Tropical Plant Biol. (2013) 6:69–84
DOI 10.1007/s12042-013-9124-2

Thermal-Stable Proteins of Fruit of Long-Living Sacred Lotus *Nelumbo nucifera* Gaertn var. China Antique

J. Shen-Miller · Petra Lindner · Yongming Xie · Sarah Villa · Kerry Wooding · Steven G. Clarke · Rachel R. O. Loo · Joseph A. Loo

Received: 1 April 2013 / Accepted: 14 June 2013 / Published online: 24 July 2013

Abstract Single-seeded fruit of the sacred lotus *Nelumbo nucifera* Gaertn var. China Antique from NE China have been shown to remain viable for as long as ~1,300 years, determined by direct radiocarbon-dating, and to have a germination rate of 84 %. The pericarp, a fruit tissue that encloses the single seeds of *Nelumbo*, is one of the major factors contributing to fruit longevity. Proteins that are heat stable and have a protective function are equally important to such centuries-long seed viability. We document proteins of *Nelumbo* fruit that are able to withstand heating, 32 % of which remained soluble in the 110 °C-treated embryo axis of a 549-year-old fruit and 76 % retained fluidity in its cotyledons. The genome of *Nelumbo* has recently been published and annotated. The amino-acid sequences of 11 "thermal proteins" (soluble at 100 °C) of modern *Nelumbo* embryo axes and cotyledons, identified by mass spectrometry, Western blot and bioassay, are assembled and aligned with those of an archaeal hyperthermophile *Methancaldococcus jannaschii* ("Mj," an anaerobic methanogen having a growth optimum of 85 °C) and with those of five mesophile angiosperms. These thermal proteins have roles in protection and repair under stress. More than half (55 %) of the durable *Nelumbo* thermal proteins are present in the archaean Mj, indicating their ancient history. One *Nelumbo* protein-repair enzyme exhibits activity at 100 °C, having a heat-tolerance higher than the comparable enzyme of Arabidopsis. A list of 30 sequenced but unassembled thermal proteins of *Nelumbo* is appended.

Keywords *Nelumbo nucifera* China antique · Heat-soluble (100 °C) Proteins · Hyperthermophile-Mesophile protein-alignments · Stress-and-repair thermoproteins

Abbreviations

AdoMet	S-adenosyl-L-methionine
CPN20/60	Chaperonine20/60
CuZn-SOD	Copper-zinc superoxide dismutase
1-CysPRX	1-Cys peroxiredoxin
ENO1	Enolase1
EF-1α	Elongation factor-1α
EST	Expressed sequence tag
HSP80	Heat-shock protein 80
LC-MS-MS	Liquid-chromatography tandem mass-spectrometry
Mj	*Methancaldococcus jannaschii*
PIMT	Protein L-isoaspartyl methyltransferase
ROS	Reactive oxygen-species

Communicated by Yuval Cohen

J. Shen-Miller for manuscript preparation, overall experimentation and coordination; P. Lindner for initiation of thermal-protein experiments; Y. Xie, K. Wooding, R.R.O. Loo, and J. A. Loo for protein mass-spectrometry; and S. Villa and S. G. Clarke for PIMT experiments.

Electronic supplementary material The online version of this article (doi:10.1007/s12042-013-9124-2) contains supplementary material, which is available to authorized users.

J. Shen-Miller (✉)
IGPP Center for the Study of Evolution and the Origin of Life, Department of Ecology and Evolutionary Biology, University of California, Los Angeles, Geology Building, Room 5676, 595 Charles E. Young Drive East, Los Angeles, CA 90095-1567, USA
e-mail: shenmiller@lifesci.ucla.edu

P. Lindner
Lehrstuhl Mikrobiologie, Regensburg University, Universitat Str. 31, 93053 Regensburg, Germany

Y. Xie · K. Wooding · R. R. O. Loo · J. A. Loo
Department of Chemistry and Biochemistry, University of California, Los Angeles, 402 Boyer Hall, Hilgard Avenue, Los Angeles, CA 90095-1569, USA

S. Villa · S. G. Clarke
Department of Chemistry and Biochemistry, University of California, Los Angeles, 640 Boyer Hall, Hilgard Avenue, Los Angeles, CA 90095-1570, USA

Introduction

Fruit of *Nelumbo nucifera* Gaertn var. China Antique hundreds of years in age, preserved in a dry lakebed in NE China at Xipaozi Village, Pulandian, Liaoning Province (Chen et al. 1965), were collected in 1952 by botanists of Academia Sinica, Beijing Institute of Botany (Chang 1978). The oldest germinated fruit was directly radiocarbon dated as being ~1,300 year-old (Shen-Miller et al. 1995). At present, a total of 16 old fruit of China Antique collected from the same lakebed have been shown to germinate within ~3 days (84 % germination; Shen-Miller et al. in press). Remarkably, this rate of germination of the old fruit is comparable not only to the modern controls (93 %) but also to those of other crops 2–3 year in age (Priestley 1986).

The pericarp (the fruit coat that encloses each seed) is the first line of defense that protects a *Nelumbo* seed and preserves its viability. This fruit coat is robust and impermeable to water, and in a centuries-old fruit has been shown to have elastic stiffness and hardness equivalent to that of the antlers of elk (Shen-Miller et al. in press). Such pericarps are wax-covered (supported further by new data on the expansion of *Nelumbo* wax-biosynthetic genes; Nelson and Schuler 2013). The fruit coats are composed of cell layers covered with water-insoluble suberin (Shaw 1929; Shen-Miller et al. in press) that are infused with latex (Esau and Kosakai 1975; Tele Images-Nature 2003; Shen-Miller et al. in press). The robust protective *Nelumbo* pericarp is superbly "architectured" — both structurally and chemically — to protect the enclosed seed.

The genome of *Nelumbo nucifera* var. China Antique has recently been annotated (Ming et al. 2013). It is the most ancient angiosperm thus far sequenced, having a remarkably slow nucleotide mutation rate. *Nelumbo* underwent a whole-genome duplication ~65 million years ago that evidently helped it weather the mass impact-extinction event at the Cretaceous-Tertiary boundary (Ming et al. 2013). Nelumbo has a genome size of ~1 Gb and coding for ~27,000 proteins. In this paper we document the presence of proteins in embryo axes and cotyledons of *Nelumbo* seeds that remain soluble after having been heated to 110 °C. The fluidity exhibited by *Nelumbo* seed-proteins under stress seems certain to have played a role in long-term maintenance of seed viability. All 11 "thermal proteins" (heat soluble at 100 °C) here identified and aligned (Supplement 1) have roles in cellular protection and restoration of plant functions under stress. One *Nelumbo* thermal protein, PIMT (protein L-isoaspartyl methyltransferase), exhibits protein-repair activity after heating to 100 °C, whereas the optimal activity of the enzymes of most mesophile plants occurs between 25 °C and 50 °C (Vieille and Zeikus 2001). Interestingly, two of the *Nelumbo* thermal enzymes, PIMT and CuZn-SOD (Supplement 1), tested here and elsewhere (Fig. 6; Chen et al. 2011), have been shown to maintain activity following higher heat-treatments than those of Arabidopsis (Villa et al. 2006) and pearl millet (Mahanty et al. 2012). Thirty additional assembled *Nelumbo* thermal proteins, identified by mass spectrometry, are listed in Supplement 2.

Ohga (1926) was the first to report discovery of old *Nelumbo* fruit at Xipaozi Village, NE China (Chen et al. 1965), the source of the seeds studied here. He documented germination of these fruits after 120 °C steam treatment (Ohga 1927). Chu et al. (2012) reported ~40 % germination after 48 h at 90 °C. Similarly, Ding et al. (2008) showed that although exposure to 100 °C for 15 min resulted in 0 % germination of maize kernels, the same treatment of *Nelumbo* fruit, for 24 h, yielded ~14 % germination, and Huang et al. (2000) also documented germination of *Nelumbo* fruit after 100 °C heating, whereas other crop seeds, such as bean, corn and peanut, did not. These data suggest that *Nelumbo* proteins active during seed germination are highly heat stable.

The stability of proteins at high temperatures is due to many factors (Unsworth et al. 2007), among others the co-occurrence of metabolites and sugars and the presence of polar amino acids. Specific polar residues (viz., Asp, Glu, Lys, Arg, Tyr: D-E-K-R-Y) can enhance the occurrence of intra-subunit ion-pair formation that confers thermal stability to the proteins. Like thermal proteins known from hyperthermophiles and used for bioengineering (Vieille and Zeikus 2001; Unsworth et al. 2007), the thermal proteins here identified in seeds of *Nelumbo* could prove useful in protein engineering, their presence in a mesophile expanding the temperature range of established enzymatic activity.

Multiple shared sequence alignments are documented for 11 *Nelumbo* thermal proteins, the archaeal hyperthermophile *Methancaldococcus jannaschii* (an anaerobic autotrophic methanogen previously known as *Methanococcus jannaschii* that has a growth optimum of 85 °C; Stetter 1996; Vieille and Zeikus 2001), and five mesophile angiosperms (Arabidopsis, corn, grape, poplar and soybean; Supplement 1).

M. jannaschii (Mj), the first archaeal hyperthermophile sequenced and having ~1,738 predicted protein-coding genes (Bult et al. 1996), is a marine single-celled methanogen occurring at abyssal depth (~2,600 m below the sea level) in the vicinity of the "white smokers" of hydrothermal vents in boiling water and extreme hydrodstatic pressure (DeWeerdt 2002). The oceans of the early Earth were likely hot (Kasting and Howard 2006; Robert and Chaussidon 2006), populated at least in part by archaeal hyperthermophiles. Carbon isotopic data indicate that methanogens, all members of the Archaea, were extant at ≥2.8 billion years ago (Schopf 1994). The proteins of Mj used in the alignment presented here have been evolutionarily highly conserved, their coding sequences occurring also in the genome of *Nelumbo* and other angiosperms. More than half (55 %) of the 11 *Nelumbo* thermal proteins identified are present in Mj, documenting their existence over billions of years.

About Sacred Lotus *Nelumbo nucifera*

Nelumbo nucifera belongs to the smallest family of angiosperms, the Nelumbonaceae that includes only a single genus, *Nelumbo*, and its two species, *N. nucifera* (Asia) and *N. lutea* (N. America). These taxa are related neither to the water lily (Nymphaeaceae; e.g., Egyptian lily-of-the-Nile, *Nymphpaea caerulea*), to *Lotus japonica* (Leguminosae), nor to the "lotos-fruits" of *Ziziphus lotus* (Rhamnaceae). Ancient Greeks used "lotus" ("*lotos*") to refer to many unrelated plants including *Nelumbo* (Shen-Miller et al. in press). To avoid such confusion, "lotus" will not be used in the following. Instead, we will refer to the plant for which data are reported here by its genus name, "Nelumbo," not encumbered by italics.

