Identification and Validation of Different HouseKeeping Genes for real time RT-PCR in Plants

By

Sitwat Aman

Department of Biochemistry
Faculty of Biological Sciences
Quaid-i-Azam University
Islamabad
2014
(We commence) with the name of Allah

The most gracious (To begin with)*

The most Merciful (To the end)**

Identification and Validation of Different HouseKeeping Genes for real time RT-PCR in Plants
A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in

Biochemistry/Molecular Biology

By

Sitwat Aman

Department of Biochemistry

Faculty of Biological Sciences

Quaid-i-Azam University

Islamabad

2014
Dedicated
To
My beloved parents,
Husband, Brothers and
Sisters
DECLARATION

I hereby declare that the work presented in the following thesis is my own effort and that this thesis is my own composition. No part of the thesis has been previously presented for any other degree.

Sitwat Aman
ACKNOWLEDGEMENTS

All praise for Almighty Allah (SWT) the most compassionate, the most beneficent and ever merciful, who gives me the power to do, the sight to observe and mind to think and judge. Peace and blessings of Almighty Allah (SWT) be upon His Prophet Hazrat Muhammad (P.B.U.H) who exhorted his followers to seek knowledge from cradle to grave.

I feel a deep sense of gratitude and indebtedness to my dignified supervisor, Dr. Samina Shakeel, Department of Biochemistry, Quaid-i-Azam University, Islamabad, for her kind supervision, useful suggestions, consistent encouragement, friendly behavior and dynamic supervision which enabled me to complete this task successfully.

I feel pride in extending my sense of gratitude and sincere feelings of reverence and regards to Prof. Dr. Bushra Mirza, Chairman Department of Biochemistry, Quaid-i-Azam University, Islamabad, for her co-operation in providing me all possible facilities in this course of study.

I am indebted to all my honorable teachers, especially Prof. Dr. Wasim Ahmed, Dr. M. Rashid Khan, Dr. M. Ansar and Prof. Dr. Asghari Bano whose teachings have brought me to this stage of academic zenith.

Furthermore, I am deeply obliged and thankful to Dr. George Eric Schaller (Dartmouth College, New Hampshire USA) because of his kind and considerate suggestions during my international research program. I am also thankful to Higher Education Commission, Pakistan who supported me by giving Indigenous and IRSIP scholarships.

Heartfelt thanks are also extended to my class fellows and co-workers, sweet seniors and juniors for their inspiring support, co-operation and help during practical work in the Lab. A lot of thanks to my friend Saima Ambreen, Beenish Gull, Madiha Amir, who gave me company during my stay in the University and helped me whenever I needed it.

Heartfelt gratitude for my beloved parents who always stood by me whenever I needed them and whose uninterrupted support in various aspects of my life and studies, kept me going with grace and honor under the shades of their prayers. I am grateful to my husband Adnan Arif who always stood at my back in every aspect and supported me in every favorable and unfavorable condition. Furthermore, I am highly obliged to my sisters Tayyaba Aman, Umber Aman, Huma Zeeshan and Wafa Muneeb whose inspiration and good wishes have always accompanied me. I extend my heartfelt gratitude to my brothers M. Ali Raza Khan, M. Ameem-ul-Ehsan and Zeeshan Arif for their prayers and best wishes.

May Almighty Allah shower His blessings and prosperity on all those who assisted me in any way during this research work.

Sitwat Aman
CONTENTS

1. Abstract 1

2. Chapter 01: Introduction 3
   1.1 Real-time PCR 3
   1.2 House keeping genes as internal controls 5
   1.3 Internal control genes in plants 6
      1.3.1 Actins 7
      1.3.2 18S rRNA 7
      1.3.3 Beta tubulins 9
      1.3.4 Ubiquitin 10
      1.3.5 Glyceraldehyde-3-Phosphate Dehydrogenase 11
      1.3.6 Elongation Factor 1 Alpha 12
      1.3.7 RNA polymerase II genes 12
      1.3.8 Cyclophilins 13
   1.4 Validation of the internal control genes 13
   1.5 Validation methods 14
      1.5.1 geNorm 15
      1.5.2 NormFinder 16
   1.6 Validations of plant internal control genes under abiotic stress 17
   1.7 Objectives of the study 19

