Theor Appl Genet (2003) 106:503–511
DOI 10.1007/s00122-002-1145-4

J.-C. Guan · X.-H. Li · Q.-F. Zhang · G. Kochert
C.-Y. Lin

Characterization of a unique genomic clone located 5′ upstream of the *Oshsp16.9B* gene on chromosome 1 in rice (*Oryza sativa* L. cv Tainung No. 67)

Received: 2 February 2002 / Accepted: 27 June 2002 / Published online: 15 November 2002

Abstract Small heat-shock proteins (sHSP) are the most abundant heat stress-induced proteins in plants. In rice, there are at least seven members of class-I sHSP. A 1.6-kb DNA fragment was isolated from the *Eco*RI-digested rice genomic library probed with the cDNA *pTS1* encoding a 16.9-kDa class-I sHSP. This fragment was composed of 365-bp tandem direct repeats (DRs) and 441-bp near perfect long terminal inverted repeats (LTIRs). The DRs contain 123-bp regions with 99% nucleotide identity to the 5′ coding region of the *Oshsp16.9B* gene. Two putative pseudogenes were deduced from the DRs. Using the LTIR as a specific probe, Southern-blotting analysis showed that there was a single copy of this 1.6-kb DNA fragment in the rice genome. By genomic walking, we located this fragment in proximity 5′-upstream of the *Oshsp16.9B* gene that was mapped on chromosome 1 with other two class-I sHSP genes, *Oshsp16.9A* and *Oshsp16.9C*. By comparative analysis of the nucleotide sequences of class-I sHSP genes clustered on chromosome 1 between Tainung No. 67 and Nipponbare cultivars, we confirmed our mapping results of these genes and only the promoter region of *Oshsp16.9B* was different. However, we found that the expression profile of *Oshsp16.9B* upon different heat stresses in Nipponbare was not significantly different relative to that in Tainung No. 67.

Communicated by R. Hagemann

J.-C. Guan · C.-Y. Lin (✉)
Department of Botany, National Taiwan University,
106 Taipei, Taiwan, R. O. C.
e-mail: chuyung@ccms.ntu.edu.tw
Tel.: (886-2-)-2363-0231 ext. 2675, Fax: (886-2-)-2363-8598

X.-H. Li · Q.-F. Zhang
National Key Laboratory of Crop Genetic Improvement,
Huazhong Agricultural University, Wuhan 430070, China

G. Kochert
Department of Botany, University of Georgia, Athens,
GA 30602, USA

Keywords Chromosomal location · *Oryza sativa* · Pseudogene · Small heat-shock proteins · Terminal inverted repeats

Introduction

In response to heat shock, plants synthesize abundant small HSPs (sHSPs) in the range of 15–30 kDa forming a diverse family that is related to the vertebrate α-crystallins (Key et al. 1981; Ingolia and Craig 1982; Waters et al. 1996). Based on the amino-acid sequence homology, immunological cross-reactivity and intracellular localization, five classes of sHSPs were identified in plants: class I and class II are in the cytosol, class III is in the chloroplast, class IV is in the endomembrane and class V is in the mitochondrion (Waters and Vierling 1999). The sHSP genes are all of nuclear origin (Waters et al. 1996). Among these classes, class-I sHSPs are the most abundant ones. Many data indicate that members of different sHSP families act as molecular chaperones in vitro (Jinn et al. 1989; Jinn et al. 1995; Yeh et al. 1997; Young et al. 1999; Lee and Vierling 2000) and in vivo (Löw et al. 2000). Properties shared by the sHSP family include high thermostability, a tendency to form multimers, and confer thermotolerance by their chaperone activity (Löw et al. 2000; van Montfort et al. 2001). Apart from heat stress, members of class-I sHSP genes are developmentally regulated in seeds, fruits and flowers in higher plants (Wehmeyer and Vierling 2000). Heavy metal ions (such as arsenite, cadmium) (Edelman et al. 1988; Tseng TS et al. 1993), chilling (Sabehat et al. 1998), oxidative stress (Pla et al. 2000), alcohol (Kuo et al. 2000) and amino-acid analogues (e.g. azetidine-2-carboxylic acid, Lee et al. 1996) were also reported to induce the expression of a subset of class-I sHSP genes in higher plants.

The highly conserved sequence of sHSP genes identified among a large number of plant species suggests that they may play an important role in coping with heat stress (Waters and Vierling 1999). In rice (*Oryza sativa* cv Tainung No. 67), five prominent hybridization bands were de-

tected in the *Eco*RI-digested genomic DNA gel-blot using rice *pTS1* cDNA encoding a 16.9-kDa class-I sHSP as a probe. We isolated and characterized six class-I sHSP genes from 9.5-, 5.9-, 3.4-, and 2.5-kb fragments of *Eco*RI-digested genomic DNA using the same probe (Tzeng SS et al. 1992; Lee et al. 1995; Chang et al. 2001). The 2.5-, 3.4- and 5.9-kb fragments were shown to contain a class-I sHSP gene, *Oshsp16.9*A (*pTS1* cDNA), *Oshsp16.9B* and *Oshsp16.9C*, respectively (Lee et al. 1995; Chang et al. 2001). In addition, another three class-I sHSP genes were found to cluster in the 9.5-kb fragment. They were *Oshsp17.3* (*pTS3* cDNA), *Oshsp17.7* and *Oshsp18.0* (*pYL* cDNA) (Guan et al. 1998). Based on two-dimensional gel analysis of in vitro hybrid-selection/translation products from the *pTS1* cDNA, we estimated that this class consisted of at least seven members in rice (Tzeng SS et al. 1993). Moreover, two-dimensional gel electrophoresis and Western-blotting analysis of the 310-kDa native heat-shock protein complex corroborated these observations (Yeh et al. 1995). Thus, we proceeded with the isolation of another member of the class-I sHSP gene family possibly localized in the 1.6-kb fragment. On the other hand, a tentative map of genes involved in the synthesis of sHSPs in cereals has been obtained by using nullisomic/tetrasomic mutants of wheat and a two-dimensional gel-electrophoretic separation of the in vivo labeled proteins (Potter et al. 1989). In an RFLP analysis of nullisomic and nullisomic/tetrasomic mutants of wheat, using the *Hvhsp17* from barley as a probe, homologous sequences were identified on chromosomes 3B and 3D (Marmiroli et al. 1993). Currently a total of six members of class-I sHSP genes have been characterized from rice, but the chromosomal locations of these genes are still unknown. As a cereal model plant, it is necessary to find out the chromosomal loci of these rice class-I sHSP genes.

Here, we report the isolation and characterization of a unique genomic clone from the 1.6-kb DNA fragment that cross hybridized with *pTS1*. We found that this fragment contains two tandem-arranged truncated sHSP pseudogenes, which were surrounded by several inverted repeats. Based on the PCR-based genomic walking analyses, this 1.6-kb fragment was located 5′ upstream to the *Oshsp16.9B* gene in close proximity to class-I sHSPs, *Oshsp16.9A* and *Oshsp16.9C* genes, which were mapped on chromosome 1.