Materials and Methods

Proteins Isolation and Heat Treatment

After pre-testing Nelumbo fruit viability using the sink/float water test ("floaters" invariably being not germanative), those deemed viable were selected for experimentation. Whole embryo axes and cotyledon tissues of equivalent dry weight (~25 mg) of a "sinker" fruit of 459-year-old China Antique OL96-60 (Table 1, Shen-Miller et al. in press) and a modern control were powdered and extracted with Tris buffer (50 mM Tris–HCl, 5 mM $MgCl_2$, 0.5 mM DTT). The extracts were spun and the supernatants estimated for protein content by absorbance measured at 595 nm and by use of the standards of bovine serum albumin (Bradford 1976). Aliquots of protein extracts in pressure–capped microfuge tubes were heated for 10 min at 30°, 50°, 80°, 100° and 110 °C in temperature controlled water-baths (30° to 100 °C) and glycerol-baths (110 °C). After chilling on ice, the tubes were spun, the protein quantity of the resultant supernatants (constituting the "crude proteins;" Table 1) were estimated, and the aliquots were frozen and stored for SDS gel-electrophoresis. Similar procedures were used for samples of trypsin digestion of in-gel protein for tandem mass-spectrometry.

Gel Electrophoresis

The method of Laemmli (1970) was employed for 1D SDS-PAGE gel-electrophoresis of the heat-treated proteins, using linear gradient gels of 5 to 25 % w/v polyacrylamide. Equal aliquots of heat-treated proteins from embryo axes and cotyledons and corresponding unheated controls were electrophoresed and compared with protein standards having masses ranging from 78 to 12.4 kDa. The resultant gels were stained with 0.1 % Coomassie Blue w/v in 30 % methanol with 10 % HAc v/v, developed at 50 °C, and photographed (Fig. 1). For MS identification of protein, 1D-SDS gels (BioRad) of 100 °C treated extracts of embryo axes and cotyledons of modern China-Antique were prepared (Fig. 2).

Protein Mass Spectrometry

Coomassie-blue stained gel-bands numbered by bins (Fig. 2) were individually excised and placed in microfuge tubes to prepare proteins for identification. The gel-embedded proteins were reduced, iodoacetaminde-alkylated and trypsin digested using Promega, a sequencing grade modified trypsin (Shevchenko et al. 1996). Product peptides were extracted with 50 % acetonitrile-0.1 % trifluoroacetic acid and the resulting extract was dried by vacuum centrifugation. Peptides were then dissolved in 10 μl of a 0.1 % formic acid (FA) solution and analyzed by liquid-chromatography tandem-mass spectrometry (LC-MS/MS) using electrospray ionization (ESI) on an Applied Biosystems OSTAR® Pulsar XL (OqTOF) mass spectrometer equipped with a nanoelectrospray interface (Protana), a Proxen (Odense) nano-bore stainless steel emitter (30 μm i.d), and using an LC packings nano-LC system. The nano-LC was equipped with a specially constructed precolumn (150 μm×5 mm) and an analytical column (75 μm×150 mm) packed with Jupiter Proteo C12.4-μm resin (Phenomenex). Typically, 6 μl samples were loaded onto the precolumn, washed with loading solvent (0.1 % FA) for 4 min and injected into the LC column. The eluents used were 0.1 % FA, aqueous (Solvent A) and 95 % CH_3CN containing 0.1 % FA (Solvent B). A column solvent-flow of 200-nL per-min was applied to the gradient: 3 % B to 6 % B in 6 s, 6 % B to 24 % B in 18 min, etc., to a final step of 36 % B to 80 % B in 2 min, where it was maintained at 80 % B for 7.9 min. At the conclusion of the run, the column was equilibrated with 3 % B for at least 15 min before the next run.

Peptide ion spectra were recorded automatically during LC-MS/MS by information-dependent analysis (IDA) software on the mass spectrometer. Argon was used as the collision gas, and collision energies for maximum fragmentation efficiencies were calculated by the instrumental software using empirical parameters based on the charge and mass-to-charge ratio of the peptide precursor ion.

Protein Sequence Identification

Nelumbo thermal proteins are identified from sequencing by liquid-chromatography tandem mass-spectrometry (Pennington and Dunn 2001) and database search utilizing the Mascot database search engine (Matrix Science). Searches were performed against a database of proteins predicted from the genome, supplemented with keratin and trypsin sequences. These protein sequence searches allowed a maximum of two missed cleavages and a

mass size of 0.3 Da for precursor and product ions. Protein hits were accepted based on >2 ascribed peptides, with at least one peptide possessing an ion score >32, such a score being defined as −10*Log (P), where P is the probability that the observed match is a random event. For the database of predicted Nelumbo proteins, ion scores >32 corresponded to <5 % probability that the peptide match was random. Correspondences between MS/MS spectra and ascribed sequences were also verified manually.

Western-Blot Protein Identification

Western blotting of graded 1D SDS-PAGE gels of the heat-treated crude proteins of modern embryo axes and cotyledons of China Antique followed the procedures of Towbin et al. (1979) and Burnett (1981). The blotted Millipore membranes (PVDF, BioRad) were stained with 0.2 % Ponceau-S to mark the protein standards. The membrane was destained with water and finally with PBS-T (4 mM KH_2PO_4, 16 mM Na_2HPO_4, 115 mM NaCl, in 0.05 % Tween-20, w/v). For protein identification, the destained membrane containing the blotted Nelumbo protein was incubated at room temperature together with dilutions of the appropriate antiserum for 1–12 h in PBS-T. The primary antibodies used included rabbit anti-spinach chloroplast chaperonin CPN20, anti-pea chloroplast CPN60 (Bersch et al. 1992; Fig. 5), anti-barley dehydrin (Close 1997; Fig. 3), and rabbit anti-Arabidopsis PIMT1 (Xu et al. 2004; Fig. 4).

After incubation, the membranes were washed 3X with TBS-T to remove unbound primary antisera and incubated for 45 min to 2 h in a 1:5000 dilution of a secondary antibody of anti-rabbit sheep IgG linked to alkaline phosphatase or horseradish peroxidase. After incubation, the membranes were washed with PBS-T to remove unbound secondary antibodies, and were then washed 3X in PBS without Tween-20. Shielded from light, color development of these washed blots (e.g., alkaline phosphatase) was carried out without shaking in a 15-ml solution of 10 mM Tris–HCl (pH 8.8), 100 mM NaCl, 5 mM $MgCl_2$ that contained 45 μl 5-bromo-4-chloro-3-indolyl phosphate (50 mg/ml, w/v) and 25 μl 4-nitro blue tetrazolium chloride (100 mg/ml, w/v) dissolved sequentially in 40 and 70 % dimethylformamide w/v. Membranes containing the stained proteins were photographed (Figs. 3, 4 and 5).

Protein Sequence Alignment

For alignment, the proteins from MS and Western blot analyses were identified from the newly annotated *Nelumbo* genome (Ming et al. 2013) using standard proteomics experimental strategies. In addition, the deduced protein sequences were further verified by EST identification (expressed sequence tag; Gouet et al. 1999), BLASTed (Basic Local Alignment Tool, by NCBI; Camacho et al. 2009) and aligned. The sequences were aligned using MUSCLE (Edgar 2004) and visualized using ESPript (Gouet et al. 1999; Supplement 1).

PIMT Bioassay

The powdered embryo axis prepared as described above was extracted on ice by mixing for 10 min with 100 mM pH 7.5 HEPES composed of 10 mM β-mercaptoethanol, 10 mM Na-hydrosulfite, 10 mM sodium-metabisulfite, 1 μM Leupeptin, 25 μM PMSF and 10 % glycerol (Villa et al. 2006). After centrifugation at 14,000×g for 10 min at 4 °C the extracts were collected. Aliquots of the extracts were then heated for 10 min in water bath at room temperature (~22 °C), 30 °C, 50 °C, 80 °C and 100 °C, and re-centrifuged. The protein concentrations of the heat-treated supernatants were assayed by the Lowry method after precipitation with 10 % trichloro-acetic acid and use of the standards of bovine serum albumin.

The heat-treated protein samples were used for Western blotting with antibodies to Arabidopsis PIMT1 (above) and for determination of methyltransferase activity as described below. Twelve μl (13–46 μg protein) of each thermally treated supernatant was incubated for 60 min at 40 °C in a reaction mixture of substrate consisting 10 μM [^{14}C]AdoMet (60 mCi/mmol-1), 625 μM Val-Tyr-Pro-(L-isoAsp)-His-Ala [VYP-(L-isoAsp)-HA], and 200 mM phosphate pH 6.7 citrate buffer and having a final volume of 40 μl. The assay tubes were immediately frozen on dry ice, subsequently thawed on wet ice for 5–10 min, and quenched with 40 μl 0.2 N NaOH in 1 % NA-dodecylsulfate for hydrolysis of methyl esters to methanol. Sixty μl aliquots of the quenched reaction mixture of each treatment were spotted onto a triple-pleated 1.5×8 cm filter paper (No. 1650962, BioRad Lab) and placed in the neck of a 20-ml scintillation vial that contained 5 ml of the counting fluor (Safety Solve High Flashpoint Cocktail; Res Prod Intl). The vials were then capped and kept at room temperature for 2 h during which volatile [^{14}C]-methanol diffused into the fluor and unreacted [^{14}C]AdoMet remained on the filter paper. The reaction vials, after removal of filter paper, were counted by liquid scintillation. Reactions were assayed in triplicate (Fig. 6).

Results and Discussions

Nelumbo Thermal Proteins

Figure 1 shows 1-D SDS gels of Nelumbo fruit heat-soluble proteins at room temperature (Rt, ~22°), 30°, 50°, 80°, 100° and 110 °C isolated from embryo axis- and cotyledon-tissues of a modern and a 549-year-old fruit of China Antique,

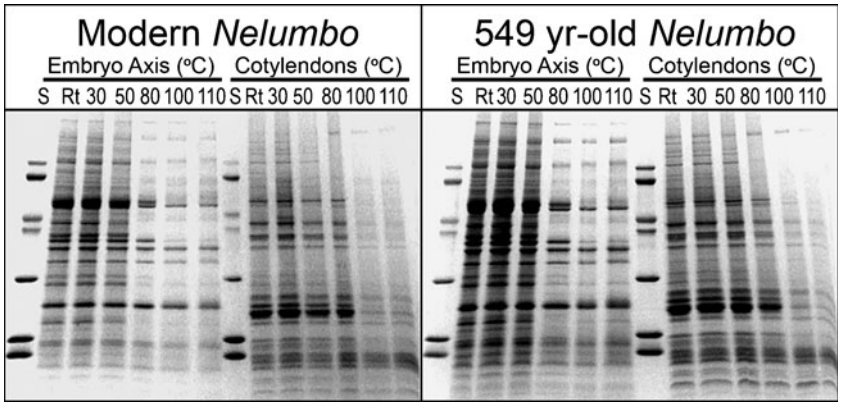

Fig. 1 1D-SDS gels of heat-soluble proteins in fruit of *Nelumbo nucifera* Gaertn var. China Antique: Embryo axes and cotyledons of a modern and a 549-year-old fruit heated for 10 min at room temperature (Rt, ~22°), 30°, 50°, 80°, 100°, and 110 °C; S, protein standards from top to bottom: 78, 67, 45, 39, 27, 18 and 12 kDa

compared with protein standards (S) having masses of 78 to 12 kDa. On the whole, these Nelumbo proteins of axes withstand heat without diminishing solubility through 50 °C and those of cotyledons through 80 °C (Fig. 1), proteins in the axes being evidently somewhat less heat stable than those of cotyledons. After five-and-a-half centuries of aging, proteins in the old fruit have remained remarkably intact (Fig. 1).

