International Journal of Botany Studies
www.botanyjournals.com
ISSN: 2455-541X
Received: 07-08-2026, Accepted: 09-09-2026, Published: 24-09-2026
Volume 11, Issue 3, 2026, Page No. 406-410

Isolation and characterization of small heat shock protein coding genes from slipper orchids: A review

Dr. Manju M George
Department of Botany, Union Christian College (Autonomous), Aluva, Ernakulam District, Kerala, India
DOI: https://doi.org/10.66856/ijbs.2026.11.3.11239

Abstract
Slipper orchids, particularly *Paphiopedilum* species, are valuable ornamental plants that frequently experience temperature fluctuations and other abiotic stresses. Small heat shock proteins (sHSPs), also called HSP20 proteins, are ATP-independent molecular chaperones that protect cellular proteins from stress-induced unfolding and aggregation. This review discusses the isolation and molecular characterization of sHSP-coding genes from slipper orchids, with emphasis on gene discovery, sequence analysis, phylogeny, subcellular localization, and stress-responsive expression. Plant sHSPs are characterized by variable amino-terminal regions and a conserved carboxyl-terminal α-crystallin domain. Their diversity reflects expansion of the sHSP gene family and specialization among cytosolic, chloroplastic, mitochondrial, endoplasmic-reticulum, and other cellular compartments. In slipper orchids, reverse-transcription PCR, degenerate-primer PCR, rapid amplification of cDNA ends, transcriptome sequencing, and quantitative RT-PCR provide complementary approaches for isolating and characterizing these genes. Although chloroplast genomes of several *Paphiopedilum* species have been analyzed, nuclear-encoded sHSP genes remain comparatively underexplored. Studying these genes could improve understanding of orchid thermotolerance, acclimatization, conservation physiology, and breeding.

Keywords: *Paphiopedilum*, slipper orchid, small heat shock protein, HSP20, gene isolation, molecular characterization, heat stress, thermotolerance

Introduction

Slipper orchids are members of the orchid subfamily Cypripedioideae and include the genera *Paphiopedilum*, *Phragmipedium*, *Cypripedium*, *Mexipedium*, and *Selenipedium*. *Paphiopedilum* is particularly important in horticulture because of its distinctive shoe-shaped labellum, attractive foliage, long-lasting flowers, and extensive species and hybrid diversity. The genus is also important for conservation research because many species are threatened by habitat destruction, illegal collection, and restricted geographical distributions.

Temperature stress is a major factor affecting orchid growth, photosynthesis, flowering, reproductive development, and survival. At high temperature, cellular proteins may unfold or aggregate, membranes may lose stability, and metabolic reactions may become disrupted. Plants respond through a coordinated heat-shock response involving heat shock transcription factors and several classes of heat shock proteins (Hsps). Among these, small heat shock proteins are particularly abundant and diverse in plants (Vierling, 1991; Sun *et al*., 2002) [5, 7].

Plant sHSPs generally range from approximately 12 to 42 kDa and possess a conserved α-crystallin domain, also known as the HSP20 domain. The domain is flanked by variable amino-terminal and carboxyl-terminal regions that contribute to substrate recognition, oligomerization, and subcellular specificity (Waters and Vierling, 2020) [8]. Isolation of sHSP-coding genes from slipper orchids can therefore provide important information about the molecular mechanisms that support survival under heat and other environmental stresses.

Biological Characteristics of Plant sHSPs

Structure and chaperone function

sHSPs are ATP-independent molecular chaperones. Rather than actively refolding proteins using ATP, they bind partially unfolded proteins and prevent irreversible aggregation. The trapped substrates can later be transferred to ATP-dependent chaperone systems, including HSP70 and HSP100, for refolding or degradation (Sun *et al*., 2002; Waters and Vierling, 2020) [5, 8].

The central structural feature of a plant sHSP is the α-crystallin domain, which contains approximately 100 amino acids and forms a conserved β-sandwich structure. This domain is flanked by variable N-terminal and C-terminal extensions (Basha *et al*., 2011) [1]. sHSP monomers commonly associate into dimers and larger oligomeric complexes. These oligomers are dynamic and can change in response to temperature, substrate binding, phosphorylation, and other cellular signals.

The N-terminal region is generally more variable than the α-crystallin domain. It contributes to substrate recognition and may explain why different sHSPs protect different proteins or operate under different physiological conditions. The short C-terminal extension also contributes to oligomer stability and inter-subunit interactions.

