ARTICLE; AGRICULTURE AND ENVIRONMENTAL BIOTECHNOLOGY

Cloning and allelic variation of two novel catalase genes (SoCAT-1 and SsCAT-1) in Saccharum officinarum L. and Saccharum spontaneum L.

Yang Liu a,b*, Yanli Yao a, Xiaowen Hu a, Shulian Xing a and Lei Xu a

aZhanjiang Experiment Station, Chinese Academy of Tropical Agriculture Science, Zhanjiang, Guangdong, P.R. China; bInstitute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Science, Haikou, Hainan, P.R. China

(Received 24 October 2014; accepted 10 February 2015)

Catalases (CAT) play important roles in plant defence mechanisms, stress response, cellular redox balance and aging delay. In order to clone CAT genes in cultivated (Saccharum officinarum L.) and wild (S. spontaneum L.) sugarcane, we designed two pairs of specific primers using a full length cDNA sequence of sorghum CAT gene (XM 002437586.1) as the probe. Two novel full-length sequences of CAT genes (SoCAT-1 in S. officinarum and SsCAT-1 in S. spontaneum) were cloned by nested polymerase chain reaction. Two allelic variants were found at SoCAT-1 (SoCAT-1a and SoCAT-1b, GenBank accession nos. KF864224-25) and SsCAT-1 (SsCAT-1a and SsCAT-1b, GenBank accession nos. KF864226-27) loci. The full lengths of SoCAT-1a, SoCAT-1b, SsCAT-1a and SsCAT-1b were 3816 bp, 3814 bp, 3777 bp and 3812 bp, respectively, and all contained eight exons and seven introns, with a similar gene structure. The length of cDNA for all four genes was 1532 bp, containing 10 bp 5'-untranslated region (UTR) and 43 bp 3'-UTR, and an open reading frame of 1479 bp encoding a polypeptide of 492 amino acids. High homology of DNA (99.0% for SoCAT-1, 98.2% for SsCAT-1) and cDNA (99.3% for SoCAT-1, 99.4% for SsCAT-1) were found. Eleven single nucleotide polymorphisms were found in cDNA SoCAT-1 and nine ones in SsCAT-1, and nine and five amino-acid mutation sites in the predicted proteins, respectively. The predicted proteins encoded by SoCAT-1 and SsCAT-1 showed high homology with CAT in sorghum, rice, corn and other species.

Keywords: Saccharum officinarum L.; Saccharum spontaneum L.; catalase; clone; PCR; allelic variation

Introduction

Sugarcane (Saccharum officinarum L.) is the most important source of sugar and energy in the world. As a tropical perennial crop with the largest distribution area in China, sugarcane is mainly planted on mountain slopes and hilly lands where the irrigation conditions are poor and there is intensive drought stress. [1] The breeding of drought-tolerant sugarcane varieties by mining the anti-drought genes is an effective means to solving this problem. Saccharum spontaneum L., a member of genus Saccharum, is a thinstalk wild variety and known for having strong adaptability, stress resistance and intensive tilling. [2,3] Saccharum spontaneum is not only the closest wild variety to sugarcane cultivars phylogenetically, but also the species of genus Saccharum with the highest sugar content. [4,5] Saccharum spontaneum has already been used in cross-breeding of sugarcane as a valuable breeding material. [6,7] Thus, it is expected that the stress resistance of sugarcane could be improved by mining the elite gene resources in S. spontaneum, using modern biotechnological methods.

Catalase (E.C.1.11.1.6; H2O2 oxidoreductase; CAT), is an iron porphyrin enzyme, widely present in animals, plants and microorganisms. CAT can scavenge the H2O2 generated by plant photosynthesis and mitochondrial electron transport by catalysing the electron transport and breaking down of H2O2 to H2O and O2. Thus, CAT plays important roles in plant defence mechanism, stress response, redox balance of cells and aging delay. [8,9] In plants, CAT enzymes are localized in peroxisomes, glyoxysomes and the cytoplasm and are encoded by multiple genes. The expression of CAT genes can be induced by non-biotic factors such as drought, low temperature, sunlight, salt stress and heavy metal stress. Different expression patterns of CAT genes have been reported under spatio-temporal changes. [10–13]

There are a number of previous studies in relation to CAT genes from sugarcane. For example, Su et al. [14] isolated one cDNA fragment of CAT genes by using homology-based cloning in sugarcane. To the best of our knowledge, however, there is no report on CAT genomic DNA in sugarcane or S. spontaneum. In this study, we used CAT gene sequences from sorghum (Sorghum bicolor L.) as a probe and homology-based cloning as the experimental strategy to explore CAT homologous genes in S. officinarum and S. spontaneum. Through molecular
techniques and bioinformatics analysis, this result may be used as the theoretical basis for breeding anti-stress sugarcane varieties with the aid of biological techniques.