After 110 °C of heating, ~32 % of the proteins of the embryo axis of this old fruit and ~76 % of its cotyledons remained soluble (Table 1), a further example of the greater heat-stability of the proteins of Nelumbo cotyledons. Surprisingly, the percentages of soluble protein remaining in the old axis and cotyledons after 110 °C heating were more than twice those of its modern counterpart. This difference in thermal stability (Table 1), also visible in the SDS gels (Fig. 1), may reflect fruit maturity. That the modern fruits require more than twice as long (~6–9 days) to germinate than the old fruits (Shen-Miller et al. in press) may evidence their relative lack of maturity, accommodating lesser amount of heat-stress proteins (Table 1). The modern Nelumbo fruit is from a 3-year-old crop, harvested at the Kenilworth Aquatic Gardens, Washington DC. This fruit is an offspring of fruit originated from the dry lakebed in Xipaozi Village, China (Shen-Miller et al. 1995) where all of the old fruit used in such studies have been collected (e.g., Ohga 1926; Shen-Miller et al. 2002; Shen-Miller et al. in press).

In Table 1 the total amount of embryo axis proteins in room temperature (Rt, ~22 °C) of the modern control is essentially identical to that after 549 year of aging (70 $\mu g \cdot mg^{-1}$ modern vs. 73 $\mu g \cdot mg^{-1}$ old). Nevertheless, the cotyledon proteins of the old fruit (15 $\mu g \cdot mg^{-1}$) are markedly less than those of the modern (63 $\mu g \cdot mg^{-1}$). This difference may be a result of long-term metabolism of the cotyledons in the old fruit over hundreds of years. Notably, the proteins remaining in the old cotyledons are evidently more thermal stable than those of the modern control (Table 1; respectively, 76 % vs. 36 %, at 110 °C), a further example that relates protein heat stability to fruit maturity (cf. Table 1, old vs. modern embryo axis at 110 °C, respectively, 31 % vs. 10 %).

Thermal Protein Identification by MS Modern Nelumbo proteins of China Antique from Xipaozi Village, NE China were treated for 10 min at 100 °C. Supernatants of the crude protein of embryo axes and cotyledons were electrophoresed using the 1-D SDS gels (Fig. 2; embryo axis). Stained protein bands were boxed into 15 bins (Bins 1–15 vs. protein standards

Table 1 Thermostable proteins in embryo axes and cotyledons of a modern and a 549-year-old fruit of *Nelumbo nucifera* Gaertn var. China Antique, heated for 10 min at room temperature (Rt, ~22°), 30°, 50°, 80°, 100° and 110 °C; ± (standard deviation; $n=5$)

Temperatures (°C)	Rt	30	50	80	100	110
Embryo axis — modern						
$\mu g \cdot mg^{-1}$	70.3±12.3	63.5±1.0	66.2±14.8	32.6±16.2	9.1±2.0	7.1±5.3
Recovery, %	100	90±2	94±7	46±13	13±3	10±5
Embryo axis — 549-year-old						
$\mu g \cdot mg^{-1}$	73.0±3.3	–	100±34	31.8±0.3	–	22.6±3.6
Recovery, %	100	–	137±34	44±2	–	31±5
Cotyledons — modern						
$\mu g \cdot mg^{-1}$	62.6±3.4	–	57.0±15.8	32.9±5.1	17.0±3.4	21.9±3.5
Recovery, %	100	–	91±9	53±10	28±6	36±6
Cotyledons — 549-year-old						
$\mu g \cdot mg^{-1}$	14.7±5.9	–	17.7±1.2	17.5±5.4	–	11.1±0.2
Recovery, %	100	–	120±40	119±18	–	76±1

~200–15 kDa). The proteins from individual bins were processed for trypsin digestion and analyzed by liquid chromatography and tandem mass-spectrometry. Product ion spectra were matched to those calculated for putative Nelumbo proteins predicted from its genome (Perkins et al. 1999).

Eight Nelumbo thermal proteins of embryo axes and cotyledons were identified by sequence. Molecular masses of the identified Nelumbo proteins match well the protein standards. The MS identified thermal proteins are listed below, from small to large (Fig. 2, Bin-15 to Bin-3):

(a) CuZn-SOD (Bin-15, ~15 kDa)
(b) 1-CysPRX (Bin-14, ~25 kDa)
(c) CPN20 (Bin-14, ~25 kDa; Western blot)
(d) Vicilin (Bin-7, ~50 kDa)
(e) ENO1 (Bin-7, ~50 kDa)
(f) EF-1α (Bin-7, ~50 kDa)
(g) HSP80 (Bin-3, ~75 kDa)
(h) Met-Synthase (Bin-3, ~75 kDa)

Thermal Protein Identification by Western Blot Western blot verifications of the thermal stability of three Nelumbo proteins from embryo axes and cotyledons — dehydrin, PIMT and CPN60 — are presented in Figs. 3, 4 and 5:

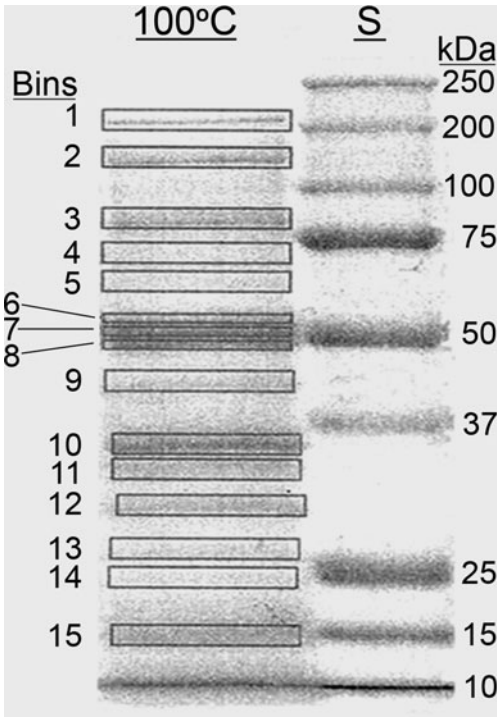

Fig. 2 Embryo-axis proteins of a modern fruit of *Nelumbo nucifera* Gaertn var. China Antique, heated for 10 min at 100 °C. S, protein standards: 250 to 10 kDa. Protein bands are boxed into Bins 1 through 15. Proteins eluted from these Bins have been identified by mass spectrometry (see Supplements 1 [aligned], 2)

Dehydrin (Fig. 3, ~22 kDa). Dehydrin is a LEA (late embryogenesis abundant) protein that responds to freezing temperature and draught (Close 1997). The ultra-high thermal stability of Nelumbo dehydrin is illustrated in the Western blot shown in Fig. 3, documenting the results for this protein in an embryo axis and cotyledons after heating for 10 min at various temperatures to 110 °C. Both in the axis and cotyledons, dehydrin remained intact to 110 °C with no diminished intensity (noted by the arrows at ~22 kDa in Fig. 3). This heat-stability is consistent with the very high concentration (55 %) in dehydrin of the thermostable polar-residues D-E-K-R-Y (Table 2, below). The Nelumbo embryo axis evidently contains a higher concentration of dehydrin than the cotyledons.

PIMT (Fig. 4, ~25 kD). PIMT recognizes L-isoaspartyl residues spontaneously produced from L-asparaginyl and L-aspartyl residues during aging and stress (Clarke 2003). These isomerized residues are methyl-esterified by PIMT at the α-carboxyl via the S-adenosyl-L-methionine-dependent reaction, followed by non-enzymatic reactions that lead to conversion of the normal L-aspartyl residues. The Western blot in Fig. 4 shows a clear band of reactivity at ~25 kDa that corresponds closely to the position noted for the Arabidopsis recombinant PIMT1 and PIMT2αω proteins (Xu et al. 2004). The 25 kDa immunoreactive species persist in the supernatant of heated fractions to 50 °C and a small amount of the enzyme is visible in extracts heated to 80 °C and 100 °C (Fig. 4). Strong staining bands of an unidentified 50 kDa species, perhaps derived from an unrelated Nelumbo protein that cross-reacted with the antibody to the Arabidopsis protein, occur in samples heated up to 50 °C.

CPN60 (Fig. 5, ~60 kDa). In the restorative folding of nascent or damaged proteins, a "lid-and-basket" structure is formed between chaperonins CPN20 and CPN60 (Thirumalia and Lorimer 2001). Thermal stability of Nelumbo CPN20 (~27 kDa; the lid) identified by sequence (above) is further verified by Western Blot of embryo axis and cotyledons through heat treatments to 110 °C (Fig. 5). In contrast, for both tissues, the CPN60 (~60 kDa; the basket) is heat-stable only through 50 °C (Fig. 5) Judging from the stain intensities of proteins in the axis and cotyledons of Nelumbo, the axis appears to contain more of both these chaperonins than the cotyledons (Fig. 5).