Subfamilies and localization

Plant sHSPs form a large multigene family. They are classified according to sequence similarity and intracellular localization. The principal groups include cytosolic/nuclear classes, chloroplast-localized sHSPs, mitochondrial sHSPs, endoplasmic-reticulum sHSPs, and other organelle-associated proteins (Scharf *et al*., 2001; Waters and Vierling, 2020) [4, 8].

The complexity of this family was demonstrated in *Arabidopsis thaliana*, where genomic analysis identified 13 sHSPs distributed among six major classes, in addition to related proteins containing α-crystallin domains (Scharf *et al*., 2001) [4]. This finding showed that plant sHSPs are not a single homogeneous group but a diversified family with potential specialization among organs, developmental stages, and stress conditions.

Subcellular localization is often predicted from the amino-terminal sequence. Chloroplast- and mitochondria-targeted sHSPs generally contain transit peptides, whereas endoplasmic-reticulum proteins may contain signal peptides. These predictions are particularly important when characterizing newly isolated slipper-orchid genes because the localization of the encoded protein can provide clues about its function.

Stress and developmental responses

Although sHSPs are best known for their response to heat, they can also be induced by drought, salinity, oxidative stress, ultraviolet radiation, high light, chilling, and pathogen infection (Sun *et al*., 2002; Waters and Vierling, 2020) [5, 8]. Some sHSP genes are stress-inducible, whereas others are constitutively expressed.

In addition to environmental stress, sHSPs participate in developmental processes such as seed maturation, embryogenesis, pollen development, flowering, and fruit ripening. This broad range of expression is relevant to slipper orchids because their vulnerable stages include seed germination, protocorm development, tissue-culture acclimatization, and floral development.

Importance of sHSP Research in Slipper Orchids

The environmental diversity of slipper orchids suggests that different species may possess distinct stress-response strategies. Some *Paphiopedilum* species occur in shaded tropical forests, whereas others occupy montane, limestone, or seasonally dry habitats. These ecological differences may be reflected in the sequence, expression, and regulation of nuclear stress-response genes.

Molecular studies of *Paphiopedilum* have already demonstrated the value of genomic data for understanding species relationships. Sun *et al*. (2022) [6] compared the complete chloroplast genomes of nine *Paphiopedilum* species and identified differences in genome structure, gene content, repetitive sequences, and variable regions. A larger plastome study analyzed 77 *Paphiopedilum* accessions and reported plastome sizes ranging from 152,130 to 164,092 base pairs (Liu *et al*., 2021) [3]. These studies provide an important genomic foundation, but chloroplast data alone cannot reveal the complete nuclear sHSP repertoire.

Isolation and characterization of nuclear sHSP genes could help clarify:

- The number and diversity of sHSP genes in slipper orchids.
- The evolutionary relationships among orchid sHSP subfamilies.
- Species-specific variation in heat-response genes.
- The relationship between gene expression and habitat adaptation.
- The molecular basis of heat tolerance during cultivation and acclimatization.
- Potential markers for conservation and germplasm management.

Methods for sHSP Gene Isolation

Plant material and heat treatment

Young leaves are often selected for gene isolation because they provide relatively high-quality RNA and contain actively expressed genes. To identify heat-responsive sHSP transcripts, plants may be divided into control and treatment groups. Heat-treated plants should be exposed to a defined temperature for a specified period, followed by sampling at several time points.

A suitable experiment may include:

1. Control plants maintained at the normal growth temperature.
2. Plants exposed to moderate heat stress.
3. Plants exposed to severe heat stress.
4. Samples collected during early stress, prolonged stress, and recovery.
5. Biological replicates for each treatment.

Leaves, roots, floral buds, flowers, developing capsules, and protocorm-like bodies should also be compared because sHSP expression may differ among tissues.

RNA extraction and cDNA preparation

RNA isolation from orchids can be difficult because tissues may contain polysaccharides, phenolic compounds, mucilage, and other substances that interfere with extraction. A modified cetyltrimethylammonium bromide method or a commercial plant RNA kit can be used, often with polyvinylpyrrolidone and reducing agents.

The RNA should be treated with DNase to remove contaminating genomic DNA. RNA integrity and purity should be assessed before cDNA synthesis. First-strand cDNA can be synthesized using oligo(dT) primers, random primers, or a combination of both.

Degenerate-primer PCR

When a complete slipper-orchid genome is unavailable, degenerate-primer PCR is a practical method for isolating partial sHSP transcripts. Conserved amino-acid sequences from known plant and orchid sHSPs are aligned, and degenerate nucleotide primers are designed against conserved regions of the HSP20 domain.