Materials and methods

Plant materials

The test material, cultivated sugarcane variety ROC22 and S. spontaneum L. 82-114 were collected and identified by our research team. ROC22 is the most widely planted sugarcane variety in China and S. spontaneum 82-114 is a very good drought-tolerance material. The experiment was carried out on potted plants at the base of Zhanjiang Experimental Station, Chinese Academy of Tropical Agriculture, in May 2013. Routine management procedures were applied, and favourable water and fertilizer conditions were maintained. Two months after seedling emergence, the young, damage-free leaves were removed with scissors and placed immediately in liquid nitrogen and subsequently stored at −70 °C. The specimens were used for total DNA extraction.

Total DNA extraction

Total genomic DNA extraction was performed using plant total DNA extraction kit (Tiangen Biotech Co., Ltd, Beijing, China), according to the manufacturer’s instructions. The extracted genomic DNA was stored at −70 °C.

Gene-cloning strategy

Using a full length cDNA sequence of the CAT gene (XM 002437586.1) from sorghum as the probe, we searched for homologous sequences against an expressed sequence tag (EST) library of genus Saccharum, using the Basic Local Alignment Search Tool (BLAST) at the National Center for Biotechnology Information (NCBI).[15] EST sequences from Saccharum with higher identity with the CAT gene from sorghum were screened preliminarily. After further screening, several EST sequences containing the complete coding region of the CAT gene from sorghum were assembled to obtain the putative complete coding reference sequences for the cloning of CAT homologous genes from S. officinarum and S. spontaneum. Two pairs of nested primers were designed according to the reference sequences including the start and stop codons. DNASTAR lasergene7.1 (http://www.dnastar.com/products/lasergene.php) was used to analyse the polymerase chain reaction (PCR) amplification products after cloning-based sequencing. Then the complete nucleotide sequences of CAT homologous genes from S. officinarum and S. spontaneum were obtained.

Primer design and synthesis

Two pairs of specific primers were designed according to the reference nucleotide sequences of CAT genes (Table 1). The primers were synthesized by Invitrogen Corporation (Shanghai, China). C1 was used for the first round of PCR amplification and C2, for the second round.

Nested PCR amplification and cloning

The extracted genomic DNA was used as a template to amplify the specific fragments. PCR procedures using primers C1 and C2 were as follows: 94 °C/5 min, 94 °C/30 s, 60 °C/30 s, 72 °C/120 s, 30 cycles, 72 °C/10 min and a final hold at 4 °C. PCR products (5 µL) mixed with 1 µL of loading dye buffer were loaded into 1% agarose gel wells for electrophoresis. The gel strips were analysed by an SC810A gel-imaging system (BIOTOP Co., Ltd, Shanghai, China) and recovered using a DNA recovery kit (TiaP., Ltd, Beijing, China), according to the manufacturer’s instructions. The recovered PCR products were cloned using the pEASY-T5 Zero Cloning Kit (TransGen Biotech (Beijing) Co., Ltd, Beijing, China) following the product manual. Two positive final clones were submitted to Invitrogen Corporation (Guangzhou, China) for sequencing. Taq 2X PCR Master Mix, DNA Marker, deoxyribonucleoside triphosphates (dNTPs) and DNA purification kit (Tiangen Biotech (Beijing) Co., Ltd, Beijing, China) were used.

