lines of petunia. *Kor J Plant Biotechnol.* **36,** 144–148 (2009).
17. Lee SY, Han BH, Kim YT, Kim JS. Resistance of SOD2 transgenic petunia line to oxidative stress. *Kor J Plant Biotechnol.* **37**, 562–566 (2010).
18. Lee SY, Cheon KS, Kim SY, Kwon OH, Lee HJ, Kim WH, Yoo BS Enhanced resistance to sulfur dioxide gas in transgenic petunia by stacking both SOD2 and NDPK2 genes. *Kor J Hortic Sci Technol.* **34**, 154–162 (2016).
19. Kumar S, Sahoo R, Ahuja PS. Isozyme of autoclavable superoxide dismutase (SOD), a process for the identification and extraction of the SOD in cosmetic, food and pharmaceutical compositions. *US Patent No. 6,485950 B1* (2002).
20. Kumar A, Dutt S, Bagler G, Ahuja PS, Kumar S. Engineering a thermo-stable superoxide dismutase functional at sub-zero to> 50 C, which also tolerates autoclaving. *Scientific reports.* **2**, 1-8 (2012).
21. Zhou, L., Xu, L., Jiang, M. M., Liu, Y., & Chen, H. Y. Structural Characterization and Expression Analysis of SmCSD1 Gene in Eggplant (Solanum melongena). *Russ J Plant Physiol.* **66**, 461-468 (2019).
22. Leng, X., Wang, P., Zhu, X., Li, X., Zheng, T., Shangguan, L., & Fang, J. Ectopic expression of CSD1 and CSD2 targeting genes of miR398 in grapevine is associated with oxidative stress tolerance. *Funct Integr Genom.* **17**, 697-710 (2017).
23. Suzuki, N., & Mittler, R. Reactive oxygen species and temperature stresses: a delicate balance between signaling and destruction. *Physiol Plant.* **126**, 45-51 (2006).
24. Gill, S. S., & Tuteja, N. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol Biochem.* **48**, 909-930 (2010).
25. Han, D., Hou, Y., Wang, Y., Ni, B., Li, Z., & Yang, G. Overexpression of a Malus baccata WRKY transcription factor gene (MbWRKY5) increases drought and salt tolerance in transgenic tobacco. *Can J Plant Sci.* **99**, 173-183 (2018).
26. Li, Y., Zhang, H., Zhang, Q., Liu, Q., Zhai, H., Zhao, N., & He, S. An AP2/ERF gene, IbRAP2-12, from sweetpotato is involved in salt and drought tolerance in transgenic Arabidopsis. *Plant Sci.* **281**, 19-30 (2019).
27. Wu, J., Jiang, Y., Liang, Y., Chen, L., Chen, W., & Cheng, B. Expression of the maize MYB transcription factor ZmMYB3R enhances drought and salt stress tolerance in transgenic plants. *Plant Physiol Biochem.* **137**, 179-188 (2019).
28. Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., & Madden, T. L. NCBI BLAST: a better web interface. *Nucleic Acids Res.* **36**, W5-W9 (2008).
29. Thompson, J. D., Gibson, T. J., & Higgins, D. G. Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinform.* 2-3 (2003).
30. Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol.* **30**, 2725-2729 (2013).
31. Adamczak, R., Porollo, A., & Meller, J. Combining prediction of secondary structure and solvent accessibility in proteins. *Proteins.* **59**, 467-475 (2005).
32. Eswar, N., John, B., Mirkovic, N., Fiser, A., Ilyin, V. A., Pieper, U., ... & Sali, A. Tools for comparative protein structure modeling and analysis. *Nucleic Acids Res.* **31**, 3375-3380 (2003).
33. Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S., & Thornton, J. M. PDBsum: Structural summaries of PDB entries. *Protein Sci.* **27**, 129-134 (2018).
34. Wiederstein, M., & Sippl, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **35**, W407-W410 (2007).
35. Maiti, R., Van Domselaar, G. H., Zhang, H., & Wishart, D. S. SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Res.* **32**, W590-W594 (2004).

36. Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., & Bairoch, A. Protein identification and analysis tools on the ExPASy server. The proteomics protocols handbook, 571-607 (2005).
37. Liu, Y., Wang, M., Xi, J., Luo, F., & Li, A. PTM-ssMP: a web server for predicting different types of post-translational modification sites using novel site-specific modification profile. *Int J Biol Sci.* **14**, 946 (2018).
38. Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., ... & Mering, C. V. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607-D613 (2019).
39. Letunic, I., Doerks, T., & Bork, P. SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res.* **40**, D302-D305 (2012).
40. De Castro, E., Sigrist, C. J., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P. S., Gasteiger, E., ... & Hulo, N. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res.* **34**, W362-W365 (2006).