Materials and methods

Plant materials

Rice (*O. sativa* L. cv Tainung No. 67) seedlings were germinated in the dark at 28 °C for 3 days in rolls of moist paper towels as described by Lin et al. (1984). Three-day old rice seedling-ablated endosperms were incubated in 5 mM of potassium phosphate buffer (pH 6.5) with 1% sucrose during treatments at control (28 °C) or heat-shock temperatures as indicated.

Preparation of rice genomic DNA and Southern-blot analysis

Total rice genomic DNA was isolated according to Malmberg et al. (1985). DNA was digested with restriction enzymes, separated on 0.8% agarose gels and transferred to Hybond-N extra membranes (Amersham, Buckinghamshire, UK). Filters were pre-hybridized in 50% formamide, 5 × SSC, 0.1% SDS, 20 mM of sodium phosphate pH 6.5, 0.1% Ficoll, 0.1% polyvinylpyrolidone, 1% glycine, 250 µg/ml of denatured salmon sperm DNA at 42 °C for at least 2 h. Hybridization was performed at 42 °C overnight in the pre-hybridized solution with ^{32}P-labeled probes (>10^7 cpm specific activity, cpm/µg DNA). The LTIR, C108 and full-length *pTS1* cDNA were labeled with (α-^{32}P)-dCTP (1,000 Ci/mmol, Amersham, Buckinghamshire, UK) using the Prime-a-Gene Labeling System (Promega, Madison, Wis., USA). Then, the filters were washed three times in 2 × SSC, 0.1% SDS at room temperature for 10 min and then two washes in 0.1 × SSC, 0.1% SDS at 56 °C for 30 min each.

Preparation of rice total RNA and Northern-blot analysis

Total rice RNA was extracted with Trizol reagent (Gibco BRL, Rockville, Maryland, USA) according to the manufacturer's protocol. Total RNA was separated on 1.0% formaldehyde gel and transferred to Hybond-N extra membranes (Amersham, Buckinghamshire, UK). The 3′-untranslated regions of *Oshsp16.9A* and *Oshsp16.9B* genes were labeled as described in the Southern-blotting analysis. Pre-hybridization, hybridization and washing were carried out as described above.

Construction and screening of the size-selected genomic library

Total rice genomic DNA was digested with the restriction enzyme *Eco*RI, separated on agarose gel, and the DNA fragments of sizes between 1 and 2 kb were eluted from the agarose gel. The size-selected genomic library was established in Lambda ZAP Express *Eco*RI/CIAP- treated vector, using the ZAP Express Gigapack II Gold cloning kit (Stratagene, La Jolla, Calif., USA) according to the manufacturer's protocol. The library was screened by hybridization with ^{32}P-labeled cDNA *pTS1* probes (>10^7 cpm) as described above. The insert from one positive clone, designated as R1.6 was in vivo excised from the ZAP Express vector and maintained in the pBK-CMV phagemid vector (Stratagene, La Jolla, Calif., USA) according to the manufacturer's protocol.

Cloning and sequence analysis of the R1.6 clone

R1.6 was subjected to sequence determination. Serial deletion was conducted for sequence analyses after sequencing with universal primers failed. The *Eco*RI-digested insert was isolated from R1.6, ligated into pGEM-7Z(+) to give the plasmid R1.6-2 for serial deletion. Nested sets of serial deletion were conducted on R1.6-2 using the Erase-a-Base System (Promega, Madison, Wis., USA) according to the technical manual. The DNA sequence was determined by using the Sequenase version 2.0 DNA sequencing kit (USB, Cleveland, Ohio, USA) according to the manufacturer's protocol. The UWGCG program suite was used to perform the sequence analysis through the NHRI of Taiwan at Nankang, Taipei.

Genome walking

The Universal GenomeWalker kit (Clontech, Palo Alto, Calif., USA) was used to conduct genome walking according to the user's manual. After rice genomic DNA has been digested with the *Hae*III restriction enzyme (Boehringer Mannheim Gmbh, Germany) overnight, a pool of uncloned, GenomeWalker adaptor-ligated genomic DNA fragments was generated. The long terminal inverted repeat-specific primer (IRSP), 5′-GAGGCTTCAGCCGGAACGAAGCGGAAA-3′, complementary to the R1.6 insert from 142 to 168 bp is used for PCR. The Advantage Genomic Polymerase Mix (Clontech, Palo Alto, Calif., USA) was used in the PCR as recommended in the protocols.