Data analysis

Sequence alignment, protein prediction and evolutionary analysis of homologous sequences were performed using DNASTAR 6.0 software. A phylogenetic tree was

| Primer | Primer sequence (5’-3’) | Amplified region | PCR fragment size (bp) |
|--------|------------------------|-----------------|-----------------------|
| C1     | forward: AGGCGTCTGCTCTGCTCC reverse: ATAGGCGTAACAACCATTCT | 98-4040 | 3943 |
| C2     | forward: ACCGGCAGCCATGGATCCG reverse: CTGTATCCTCGGACACCGGC | 158-3988 | 3831 |

Note: The length of the amplified region was determined based on the cDNA sequence of SsCAT-1a.
constructed. The gene structure was drawn by DNAMAN version 6.0 software (http://www.lynnon.com/). Alignments of the obtained sequences and the predicted protein sequences were done by Blast search after accessing the NCBI website.[15] The molecular formula, molecular weight and isoelectric points (PI) of amino-acid residues were determined using ProtParam software (http://web.expasy.org/protparam/).[16] The hydrophobicity of proteins was analysed by ProtScale; the signal peptides and transmembrane domains were predicted by Interpro. The protein functions and 3D structures were predicted using the assumed protein sequences on SWISS-MODEL Repository (http://swissmodel.expasy.org).[17]

Results and discussion
Searching EST library of Saccharum with CAT gene from sorghum as a probe
The EST library of Saccharum was searched on the NCBI website [18] with the cDNA sequence of the CAT gene (XM 002437586.1) from sorghum. A number of Saccharum EST sequences related to the CAT gene were found, and five of them showed higher identity with the cDNA sequence of the CAT gene from sorghum (Table 2). These contained the complete coding region of the CAT gene (Figure 1). These five sequences were spliced and assembled to obtain a 2047 bp nucleotide sequence, which was used as the putative reference sequence for primer design in the cloning of CAT-1 homologous gene in S. officinarum and S. spontaneum.

The common techniques to obtain the target genes include map-based cloning, transposon tagging, differential display and homology cloning.[19] Although homology cloning is faster and has higher efficiency, there are stricter requirements on the amount and accuracy of information available on DNAs. Given the fact that sorghum is phylogenetically close to sugarcane,[20] richer DNA information can be harnessed using CAT genes from sorghum.

In this study, we used the cDNA sequence of the CAT gene from sorghum as a probe and identified five EST sequences from Saccharum EST library that had very high identity with the query sequence. These EST sequences were the putative sequences of CAT homologous genes used for the purpose of primer design. Generally, it is difficult to obtain ideal results for the PCR primer design if there is single-nucleotide polymorphism (SNP) or DNA variation. Here we designed the primers based on the EST sequences of the species close to our plant model species, S. officinarum and S. spontaneum, and obtained much more precise results.

PCR verification of CAT-1 and the sequence structure
When nested PCR amplification of genomic DNAs of S. officinarum ROC22 and S. spontaneum 82-114 was performed with two pairs of primers, C1 and C2 (Table 1), in the second round of PCR amplification, a single band appeared with a size of about 3800 bp (Figure 2). The PCR products with primer C2 were recovered, cloned and sequenced, and four DNA sequences were obtained. These sequences were further aligned with those downloaded from NCBI,[18] and these four DNA sequences including the start and stop codons were assembled. They were denoted as SoCAT-1a, SoCAT-1b, SsCAT-1a and SsCAT-1b, respectively. These four sequences were submitted to GenBank under accession numbers KF864224–27.

Table 2. Characterization of five sugarcane ESTs.

| GenBank accession number | Length (bp) | Position | Similarity (%) | Expected value (E) |
|-------------------------|------------|----------|----------------|-------------------|
| CN606452.1              | 405        | 37–441   | 92             | 6e–163            |
| CA160849.1              | 711        | 143–842  | 96             | 0                 |
| CN606452.1              | 801        | 529–1325 | 96             | 0                 |
| CA131322.1              | 886        | 873–1768 | 94             | 0                 |
| CN606452.1              | 692        | 1385–2072| 91             | 0                 |
The full lengths of SoCAT-1a, SoCAT-1b, SsCAT-1a and SsCAT-1 were 3816 bp, 3814 bp, 3777 bp and 3812 bp, respectively. They all contained eight exons and seven introns, with a similar gene structure (Figure 3).

The cDNA sequence length of SoCAT-1a, SoCAT-1b, SsCAT-1a and SsCAT-1 was 1532 bp, including 10 bp of 5’-untranslated region (UTR), 43 bp of 3’-UTR and a 1479 bp open reading frame (ORF) that encoded 492 amino acids.