Figures

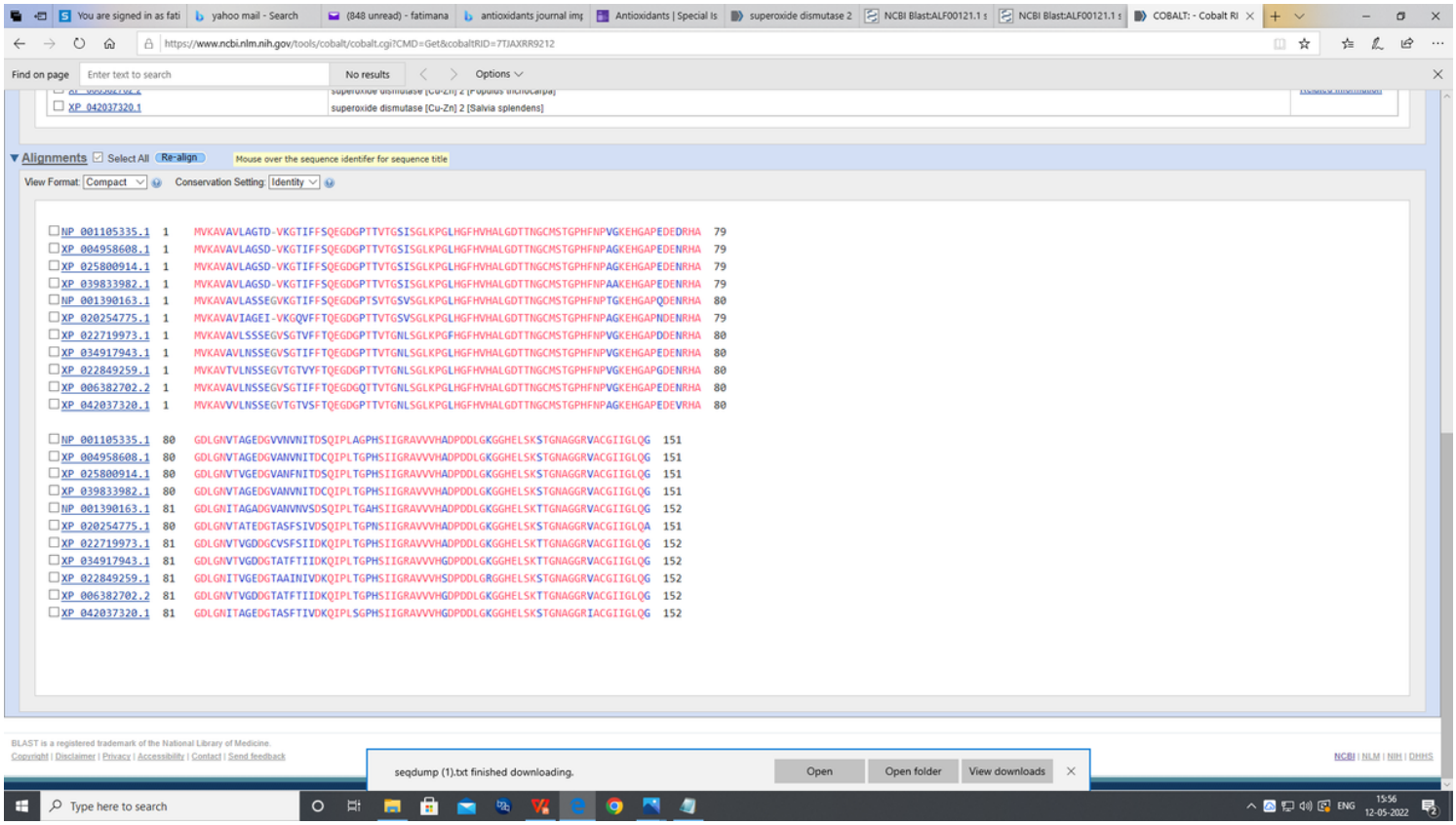

Figure 1

Multiple Sequences Alignment using Clustal W from different plant species for SOD2 sequences. Blue color letters indicate that there is substitution of amino acid at that particular position.