3. Chapter 02: Material and Methods 20
   2.1 Plant materials and growth conditions 20
   2.2 Treatments or stress conditions 21
2.3 Primer designing and genomic DNA amplification 23
2.4 Sequencing and analysis of partial sequence of selected internal control genes 24
2.5 RNA isolation, cDNA synthesis and RT-PCR 24
2.6 Real-time PCR 25
2.7 Statistical analysis for stability of reference gene 25

4. Chapter 03: Results

3.1 Internal control gene identification and characterization of different internal control genes of selected plants 27
3.2 Validation of C.album internal control genes by using real-time PCR 32
3.3 Measurement of expression stability of C.album internal control genes 33
3.4 Validations of F.carica internal controls by using real-time PCR 89
3.5 Validations of M.spicata internal controls by using real-time PCR 99

5. Discussion 107
6. Conclusion 117
7. References 118
LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ºC</td>
<td>Degree centigrade</td>
</tr>
<tr>
<td>Cat</td>
<td>Catalog</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complimentary DNA</td>
</tr>
<tr>
<td>CTAB</td>
<td>Cetyl tri-methyl ammonium bromide</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxyribonucleotide tri-phosphate</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>HCl</td>
<td>Hydrochloric acid</td>
</tr>
<tr>
<td>Hrs</td>
<td>Hours</td>
</tr>
<tr>
<td>Min</td>
<td>Minutes</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>Magnesium chloride</td>
</tr>
<tr>
<td>Mg</td>
<td>Miligram</td>
</tr>
<tr>
<td>Ml</td>
<td>Milliliter</td>
</tr>
<tr>
<td>Mm</td>
<td>Milimolar</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger Ribonucleic acid</td>
</tr>
<tr>
<td>MRC</td>
<td>Molecular Research Centre</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Centre for Biotechnology Information</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PVP</td>
<td>Polyvinylpyrrolidone</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Quantitative Real-time PCR</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal RNA</td>
</tr>
<tr>
<td>RPM</td>
<td>Revolution per minute</td>
</tr>
<tr>
<td>RQ</td>
<td>Relative Quantification</td>
</tr>
</tbody>
</table>
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT</td>
<td>Reverse transcription</td>
</tr>
<tr>
<td>Sec</td>
<td>Seconds</td>
</tr>
<tr>
<td>TE</td>
<td>Tris ethylene diamine tetra acetic acid</td>
</tr>
<tr>
<td>T.HCl</td>
<td>Tris hydrochloric acid</td>
</tr>
<tr>
<td>UV</td>
<td>Ultra violet</td>
</tr>
<tr>
<td>µl</td>
<td>Microliter</td>
</tr>
<tr>
<td>Ct</td>
<td>Cycle threshold</td>
</tr>
<tr>
<td>HKGs</td>
<td>Housekeeping genes</td>
</tr>
<tr>
<td>ABC</td>
<td>ATP-binding cassette</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>EFα1</td>
<td>Elongation factor alpha 1</td>
</tr>
<tr>
<td>Ubq</td>
<td>Ubiquitin</td>
</tr>
<tr>
<td>Act-α</td>
<td>Actin alpha</td>
</tr>
<tr>
<td>Act</td>
<td>Actin</td>
</tr>
<tr>
<td>β-act</td>
<td>Beta actin</td>
</tr>
<tr>
<td>18S rRNA</td>
<td>18S ribosomal RNA</td>
</tr>
<tr>
<td>β-tub</td>
<td>Beta-tubulin</td>
</tr>
<tr>
<td>β-tub-1</td>
<td>Beta-tubulin-1</td>
</tr>
<tr>
<td>NF</td>
<td>Normalization factor</td>
</tr>
<tr>
<td>V</td>
<td>Pairwise variation</td>
</tr>
<tr>
<td>PPFD</td>
<td>Photosynthetic photon flux density</td>
</tr>
<tr>
<td>M</td>
<td>Average expression stability</td>
</tr>
<tr>
<td>UDG</td>
<td>Uracil DNA Glycosylase</td>
</tr>
<tr>
<td>IAA</td>
<td>Indole acetic acid</td>
</tr>
</tbody>
</table>
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAA</td>
<td>1-Naphthaleneacetic acid</td>
</tr>
<tr>
<td>2,4 D</td>
<td>2,4-dichlorophenoxy acetic acid</td>
</tr>
</tbody>
</table>
### List of Tables