The validity of this approach is confirmed by the PCR results. As shown by the search results, several Saccharum EST sequences having high identity with the sequences of CAT genes from sorghum were obtained. The homologous similarity of these Saccharum EST sequences with the sequences of CAT genes from sorghum was over 90%. With the sequences assembled into a full length coding sequence of the CAT gene, the reliability of the experiment was confirmed. Sequence alignment was done with the sequences of CAT genes from sorghum and other related species. Based on the structure of introns, the positions of exons in DNA sequences of CAT genes were predicted, and the primers were further designed, with each primer bounded by complete exons.

The alignment of the cDNA sequences of the four genes made using MegAlign software revealed 11 SNP sites in the coding regions of SoCAT-1a and SoCAT-1b. There were six C/T mutations, four A/G mutations and an A/T mutation (Table 5). The SNP sites in the coding regions of SsCAT-1a and SsCAT-1b were nine. Of these nine sites, six ones showed C/T mutations, two ones A/G mutations and a single one an A/C mutation (Table 6).

The phenomenon of allelic variation is common in crops such as rice, wheat and maize. Allelic variation occurs as a result of ancestral gene evolution and mutation. The direct effect of allelic variation is a change in the corresponding protein coding sequence, which may result in alterations in phenotype and function. However, some allelic variations occur in introns and have no influence on the protein-coding sequence of the gene. Introns may affect the gene transcript by modulating alternative splicing. In some genes, insertion/deletion changes in introns may result in changes in the enzymatic activity and thus, in the phenotype. Yau and Simon observed that the insertion of a 2.5 kb sequence in an intron of the soluble acid invertase (SAI) gene altered the transcription and resulted in a significant increase in sucrose content in carrots. Liu et al. also reported that a domain of ATTGA repeat sequences in the second intron of the Sai-1 gene in sweet sorghum might be an

### Table 3. Percent similarity of CAT-1 DNA sequences and cDNA sequences.

| DNA     | SoCAT-1a | SoCAT-1b | SsCAT-1a | SsCAT-1b |
|---------|----------|----------|----------|----------|
| cDNA    | 99.0     | 97.8     | 97.9     |          |
| SoCAT-1a| 99.3     | 97.4     | 97.7     |          |
| SoCAT-1b| 99.0     | 98.8     | 98.2     |          |
| SsCAT-1a| 99.0     | 98.8     | 99.4     |          |
| SsCAT-1b| 99.0     | 98.8     | 99.4     |          |
important regulatory site, and that the functional marker SBXI was closely related to the Brix value.

To the best of our knowledge, there are no reports on the relationship between allelic variants of the CAT gene and agronomic traits. In this paper, we found two insertion/deletion sites in SoCAT-1 and 53 ones in SsCAT-1; especially, there was a 21 bp deletion in the first intron of SsCAT-1. Analysis of AFE (alternative first exons)-containing genes in rice and Arabidopsis indicates that AFEs have multiple functions, from regulating gene expression to generating protein diversity.[27] However, whether functional markers exist in CAT genes needs further investigation.

Protein variation analysis and function prediction
The four sequences encoded proteins that were composed of 492 amino acids. The predicted molecular weights of the four proteins were 56.80, 56.63, 56.81 and 56.77 kDa, and their pI values were 6.92, 6.70, 7.14 and 7.02, respectively.

Further sequence alignment showed that the amino-acid sequences of the proteins encoded by SoCAT-1a and SoCAT-1b both had nine variation sites (Figure 5). There was an S/P mutation in position 24, a Y/H mutation in position 45, a V/A mutation in position 64, an L/F mutation in position 143, an I/V mutation in position 195, another V/A mutation in position 222, a Q/L mutation in position 413, an R/G mutation in position 442 and a K/E mutation in position 481. Of these, a leucine to phenylalanine substitution at residue 143 occurs in the predicted amino-acid sequence of the protein encoded by SoCAT-1b. Since both these amino acids are non-polar and hydrophobic, with similar physical and chemical properties, this mutation could be expected to have little effect on the enzymatic activity of the protein. However, the mutation leading to a serine to proline substitution at residue 24 changes a polarized neutral amino acid to a non-polar alkaline one. Five variation sites were detected in the amino-acid sequences of the proteins encoded by SsCAT-1a and SsCAT-1b (Figure 6). There was a T/A mutation in position 97, an L/M mutation in position 149, a P/S mutation in position 219, an R/C mutation in position 247 and an L/F mutation in position 380. These mutational changes may alter the protein structure, but it is inconclusive whether the enzymatic activity would also change.