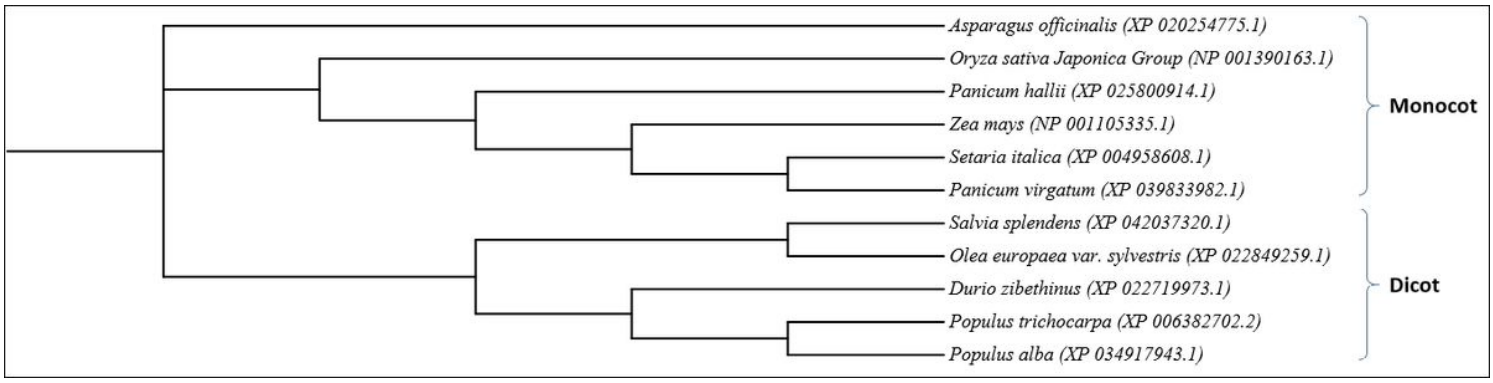

Figure 2

Phylogenetic analysis showing *Zea mays* SOD2 with other plants species SOD2 proteins from dicots and monocots.

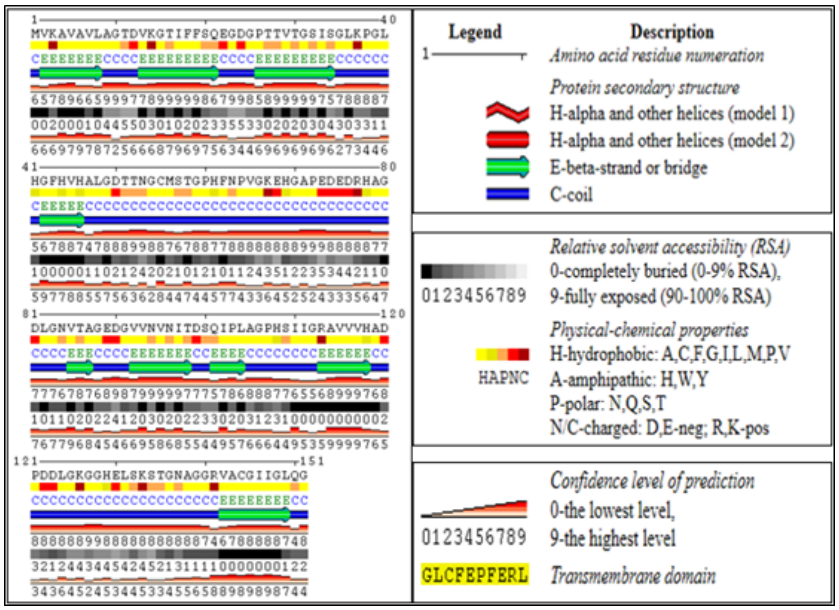

Figure 3

Secondary structure predictions for ZmSOD2 protein from drought tolerant inbred HKI 335. The colour coded onto the sequence are according to the sequence feature.