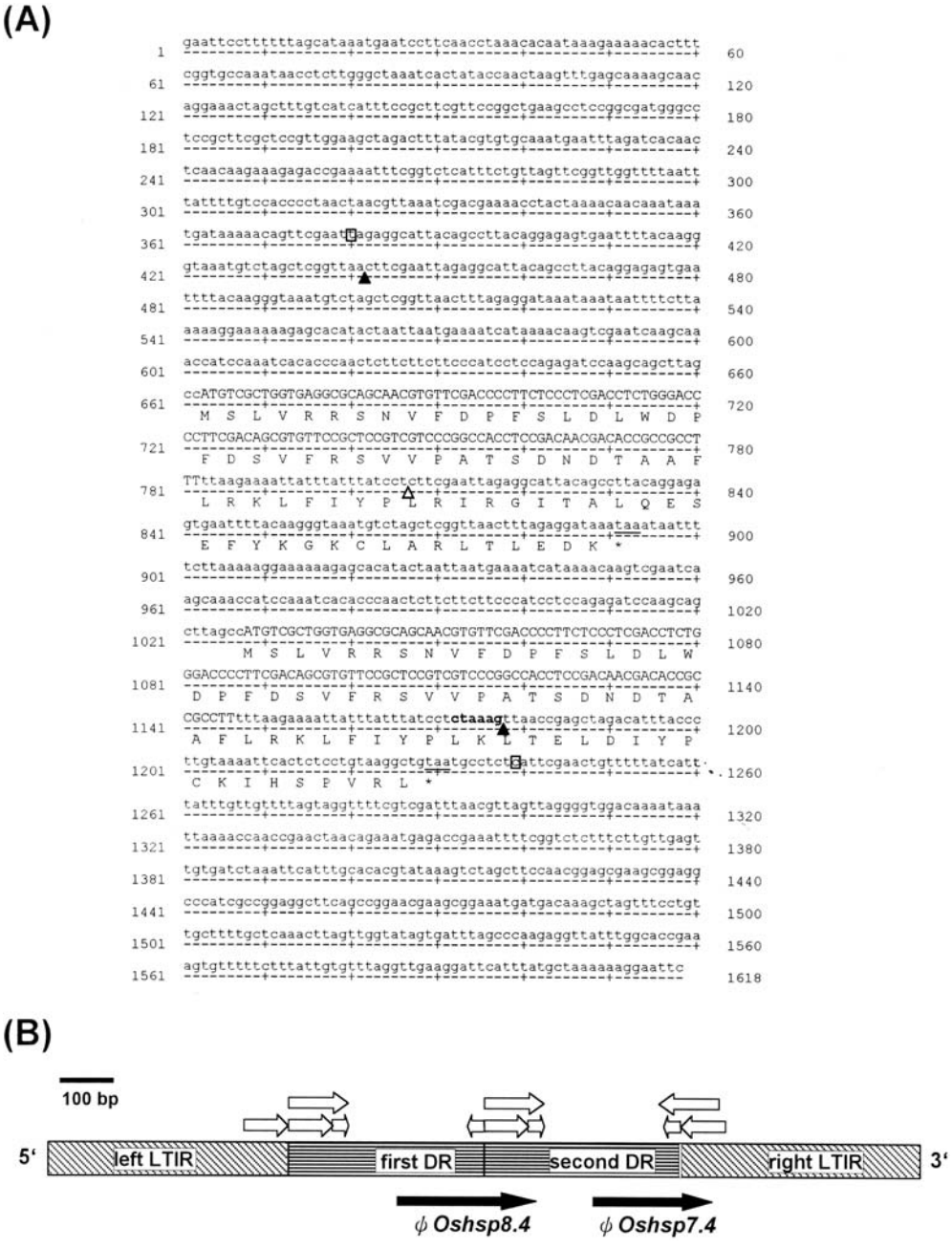

Fig. 1A, B The nucleotide sequence and schematic map of the R1.6 fragment (GenBank accession No. AF104864). (**A**) The two deduced amino-acid sequences were presented along with the putative open reading frames. The stop codons were *underlined*. The capitalized characters represented the sequence identical to the rice class-I sHSP gene, *Oshsp16.9B*. The *filled triangles* (▲) indicated the sites separated the long terminal inverted repeats (LTIRs) from the direct repeats (DRs). The *blank triangle* (△) separated the direct repeats. The *bold characters* indicated the spacer between the second DR and right LTIRs. The nucleotides that were not complementary with each other in the LTIRs were indicated by *square boxes*. (**B**) Schematic map of the R1.6 fragment. The LTIRs (*hatched boxes*) and DRs (*lined boxes*) were indicated. The *hollow arrows* indicated the differenst lengths of inverted or direct subrepeats. The *black arrows* represented the putative pseudogenes (ψ*Oshsp8.4* and ψ*Oshsp7.4*). The *blank box* was the spacer between the second DR and right LTIR

Chromosome mapping

A population consisting of 235 individuals derived from a three-way cross was used to map the sequences to the chromosomes with the 3′-UTRs of *Oshsp16.9A*, *Oshsp16.9B* and *Oshsp16.9C* as hybridization probes. The length of the probes is 209 bp, 155 bp and 190 bp for *Oshsp16.9A*, *Oshsp16.9B* and *Oshsp16.9C*, respectively. The three-way cross was conducted by Balilla (*O.sativa* ssp. Japonica)/Dular (*O.sativa* ssp. Indica)//NanJing11 (*O.sativa* ssp. Indica). A molecular linkage map containing 196 RFLP loci was developed with the above population (Wang et al. 1998). DNA digestion, Southern blotting and hybridization were all conducted essentially as described by Liu et al. (1997). Then, the data were scored as described by Wang et al. (1998). Chromosomal locations of the three clones on the marker-linkage map were determined by using MAPMAKER/EXP 3.0 (Lincoln et al. 1992).

Results

Isolation and sequence analysis of a 1.6-kb genomic clone

One 1.6-kb *Eco*RI-digested DNA fragment which showed a positive hybridization signal with *pTS1* cDNA was isolated and subcloned into a pGEM-7Zf(+) vector (Promega, Madison, Wis., USA), designated as clone R1.6. According to DNA sequence analyses, the R1.6 clone is 1,618 bp in length as shown in Fig. 1A along with the predicted open reading frames (ORFs). A pair of 441-bp long terminal inverted repeats (LTIRs) at nucleotides 1–441 (left LTIR) and 1,178–1,618 (right

LTIR) was found in this genomic clone (Fig. 1B). The 365-bp tandem DR located in nucleotides 442–806 (the first DR) and 807–1,171 (the second DR) made up the internal part of the 1.6-kb fragment. There is a spacer of six nucleotides between the second DR and the right LTIR. Each DR contains a stretch of 123 bp identical to the 5′ coding regions of two rice class-I sHSP genes, *Oshsp16.9A* and *Oshsp16.9B*, which only differ with a single nucleotide within this region. Two open reading frames were deduced from both DRs. One of the ORFs with 76 amino acids is located from nucleotides 663 to 893, and the other with 67 amino acids is located from nucleotides 1,028 and 1,231 (Fig. 1A). The N-terminal 40 amino acids of both ORFs are identical to the N-termini of either Oshsp16.9A or Oshsp16.9B, but the other residues show no similarity to any known proteins in the databases. For the test of the possible existence of transcripts of the ORFs, we performed primer extension using a primer specifically annealing to the DR. The direct repeat-specific primer (DRSP), 5′-GGCTAAGCTGCTTGGATC-3′, was complementary to nucleotides 645–662 and 1,010–1,027 which are located at the immediate 5′ to the start codons of the ORFs. No primer extension products were detected (data not shown). Additionally, we could not define any minimum promoters in the 5′ upstream regions of the ORFs, such as the canonical TATA box or CCAAT box, as well as HSEs that are a typical consensus elements of the sHSP genes (Schöffl et al. 1998). Hence, the ORFs in the R1.6 clone are pseudogenes designated ψ*Oshsp8.4* and ψ*Oshsp7.4*, respectively (Fig. 1B). The N-terminal 49 amino acids in ψ*Oshsp8.4* and ψ*Oshsp7.4* are identical, but the C-terminal residues are totally different. In addition, there were three different lengths of direct or inverted subrepeats in the DRs and LTIRs. The inverted subrepeats of 31 bp, 70 bp and 104 bp are different in length and orientation, as shown in Fig. 1B. One of the 104-bp subrepeats located in nucleotides 1,143–1,246 contained the spacer between the second DR and the right LTIR. In the DRs, the 123-bp motif identical to *Oshsp16.9A* and *Oshsp16.9B* is surrounded by the AT-rich 31-bp inverted subrepeats.