The general procedure involves:

1. Aligning available sHSP sequences.
2. Identifying conserved nucleotide or amino-acid motifs.
3. Designing degenerate primers.
4. Amplifying partial cDNA fragments by RT-PCR.
5. Cloning and sequencing the amplified products.
6. Confirming the presence of the conserved HSP20 domain.

Because highly expressed genes are more likely to be amplified, degenerate PCR may preferentially recover abundant cytosolic sHSPs. Less abundant, divergent, or organelle-specific genes may require transcriptome sequencing or gene-specific approaches.

Rapid amplification of cDNA ends

Rapid amplification of cDNA ends is used to obtain the complete 5′ and 3′ regions of a partial sHSP transcript. Gene-specific primers are designed from the known internal sequence, and separate 5′-RACE and 3′-RACE reactions are performed.

A complete sHSP cDNA normally contains:

- A 5′ untranslated region.
- An open reading frame.
- A 3′ untranslated region.
- A polyadenylation region.

Full-length sequences are necessary for predicting the complete protein, identifying organelle-targeting peptides, designing expression primers, and distinguishing closely related paralogues.

Transcriptome and genome sequencing

RNA sequencing can identify many sHSP transcripts simultaneously and is particularly useful when no reference genome is available. Transcriptomes from control and heat-treated plants can be assembled *de novo*, and candidate HSP20 sequences can be identified through similarity searches and domain analysis.

Genome sequencing provides additional information about:

- Exon–intron organization.
- Promoter sequences.
- Tandem gene duplication.
- Gene copy number.
- Regulatory regions.
- Possible pseudogenes.

Because HSP20 domains can also occur in multidomain proteins, candidate sequences should be confirmed using conserved-domain analysis, phylogenetic placement, subcellular-targeting prediction, and expression evidence.

Molecular Characterization of Isolated Genes

Nucleotide and protein sequence analysis

The nucleotide sequence of each isolated cDNA should be analyzed to identify the open reading frame, untranslated regions, and predicted polyadenylation site. The deduced protein sequence can then be used to calculate:

- Amino-acid length.
- Predicted molecular mass.
- Theoretical isoelectric point.
- Amino-acid composition.
- Conserved residues.
- Hydrophobicity.
- Potential targeting peptides.

The presence of the HSP20/α-crystallin domain is the principal criterion for assigning a sequence to the sHSP family. However, the complete sequence should also be compared with known orchid and monocot proteins.

Conserved domain analysis

The conserved α-crystallin domain should be identified using databases such as Pfam, SMART, and the NCBI Conserved Domain Database. The domain generally occupies the C-terminal portion of the protein, while the N-terminal region is more variable.

Conserved structural motifs may support the predicted ability of the protein to form dimers or oligomers. Nevertheless, sequence similarity alone does not prove chaperone activity. Biochemical assays are required to determine whether the encoded protein prevents thermal aggregation or assists substrate refolding.

Subcellular localization

Localization prediction should be performed for every candidate gene. Chloroplast transit peptides, mitochondrial targeting peptides, signal peptides, and peroxisomal targeting motifs may indicate specialized cellular functions.

Predictions can be experimentally tested using green fluorescent protein fusions. For example, the N-terminal region of a candidate slipper-orchid sHSP can be fused to GFP and expressed in plant cells. Fluorescence microscopy can then determine whether the protein is cytosolic, nuclear, chloroplastic, mitochondrial, or associated with another organelle.

Phylogenetic analysis

Phylogenetic analysis is useful for assigning isolated genes to sHSP subfamilies. The analysis should include sequences from:

- The target slipper-orchid species.
- Other *Paphiopedilum* species.
- Related orchid genera.
- Representative monocots and eudicots.
- Appropriate outgroup proteins.

Multiple sequence alignment should include both the conserved α-crystallin domain and informative N-terminal regions. Maximum-likelihood, neighbor-joining, or Bayesian methods may be used, with bootstrap support or posterior probabilities reported.

A sequence clustering with cytosolic class I or class II sHSPs is likely to have a different expression pattern and function from a sequence clustering with chloroplast or mitochondrial sHSPs. Phylogenetic assignment should, however, be combined with localization prediction and expression data.

Expression Analysis

Quantitative RT-PCR

Quantitative RT-PCR is a widely used method for studying sHSP expression in response to heat. Primers should be designed in regions that distinguish closely related genes. Primer specificity can be evaluated by melting-curve analysis, gel electrophoresis, and sequencing of representative amplicons.

Reference genes must be validated under heat-treatment conditions because commonly used housekeeping genes may change in abundance during stress. At least two stable reference genes should preferably be used.