In terms of natural selection, genes tend to be conserved in many species and exist as members of a specific

Table 4. Type and number of variation sites in SoCAT-1a and SoCAT-1b and in SsCAT-1a and SsCAT-1b.

| Type of variation sites                                      | Numbers of variation sites |
|--------------------------------------------------------------|----------------------------|
| SoCAT-1a/SoCAT-1b                                           | SsCAT-1a/SsCAT-1b          |
| Total number of sites (excluding sites with gaps/missing data) | 3814                       | 3768                       |
| Sites with alignment gaps or missing data                    | 2                         | 53                        |
| Invariable (monomorphic) sites                               | 3776                       | 3701                       |
| Singleton variable sites                                     | 38                         | 67                         |
family. CATs encompass a large family of proteins, which can be found in almost every oxygen-requiring microorganism, animal and plant. There are three members identified in *Arabidopsis thaliana* [28], *Nicotiana tabacum* [29] and *Zea mays* [30]; and two in *Hordeum vulgare* [31] and *Solanum lycopersicum*. In the catalase gene family, different members encode distinct catalase enzymes. In this context, *SoCAT-1a*, *SoCAT-1b*, *SsCAT-1a* and *SsCAT-1b* could most likely be considered as two pairs of alleles from a member of the CAT gene family rather than from two members. In other words, there are at least two alleles in *S. officinarum* and *S. spontaneum*.

BLASTP search and protein function prediction performed for the four protein sequences through NCBI showed that the predicted protein contained a conserved structural domain of CAT binding to protoheme (16–486 amino acids) (Figure 7). The presence of such a domain indicates that this protein probably has the function of breaking down hydrogen peroxide.

Based on gene structure and sequence, Zmocky and Koller [34] divided CATs into three subgroups, namely,
mono-functional catalase, catalase—peroxidase (CAT—POD) and Mn-catalase. According to the structure of the catalytic centre, there are two CAT categories: (1) ferricoporphyrin-containing CAT, or Fe-CAT; (2) Mn-containing CAT, or Mn-CAT.[35] Existing studies revealed that most CATs in the organisms are Fe-CAT, while Mn-CAT only occurs in very few organisms. Sequence alignment showed that SoCAT-1a, SoCAT-1b, SsCAT-1a and SsCAT-1b obtained in our experiment were very likely to be Fe-CAT.

The functions and three-dimensional (3D) structure of the proteins encoded by SoCAT-1 and SsCAT-1 were predicted by SWISS-MODEL Repository (http://swissmodel.expasy.org).[17] The results showed that the protein sequences were aligned with the target protein 1m7sA from amino acid 14 to 491, and the sequence similarities of the two genes were 50.11% and 50.54%, respectively; the expected value and QMEAN Z-Score were -4.59 and -4.51, respectively. That is, the two proteins were predicted to be exactly identical in 3D structure, as shown in Figure 8.

Figure 6. Alignment of the predicted protein sequences of SsCAT-1a and SsCAT-1b.

Figure 7. Size and location of conserved protein domains encoded by SoCAT-1 and SsCAT-1.

Figure 8. Predicted tertiary structure of the proteins encoded by SoCAT-1 and SsCAT-1. Subunit A, B, C and D are shown in blue, green, yellow and red, respectively. Subunits A and C are associated so that their N-termini are overlapped. Subunits B and D form a similar intertwined dimmer.
Figure 9. Phylogenetic tree based on predicted amino-acid sequences of the proteins encoded by SoCAT-1a, SoCAT-1b, SoCAT-1a, ScCAT-1b and other orthologues.
Evolutionary analysis of CAT proteins

To perform evolutionary analysis of the predicted CAT proteins, BLASTP search was done at the NCBI website. [15] The results showed that the predicted amino-acid sequences of the proteins encoded by SoCAT-1a, SoCAT-1b, SsCAT-1a and SsCAT-1b had high homology (83%–98%) with CAT proteins from sorghum, maize, rice and wheat. From the phylogenetic tree, the four predicted proteins encoded showed closest relationship to their orthologue(s) from sorghum, but were distant from those in Picea glauca, cotton and other species. The species were grouped into two clusters based on CAT sequences. The proteins encoded by SoCAT-1a, SoCAT-1b, SsCAT-1a and SsCAT-1b belonged to the second cluster and showed high similarity (Figure 9).