Identification of the R1.6 fragment by genomic southern blotting

Genomic DNA gel-blot analysis was carried out to investigate the genomic organization of the R1.6 fragment using the LTIR fragment as a probe. Only one prominent band was detected under high-stringency conditions (Fig. 2A). This suggested that the R1.6 fragment is present as a single-copy in the rice genome. On the other hand, Fig. 2B shows the hybridization bands detected by the C108 insert which made up the most conserved region of rice class-I sHSP genes (Chang et al. 2001). There are four prominent hybridization signals in the *Eco*RI-digested lane as expected. For comparison, Fig. 2C shows the five prominent hybridization bands in the *Eco*RI-digested lane using the *pTS1* cDNA as a probe as described previously (Tzeng SS et al. 1993). From these Southern-blot analyses we speculated that the 1.6-kb signal detected in the *Eco*RI-digested lane appeared to consist the truncated sHSP genes.

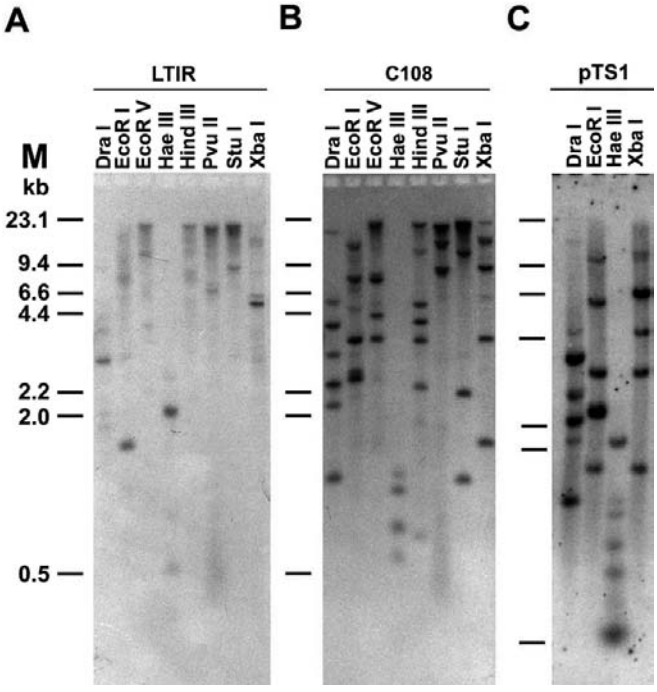

Fig. 2A–C Southern-blotting analyses of rice genomic DNA probed with the insert of LTIR (**A**), the C108 clone (**B**), and the coding region of *pTS1* cDNA (**C**). Equal amounts (8 μg per lane) of rice genomic DNA were digested with restriction enzymes as indicated at 37 °C for overnight and separated on a 0.8% agarose gel in 1 × TBE buffer. The size marker is λ/*Hind*III marker (*M*)

Extending the R1.6 fragment by genome walking

Pseudogenes that represent inactive versions of currently active genes were reported in many gene families that were interspersed in the vicinity of the clustered active genes, including hsc70 (Heschl and Baillie 1989; Rothermel et al. 1995), hsp47 (Nagai et al. 1999), hsp27 (Hickey et al. 1986) and hsp25 (Frohli et al. 1993). In order to examine whether there were any active rice class-I sHSP genes adjacent to the R1.6 fragment, we carried out PCR-based genomic walking with a LTIR-specific primer (IRSP). We obtained a 2,003-bp long clone, designated H10 (Fig. 3A). In our previous work, the 5′ upstream and coding regions of the *Oshsp16.9A* and *Oshsp16.9B* genes from R2.5 (EMBL accession No. M80938) and R3.4 (EMBL accession No. M80939) genomic clones were shown to be almost identical, except for a 41-bp insertion located between nucleotides –1,595 and –1,635 in the R2.5 clone relative to the transcription start site (Tzeng SS et al. 1993). Alignment analyses revealed that the sequence of H10 located between nucleotides 163 and 2,003 was completely overlapping with the sequence of the R3.4 fragment

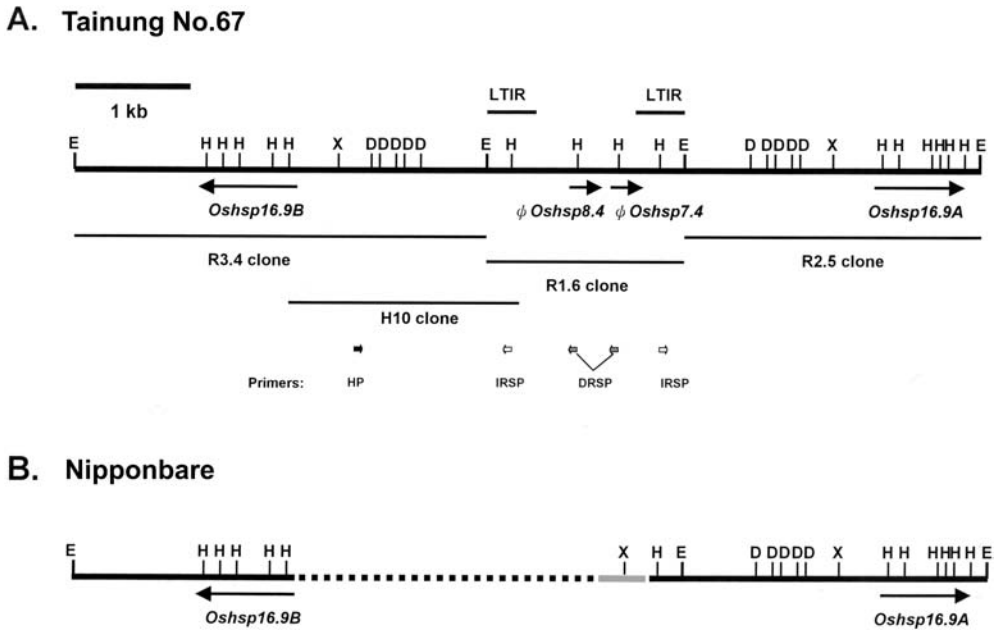

Fig. 3A, B Restriction maps of the rice genomic DNA fragments clustering with class-I sHSP genes in two rice cultivars. (**A**) The DNA fragment from Tainung No. 67 was deduced by Southern-blotting analyses as described in the result. *Arrows* represent the sHSP genes and pseudogenes. H10 was the genomic clone from rice *Hae*III-digested genome-walking library in Tainung No. 67. The R2.5 and R3.4 genomic clones containing the class-I sHSP *Oshsp16.9A* and *Oshsp16.9B* genes, respectively, were characterized previously. The length of the overlapped region between H10 and R3.4 was 1,835-bp long. The locations of the LTIR used as a probe in Southern blotting analyses are indicated. The primers used to perform PCRs are also shown. The inverted repeat-specific primer (*IRSP*) was used for performing PCR-based genome walking. The H10 primer (*HP*) and the DR-specific primer (*DRSP*) were used to conduct genomic PCR for confirming that the H10 fragment was linked to the left LTIR of the R1.6 clone. (**B**) The DNA fragment from Nipponbare was obtained from the PAC clone P0443D08. Only the fragment containing *Oshsp16.9A* and *Oshsp16.9B* was shown. The *dashed line* indicates the fragment is absent in the 5′ upstream of *Oshsp16.9B* in Nipponbare. The *gray line* indicates the putative minimum promoter region of *Oshsp16.9B* in Nipponbare, which is absent in the Tainung No. 67. *D Dra*I, *E Eco*RI, *H Hae*III, *X Xba*I