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Title</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Selected internal control genes used for gene expression studies of <em>C. album</em>, <em>F.carica</em> and <em>M.spicata</em></td>
<td>36</td>
</tr>
<tr>
<td>3.2</td>
<td>Candidate internal control gene ranking of <em>C.album</em> under different types of abiotic stresses according to NormFinder</td>
<td>88</td>
</tr>
<tr>
<td>3.3</td>
<td>Candidate internal control gene ranking of <em>F.carica</em> with different abiotic stresses according to NormFinder</td>
<td>97</td>
</tr>
<tr>
<td>3.4</td>
<td>Candidate internal control gene ranking of <em>M.spicata</em> with different abiotic stresses according to NormFinder</td>
<td>106</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Title</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 3.1</td>
<td>Genomic DNA amplification of nine candidate internal control genes under different abiotic conditions in (a) <em>C. album</em>, (b) <em>F. carica</em>, (c) <em>M. spicata</em></td>
<td>37</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>Sequencing of <em>C.album</em>, <em>F.carica</em> and <em>M.spicata</em> internal control genes</td>
<td>60</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>Sequence alignments of newly identified partial sequence of 18S rRNA gene of <em>C.album</em>, <em>F.carica</em> and <em>M.spicata</em></td>
<td>61</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>Phylogenetic relationships of 18S rRNA gene of <em>C. album</em>, <em>F. carica</em> and <em>M. spicata</em> constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI genebank.</td>
<td>62</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>Sequence alignment of newly identified partial <em>Actin</em> gene of <em>C.album</em>, <em>F.carica</em> and <em>M.spicata</em></td>
<td>64</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>Phylogenetic tree of partial <em>Actin</em> gene of <em>C. album</em>, <em>F. carica</em> and <em>M. spicata</em> constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI genebank</td>
<td>65</td>
</tr>
<tr>
<td>Figure 3.7</td>
<td>Sequence alignments of newly identified <em>Actin-α</em> gene of <em>C.album</em>, <em>F.carica</em> and <em>M.spicata</em></td>
<td>66</td>
</tr>
<tr>
<td>Figure 3.8</td>
<td>Phylogenetic tree of newly identified partial <em>Actin-α</em> gene of <em>C. album</em>, <em>F. carica</em> and <em>M. spicata</em> constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI genebank</td>
<td>67</td>
</tr>
<tr>
<td>Figure 3.9</td>
<td>Sequence alignments of newly identified <em>β-actin</em> gene of <em>F. carica</em> done by BioEdit software</td>
<td>69</td>
</tr>
<tr>
<td>Figure 3.10</td>
<td>Phylogenetic tree of β-actin of <em>F. carica</em> constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI database</td>
<td>70</td>
</tr>
<tr>
<td>Figure 3.11</td>
<td>Sequence alignments of newly identified β-tubulin gene of <em>C. album, F. carica</em> and <em>M. spicata</em> done by the BioEdit software</td>
<td>71</td>
</tr>
<tr>
<td>Figure 3.12</td>
<td>Phylogenetic tree of newly identified partial β-tubulin gene of <em>C. album, F. carica</em> and <em>M. spicata</em> constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI genebank</td>