Alignment of Thermal-Protein Sequences of Nelumbo Embryo-Axes and Cotyledons

Data of 11 Nelumbo thermal protein alignments with those of the archaeal hyperthermophile *Methancaldococcus jannaschii* (Mj) and five mesophile angiosperms (Arabidopsis, corn,

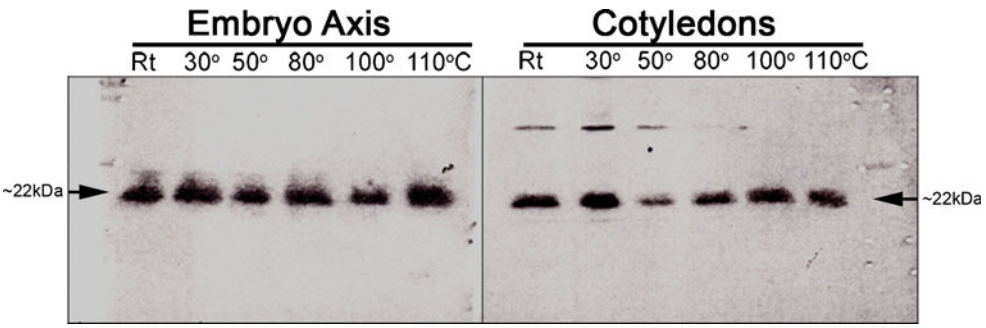

Fig. 3 Western blots of dehydrin antibodies of barley to heat-soluble proteins of embryo axis and cotyledons of modern fruit of *Nelumbo nucifera* Gaertn var. China Antique. Crude protein heated for 10 min at room temperature (RT, ~22°), 30°, 50°, 80°, 100°, 110 °C; *arrows* indicate the location of dehydrin, ~22 kDa

grape, poplar and soybean) are presented in the Supplementary Material 1 (Supplemental Figs S-1a to S-1k):

(a) **CuZn-SOD** (cotyledons; NNU_001676; Fig. 2 Bin15, ~17 kDa; Supplemental, Fig. S-1a). As a result of the mid-Precambrian Great Oxidation Event ~2.4 billion years ago (Holland 2002), when free oxygen first became a significant component of Earth's atmosphere, aqueous Fe^{2+} became scarce and soluble Cu^{2+} became the dominant cofactor at the active site of dismutases (Alscher et al. 2002). CuZn-SOD, a superoxide dismutase of aerobic bacteria and eukaryotes (Alscher et al. 2002) thus appears to be a relatively late evolutionary innovation (i.e., not present for the earliest half of Earth history) that has no evident sequence-similarity with the earlier-evolved, much more ancient Fe-SOD and Mn-SOD of other organisms or with that of the archaeal hyperthermophile Mj. Hence, the antioxidative plant enzyme CuZn-SOD, not present in Mj (Fig. S-1a) appears to have become prevalent at ~2.4-Byr when it first played an important role in catalyzing the dismutation of superoxide (O^-, a reactive oxygen species, ROS) and thereby protecting cellular integrity.

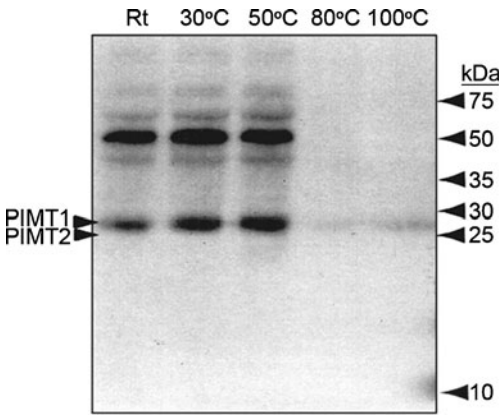

Fig. 4 Western blot of PIMT1 (protein L-isospartyl methyltransferase) antibodies of Arabidopsis to heat soluble-proteins of modern fruit of *Nelumbo nucifera* China Antique embryo-axis extracts heated for 10 min at different temperatures; PIMT1 and PIMT2 arrows at ~25 kDa denote PIMTs of Arabidopsis (Xu et al. 2004)

The CuZn-SOD of Nelumbo cotyledons, NNU_001676, is highly conserved among such angiosperms as Arabidopsis, corn, grape, soybean and poplar (Supplemental Fig. S-1a), composed of 152 residues (~16.7 kDa) and having two cysteines, Cys56 and Cys145, that form a disulfide bridge identical to that of the CuZn-SOD in the halophyte *Polygonum sibirium* Laxm (Qu et al. 2010) and harsh-arid monocot pearl-millet *Pennisetum glaucum* (Mahanty et al. 2012). In *P. sibirium*, the two CuZn-SOD catalytic sites (GFHVHALGD, the first site at amino acid position 43–51, and GNAGGRI*ACGII at 137–148) are essentially identical both in their composition and sequence location to those of Nelumbo and other plants (Supplemental Fig. S-1a). The first active site in Nelumbo is totally conserved, and the second differs only by having a "V" in place of "I*" (Supplemental Fig. S-1a). NNU_001676 is composed of ~18 % polar residues that form thermally stable intramolecular ion-pairs (D-E-K-R-Y, Unsworth et al. 2007), with Asp, Lys and Glu being the most abundant.

(b) **1-CysPRX** (NNU_016172; Fig. 2 Bin14, ~25 kDa; Supplemental Fig. S-1b). In animals, 1-Cys-peroxiredoxin is bifunctional, exhibiting both glutathione peroxidase and phospholipase A_2 activities (Chen et al. 2000). It has the simultaneous catalytic role of regulating phospholipid turnover and of providing oxidative protection by detoxifying reactive oxygen ROSs. The two active sites performing these functions have, respectively, the consensus sequences of $GDS_{32}WG$ and $PVC_{47}TTE$ (Chen et al. 2000). In plants as well as Mj, the active site of peroxisdase $PVC_{46}TTE$ is highly conserved (Supplemental Fig. S-1b), but the phospholipase sequence (in which Ser_{32} is the catalytic nucleophile) is not present in the Mj-angiosperm alignments (Supplemental Fig. S-1b). Nelumbo has two 1-CysPRXs: NNU_016172 is predominantly conserved among the plants as well as in the archaean Mj (Supplemental Fig. S-1b), consisting of ~213 residues (~23.4 kDa). A Nelumbo PER 12 peroxidase is a thermostable protein listed in Supplement 2.

(c) **CPN20** (NNU_010559; Fig. 2 Bin14, ~27 kDa; Supp1emental Fig. S-1c). Chaperonin20 is a

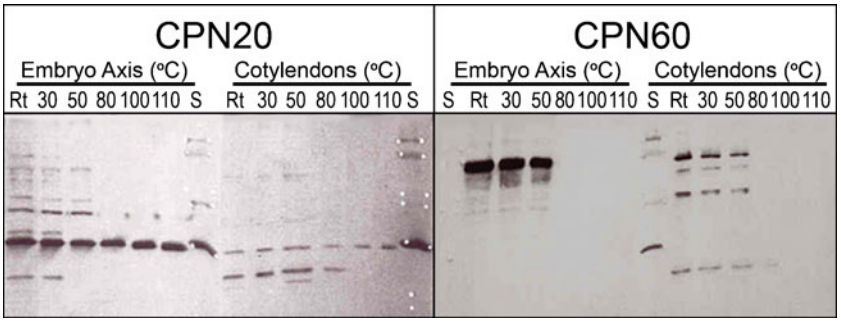

Fig. 5 Western blots of CPN20 (~27 kDa) and CPN60 (~60 kDa) antibodies, respectively, of spinach and pea to heat-soluble proteins of embryo axis and cotyledons of modern fruit of *Nelumbo nucifera* China Antique heated for 10 min at different temperatures; S, protein standards from top to bottom: 78, 67, 45, 39, 27, 18 and 12 kDa (cf. under CPN20 cotyledon, S, "*white-dots*")

molecular chaperone that together with its companion, chaperonin CPN60 (Supplemental Fig. S-1k), forms protein-folding machinery powered by ATP serves to restore nascent or misfolded proteins to functional conformation (Thirumalia and Lorimer 2001). CPN20 is a highly heat-stable protein through 110 °C both in the Nelumbo embryo axis and cotyledons (Fig. 5). NNU_010559 and NNU_010102 are dimers. Of these, NNU_010102 has sequence similarity more closely similar to a dimer of *Homo sapiens* than to that either of grape or rice whereas two other dimers, including NNU_010559, are similar to those of Arabidopsis. CPN20 (NNU_010559, ~27.1 kDa) is composed of ~246 residues. With the exception of the first ~40 residues at the N-terminus, the remaining ~200 amino acids are highly conserved among the aligned angiosperms (Supplemental Fig. S-1c). The "MASI" motif at the N-terminus of NNU_010559 reflects its localization in chloroplasts as is evidenced in an Arabidopsis chloroplastic CPN10 (Hill and Hemmingsen 2001). In the Nelumbo protein, amino acids 54–63 (SIK**PLGDRVL),** 76–80 **(GILLP)** and three shorter segments up to amino acid 96 show similarity or identity (shown by pink or red type in Supplemental Fig. S-1c) to the Arabidopsis CPN10. CPN20, per se, has not been identified in the archaean Mj; however, the larger companion chaperonin CPN60 is present (Supplemental Fig. S-1k).

(d) **Vicilin** (NNU_007171, Fig. 2 Bin7, ~54 kDa; Supplemental Fig. S-1d). Vicilin, a component of the plant seed-storage protein 7S-globulin, has the dual function of providing nutritional reserve and microbial defense (De Souza et al. 2011). The vicilin NNU_007171 of Nelumbo, similar to that of *Macademia integrifolia*, contains a short hydrophobic N-terminus (Marcus et al. 1999). During seed germination, two similar motifs of *Macademia* vicilin (Cys-X-X-X-Cys) have been identified as promoting the release of exudates that exhibit antimicrobial activity. Nelumbo vicilin NNU_007171 contains three Cys residues, at positions 10, 150, and 348, with long stretches of Xs between. Vicilin is not found in the archaean Mj, and the aligned vicilin proteins show similarity but less identity among the aligned plants (Supplemental Fig. S-1d), differing especially at their N- and C-termini. In all of the aligned plant sequences compared with Nelumbo vicilin, most identities occur in two sets of amino acid pairs, PY and AG, between positions 60–70 and 420–430. In these alignments, the longest sequence of similarity to Nelumbo is EQIKAMS at position 243–249. Vicilin NNU_007171 (~53.5 kDa) is composed of ~486 residues, ~30 % of which are the thermal-stable polar amino acids (Unsworth et al. 2007) from the N-terminus to the mid-region of the protein.

(e) **ENO1** (NNU_002362, NNU_020386, NNU_021001; Fig. 2 Bin7, ~48–53 kDa; Supplemental Fig. S-1e). Three Nelumbo enolase1s have been aligned (Supplemental Fig. S-1e). All of these show high conservation with the ENO of Mj and with those of the angiosperms analyzed, including corn (Lal et al. 1998) that exhibits similarity at the N-terminus of the protein (viz., MAVT**I**TWVK**AR**Q, in which the amino acid differences are underlined and identities in bold). This same terminal sequence occurs also in the maize enolase isozymes ANP45A and ANP45B (Lal et al. 1998). Activity of ENO1 in corn roots is induced 10-fold under anaerobic conditions and the two isozymes appear to remain functional during long-term anoxia when the activity of other enolases declines to a nondetectable level (Lal et al. 1998). The longest of these, NNU_002362 (~53.4 kDa), is composed of ~485 amino acids. In Nelumbo, the ENO1 protein has a Tyr53 residue (similar to the Tyr46 in vertebrate ENO1) that has been shown to be phosphorylated (Lal et al. 1998). The regions adjacent to the Tyr(Y) residue in Nelumbo are highly conserved among the aligned species except Mj (Supplemental Fig. S-1e). In animals, dephosphorylation at this site is reported to occur under anoxia that increases activity of the enzyme (Lal et al. 1998).