located between nucleotides –1,620–222 relative to the transcription start site. Furthermore, in order to verify from which LTIR was extended in the PCR-based genome walking, we performed the genomic PCR using the DR-specific primer (DRSP) and the H10-primer (HP) (Fig. 3A). The HP, 5′-GATTAATCAGCGACCCACTG-3′, was complementary to the nucleotides 1,455–1,474 of the H10 clone. Two PCR products of different sizes were detected (data not shown). The result suggested that the H10 fragment was extended from the left LTIR of the R1.6 fragment (Fig. 1B). Therefore, it was plausible that the R1.6 fragment might be linked to the 5′-upstream region of the *Oshsp16.9B* gene. Previous sequence data from clones R1.6, R2.5 and R3.4 also supported this hypothesis. Combining the result of restriction mapping of R1.6, R2.5 and R3.4, and Southern-blotting analysis in Fig. 2A, we could link these genomic clones together into an approximate 7.5-kb DNA fragment (Fig. 3A). For example, in the *Dra*I-digested lane the size of an approximate 3-kb signal was detected using the LTIR as a probe (Fig. 2A), and this fragment could result from 680 bp in the R2.5 clone, 639 bp from the R3.4 clone and 1,618 bp from the R1.6 clone. There are two hybridization signals about 2 kb and 0.6 kb at the *Hae*III-digested lane in the blots hybridizing with LTIR (Fig. 2A) and the full-length *pTS1* cDNA (Fig. 2C), but these bands are not present in the blot probing with C108 (Fig. 2B). There are four cutting sites for *Hae*III (5′-GGCC-3′) in the R1.6 clone, and one of them is located at the region of probe LTIR. So, using the LTIR as a probe we could detected a signal of about 0.6 kb from the R1.6 clone and a 2-kb signal from part of R1.6 (178 bp), and part of the R2.5 clone (1,885 bp) or R3.4 clone (1,844 bp). Using the full-length *pTS1* cDNA as a probe these signals could be detected, because these fragments also contain the 123-bp 5′ coding region of *pTS1* cDNA. Attempts in amplifying this 7.5-kb fragment with PCR reactions was unsuccessful possibly due to the formation of secondary structures.

Chromosome mapping of three rice class-I sHSP genes

The coding sequences of class-I sHSP genes are conserved with homology over 80% of the members, but the homologies of the 3′-untranslated regions (3′-UTRs) are below 40%. In our previous studies, cross-hybridization experiments and RNA gel-blot analyses had shown that the 3′-UTRs were suitable to be used as gene-specific probes (Lin 1998). Thus, we used 3′-UTRs of each class-I sHSP gene as a probe for chromosome mapping. The chromosomal locations of the class-I sHSP genes were determined by the RFLP mapping approach using a three-way cross population and its molecular linkage maps

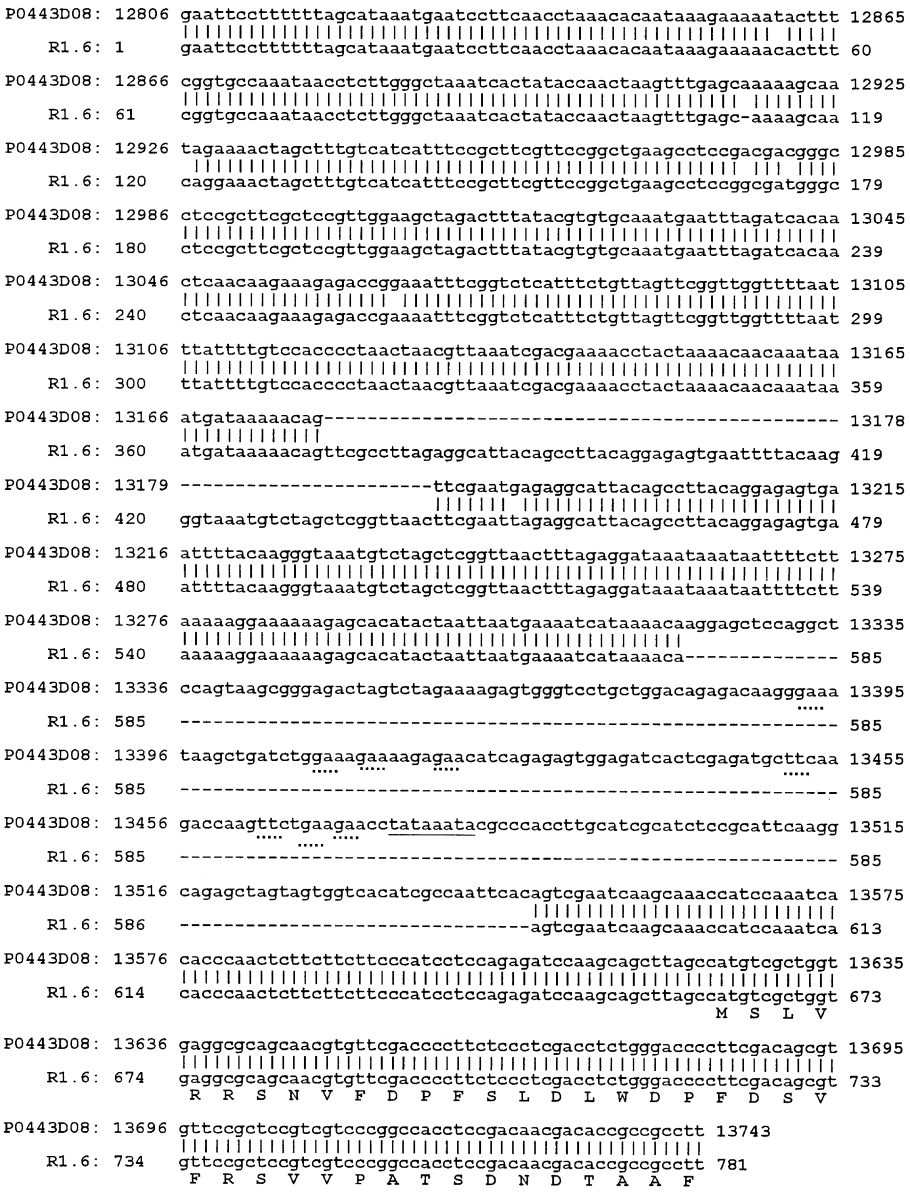

Fig. 4 Alignment analysis between the promoter region of *Oshsp16.9B* in Nipponbare and the R1.6 clone in Tainung No. 67. The putative promoter region of the *Oshsp16.9B* gene in Nipponbare cultivar (P0443D08 PAC clone) was aligned with clone R1.6. They shared high homology except for a 70-bp insertion in the LTIR and a 226-bp deletion in the DR of the R1.6 fragment. The dashed lines indicated the deletions in the R1.6 clone or the P0443D08 clone. The putative TATA box was *underlined*. The *dotted underlines* indicated the putative HSEs

(Wang et al. 1998). The mapped positions of *Oshsp16.9A*, *Oshsp16.9B* and *Oshsp16.9C* are on the short arm of chromosome 1 (data not shown). *Oshsp16.9B* is 0.4 centimorgan (cM) apart from *Oshsp16.9C*, which shows that rice class-I sHSP genes, *Oshsp16.9A*, *Oshsp16.9B*, and *Oshsp16.9C*, are closely clustered.