<td>72</td>
</tr>
<tr>
<td>Figure 3.13</td>
<td>Sequence alignments of newly identified and sequenced partial β-tubulin-1 gene of <em>F. carica</em> and <em>M. spicata</em> done by the BioEdit software</td>
<td>73</td>
</tr>
<tr>
<td>Figure 3.14</td>
<td>Phylogenetic tree of <em>F. carica</em> and <em>M. spicata</em> partial β-tubulin-1 gene constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI database</td>
<td>74</td>
</tr>
<tr>
<td>Figure 3.15</td>
<td>Sequence alignments of newly identified and sequenced, partial GAPDH genes of <em>C. album, F. carica</em> and <em>M. spicata</em> by using BioEdit software</td>
<td>75</td>
</tr>
<tr>
<td>Figure 3.16</td>
<td>Phylogenetic relationships <em>C. album, F. carica</em> and <em>M. spicata</em> partial GAPDH genes constructed by ClustalW multiple alignments with other known gene sequences downloaded from genebank</td>
<td>76</td>
</tr>
<tr>
<td>Figure 3.17</td>
<td>Sequence alignments of newly identified Ubiquitin gene of <em>C. album, F. carica</em> and <em>M. spicata</em> done by BioEdit software</td>
<td>78</td>
</tr>
<tr>
<td>Figure 3.18</td>
<td>Phylogenetic tree of Ubiquitin gene of <em>C. album, F. carica</em> and <em>M. spicata</em> constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI</td>
<td>79</td>
</tr>
<tr>
<td>Figure 3.19</td>
<td>Sequence alignments of newly identified EFα1 gene of <em>C. album, F. carica</em> and <em>M. spicata</em> done by BioEdit software</td>
<td>80</td>
</tr>
<tr>
<td>Figure 3.20</td>
<td>Phylogenetic tree of EFα1 gene of <em>C. album, F. carica</em> and <em>M. spicata</em> constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI genebank</td>
<td>80</td>
</tr>
<tr>
<td>Figure 3.21</td>
<td>Dissociation curves of different putative internal control gene primers used for real time PCR in <em>C. album, F. carica</em> and <em>M. spicata</em></td>
<td>82</td>
</tr>
<tr>
<td>Figure 3.22</td>
<td>Expression levels of seven candidate reference genes of <em>C. album</em></td>
<td>83</td>
</tr>
<tr>
<td>Figure 3.23</td>
<td>Gene expression stabilities of the seven candidate genes of <em>C. album</em> as given by geNorm</td>
<td>86</td>
</tr>
<tr>
<td>Figure 3.24</td>
<td>Determining the optimal number of reference genes of <em>C. album</em></td>
<td>87</td>
</tr>
<tr>
<td>Figure 3.25</td>
<td>Expression levels of nine candidate internal control genes in <em>F. carica</em></td>
<td>93</td>
</tr>
<tr>
<td>Figure 3.26</td>
<td>Gene expression stabilities of the nine candidate genes of <em>F. carica</em> as predicted by geNorm</td>
<td>95</td>
</tr>
<tr>
<td>Figure 3.27</td>
<td>Determination of the optimal number of reference genes of <em>F. carica</em></td>
<td>96</td>
</tr>
<tr>
<td>Figure 3.28</td>
<td>Expression levels of eight candidate internal control genes in <em>M. spicata</em></td>
<td>102</td>
</tr>
<tr>
<td>Figure 3.29</td>
<td>Gene expression stability index of the eight candidate genes of <em>M. spicata</em> as predicted by geNorm</td>
<td>104</td>
</tr>
<tr>
<td>Figure 3.30</td>
<td>Determination of the optimal number of reference genes of <em>M. spicata</em></td>
<td>105</td>
</tr>
</tbody>
</table>
Abstract