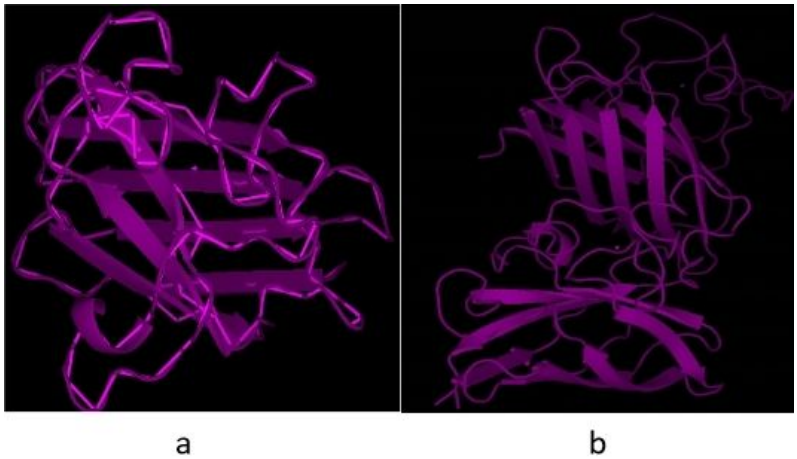

Figure 4

(a) Target structure of SOD2 from HKI 335 by Swiss model, (b) template SOD pdb structure (2Q2L).

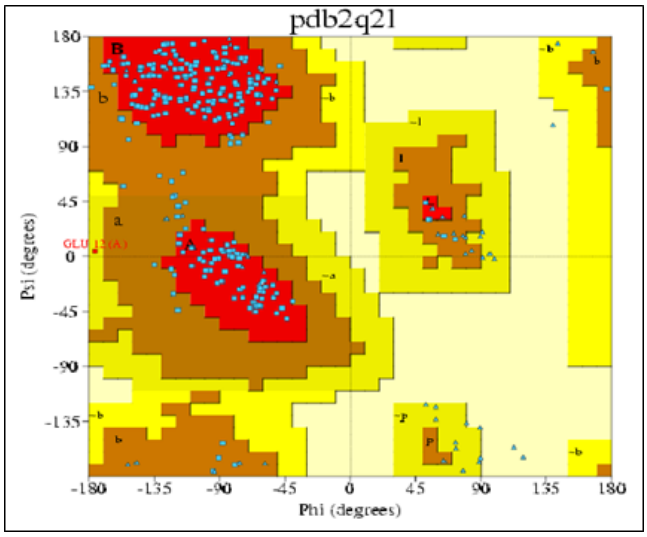

Figure 5

Ramachandran plot for phi and psi bond length for model validation for SOD2 (from drought tolerent HKI 335 maize inbred) designed by PDBsum program. All residures are in most favoured region.

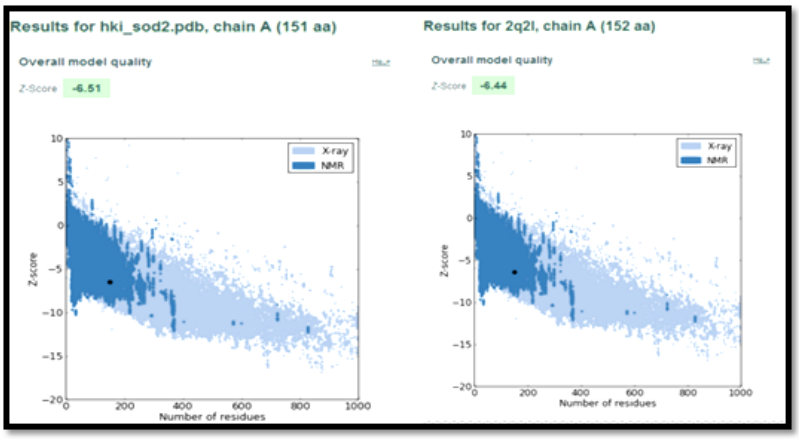

Figure 6

ProSA-web used for evaluation of Z-score value of both target and template structure.