Comparative analysis of the class-I sHSP gene cluster on chromosome 1 of two rice cultivars

Recently, from the databases search, we noticed that in the PAC clone P0443D08 (AP003250), located on chromosome 1 of *O. sativa* cv Nipponbare, four class-I sHSP genes were deduced within the first 15-kb region. Three of the genes correspond to *Oshsp16.9A*, *Oshsp16.9B* and *Oshsp16.9C* in Tainung No. 67. This result agreed with the chromosomal assignment of these genes by 3′-UTRs, as described above. *Oshsp16.9A* and *Oshsp16.9B* are arranged in a head-to-head orientation as in the Nipponbare cultivar. The restriction map of the DNA fragment from Nipponbare was constructed according to the nucleotide sequence of the PAC clone P0443D08 (Fig. 3B). The *Oshsp16.9B* gene in the R3.5 clone shares homology with the *Oshsp16.9A* gene of the R2.5 clone in the 5′ upstream region, as described above; however, the 5′ region beyond the coding region of the *Oshsp16.9B* gene in the P0443D08 clone did not share homology with the 5′ upstream region of the *Oshsp16.9A* gene. Moreover, by alignment of the promoter region of *Oshsp16.9B* in Nipponbare with the R1.6 clone, we found that the promoter region shares high homology with the LTIR (97%) and DR (99%), except for a 70-bp deletion and a 226-bp insertion in the P0443D08 clone (Fig. 4). The 226-bp fragment absent in the DRs of R1.6 contains one putative TATA box and several putative HSEs (Fig. 4).

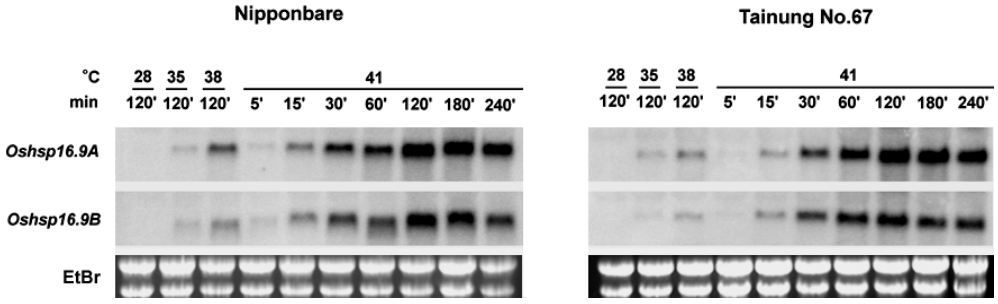

Fig. 5 Northern-blotting analyses of *Oshsp16.9A* and *Oshsp16.9B* under different heat stresses in rice Nipponbare and Tainung No. 67 cultivar. Total RNA (15 μg) from rice seedlings incubated at various temperatures was loaded in each lane and was fractionated on a 1.0% formaldehyde/agarose gel. The 3′-UTR of *Oshsp16.9A* and *Oshsp16.9B* were used as probes, separately. The specific activity was about 6.0×10^8 cpm/μg DNA for each probe. After washing under high stringency, the blots were exposed on Phosphor Screen (Molecular Dynamics) for 24 h. The EtBr staining gel was shown as a RNA loading reference. These experiments were conducted twice for confirmation of the results

It is interesting to note that the promoter region of *Oshsp16.9B* in the Nipponbare cultivar is different from that in Tainung No. 67; however, the nucleotide sequence identity is more than 99% in the other regions (including the coding region and 3′-downstream from the coding region). The promoter region of *Oshsp16.9B* in Tainung No. 67 is almost identical to the promoter region of *Oshs16.9A* in both cultivars. Thus, it is interesting to investigate the expression profiles of *Oshsp16.9A* and *Oshsp16.9B* under different heat stresses in these two cultivars. We performed Northern-blot analyses at different heat-shock (HS) conditions between these two cultivars, using gene-specific 3′-UTRs of *Oshsp16.9A* and *Oshsp16.9B* as probes. Three-day old etiolated rice seedlings were incubated at 28 °C, 35 °C, 38 °C and 41 °C for 2 h to evaluate the levels of expression of *Oshsp16.9A* and *Oshsp16.9B* in the rice cultivars, Nipponbare and Tainung No. 67. The transcripts began to accumulate at 35 °C, and reached a maximal level at 41 °C in both genes from the two cultivars during a 2-h incubation (Fig. 5). The Nipponbare cultivar accumulated more transcripts of both genes under 35 °C and 38 °C than the Tainung No. 67 cultivar. The time course of gene expression was also performed at 41 °C. The transcripts were detected in the first 5-min incubation and gradually increased to the highest level after the 120-min incubation at 41 °C in both cultivars. Nipponbare accumulated more transcripts after 5-min and 15-min heat-shock treatments than Tainung No. 67; however, the transcript level of Tainung No. 67 reached almost the same level as Nipponbare after a 60-min heat-shock treatment. The transcripts levels decreased gradually in a prolonged incubation time at 41 °C in both cultivars. The *Oshsp16.9A* transcripts were found to accumulate significantly higher levels than *Oshsp16.9B* transcripts in both cultivars. In our previous study on Tainung No. 67, a similar result was demonstrated (Lin 1998).

Discussion

In this study we report the isolation and characterization of a unique rice genomic clone containing two tandem truncated class-I sHSP genes with almost perfect long terminal inverted repeats. Although many plant sHSP genes have been isolated and characterized from different plant species (Waters and Vierling 1999), only one pseudogene of the plant sHSP gene was reported, which is a putative prematurely truncated class-I *HSP17.7b* pseudogene of *Brassica oleracea* (AJ243566). There were two human *HSP27* pseudogenes and one mouse *HSP25* pseudogene (McGuire et al. 1989; Frohli et al. 1993). Typical pseudogenes (ψ) were defined by the possession of sequences related to those of the functional genes, but cannot be transcribed or translated into a functional protein. We designated the truncated class-I sHSP genes as pseudogenes based on gene structure and the lack of detectable transcripts from the primer extension study. The high sequence conservation between these pseudogenes and *Oshsp16.9B* is very high (>99%), indicating that these pseudogenes have emerged very recently.