Expression analysis should include:

- Control and heat-treated plants.
- Several treatment temperatures or durations.
- Recovery after heat stress.
- Different organs.
- Independent biological replicates.
- Technical replicates.

Expected expression patterns

Heat-responsive sHSP genes commonly show rapid transcript accumulation after exposure to high temperature, followed by either sustained expression or decline during recovery (Sun *et al*., 2002; Waters and Vierling, 2020) [5, 8]. However, individual genes may differ substantially in timing and magnitude of induction.

Some genes may be strongly induced in leaves but not roots, whereas others may be associated with seeds, reproductive

tissues, or protocorm development. Therefore, the expression of each gene should be analyzed independently rather than assuming that the entire sHSP family responds uniformly.

Transcriptomics and co-expression

Transcriptome analysis can place sHSP expression within a broader stress-response network. sHSP induction may occur alongside changes in genes encoding:

- Heat shock transcription factors.
- HSP70 and HSP90 proteins.
- Antioxidant enzymes.
- Osmoprotectant biosynthetic enzymes.
- Membrane-stabilizing proteins.
- Abscisic-acid signaling components.
- Proteases and autophagy-related proteins.

Co-expression analysis can identify candidate genes associated with acquired thermotolerance. However, RNA abundance does not necessarily reflect protein accumulation or activity, and transcriptomic results should be complemented with protein and physiological measurements.

Functional Significance

Sequence–function relationships

The α-crystallin domain indicates that an isolated protein belongs to the HSP20 family, but functional activity should be experimentally demonstrated. Recombinant slipper-orchid sHSPs can be expressed in *Escherichia coli* and purified for *in vitro* assays.

Potential assays include:

- Measurement of thermal aggregation by light scattering.
- Protection of enzyme activity during heating.
- Substrate-refolding assays with HSP70 or HSP100.
- Size-exclusion chromatography to examine oligomerization.
- Native PAGE or blue-native PAGE.
- Thermal-stability assays.
- Analysis of intracellular aggregation using fluorescent reporters.

A candidate that is highly heat-induced but inactive in a standard chaperone assay may have a specialized developmental or regulatory function. Conversely, a constitutively expressed sHSP may provide basal protection under normal conditions.

Promoter analysis

When genomic sequence is available, promoter regions can be examined for heat shock elements and other regulatory motifs. Heat shock factors recognize repeated nGAAn motifs in heat shock elements, while additional promoter sequences may connect sHSP transcription with drought, oxidative stress, abscisic acid, light, or developmental pathways.

Comparing promoters among slipper-orchid species may reveal whether differences in heat responsiveness are associated with regulatory sequence variation. This approach could be particularly informative for species occupying contrasting habitats.

Gene duplication and diversification

Plant sHSP families have expanded through gene duplication and subsequent functional diversification. Duplicated genes may become specialized for different tissues, organelles, developmental stages, or stress conditions (Scharf *et al*., 2001; Waters and Vierling, 2020) [4, 8].

In slipper orchids, one paralogue might be heat-inducible in leaves, another might function during seed maturation, and a third might protect chloroplast proteins. Consequently, isolation of a single sHSP gene should not be interpreted as a complete representation of the species' sHSP family.

Applications in Conservation and Horticulture

Conservation physiology

sHSP expression may serve as a molecular indicator of stress sensitivity. Populations or species that rapidly induce protective sHSPs may tolerate short-term heat episodes more effectively than those with weaker responses. Expression profiles could therefore complement measurements of survival, photosynthetic performance, membrane stability, and recovery.

Tissue culture and acclimatization

Plantlets produced *in vitro* often experience physiological stress when transferred to greenhouse or natural conditions. Monitoring sHSP expression during acclimatization may identify sensitive stages and help optimize temperature, humidity, light, and watering regimes.

Breeding and hybrid selection

sHSP genes could serve as candidate markers for selecting parents or hybrids with improved heat responses. However, sHSP expression should not be used as the sole measure of thermotolerance because heat tolerance is controlled by multiple mechanisms, including antioxidant defense, membrane composition, water relations, photosynthetic stability, and developmental timing.

Molecular identification

Chloroplast genomic studies have shown that *Paphiopedilum* species contain informative variable regions that can assist phylogenetic and taxonomic analyses (Liu *et al*., 2021; Sun *et al*., 2022) [3, 6]. Nuclear sHSP sequences could provide complementary markers for germplasm identification, population studies, and conservation genetics, particularly when combined with plastid and nuclear ribosomal markers.