Investigations of gene expression patterns of a gene or set of different genes in eukaryotic cells at particular stage or conditions require appropriate set of internal controls for accurate quantitative data analysis. These internal controls or reference genes are essential for ensuring the cell viability. Housekeeping gene/s (HKG) has been commonly used as reference genes for the normalization of gene expression data due to their presumed stable and constitutive expression in living organisms. The selection of appropriate internal control gene is a critical step for the gene expression studies by qRT-PCR due to variability in the stable expression of candidate HKGs in different organisms, different organs/tissues of same organism or most importantly due to different experimental conditions or environmental conditions for the same sample.

In this study, we challenged the consensual thinking that all HKGs are reliable controls for expression studies through detailed investigation of set of potential reference genes suitable for gene expression analysis of *Ficus carica* (dicot), *Chenopodium album* (dicot) and *Mentha spicata* (dicot) after treating with different experimental conditions of abiotic stresses. We identified and sequenced three isoforms of *actin* (*actin, β-actin* and *actin α*), two isoforms of tubulin gene (*β-tubulin* and *β-tubulin-1*), *GAPDH, EFα1, ubiquitin* and *18S rRNA* genes from these three plants and checked their validity as good internal control genes. Gene specific primers of above mentioned genes were designed from the conserved regions of similar genes from other plants to amplify genomic DNA followed by sequencing. After confirmation of right products from each plant using several bioinformatics tools, all of these sequences were submitted to genebank. Initially the similarities of these newly isolated genes with other known homologues or orthologues were determined. Alignments and phylogenetic trees have shown high levels of conservation in the genes among diverse set of selected monocots and dicots, as expected.
In order to validate the utility of these newly isolated genes as internal controls, we treated the above mentioned plants with various abiotic stresses including heat, metal, cold, drought, salt and growth hormones. The respective transcripts were amplified and sequenced for further confirmations. We could hardly observed significant difference in the expression of these genes by semi-quantitative RT-PCR as expected. For quantitative validations, we used all of the above candidate internal controls to study the differences in their expression by relative quantitative real-time to validate the best gene or set of genes for the normalization of target genes in these plants. Different reference genes of *C. album, F. carica and M. spicata* showed somewhat variable expression under one or the other type of abiotic stress treatment suggesting; 1) Different experimental conditions can variably affect the stability of these internal control genes. 2) β-tub, EFα1, Ubq and Act-a were identified as the most stable genes under cold, drought, heat, metal and salt stress in *C. album*. In *F. carica, 18S rRNA, β-tub* and Act were three stable genes under cold, drought, salt and growth hormone stress. Similarly, *GAPDH, β-tub, β-tub-1* and *EFα1* were the stable internal control genes under drought, cold, heat and salt stress in *M. spicata* respectively.

Furthermore, the gene stability index was generated for these selected HKGs which facilitated the selection of more than one internal control gene for further verifications of gene expression studies of these or any other related plant species.
Introduction

Gene expression analysis is very important to gain insight knowledge of the functions of important genes of different living organisms. When the organism that is adaptive to the change responses to the environment, it always brings differences in gene expression levels. For the analysis of gene expression patterns certain tools are required that enables the sensitive, more precise, and reproducible quantification of specific mRNAs (Bustin et al., 2009). Now a days, many high throughput techniques can be used for the accurate quantification of gene expression levels (transcript abundance) of hundreds or thousands of genes (Lee et al., 2005). The commonly used techniques are cDNA- and oligonucleotide based microarrays that can measure transcripts at a genome-wide scale (Lee et al., 2005). However, these techniques are not useful all the times as they have limited sensitivity and accuracy especially for the transcripts that are in low abundance. For this reason, quantitative reverse transcription polymerase chain reaction (qRT-PCR) can be used as it quantifies even very weakly expressed genes accurately (Pfaffl et al., 2002).

1.1 Real-time PCR

There are several choices for gene expression technologies depending upon the specific objectives, but quantitative real-time PCR (qRT-PCR) is important in quantifying gene expression because it is sensitive, specific and have broad dynamic range when compared with other analytical techniques like Northern blotting, quantitative competitive polymerase chain reaction (PCR), and semi quantitative reverse transcription polymerase chain reaction (RT-PCR) (Freeman et al., 1999; Bustin, 2002; Gachon et al., 2004). qRT-PCR has become the most common technique for expression analysis of different genes for molecular diagnostics and microarray data validations (Giulietti et al., 2001; Chuaqui et al., 2002; Czechowski et al., 2004; Jain et al., 2006; Jain and Khurana, 2009). Normalization of qRT-PCR data with one or more internal control genes, is extremely important for the accurate and reliable analysis of expression data.
Chapter 01

Introduction

Identification and validation of different housekeeping genes for real time RT-PCR in plants

(Thellin et al., 1999; Guenin et al., 2009). qRT-PCR is now recognized as the most accurate, sensitive and exact method of quantifying mRNA transcripts (Bustin, 2000; Bustin, 2002). This method allows the detection of amplicons from the start of target amplification using different fluorogenic probes or intercalating dyes such as SYBR Green I and does not involves the conventional endpoint analysis. This method is more reliable and quicker than regular RT-PCR due to no need of post-amplification procedures. PCR efficiency of qRT-PCR can also be measured and permits the selection of more sensitive and appropriate assay with excellent reproducibility and broad dynamic range (Gibson et al., 1996). Normalization of qRT-PCR data is a key factor in gene quantification analysis (Bustin, 2000; Pfaffl et al., 2004). The reliability of qRT-PCR experiment can be improved by including an invariant and stable endogenous control in the assay to normalize the expression of target genes to the expression of endogenous controls. By using such approach we can reduce the variations of expression in a sample and increase qRT-PCR efficiency by minimizing the related errors during quantifications. qRT-PCR is commonly used technique to measure the transcript abundance of target genes. These approaches are high throughput and are easy to get good results by the use of appropriate internal control genes to check the expression of several samples simultaneously. However, the evaluation strategy can be affected by experimental variations and correct quantification of target gene requires careful normalization. The experimental variability of the target gene is usually calibrated with some kind of housekeeping genes or internal control genes. For normalization, the expression of a target gene under given conditions is checked and housekeeping genes are used as common denominator in biological fractions (Thellin et al., 1999).