Found extensively in prokaryotes and eukaryotes, CATs catalyse the decomposition of hydrogen peroxide into water and oxygen.[8,9] As it contains four small subunits (55–69 kDa) and binds to four protoheme IX (protoheme b) groups, CAT forms a tetramer complex.[34] In eukaryotes, CATs are localized in peroxisomes. In our study, the predicted structures and functions of the proteins encoded by SoCAT-1a, SoCAT-1b, SsCAT-1a and SsCAT-1b are in accordance with their expected biological role.

It is well known that CATs are related to stress resistance in plants, and their expression is induced by factors such as sunlight, temperature, cytokinin, salicylic acid and sulphur dioxide.[36] Therefore, more experiments need to be done to understand the function of the proteins encoded by SoCAT-1a, SoCAT-1b, SsCAT-1a and SsCAT-1b and to explore their potential for inclusion in breeding programmes focused on improving the drought tolerance in sugarcane.

Conclusions

In this study, two novel full-length genomic DNA sequences of CAT genes (SoCAT-1 in S. officinarum L. and SsCAT-1 in S. spontaneum L.) were cloned by nested PCR. Two allelic variants were found in SoCAT-1 and SsCAT-1, especially, there was a CTTCAAC-GATTCTGCCCTACT deletion in the first intron of SsCAT-1. The obtained results will help to gain deeper understanding of sugarcane catalase genes, and could contribute to the development of molecular markers in sugarcane.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was partially supported by the National Non-profit Institute Research Grant of CATAS-ITBB [grant number ITBB120504], [grant number ITBB140103].

References

[1] Chen RK, Xu LP, Lin YQ. Modern sugarcane genetic breeding. Beijing: China Agriculture Press; 2011. Chinese.
[2] Tai PYP, Miller JD. A core collection for Saccharum spontaneum. From the world collection of sugarcane. Crop Sci. 2001;41:879–885.
[3] Bremer G. Problems in breeding and cytology of sugarcane. Euphytica. 1961;10:59–78.
[4] Glaszmann JC, Fautret A, Noyer JL, Feldmann P, Lanaud C. Biochemical genetic markers in sugarcane. Theor Appl Genet. 1989;78:537–543.
[5] D’Hont A, Lu YH, Feldmann P, Glaszmann JC. Cytoplasmic diversity in sugarcane revealed by heterologous probes. Sugarcane. 1993;1:12–15.
[6] Mary S, Nair N, Chaturvedi PK, Selvi A. Analysis of genetic diversity among Saccharum spontaneum from four geographical regions of India, using molecular markers. Genet Resour Crop Evol. 2006;53:1221–1231.
[7] Pan YB. Highly polymorphic microsatellite DNA markers for sugarcane germplasm evaluation and variety identity testing. Sugar Tech. 2006;8:246–256.
[8] Schroder M, Fahimi HD. Peroxisomes and oxidative stress. Biochimica et Biophysica Acta. 2006;1763 (12):1755–1766.
[9] Zamocky M, Furtmüller PG, Obinger C. Evolution of catalases from bacteria to humans. Antioxid Redox Signaling. 2008;10(9):1527–1548.
[10] Purev M, Kim YJ, Kim MK, Pulla RK, Yang DC. Isolation of a novel catalase (Cat1) gene from Panax ginseng and analysis of the response of this gene to various stresses. Plant Physiol Bioch. 2010;48:451–460.
[11] Chen HJ, Wu SD, Huang GJ, Shen CY, Afifyanti M, Li WJ, Lin YH. Expression of a cloned sweet potato catalase SP/CAT1 eleviates ethephon-mediated leaf senescence and H2O2 elevation. J Plant Physiol. 2012;169:86–97.
[12] Du YY, Wang PC, Chen J, Song CP. Comprehensive functional analysis of the catalase gene family in Arabidopsis thaliana. J Integr Plant Biol. 2008;50:1318–1326.
[13] Guan ZQ, Chai TY, Zhang YX, Xu J, Wei W. Enhancement of Cd-tolerance in transgenic tobacco plants overexpressing a Cd-induced catalase cDNA. Chemosphere. 2009;76:623–630.
[14] Su YC, Guo JL, Ling H, Chen SS, Wang SS, Xu LP, Allan AC, Que YX. Isolation of a novel peroxisomal catalase gene from sugarcane, which is responsive to biotic and abiotic stresses. PLOS One. 2014;9(1):e84426.
[15] NCBI BLAST [Internet]. Rockville Pike: National Center for Biotechnology Information; c1993-2014 [cited 2014 Aug 12]. Available from: http://blast.ncbi.nlm.nih.gov/Blast.cgi/
[16] Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics. 2006;22(2):195–201.
[17] Rost B, Yachdav G, Liu J. The Predict Protein server. Nucleic Acids Res. 2004;32:W321–W326.
[18] NCBI [Internet]. Rockville Pike: National Center for Biotechnology Information; c1993-2014 [cited 2014 Oct 7]. Available from: http://www.ncbi.nlm.nih.gov/BLAST.cgi/
[19] Zhou GL, Guang S, Fu TD. [Gene cloning techniques]. J Huazhong Agric Univ. 2001;20(6):584–592. Chinese.
[20] Michael DG, Katrien MD. Comparative genetics in the grasses. Proc Natl Acad Sci. 1998;95:1971–1974.
[21] Jeppe RA, Thomas L. Functional markers in plants. Trends Plant Sci. 2003;8(11):554–560.
[22] Zhang W, Dubcovsky J. Association between allelic variation at the Phytoene synthase 1 gene and yellow pigment content in the wheat grain. Theor Appl Genet. 2008;116 (5):635–645.