Enolase1 is a highly conserved polypeptide. The consensus sequence of its anaerobic responsive element, **AR**E at the 5′-terminus, is present in all organisms examined (Van Der Straeten et al. 1991) including Nelumbo and the others analyzed (Supplemental Fig. S-1e, with identities to Nelumbo shown in bold type) in which in plants, the E is replaced by Q at position 18. Six active sites that are homologous in the enolases of plants and vertebrates (E-170, D-248, E-289, DD-314-5, K-339 and K-402; Van Der Straeten et al. 1991) are fully conserved among Nelumbo, the archaean Mj, and the aligned angiosperms (red columns in Supplemental Fig. S-1e).

(f) **EF-1α** (NNU_024576, NNU_024577, NNU_026673; Fig. 2 Bin7, ~42–49 kDa; Supplemental Fig. S-1f). Three Nelumbo elongation-factor EF-1α proteins have been identified, two of which, NNU_024576 and NNU_024577 are in tandem and exhibit 100 % sequence identity (Fig. S-1f). The third protein, NNU_026673, shows sequence similarity with this tandem pair except for having fewer residues toward the C-terminus and containing a total of 377 amino acids (~41.5 kDa) rather than 443 (48.7 kDa; Supplemental Fig. S-1f). All three Nelumbo EF-1αs exhibit sequence conservation with the other aligned plant sequences and the archaean Mj (Supplemental Fig. S-1f). The Nelumbo EF-1α proteins have sequence similarities also with those of carrots, wheat and barley in the regions at positions 6–30 (VHI**N**IVV**IGHVD** S**GKSTT**T**G**HLI**Y**K), 167–174 (EV**S.**SYL**K**), 274–294 (S**V**E**MHHE**TLQE**AL**P**GDN**V**GFN**) and 297–304 (N**V**AV**KD**L**K**; in which bold type indicates identity with Nelumbo; Durso and Cyr 1994). The "lack of alignment" (the lack of residue alignments in pink or red) at the C-termini of these proteins in Supplementary Figure S-1f is due to the absence of sequences in the relatively short Nelumbo protein NNU_026673; otherwise their C-terminal sequences are of high similarity.

(g) **HSP80** (NNU_010290, Fig. 2 Bin3, ~77 kDa; Supplemental Fig. S-1g). Heat-shock protein 80 falls in the HSP70-90 group of the evolutionarily conserved hsps. The Nelumbo HSP80, NNU_010290 (~76.9 kDa), composed of 699 residues, shows a very high degree of conservation in comparison with those of the other aligned angiosperms (Supplemental Fig. S-1g). Of these, the Nelumbo N-terminus exhibits the longest sequences identical to those of the other species, beginning from amino acid 5 followed by 34 residues and flanked by an identical amino acid A at each end: (AETFAFQAEINQLLSLIINTFYSNKEIFLRELISNA; Supplemental Fig. S-1g). This long sequence and many more other identical residues are present also in the HSP80 both of tomato (Koning et al. 1992) and *Triticum aestivum* (GeneBank CAA67191.1). Interestingly, the HSP80 protein is absent in the hyperthermophile Mj. The Nelumbo HSP80 contains a large number (264) of the heat-stable polar residues D-E-K-R-Y that comprise ~37.8 % of this protein (Unsworth et al. 2007). These polar amino acids are most abundant in the 170–290 and 410–585 regions (Supplemental Fig. S-1g).

(h) **Met-Synthase** (NNU_013651; Fig. 2 Bin3, ~84 kDa; Supplemental Fig. S-1h). The methionine synthase of Nelumbo, comprised of 765 residues (~84.2 kDa), is another highly conserved protein among the angiosperms aligned (Supplemental Fig. S-1h). The Arabidopsis protein present in the alignment (Supplemental Fig. S-1h) is a cytosolic cobalamin-independent Met-synthase (GenBank U97200). The archaean Mj has a small Met-Synthase composed of ~308 amino acids, fewer than half those of the plant homologues (Mj alignment data not included). However, the first ~70 residues at the N-terminus of the plant proteins (viz., **M**AS.I**VG**.**YP**RM..**K**R**E**.....**ESF**.**D**.....**L**QKV etc. to ~70 residues) are identical or similar to those of the Mj. In Mj, all of the residues occurring after amino acid position 83 are highly conserved in the last third of the plant sequences through the C-terminus.

(i) **Dehydrin** (NNU_006332, NNU_013652, NNU_013851; Fig. 3, ~22 kDa; Supplemental Fig. S-1i). Thermal stability of Nelumbo dehydrin is shown by the Western blot (Fig. 3). The dehydrins of the plants analyzed do not evidence strong similarity in their alignment (Supplemental Fig. S-1i). It is mainly a result of missing sequence-regions in grape, corn1, Arabidopsis4, 5, soy1 and Nelumbo. Nevertheless, all plants, including those for which data are presented in Supplemental Figure S-1i, have been shown to typically exhibit conserved and repetitive dehydrin sequences, in which "k-segments" (EKKGIMDKIKEKLPG) are of common occurrence (Close 1997).

Supplemental Figure S-1i compares the alignments of three Nelumbo dehydrins (NNU_006332, _013652, _013851) with other angiosperms. Nelumbo dehydrin NNU_013851 (the longest of its three) has a K-like segment beginning at amino acid 85 (**EKKGI**LE**K IKEKLPG;** bold-type indicating identity, the underlined residues being similar to those of the above K-segment), and the occurrence of protonatable His-residues at either end (Eriksson et al. 2011) not immediately adjacent to the segment. The dehydrins of Arabidopsis1 and 2 have identical K-like segments at this same location and His-residues situated similarly to those in Nelumbo dehydrin NNU_013851. This dehydrin is composed of ~197 amino acids and has an average molecular mass of ~21.7 kDa, a mass similar to that of thermal proteins shown on the Western blot (Fig. 3; arrows, ~22 kDa). Of all Nelumbo thermal proteins thus far assembled,

dehydrin has the richest concentration (54.6 %, Table 2) of D-E-K-R-Y heat-stable polar amino acids, as well as demonstrating a high heat-stability (Fig. 3).

A Nelumbo medium-length dehydrin protein, NNU_006332, has a K-segment (**EKKG**MMD **KIKEKLPG)** that differs at the three underlined residues from NNU_013851 but is identical to that of Arabidopsis5 having only one His-flanking residue close to its lead-end. The third and shortest Nelumbo dehydrin (NNU_013652) is composed of a hefty ~38 % of thermostable amino acids and exhibits a slightly different K-segment (KKKGEH.....G) at the same position of the K-segments of the other two dehydrins of Nelumbo and the other analyzed plants. This dehydrin has an additional His-residue within the segment. The plant dehydrin proteins have no equivalent in the archaean Mj.

(j) **PIMT** (NNU_002938, NNU_004234; Fig. 4, ~25–33 kDa; Supplemental Fig. S-1j). Protein repair L-isoaspartyl methyltransferase (PIMT) is a Nelumbo protein for which thermal stability has been verified by use of Western blot (Fig. 4). Similar to Arabidopsis (Xu et al. 2004), Nelumbo has two PIMTs in its genome, one long (NNU_002938, composed of 301 amino acid residues) and one short (NNU_004234, 231 residues). Over 229 of these residues, the amino acid sequences of these two Nelumbo proteins are 89 % identical. The longer one has an additional 72 residues at its N-terminus not present in the shorter form (Supplemental Fig. S-1j).

The longer Nelumbo protein (NNU_002938) shows 63 % identity over 286 residues to the PIMT2 of Arabidopsis (Villa et al. 2006), including some identities in the N-terminal 72 sequences of Nelumbo. In Arabidopsis, PIMT2 is located in the nucleus, suggesting that the Nelumbo 002938 species may also be nuclear. The shorter form of the Nelumbo PIMT shows 69 % identity with PIMT1 of Arabidopsis over 230 residues.

The alignment of PIMTs in higher plants with that of the archaean Mj demonstrates a very high degree of conservation among these enzymes, indicating their ancient origin. For example, the L-isoasparate substrate-binding motifs at residues 59–67 (**TISA**P**HM**HA) and 202–209 (**VRY**V**PL**TS) are similar or identical in all plants sequences studied (Griffith et al. 2001) as well as the archaean Mj enzyme (Supplemental Fig. S-1j). The nuclear location signal "NLS" (KIIKKRKKKMR) of the lead sequences of Arabidopsis PIMT2 (Xu et al. 2004) is present in the lead sequences of Soy2 PIMT as KRKSEKKKKMR (with both signals having been confirmed by the NLStradamus NLS prediction program; Nguyen et al. 2009). The nuclear location of the forms of PIMT in corn2, Nelumbo NNU_002938, and poplar2 have not been shown, nor have the Arabidopsis PIMT2 and Soy2 PIMT shown to be plastid-, mitochondrial- or ER-localized (Predotar Program; Small et al. 2004). The PIMTs of corn1 and poplar2 could be situated in plastids, mitochondria and/or ER, but no clear Predotar-based data are available to evidence the location of the long Nelumbo PIMT NNU_002938 in these cellular components. We detected no immunoreactive species corresponding to the expected size of NNU_002938 ~33 kDa on the Western blot (Fig. 4), suggesting that it may be synthesized in lower amounts than the shorter form NNU_004234.

(k) **CPN60** (NNU_011934, NNU_023642; Fig. 5, ~60 kDa; Supplemental Fig. S-1k). Chaperonin 60 is the third Nelumbo protein for which thermal stability is evidenced by use of Western blot (Fig. 5). CPN60 (cpn60β) is a homologue of *E coli* GroEL (Viitanen et al. 1995). Two CPN60s are represented in the Nelumbo genome. Both are conserved among the aligned plants and the archean Mj, and are ~96 % identical, composed of ~542 amino acids, and have a molecular mass of ~59.6 kDa.

High sequence similarity to the other taxa analyzed is exhibited by Nelumbo CPN60 in which all aligned sequences in various regions are mostly similar, some identical (Supplemental Fig. S-1k). The Mj sequences, at positions 75–107 and 152–170 Supplemental Fig. S-1k) the so-called equatorial domains, exhibit high conservation with the ATP-dependent thermosome Mj protein MJ0999 (~60 kDa), that has been shown to be up-regulated at 95 °C (Kowalski et al. 1998; Boonyaratanakornkit et al. 2005). The highly conserved motif of Nelumbo **GDGTT** beginning at position 93 (Supplemental Fig. S-1k) has been shown to form the ATP-binding site for GroEL (Kowalski et al. 1998). This conserved motif of CPN60 is found in all of the aligned angiosperms and the archean Mj (Supplemental Fig. S-1k), as well as in *Methanopyrus kandleri* (98 °C growth optimum), *Pyrococcus* sp. (100 °C growth optimum) and yeast (Kowalski et al. 1998).