Two common methods of gene quantifications can be used by using these reference genes; absolute quantification and relative quantification. The principle of absolute quantification requires a suitable standard whose concentration is
known to make dilutions in a serial, resulting in the formation of a standard curve. This standard curve gives a linear relationship between the initial amount of starting template and cycle threshold (Ct) values, that is useful in determining the concentration of unknown samples based on their (Ct) values (Bustin, 2000; Wong and Medrano, 2005). By using this method the number of template copies in a known amount of starting sample can be quantified. In relative quantification, target gene expression is compared with one or more reference genes within the same sample set. The expression of these reference genes should be ideally constant across the samples of survey or study (e.g. developmental series, tissue sets, and experimental regimes). Internal control genes such as β-actin, 18S rRNA, GAPDH and elongation factor-1 are commonly used as reference genes in qRT-PCR experiments (Bustin, 2000; Olsvik et al., 2005).

1.2 Housekeeping genes as internal controls

Housekeeping genes are constitutively expressed genes in the living cells with minimal variations in their expression in different tissues. HKGs are usually GC rich genes and are present in plants and animals. These genes have less number of small introns (Carels and Bernardi, 2000). These features give advantage to these genes to be transcribed constitutively. Housekeeping genes are mostly conserved with slow mutation rate and mutations in these genes are usually not tolerated. The consequences of mutations in housekeeping genes are observed in Arabidopsis thaliana, where the replacement of lysine amino acid at position 40 by alanine or glutamine in tubulin-α produces a dwarf plant (Xiong et al., 2012). Housekeeping genes can be used as good internal controls due to stable expression in a variety of stresses and changing environmental factors.

These genes are used as internal controls because they maintain their constant basal level of expression that is consistent, non-regulated and does not depend upon the cell cycle, and non-responsive to external treatments or developing stages. Therefore, it is necessary to know which housekeeping gene is suitable to
serve as an internal RNA control under particular experimental conditions (DeRisi et al., 1997). The selection of right and ideal housekeeping gene for every experiment is very crucial in order to ensure the credibility of the results (Foss et al., 1998).

1.3 Internal control genes in plants

Housekeeping genes have been extensively studied in animals. She et al., 2009 identified 1522 housekeeping genes by high density microarrays out of 18,149 genes in 42 normal human tissues. In contrast, relatively few sets of plant reference genes have been described to the animals (Gutierrez et al., 2008).

Expression of many plant housekeeping genes is studied under heat stress (Volkov et al., 2003; Aman et al., 2012; Haq et al., 2012), other environmental stresses (Haq et al., 2012) and developmental stages (Czechowski et al., 2005; Remans et al., 2008; Goncalves et al., 2005).

Housekeeping genes like ATP-binding cassette (ABC) transporter, F-box protein family, metallo-protease and CDPK-related protein kinase have been confirmed in soybean (Glycine max) studied under 130 different conditions (Brechenmacher et al., 2008). Ubiquitin, beta tubulin and glyceraldehyde-3-phosphate dehydrogenase showed stable expression under herbicide stress in plants (Petit et al., 2012). Most of the work on rice internal control gene validations is under the hormonal, drought and salt stress (Kim et al., 2003). Similarly, in case of potatoes, the reference genes have been explored under biotic and abiotic stresses (Nicot et al., 2005). Housekeeping genes are used to study the genes with changed expression under variable environmental conditions during different stages of the living organisms (Jin et al., 2004; Szabo et al., 2004).