[23] He XY, Zhang YL, He ZH, Wu YP, Xiao YG, Ma CX, Xia XC. Characterization of phytoene synthase 1 gene (Psy1) located on common wheat chromosome 7A and development of a functional marker. Theor Appl Genet. 2008;116 (2):213–221.

[24] Fedorova L, Fedorov A. Introns in gene evolution. Genetica. 2003;118:123–131.

[25] Yau YY, Simon PW. A 2.5 kb insert eliminates acid soluble invertase isozyme II transcript in carrot (Daucus carota L.) roots, causing high sucrose accumulation. Plant Mol Biol. 2003;53:151–162.

[26] Liu Y, Nie YD, Han FX, Zhao XN, Dun BQ, Lu M, Li GY. Allelic variation of a soluble acid invertase gene (SAI-1) and development of a functional marker in sweet sorghum [Sorghum bicolor (L.) Moench]. Mol Breeding. 2014;33:721–730.

[27] Chen WH, Lv GT, Lv CY, Zeng CQ, Hu SN. Systematic analysis of alternative first exons in plant genomes, BMC Plant Biol. 2007;7:55.

[28] Frugoli JA, Zhong HH, Nuccio ML, McCourt P, McPeek MA, Thomas TL, McClung CR. Catalase is encoded by a multigene family in Arabidopsis thaliana (L.) Heynh. Plant Physiol. 1996;112:327–336.

[29] Willekens H, Villarroel R, Van-Montagu M, Inze D, Van-Camp W. Molecular identification of catalases from Nicotiana plumbaginifolia (L.). FEBS Lett. 1994;352:79–83.

[30] Guan L, Scandalios JG. Developmentally related responses of maize catalase genes to salicylic acid. Proc Natl Acad Sci. 1995;92:5930–5934.

[31] Skadsen RW, Schulze-Lefert P, Herbst JM. Molecular cloning, characterization and expression analysis of two catalase isozyme genes in barley. Plant Mol Biol. 1995;29:1005–1014.

[32] Drory A, Woodson WR. Molecular cloning and nucleotide sequence of a cDNA encoding catalase from tomato. Plant Physiol. 1992;100:1605–1606.

[33] Kwon SI, An CS. Molecular cloning, characterization and expression analysis of a catalase cDNA from hot pepper (Capsicum annuum L.). Plant Sci. 2001;160:961–969.

[34] Zmocky M, Koller F. Understanding the structure and function of catalase: Clues from molecular evolution and in vitro mutagenesis. Prog Biophys Mol Bio. 1999;72 (1):19–66.

[35] Wang ZH, Shen L. Manganese catalase and their analogue study progress. J Hangzhou Teach Coll (Nat Sci Edition). 2006;5(6):465–468.

[36] Nan ZR, Fan YX. Advance of researches on catalase in plants. Anhui Agri Sci Bull. 2008;14(5):27–29. Chinese.