In addition to the 11 Nelumbo thermal proteins discussed above, 30 other heat-stable proteins of Nelumbo China Antique, identified by mass spectrometry but not yet aligned, are listed in Supplementary Material 2.

Thermostable Polar Amino-Acids in Nelumbo Proteins

Table 2 summarizes the distribution of thermostable, polar amino-acids (D-E-K-R-Y) present in the 11 Nelumbo thermal proteins, listed in the descending order of their abundance. Of these, amino acids capable of forming intramolecular thermally stable ion-pairs (Unsworth et al. 2007) are dominant: Lys, Glu and Asp (Table 2). The dehydrin of Nelumbo particularly

Table 2 Heat-stable protein polar amino-acid residues (D-E-K-R-Y; Unsworth et al. 2007) in fruit of *Nelumbo nucifera* var. China Antique, and ± protein sequence alignments of Nelumbo with those of Mj, *Methancaldococcus jannaschii* (85 °C optimal growth), aa, amino acid

Thermoprotein	Most abundant residues	% Total aa	Mj
Dehydrin	Glu, Lys	55	
HSP80	Lys/Glu, Asp	38	
CPN60	Lys, Glu	33	+
Vicilin	Glu, Lys, Arg/Asp	30	
EF-1α	Lys, Asp, Glu	29	+
Met-Synthase	Glu, Lys, Asp	26	+
1-CysPRX	Lys, Asp	26	+
ENO1	Lys, Glu	24	+
CPN20	Lys, Glu, Asp	22	
PIMT	Glu, Lys, Asp	21	+
CuZn-SOD	Asp, Lys, Glu	18	

stands out, being composed of 55 % of such heat-stable residues, as well as showing high heat-stability (Fig. 3).

PIMT contains 21 % thermostable polar amino acids (Table 2). This protein-repair enzyme retains 100 % solubility and activity through 50 °C (Figs. 3 and 6). Chaperonin CPN60 is similarly heat soluble through 50 °C (Fig. 4) but contains a much higher percentage (33 %) of such polar residues than PIMT (Table 2). On the other hand, although PIMT and Chaperonin CPN20 differ substantially in thermal solubility (Figs. 4 and 5), the abundance of polar amino acids in these enzymes is essentially identical (~22 % and 21 %, respectively; Table 2). Protein thermal stability is not determined by the presence of polar residues alone. Vieille and Zeikus (2001) and Unsworth et al. (2007) have reported that in addition to the occurrence of such amino acids, the presence of other substances (e.g., salts, polyamines, soluble sugars) and post-transcriptional modifications (e.g., glycosylation) can also contribute to protein heat stability.

Notably, 55 % of the Nelumbo thermal proteins are present in the archaean hyperthermophile *Methancaldococcus jannaschii* (85 °C growth optimum; Table 2), an indication of the high thermal stability and antiquity of these proteins.

PIMT Bioassay

PIMT is the only Nelumbo thermal protein for which enzymatic activity has been tested at differing temperatures. The methyltransferase activity of three such supernatant PIMT-containing fractions in the presence and absence of a peptide substrate is illustrated in Fig. 6. The data presented show that the heated extracts of Nelumbo are able to catalyze the methylation of a synthetic substrate L-isoaspartyl-containing peptide. A robust peptide-dependent activity (red curve in Fig. 6) is shown in samples at room temperature (~22 °C) as well as those heated to 30° and 50 °C, whereas those heated to 80° and 100 °C show smaller but readily detectable peptide-dependent activity (~9 and ~6 %, respectively). Significantly, these data show that although the concentration of Nelumbo PIMT decreases with heat treatment (Fig. 4), the proteins remaining at 80° and 100 °C, present in lower concentrations, remain biologically functional (Fig. 6).

Concluding Discussion

The guanine-cytosine (G-C) content of RNA is evidently correlated directly with the optimal growth temperature of many organisms (Unsworth et al. 2007). Nelumbo, a basal eudicot, has a G-C content of 38 % (Ming et al. 2013), higher than that of such core-eudicot angiosperms as Arabidopsis, poplar and grape, but lower than that of corn. The optimum and/or maximum growth temperature of plants may be related to the maintenance-temperature of their protein solubility/functional activity. In hyperthermophiles, the optimal activities of most enzymes coincide with optimal temperatures of growth (cf.

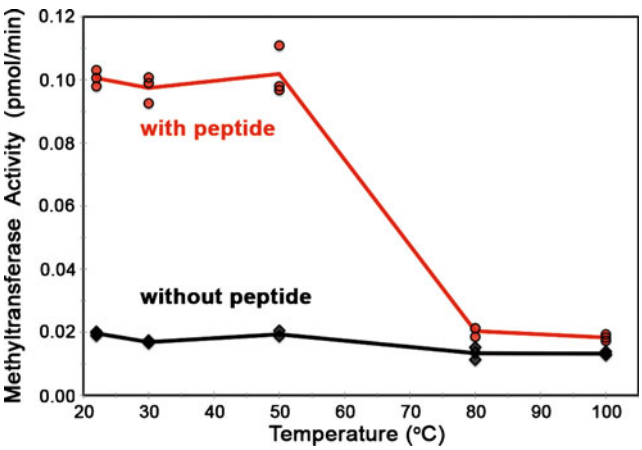

Fig. 6 PIMT bioassay of heat stability of the activity of protein L-isoaspartyl methyltransferase in soluble extracts of embryo axis of modern *Nelumbo nucifera* Gaertn var. China Antique, heated for 10 min at ~22°, 30°, 50°, 80°, 100 °C. The enzyme activity is shown in the presence (*red curve*) and absence (*black curve*) of a synthetic substrate L-isoaspartyl-containing peptide in triplicate assays

Vieille and Zeikus 2001). Several proteins of the anaerobic methanogen hyperthermophile *Methancaldococcus jannaschii* (85 °C growth optimum) are highly conserved in Nelumbo (Table 2, Supplment 1). Under anaerobic conditions, such as that within the pericarp of Nelumbo, such proteins can maintain solubility and remain enzymatically active.

The functions of the Nelumbo fruit thermoproteins are numerous, playing a role in antioxidation, membrane-maintenance, chaperonine/chaperon protein-restoration, anaerobic glycolysis, microbial defense, provision of nutrients, cellular repair and abatement of environmental stress. Taken together, these factors seem certain to have contributed to the centuries-long viability of Nelumbo fruit.

The 11 Nelumbo thermal proteins studied can be grouped into the four functionally defined categories discussed below.

- Antioxidant and Membrane Maintenance (CuZn-SOD, 1-CysPRX, dehydrin).

  (1) **CuZn-SOD** Nelumbo is a widespread crop throughout Asia. In China its habitat extends from the near-tropics to the frigid northern reaches of Heilongjiang Province, bordering Siberia (Huang 1987). Gill et al. (2010) found versatility in the CuZn superoxide dismutases of alpine plants collected at high altitudes in the Himalayas, having "an autoclavable property and a reaction maximum at 0 °C." CuZn-SOD is one of four SODs identified in plants that not only have antioxidative properties but are regarded as important anti-aging and heat-tolerant enzymes (Qu et al. 2010; Chen et al. 2011). Ninety percent of the CuZn-SOD activity of Nelumbo remains intact at 65 °C, and 20 % after 95 °C for 30 min (Chen et al. 2011). Such enzyme heat stability in the Nelumbo is much higher than that of the C_4 pearl-millet (Mahanty et al. 2012).

  Plant CuZn-SOD functions mainly in the cytoplasm and chloroplasts, and can initiate activity under mild water-stress (Alscher et al. 2002) and salt-stress (Qu et al. 2010) with increased expression of SOD activity being correlated with an increased tolerance of membrane damage (Kwon et al. 2002). Analyses of the cotyledons of a viable ^{14}C-dated 466-year-old Nelumbo China Antique fruit and two viable undated fruit from the same Holocene lakebed at Xipaozi Village showed high concentrations of polyunsaturated lipids, indicating the maintenance of membrane fluidity likely due to a lack of extensive auto-oxidation (Priestley and Posthumus 1982).

  (2) **1-CysPRX** (1-Cys-peroxiredoxin) is regarded to have evolved during the early history of life (Dietz 2011), serving as a detoxifier of ROS released by early-evolved oxygenic photosynthesizers and the nonbiologic photodissociation of water and CO_2 (Schopf 2011). Workers on circadian rhythms go so far as to suggest that the redox homeostatic mechanisms of PRX and associated cellular time-keeping biology may have co-evolved after the Great Oxidation Event (GOE) 2.4 billion-year ago (Edgar, et al. 2012). That this Nelumbo protein is present in *Methancaldococcus jannaschii* suggests an even more ancient evolutionary history as methanogens seem certain that have been extant ~2.8 billion years ago and probably earlier (Schopf 1994). The gene coding for 1-CysPRX is strongly expressed during seed germination (rice; Dietz 2011). 1-CysPRX in the hermetically enclosed Nelumbo embryo axis could provide antioxidative protection during centuries of seed-aging. The notable expansion of redoxin genes of the Nelumbo genome (Ming et al. 2013), coding for enzymes involved in antioxidation and cellular O_2-damage repair, may have played a crucial role enhancing long-term Nelumbo seed viability.

  (3) **Dehydrin** a late embryogenesis abundant protein, is up-regulated during seed maturation when it functions in combating freezing-, drought- and salinity-stress (Close 1997). It is understandable that the archaean Mj, having an 85 °C growth optimum and inhabiting the vicinity of submarine thermal vents, has no need for such a protein (Supplemental Fig. S-1i). Given the functions of dehydrin for detoxification and protection from cold stress, this protein presumably evolved later than the early Precambrian origin of oxygenic photosynthesis but before the first major glacial epoch ~2.7 billion years ago.

  All three of the dehydrin-related Nelumbo proteins contain its characteristic K-segment (see above). Protonation of the His residues flanking such K-segments induces membrane binding that by stabilizing membrane topology may modulate lipid fluidity (Eriksson, et al. 2011). Such potential of membrane fluidity has been documented by Priestley and Posthumus (1982) in old China Antique Nelumbo fruit in which dehydrin may have played a role in the viability of the centuries-old fruit.

- Chaperonins, Chaperone, and Stress Proteins (Cpn20, Cpn60, HSP80, EF-1α).

  (1 and 2) **CPN20 and CPN60** are molecular chaperones, responsible for such cellular processes as protein folding, assembly, translocation and degradation. CPN20 is a dimer of CPN10, an equivalent of the bacterial GroES of *E. coli* (Viitanen et al. 1995). With coordinated ATP hydrolyses between CPN60

and the co-chaperonin CPN20, mal-folded proteins can be rearranged to functional conformation states (Thirumalia and Lorimer 2001). The genome of Nelumbo is coded for at least seven CPN10s and five CPN20s (Ming et al. 2013). It would be interesting to know whether the greater heat stability of Nelumbo CPN20 (the "lid;" Fig. 5) serves to stabilize the companion component that forms the underlying "basket" (CPN60) after conjugation for protein folding (Fig. 5). A thermoprotein yajl Chaperone protein is listed in Supplement 2.