In our previous study, we have shown that the promoter regions of *Oshsp16.9A* and *Oshsp16.9B* in Tainung No. 67 were almost identical, except for a 41-bp insertion in the *Oshsp16.9A* gene (Tzeng SS et al. 1993). Comparative analysis of the class-I sHSP genes clustering on chromosome 1 revealed that the putative promoter region of *Oshsp16.9B* in Nipponbare was about 2-kb shorter than that of *Oshsp16.9B* in Tainung No. 67 (Fig. 3). Hence, *Oshsp16.9B* could be under different regulation in these cultivars. It is unclear, at this point, why and how a different promoter was generated for the same gene in these two cultivars. We assume that the expressions of *Oshsp16.9A* and *Oshsp16.9B* under the similar promoter may reflect the synchronous accumulation of both transcripts in response to heat stress. In our previous studies, we have shown that the transcriptional rate and stability of *Oshsp16.9A* is higher than that of *Oshsp16.9B* in Tainung No. 67 (Lin 1998). In this study, the comparative Northern-blot analyses between Tainung No. 67 and Nipponbare cultivars upon different HS conditions indicated that the induction of *Oshsp16.9A* and *Oshsp16.9B* in Nipponbare was similar to that of Tainung No. 67 (Fig. 5). The transcript level of *Oshsp16.9A* was higher than that of *Oshsp16.9B* in Tainung No. 67, as described previously (Lin 1998), as well as in Nippon-

bare. Thus, the promoter region of *Oshsp16.9B* in the Nipponbare cultivar does not alter the expression pattern in response to HS conditions. In addition, whether the promoter rearrangement introduces the diverse inducibility of sHSP genes in response to developmental regulation and different stresses in these two cultivars still needs to be further characterized in the future.

On the other hand, it also remains unclear how the complex structure of the R1.6 fragment was generated in the rice genome. We know that plant genomes can change rapidly within a short evolutionary time span. Gene duplication, conversion and inversion events have been discussed to be the major mechanism for the origin of sHSP genes in plants (Waters and Vierling 1999). Thus, the complex structure of the R1.6 fragment might be just reminiscent of the recombination of the rice sHSP gene family or it might be induced by human breeding. The distinctive feature of the R1.6 fragment is characterized by its LTIRs and inverted subrepeats of the DRs associated with the truncated class-I sHSP genes (Fig. 1B), because the terminal inverted repeats (TIRs) are characters of DNA transposable elements. The DNA transposons were characterized by terminal inverted repeats (TIRs) and moved directly through a DNA form by a "cut and paste" mechanism (Le et al. 2001). Transposable elements are an integral component of most genomes and may play an important role in the evolution of developmental processes (Bureau et al. 1996). Structural features such as sequence similarity, and the size and sequence of the target site duplication generated upon insertion, serve to further distinguish transposon superfamilies (Le et al. 2001). Several *Mutator*-like elements (MULEs) including rice elements that harbor truncated host cellular genes were identified (Le et al. 2001; Turcotte et al. 2001). Five MULE-I representatives contain a 105-bp region with high nucleotide similarity (>90%) to a portion of the rice 5S rDNA gene, and one MULE-IV element contains a 149-bp region showing high nucleotide similarity (>83%) to a region of the rice PCF2 mRNA sequence. The acquisition of truncated cellular genes perhaps resulted from illegitimate recombination and repair events (Turcotte et al. 2001). Computer-based systematic surveys revealed the predominance of miniature inverted-repeat transposable elements (MITEs), which might facilitate rapid sequence evolution in the rice genome (Bureau et al. 1996; Bennetzen 2000; Lee et al. 1995). Two MITEs, *ditto* and *explorer* (Bureau et al. 1996), were identified in the 1.6-kb promoter regions of *Oshsp16.9A* and *Oshsp16.9B*. Therefore, the transposable elements may play an important role in the evolution of the class-I sHSP gene family in rice.

Acknowledgements This work was supported by the National Science Council, Taiwan, Republic of China (R. O. C.), under the grant NSC89-2311-B-002-006 to C.Y. Lin. We thank Dr. Fritz Schöffl and Dr. T.S. Tseng for helpful comments on this manuscript.

References

Bureau E, Ronald PC, Wessler SR (1996) A computer-based systematic survey reveals the predominance of small inverted-repeat elements in wild-type rice genes. Proc Natl Acad Sci USA 93:8524–8529

Bennetzen JL (2000) Comparative sequence analysis of plant nuclear genomes: microlinearity and 75 many exceptions. Plant Cell 12:1021–1029

Chang PFL, Huang CI, Chang FC, Tseng TS, Lin WC, Lin CY (2001) Isolation and characterization of the third gene encoding a 16.9-kDa class I low-molecular-mass heat shock protein, Oshsp16.9C, in rice. Bot Bull Acad Sinica 42:85–92

Edelman L, Czarnecka E, Key JL (1988) Induction and accumulation of heat shock-specfic poly (A^+) RNAs and proteins in soybean seedlings during arsenite and cadmium treatments. Plant Physiol 86:1048–1056

Frohli E, Aoyama A, Klemenz R (1993) Cloning of the mouse hsp25 gene and an extremely conserved hsp25 pseudogene. Gene 128:273–277

Guan JC, Chang FC, Tseng TS, Chang PFL, Yeh KW, Chen YM, Lin CY (1998) Structure of rice genes encoding three class-I low-molecular-mass heat-shock proteins (accession Nos. U83669, U83670, U83671) (PGR 98–178). Plant Physiol 118:1101

Heschl MF, Baillie DL (1989) Identification of a heat-shock pseudogene from *Caenorhabditis elegans*. Genome 32:190–195

Hickey E, Brandon SE, Potter R, Stein G, Stein J, Weber LA (1986) Sequence and organization of genes encoding the human 27-kDa heat shock protein. Nucleic Acids Res 14:4127–4145

Ingolia TD, Craig EA (1982) Four small *Drosophila* heat shock proteins are related to each other and to mammalian α-crystallin. Proc Natl Acad Sci USA 79:2360–2364

Jinn TL, Yeh YC, Chen YM, Lin CY (1989) Stabilization of soluble proteins in vitro by the heat shock protein-enriched ammonium sulfate fraction from soybean seedlings. Plant Cell Physiol 30:463–469

Jinn TL, Chen YM, Lin CY (1995) Characterization and physiological function of class I low-molecular-mass, heat-shock protein complex in soybean. Plant Physiol 108:693–701

Key JL, Lin CY, Chen YM (1981) Heat shock proteins of higher plants. Proc Natl Acad Sci USA 78:3526–3530