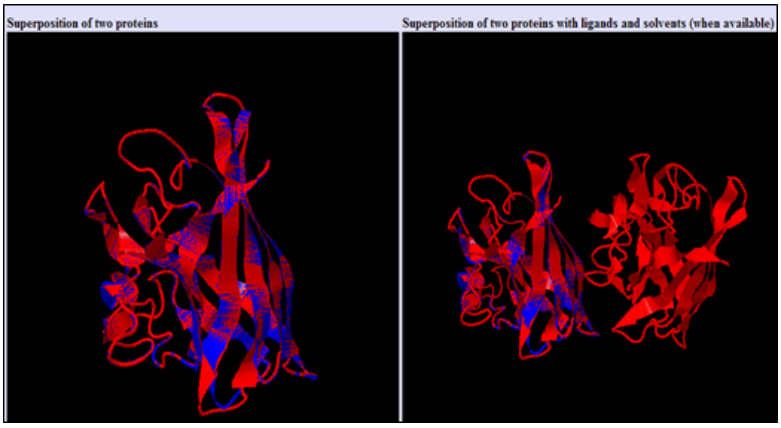

Figure 7

Superposed structure of SOD2 from ZmHKI 335 by swiss model modeling and crystal structure of putative SOD2 from *Potentilla atrosanguinea* by PDB. Superposition of Cα backbone of the target and the template was represented by blue and red colours. [Name of Chain_1: A22472, Name of Chain_2: B22472, Length of Chain_1: 151 residues; Length of Chain_2: 152 residues, Aligned length= 151, RMSD= 0.27; Seq ID = n_identical/n_aligned= 0.834; TM-score = 0.99714 (if normalized by length of Chain_1); TM-score = 0.99059 (if normalized by length of Chain_2); (should use TM-score normalized by length of the reference protein)]

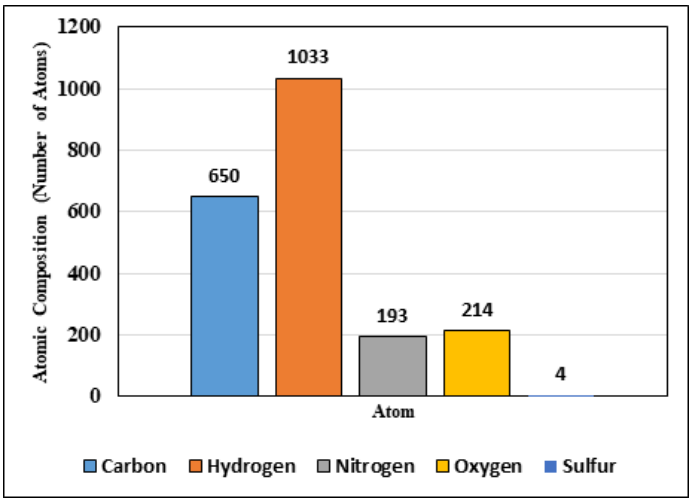

Figure 8

Atomic composition (i.e, number of atoms) of protein SOD2.

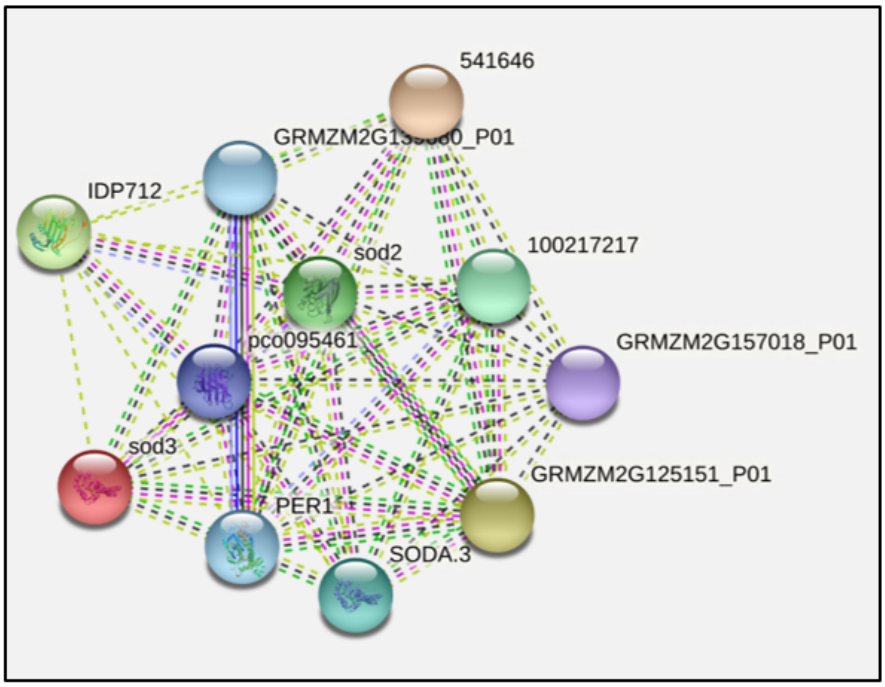

Figure 9

Protein-Protein Interaction network of SOD2 with other proteins

Supplementary Files

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

- SupplementaryFiguresGautametal2022.pdf
- SupplementaryTables.xlsx