(3) The **HSP80** of Nelumbo, a stress protein, is highly conserved with those of the other angiosperms here aligned (Supplemental Fig. S-1g). Because this hsp is absent from the hyperthermophile Mj (Supplemental Fig. S-1g), a member of an especially ancient lineage, this heat-shock protein seems likely to be a later evolved. Interestingly, however, Mj has small hsps that are induced under stress. The Mj HSP16.5 protein forms a large homomeric complex of 24 subunits having a molecular mass of ~400 kDa (Kim et al. 2003). This huge complex of small hsps functions to prevent protein aggregation at elevated temperature and, thus, may serve the same purpose as the chaperones composed of the larger hsps in plants. Supplement 2 lists two additional thermally stable Nelumbo heat-shock species, HSP18.2 and HSP70.

(4) **EF-1α** elongation factors are regarded to be among the slowest evolving enzymes; they exhibit highly conserved sequences (Gaucher et al. 2001) and have diverse functions in protein biosynthesis, chaperone activity (cf. HSP45), disulfide bond formation, protection from abiotic and biotic stresses, and protein degradation (Fu et al. 2012). The occurrence of a tandem pair of EF-1α proteins in Nelumbo reinforces their functional effectiveness. The protein elongation factors in plant cytoplasm are homologs of the EF-Tu in plastids, mitochondria, and bacteria (Fu et al. 2012). Transgenic crop plants containing EF-Tu developed heat stress tolerance, and the induced expression of EF-1α has been shown to prolong the lifespan of *Drosophila* (Tatsuka et al. 1992).

- Anaerobic Glycolytic Protein (Enolase1).

ENO1 is an ancient glycolytic enzyme originating early in Earth history when the environment was anoxic. In the process of glycolysis — present in all known organisms — ENO1 converts 2-phosophoglycerate to phosphoenopyruvate, PEP (Van Der Straeten et al. 1991). For survival under anaerobic conditions, most plant cells shift metabolism from the oxidative to the fermentative mode and ENO1 comes into action, exhibiting intense activity particularly in non-green tissues. The glycolytic pathway of anaerobic respiration evidently provided the cellular energy required to sustain viability of the hermetically sealed old Nelumbo fruit, particularly the non-green tissues in cotyledons and the embryonic radicle. This may explain the apparent diminution of protein content in the cotyledons of the 549-yr-old fruit in comparison with its modern counterparts (Table 1).

- Defense, Food and Repair (vicilin, Met-Synthase, PIMT).

(1) **Vicilins** and cruciferins (legumins) are major members of 7S-globulin seed-storage proteins (Tiedmann et al. 2000; De Souza et al. 2011). In addition to seven vicilins, Nelumbo has 19 cruciferins and three cruciferin-like proteins (Ming et al. 2013). In addition to storage, these proteins have antimicrobial functions (discussed above). The heaviest staining band on the SDS-PAGE gels in Figs. 1 and 2, corresponding to a mass of ~50 kDa, shows the location of these most abundant Nelumbo seed-storage proteins. Although not present in the archaean Mj, the vicilin protein has an ancestral chimeric genome derived both from a gram-negative bacterium and most probably an archaean (Dunwell et al. 2000). The vicilins are desiccation resistant and by their storage-function provide energy for germinating embryos. The vicilin of peanut, Ara h1, is highly resistant to harsh environments (like that of highly acidic mammalian intestines and factory food-processing facilities; Dunwell et al. 2000). The Nelumbo Ara h1 has been identified as being thermally stable (Supplement 2) as have been Nelumbo 11S-globulin β, FA02 13S-globulin and legumin proteins LEGA and LEGB.

(2) **Met-Synthase** catalyzes the biosynthesis of methionine via the cobalamine independent pathway present in plants and some prokaryotes including archaeans. In addition to providing monomeric blocks for protein building, Met is the precursor of AdoMet (S-adenosyl-L-methionine), a major methyl donor in numerous cellular reactions (cf. PIMT protein repair). In plants, Met plays a central role in the cellular metabolism of protein synthesis, methyl-group transfer, and the biosynthesis of ethylene and polyamines (Ravanel et al. 1998), all processes of relevance to aging-repair, stress protection and protein heat-stability. AdoMet is an immediate intermediate in the biosynthesis of the

plant stress-hormone ethylene (Ravanel et al. 1998). The Met-mediated production of ethylene is typically promoted by such abiotic stresses as heat, water, cold, salinity, hypoxia and water deprivation (Morgan and Drew 1997). The Met-Synthase in plants appears to have undergone major changes following its early origination in archaeans such as *Methancaldococcus jannaschii* (Supplemental Fig. S-1h, discussed above).

(3) **PIMT** an important widespread protein-repair enzyme, was tested to determine its capability after heating to catalyze the methylation of a synthetic L-isoaspartyl-containing polypeptide (discussed above). Our results show PIMT activity at 100 °C (Fig. 6) indicating that the Nelumbo protein exhibits greater thermostability than the recombinant PIMT enzymes of humans and Arabidopsis heated to ~65 °C (Villa et al. 2006). Interestingly, all of these PIMT enzymes exhibit enzymatic activity at elevated temperatures (Thapar et al. 2001), suggesting that the repair of isomerization damage is significant in resistance to heat stress and that enzyme-repair by PIMT is likely to be an important factor to the maintenance of viability of the centuries-old Nelumbo fruit. Supplementary Material 2 lists an additional Nelumbo thermostable methyltransferase (METE) identified by mass spectrometry.

Summary and Future Research

All the assembled 11 thermostable proteins identified in Nelumbo fruit function in the protection and repair of cellular processes, properties reflected in the high rate of germination of their centuries-old seeds and the notable heat-tolerance of their enzymatic activity that have evidently promoted healthy aging and long-term viability. Fifty-five percent of these proteins occur also in the microbe *Methancaldococcus jannaschii*, a modern member of the Archaea that dates to ~3.0 billion years ago. Thus, these thermostable proteins are exceedingly ancient and evolutionarily well-tested, enzymes that have changed slowly over geological time and that today serve to maintain the viability of Nelumbo fruit for as long as 1,300 years (Shen-Miller et al. 1995). The remarkable long-term existence of these proteins provides firm evidence of the adaptive value they have played in the history of life.

Of these 11 Nelumbo thermostable proteins, only one, the protein-repair enzyme PIMT, has been tested for functionality after having been heated to 100 °C (Fig. 6). Bioassays of the other thermal proteins here studied, and those listed in Supplementary Material 2, need to be performed. The purified CuZn-SOD protein of Nelumbo has been shown to have heat tolerance and to maintain activity at high temperatures (Chen et al. 2011; Ding et al. 2008). Another Nelumbo heat-tolerant protein, annexin, is up-regulated at 90 °C (Chu et al. 2012). Two additional such 90 °C up-regulated thermal proteins are listed in Supplement 2: aldose reductase and phosphoglycerate kinase. Plant annexin is a multifunctional stress protein, its binding to lipids having a membrane protective role as evidenced by the membrane fluidity in centuries-old Nelumbo seeds (Priestley and Posthumus 1982). The Nelumbo genome is now available for broad-range research (Ming et al. 2013), a recent advance that may prove to be a basis for increased understanding of seed aging and longevity, and the betterment of agricultural crops.

From the data presented above, we suggest below some "fruit-for-thought" questions for future studies of the Nelumbo fruit:

1) How long a maturation period is required for accumulation of heat-stabled proteins?
2) How does protein heat-solubility relate to protein activity?
3) What factors govern protein heat-stability?
4) In addition to the 11 enzymes discussed above and the 30 thermostable proteins identified in Supplement 2, what other heat resistant proteins occur and how might they be related to stress and the exceptional long-term viability of the Nelumbo fruit?

Acknowledgments We thank the reviewers for valuable suggestions, C. Haas-Blaby for bioinformatics and protein alignments, J.W. Schopf for constructive comments, J. Lowenson for helpful discussions, and H. Nguyen for compilation assistance of the MS data. For antisera, we are grateful to P. Viitanen for CPN20 and CPN60, T. Close for dehydrin, and B. Downie and D. Martin for PIMT1. Work in SGC laboratory is supported by NIH grant GM206020. We also thank K.O. Stetter whose seminal research on hyperthermophiles stimulated this 'hot protein' study of Nelumbo enzymes.

References

Alscher RG, Erturk N, Heath LS (2002) Role of superoxide dismutases (SODs) in controlling oxidative stress in plants. J Exp Bot 53(327):1331–1341

Bersch U, Soll J, Seetharam R, Viitanen P (1992) Identification, characteristic and DNA sequencing of a functional "double" GroES-like chaperonine from chloroplasts of higher plants. Proc Natl Acad Sci (USA) 89:8696–8700

Boonyaratanakornkit BB, Simpson AJ, Whitehead TA et al (2005) Transcriptional profiling of the hyperthermophilic methanoarchaeon *Methanococcus jannaschii* in response to lethal heat and non-lethal cold shock. Environ Microbiol 7(6):789–797

Bradford MM (1976) A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 27:248–254

Bult CJ, White O, Olsen GJ, et al (1996) Complete genome sequence of the methanogenic archaeon *Methanococcus jannaschii*. Science 273:1058–1073

Burnett WN (1981) "Western blotting" electrophoretic transfer of proteins from sodium dodecylsulfate-polyacrylamide gels to unmodified nitrocellulose, and radiographic detection with antibody and radio-iodinated protein A. Anal Biochem 112:195–203

Camacho C, Coulouris G, Avagyan V et al (2009) BLAST+: architecture and applications. BNC Bioinforma 10:421

Chang YJ (1978) Thousand-year-old *Nelumbo* has awakened. Fossil 1:22–23 (in Chinese)

Chen CH, Chen SM, Zhou KS (1965) Palynological analysis of the Holocene Nymphaea seed-bearing deposits at the vincinity in Liaoning Peninsula. Quaternaria Sin 4:167–173 (in Chinese)

Chen JW, Dodia C, Feinstein SI et al (2000) 1-Cys peroxiredoxin, a Bifunctional enzyme with glutathione peroxidase and phospholipase A_2 activities. J Biol Chem 275(37):28421–28427

Chen D, Zheng X, Li G et al (2011) Molecular cloning and expression of two cytosolic copper-zinc superoxide dismutases genes from *Nelumbo nucifera*. Appl Biochem Biotechnol 163:679–691

Chu P, Chen H, Zhou Y et al (2012) Proteomic and functional analyses of *Nelumbo nucifera* annexins involved in seed thermotolerance and germination vigor. Planta 235:1271–1288

Clarke S (2003) Aging as war between chemical and biochemical processes: protein methylation and the recognition of age-damage proteins for repair. Ageing Res Rev 2:263–285