Kuo HF, Tsai YF, Young LS, Lin CY (2000) Ethanol treatment triggers a heat shock-like response but no thermotolerance in soybean (*Glycine max* cv Kaohsiung No. 8) seedlings. Plant Cell Environ 23:1099–1108

Le QH, Wright S, Yu ZH, Bureau T (2001) Transposon diversity in *Arabidopsis thaliana* (2000). Proc Natl Acad Sci USA 97:7376–7381

Lee GJ, Vierling E (2000) A small heat shock protein cooperates with heat shock protein 70 systems to reactivate a heat-denatured protein. Plant Physiol 122:189–197

Lee YL, Chang PFL, Yeh KW, Jinn TL, Kung CCS, Lin WC, Chen YM, Lin CY (1995) Cloning and characterization of a cDNA encoding an 18.0-kDa class-I low-molecular-weight heat-shock protein from rice. Gene 165:223–227

Lee YRJ, Nagao RT, Lin CY, Key JL (1996) Induction and regulation of heat-shock gene expression by an amino-acid analog in soybean seedlings. Plant Physiol 110:241–248

Lin CY, Roberts JK, Key JL (1984) Acquisition of thermotolerance in soybean seedlings 1.Synthesis and accumulation of heat shock proteins and their cellular localization. Plant Physiol 74:152–160

Lin WC (1998) Gene regulation and mRNA stability of rice 16.9-kDa heat shock protein genes at high temperature. Ph.D. dissertation, Institute of Botany, National Taiwan University, Taiwan, R.O.C.

Lincoln S, Daly M, Lander E (1992) Constructing genetic maps with MAPMAKER/EXP 3.0 (3rd edn). Whitehead Institute. Cambridge, Massachusetts

Liu KD, Wang J, Li HB, Xu CG, Liu AM, Li XH, Zhang Q (1997) A genome-wide analysis of wide compatibility in rice and the precise location of the S5 locus in the molecular map. Theor Appl Genet 95:809–814

Löw D, Brändle K, Nover L, Forreiter C (2000) Cytosolic heat-stress proteins Hsp17.7 class I and Hsp17.3 class II of tomato act as molecular chaperones in vivo. Planta 211:575–582

Malmberg R, Messing J, Sussex I (1985) Molecular biology of plants: a laboratory course manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Marmiroli N, Pavesi A, Di Cola G, Hartings H, Raho G, Conte MR, Perrotta C (1993) Identification, characterization, and analysis of cDNA and genomic sequences encoding two different small heat shock proteins in *Hordeum vulgare*. Genome 36:1111–1118

McGuire SE, Fuqua SA, Naylor SL, Helin DD, McGuire WL (1989) Chromosomal assignments of human 27-kDa heat shock protein gene family. Somatic cell Mol genet 15:167–171

Montfort RLM van, Basha E, Friedrich KL, Slingsby C, Vierling E (2001) Crystal structure and assembly of a eukaryotic small heat shock protein. Nature Struct Biol 8:1025–1030

Nagai N, Tetuya Y, Hosokawa N, Nagata K (1999) The human genome has only one hsp47 gene (CBP2) and a pseudogene (pshsp47). Gene 227:241–248

Nagao RT, Czarnecka E, Gurley WB, Schöffl F, Key JL (1985) Genes for low-molecular-weight heat shock proteins of soybeans: sequence analysis of a multigene family. Mol Cell Biol 5:3417–3428

Pla M, Jofre A, Martell M, Molinas M, Gomez J (2000) Large accumulation of mRNA and DNA point modifications in a plant senescent tissue. FEBS Lett 472:14–16

Potter DR, Nguyen HT, Burke JJ (1989) Chromosomal and location of genes controlling heat shock proteins in hexaploid wheat. Theor Appl Genet 78:873–878

Rothermel E, Detzler E, Walter L, Levan G, Gunther E (1995) Characterization and mapping of a highly conserved processed pseudogene and an intron-carrying gene of the heat shock cognate protein 70 (Hsc70) gene family in the rat. Mammal Genome 6:602–606

Sabehat A, Lurie S, Weiss D (1998) Expression of small heat-shock proteins at low temperatures. Plant Physiol 117:651–658

Schöffl F, Prändl R, Reindl A (1998) Regulation of the heat-shock response. Plant Physiol 117:1135–1141

Tseng TS, Yeh KW, Yeh CH, Chang FC, Chen YM, Lin CY (1992) Two rice (*Oryza sativa*) full-length cDNA clones encoding low-molecular-weight heat-shock proteins. Plant Mol Biol 18:963–965

Tseng TS, Tzeng SS, Yeh KW, Yeh CH, Chang FC, Chen YM, Lin CY (1993) The heat-shock response in rice seedlings: isolation and expression of cDNAs that encode class I low-molecular-weight heat-shock proteins. Plant Cell Physiol 34:165–168

Tzeng SS, Yeh KW, Chen YM, Lin CY (1992) Two *Oryza sativa* genomic DNA clones encoding 16.9-kilodalton heat-shock proteins. Plant Physiol 99:1723–1725

Tzeng SS, Chen YM, Lin CY (1993) Isolation and characterization of genes encoding 16.9-kDa heat shock proteins in *Oryza sativa*. Bot Bull Acad Sinica 34:133–142

Turcotte K, Srinivasan S, Bureau T (2001) Survey of transposable elements from rice genomic sequences. Plant J 25:169–179

Wang J, Liu KD, Xu CG, Li XH, Zhang Q (1998) The high level of wide-compatibility of Dular has a complex genetic basis. Theor Appl Genet 97:407–412

Waters ER, Vierling E (1999) The diversification of plant cytosolic small heat shock proteins preceded the divergence of mosses. Mol Biol Evol 16:127–139

Waters ER, Lee GJ, Vierling E (1996) Evolution, structure and function of the small heat shock proteins in plants. J Exp Bot 47:325–338

Wehmeyer N, Vierling E (2000) The expression of small heat shock proteins in seeds responds to discrete developmental signals and suggests general protective role in desiccation tolerance. Plant Physiol 122:1099–1108

Yeh CH, Yeh KW, Shu-Hsing Wu, Chang PFL,Chen YM, Lin CY (1995) A recombinant rice 16.9-kDa heat shock protein can provide thermoprotection in vitro. Plant Cell Physiol 36:1341–1348

Yeh CH, Chang PFL, Yeh KW, Lin WC, Chen YM, Lin CY (1997) Expression of a gene encoding a 16.9-kDa heat-shock protein, Oshsp16.9, in *Escherichia coli* enhances thermotolerance. Proc Natl Acad Sci USA 94:10,967–10,972

Young LS, Yeh CH, Chen YM, Lin CY (1999) Molecular characterization of *Oryza sativa* 16.9-kDa heat shock protein. Biochem J 94:10,967–10,972