Limitations and Future Directions

Several limitations should be considered in slipper-orchid sHSP research. Partial cDNA sequences may be incorrectly interpreted as complete genes, and degenerate PCR may fail to recover low-abundance or divergent family members. In addition, qRT-PCR measures transcript accumulation but does not establish protein abundance, subcellular localization, oligomerization, or chaperone activity.

A further challenge is distinguishing effective acclimation from general stress injury. A large increase in sHSP transcripts may indicate a protective response, but it may also reflect severe cellular damage. Molecular results should therefore be integrated with chlorophyll fluorescence, electrolyte leakage, membrane stability, antioxidant activity, water status, and survival after recovery.

Future studies should focus on:

- Genome-wide identification of HSP20 genes in representative slipper orchids.
- Comparative analysis among *Paphiopedilum*, *Phragmipedium*, and *Cypripedium*.
- Tissue-specific expression during heat, drought, chilling, and oxidative stress.
- Promoter sequencing and heat shock element analysis.
- Proteomic validation of transcriptional responses.
- GFP-based localization studies.
- Recombinant-protein chaperone assays.
- Population-level analysis of sHSP sequence variation.
- Integration of molecular responses with physiological thermotolerance.

Long-read genome sequencing combined with RNA sequencing would be especially useful for resolving duplicated sHSP genes, alternative transcripts, promoter regions, and species-specific paralogues.

Conclusion

The isolation and characterization of sHSP-coding genes from slipper orchids can provide valuable insight into the molecular mechanisms underlying heat tolerance and environmental adaptation. sHSPs are ATP-independent chaperones with a conserved α-crystallin domain but substantial diversity in their amino-terminal regions, localization, expression, and function.

A reliable research strategy should combine heat treatment, high-quality RNA extraction, degenerate PCR or transcriptome-based gene discovery, RACE-based recovery of full-length cDNAs, sequence analysis, phylogenetic reconstruction, localization prediction, and quantitative expression profiling. Functional studies and physiological measurements are necessary to establish whether individual genes directly contribute to thermotolerance.

For *Paphiopedilum* and other slipper orchids, this research has potential applications in conservation physiology, tissue-culture acclimatization, hybrid selection, and molecular germplasm management. Existing plastome studies provide a valuable genomic context, but more research is needed on the nuclear sHSP gene family and its role in the survival of these ecologically and horticulturally important orchids.

References

1. Basha E, O'Neill H, Vierling E. Small heat shock proteins and α-crystallins: dynamic proteins with flexible functions. Trends in Biochemical Sciences. 2012;37:106-117. https://doi.org/10.1016/j.tibs.2011.11.005
2. Liu Y, Zhang H, Ma N, Wang J, Hu J, *et al*. Comparative analyses of chloroplast genomes provide comprehensive insights into the adaptive evolution of Paphiopedilum (Orchidaceae). Horticulturae. 2022;8:391. https://doi.org/10.3390/horticulturae8050391
3. Liu ZJ, *et al*. The chloroplast genome evolution of Venus slipper orchids (Paphiopedilum). Frontiers in Plant Science. 2021;12:682623. https://doi.org/10.3389/fpls.2021.682623
4. Scharf KD, Siddique M, Vierling E. The expanding family of *Arabidopsis thaliana* small heat stress proteins and a new family of proteins containing α-crystallin domains. Cell Stress & Chaperones. 2001;6:225-237. https://doi.org/10.1379/1466-1268(2001)0062.0.CO;2
5. Sun W, Montagu MV, Verbruggen N. Small heat shock proteins and stress tolerance in plants. Biochimica et Biophysica Acta—Gene Structure and Expression. 2002;1577:1-9. https://doi.org/10.1016/S0167-4781(02)00417-7
6. Sun Y, Zou P, Jiang N, Fang Y, Liu, *et al*. Comparative analysis of the complete chloroplast genomes of nine Paphiopedilum species. Frontiers in Genetics. 2022;12:772415. https://doi.org/10.3389/fgene.2021.772415
7. Vierling E. The roles of heat shock proteins in plants. Annual Review of Plant Physiology and Plant Molecular Biology. 1991;42:579-620. https://doi.org/10.1146/annurev.42.060191.003051
8. Waters ER, Vierling E. Plant small heat shock proteins—evolutionary and functional diversity. New Phytologist. 2020;227:24-37. https://doi.org/10.1111/nph.16536
9. Waters ER, Lee GJ, Vierling E. Evolution, structure and function of the small heat shock proteins in plants. Journal of Experimental Botany. 1996;47:325-338. https://doi.org/10.1093/jxb/47.3.325