Close TJ (1997) Dehydrin, a commonality in response of plants to dehydration and low temperature. Physiol Plant 100:291–296

De Souza CE, Grossi-De-Sa MF, Lima TB et al (2011) Plant storage proteins with antimibrobial avtivity novel insights into plant defense mechanisms. FASEB 25(10):3290–3305

DeWeerdt SE (2002) The first sequenced extremophile, what scientists have learned from the *M. jannaschii* genome. Genome News Network, Feb 1, 2002 www.genomenewsnetwor.org/article/02_03/extremo.shtml

Dietz KJ (2011) Peroxiredoxins in plants and cyanobacteria. Antioxid Redox Signa 15(4):1129–1159

Ding YF, Cheng HY, Song SQ (2008) Changes in extreme high-temperature tolerance and activities of antioxidant enzymes of sacred lotus seeds. Sci China Ser C: Life Sci 51(9):824–853

Dunwell JM, Kuri S, Gane PJ (2000) Microbial relatives of the seed storage proteins of higher plants: conservation of structure and diversification of function during evolution of the cupin superfamily. Microbiol Mol Biol Rev 64(1):153–179

Durso NA, Cyr RJ (1994) A calmodulin-sensitive interaction between microtubules and a higher plant homolog of elongation factor 1-α. Plant Cell 6:893–905

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32(5):1792–1797

Edgar RS, Green EW, Zhao Y et al (2012) Peroxiredoxins are conserved markers of circadian rhythms. Nature 485:459–467

Eriksson SK, Kutzer M, Procek J et al (2011) Tunable membrane binding of the intrinsically disordered dehydrin Lti30, a cold induced plant stress protein. Plant Cell 23:2391–2404

Esau K, Kosakai H (1975) Laticifers in Nelumbo nucifera Gaertn.: distribution and structure. Am Bot 39:713–719

Fu J, Momcilovic I, Prasad PVV (2012) Roles of protein synthesis elongation factor EF-Tu in heat tolerance in plants. J Bot ID 835836:8pp

Gaucher EA, Miyamoto MM, Benner SA (2001) Function-structure analysis of proteins using covarion-based evolutionary approaches elongation factors. USA PNAS 98:548–552

Gill T, Kumar A, Ahuja PS, Sreenivasulu (2010) Over-expression of *Potentilla* superoxide dismutase improve salt stress tolerance during germination and growth in *Arabidopsis thaliana*. J Plant Genet Trangenics 1:1–10

Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: multisequence alignments in postscript. Bioinformatics 15:305–308

Griffith SC, Sawaya MR, Boytz DR et al (2001) Crystal structures of a repair methyltransferase from *Pyrococcus furiosus* with its L-isoaspartyl peptide substrate. J Mol Biol 313(5):1103–1116

Hill JE, Hemmingsen SM (2001) *Arabidopsis thaliana* type I and II chaperonins. Cell Stress Chaperones 6(3):190–200

Holland HD (2002) Volcanic gases, black smokers, and the great oxidation event. Geochim Cosmochim Acta 66(21):3811–3826

Huang GH (1987) Systematic and distribution of *Nelumbo nucifera* Gaertn. In: China Nelumbo, Chp 2, Academia Sinica Wuhan Bot Inst, Science Publ pp 9–12 (in Chinese)

Huang SZ, Tang XJ, Lu CB et al (2000) Characteristic of superoxide dismutase in lotus seeds. Acta Physiol Sin 26(6):492–496, English abstract

Kasting JF, Howard MT (2006) Atmospheric composition and climate on the early earth. Philos Trans R Soc Lond Bio Sci 361(1474):1733–1742

Kim R, Lai L, Lee HH et al (2003) On the mechanism of chaperone activity of the small heat-shock protein of *Methanococcus jannaschii*. PNAS-USA 100(14):8151–8155

Koning AJ, Rose R, Comai L (1992) Developmental expression of tomato heat-shock cognate protein 80. Pl Physiol 100:801–811

Kowalski JM, Kelly RM, Konisky J et al (1998) Purification and functional characterization of a charperone from *Methanococcus jannaschii*. Syst Appl Microbiol 21:173–178

Kwon SY, Jeong YJ, Lee HS et al (2002) Enhanced tolerances of transgenic tobacco plants expressing both superoxide dismutase and ascorbic peroxidase in chloroplasts against methyl viologen-mediated oxidative stress. Plant Cell Environ 25:873–882

Laemmli UK (1970) Cleavage of structural proteins during assembly of the head of the bacteriophage. Nature 227:680–685

Lal SK, Lee C, Sachs MM (1998) Differential regulation of enolase during anaerobiosis in maize. Pl Physiol 118:1285–1293

Mahanty S, Kaul T, Pandey P et al (2012) Biochenical and molecular analyses of copper-zinc superoxide dismutase from a C_4 plant *Pennisetum glaucum* reveals an adaptive role in response to oxidative stress. Gene 505:309–317

Marcus JP, Green JL, Goulter KC, Manners JM (1999) A family of antimicrobial peptides is produced by processing of a 7S globulin protein in *Macadamia integrifolia* kernels. Plant J 19(6):699–710

Ming R, VanBuren R, Liu Y et al (2013) Genome of the long-living sacred lotus (*Nelumbo nucifera* Gaertn.). Genome Biol 14(5):R41

Morgan PW, Drew MC (1997) Ethylene and plant responses to stress. Physiol Plant 100:620–630

Nelson DR, Schuler MA (2013) Cytochrome P450 genes from the sacred lotus Genome. Trop Plant Biol. doi:10.1007/s12042-013-9119-z

Nguyen BA, Pogoutse A, Provart N, Moses AM (2009) NLStradamus: a simple hidden Markov model for nuclear signal prediction. BMC Bioinformatics 10:202. doi:10.1186/1471-2105-10-202

Ohga I (1926) On structure of some ancient, still viable fruits of Indian lotus, with special reference to their prolonged dormancy. Jpn J Bot 3:1–20

Ohga I (1927) Supramaximal temperature and life duration of the ancient fruits of Indian lotus. Bot Mag 41:161–172

Pennington SR, Dunn MJ (2001) Proteomics, from protein sequence to function. BIOS Sci Publ Ltd, UK, p 313

Perkins DN, Pappin DJC, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551–3567

Priestley DA (1986) Seed aging, implication of seed storage and persistence in the soil. Comstock Publ Assoc, Ithaca, 304 pp

Priestley DA, Posthumus MA (1982) Extreme longevity of lotus seeds from Pulantien. Nature 299(9):148–149

Qu CP, Xu ZR, Liu GJ et al (2010) Differential expression of copper-zinc superoxide dismutase gene of *Polygonum sibiricum* leaves, stems and underground stems, subjected to high-salt stress. Int J Mol Sci 11:5234–5245

Ravanel S, Gakiere B, Job D, Douce R (1998) The specific features of methionine biosynthesis and metabolism in plants. Proc Natl Acad Sci USA 95:7805–7812

Robert F, Chaussidon M (2006) A paleotemperature curve for the Precambrian oceans based on silicon isotopes in chert. Nature 443:969–972

Schopf JW (1994) The oldest known records of life: stromatolites, microfossils, and organic matter from the early Archaean of South Africa and Western Australia. In: Bengtsen S (ed) Early life on earth. Columbia Univ Press, NY, pp 193–206

Schopf JW (2011) The paleobiological record of photosynthesis. Photosynth Res 107:87–101

Shaw MF (1929) A microchemical analysis study of the fruit coat of *Nelumbo lutea*. Am J Bot 16:259–276

Shen-Miller J, Mudgett MB, Schopf JW et al (1995) Exceptional seed longevity and robust growth: ancient sacred lotus from China. Am J Bot 82(11):1367–1380

Shen-Miller J, Schopf JW, Harbottle G et al (2002) Long-living lotus: germination and soil γ-irradiation of centuries-old fruits, and cultivation, growth, and phenotypic abnormality of offspring. Am J Bot 89(2):236–247

Shen-Miller J, Aung LH, Turek J, Schopf JW, Tholandi M, Yang M, Czaja A (in press) Centuries-old viable fruit of Sacred Lotus *Nelumbo nucifera* Gaertn var. China Antique. Trop Plant Biol

Shevchenko A, Wilm M, Vorm, Mann M (1996) Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gells. Annal Chem 68:850–858

Small T, Peeters N, Legeai F, Lurin C (2004) Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics 4(6):1581–1590

Stetter KO (1996) Hyperthermophilic prokaryotes. FEMS Microbiol Rev 18:149–158

Tatsuka M, Mitsui H, Wada M et al (1992) Elongation factor-1 alpha gene determines susceptibility to transformation. Nature 359(6393):333–336

Tele Images-Nature (2003) *Eternal Seeds*, a documentary of *Nelumbo nucifera*: Xipaozi, China, USA, Japan & Africa. In: *Power Plants* (Series #4). Dir L. Frapat, Assist Dir I. Han, Camera P. Moreau, Sound P. Fleurant. Paris, France, 50 min

Thapar N, Kim AK, Clarke S (2001) Distinct patterns of expression but similar biochemical properties of protein L-isoaspartyl methyltransferase in higher plants. Plant Physiol 125(2):1023–1035

Thirumalia D, Lorimer G (2001) Chaperonine-mediated protein folding. Ann Rev Biophys Biochem Struc 30:245–269

Tiedmann J, Neubohn B, Muntz K (2000) Different functions of vicilin and legumin are reflected in the histopattern of globulin mobilization during germination of vetch (*Vicia sativa* L.). Planta 211:1–12

Towbin H, Staehlin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedures ans some applications. Proc Nat Acad Sci (USA) 76:4350–4354

Unsworth LD, Van Der Ost J, Koutsopoulos S (2007) Hyperthermopholic enzymes – stability, activity and implementation strategies for high temperature applications. FESB J 274: 4044–4056

Van Der Straeten D, Rodrigues-Pousada A, Goodman HM, Van Montagu M (1991) Plant enolase: gene structure, experession, and evolution. Plant Cell 3:719–735

Vieille C, Zeikus GJ (2001) Hyperthermophlic enzyme sources, uses, and molecular mechanisms for thermostability. Microbiol Mole Biol Rev 65(1):1–43

Viitanen PV, Schmidt M, Bucher J et al (1995) Functional characterization of the higher plant chloroplast chaperonins. J Biol Chem 270:18158–18164

Villa ST, Xu Q, Downie AB, Clarke SG (2006) Arabidopsis protein repair L-isoaspartyl methyltrransferase: predominant activities at lethal temperatures. Physiol Plant 128:581–592

Xu Q, Belcastrp MP, Villa ST et al (2004) A second protein L-isoaspartyl transferase gene in Arabidopsis produces two transcripts whose products are sequestered in the nucleus. Pl Physiol 136:2652–2664