There are several housekeeping genes which can be used as internal controls for the quantitation of the target genes such as, Actins, 18S rRNA, beta tubulin, Ubiquitin, GAPDH, EF-1a, RPII and cyclophilins etc. (Choi et al., 1999; Finnegan et al., 1993; Foss et al., 1998; Goldsworthy et al., 1993; Lemay et al.,
1996; Marten et al., 1994; Serels et al., 1998; Tang et al., 1996; Wu et al., 1999). But most of the information is available for the model plants or the economically important plants as explained below.

1.3.1 Actins

Multigene families are very helpful to study evolution of eukaryotic genes to find out their functions and regulations. It is highly conserved and most abundant protein in eukaryotes. These genes sometimes arise as a result of duplication of ancestral genes. Actin belongs to a multigene family in all animals, plants and protozoa but single gene exists in yeast. There are three main groups of isoforms of Actin gene are α, β, γ. Actin proteins are of great importance and have crucial role in cell motility and cytoskeleton maintenance (Hunter and Garrels, 1977). Actin is not only involved in cell motility and chemotaxis, it can also be involved in cellular functions like signaling, cellular shape, cell division, cellular volume regulation, movement and phagocytosis (De Loof et al., 1996; Sutherland and Witke, 1999).

Actin was validated as an internal control in the gene expression studies of Anoectochilus roxburghii (Zhang et al., 2012), Cichorium intybus (Maroufi et al., 2010). β-actin was used as a validated internal control for the expression studies by real-time PCR in chimeras detection of transgenic tobacco and apricot (Faize et al., 2010). Actin was ranked the most stable in the study of reference genes characterization by quantitative real-time PCR analysis in various tissues of Salvia miltiorrhiza (Yang et al., 2010). ACTB was used as an internal control gene for the normalization of mRNA levels in qRT-PCR analysis of diverse cultivars and tissues of citrus (Jiawen et al., 2012).

1.3.2 18S rRNA gene

Proteins are functional molecules of the cell. A cell may require thousands of different proteins to perform several functions. Cells have to synthesize proteins in order to cope with different requirements of life. However, most proteins are
short life and need to be replaced or recycled. After synthesis, each protein has to perform specific role/s followed by ubiquitination for degradation and can be recycled depending upon the cell demands in the given set of environmental conditions. Ribosomes are protein factories in all organisms and are usually made up of about equal quantity of proteins and rRNAs. 18S rRNA is one of the constituent parts of the ribosome structure. It is found in 40S smaller subunit of ribosomes. 16S rRNA in smaller subunit of prokaryotic and mitochondrial ribosomes is homologous to 18S rRNA. Clinger segments of this 18S rRNA provide binding site to mRNA in smaller subunit of ribosomes during the translation process (Torczynski et al., 1983; Chan et al., 1984; Connaughton et al., 1984; Nelles et al., 1984; Noller, 1984; Raynal, 1984). 18S rRNA gene is a high copy number gene and can be used in evolutionary biology to study phylogenetic relationships of several plants (Hillis and Dixon, 1991; Banaras et al., 2012). The use of 18S rRNA gene has helped to rearrange many animals and plants taxons in their accurate groups, which were placed incorrectly in past like metazoan ancestry.

18S rRNA gene is highly conserved in plants and except a few regions (Soltis and Soltis, 1997). 18S rRNA can also be used as an internal control gene to normalize mRNA levels in real-time PCR analysis of citrus plant (Jiawen et al., 2012). Ribosomal RNA is not polyadenylated so can be distinguished from mRNA and are widely used for RT-PCR and relative qRT-PCR studies for normalizations. (Jain et al., 2006) assessed the gene expression of ten frequently used housekeeping genes in rice, including 18S rRNA in a diverse set of experimental conditions on twenty five different rice samples. 18S rRNA gene expression of 18S rRNA was suggested to be used for normalization of qRT-PCR results for gene expression studies in a variety of samples. Similarly, 18S rRNA together with other internal control genes was used for validations as internal control for gene expression studies in rice (Hedtke et al., 2002). 18S rRNA is
identified and validated as a stable internal control for gene expression study under heat stress in *Agave americana* (Aman et al., 2012).

### 1.3.3 Beta Tubulins

Tubulin is a globular protein which polymerizes to form microtubules. A tubulin molecule is a heterodimer of two major classes, α tubulin and β tubulin. Isoforms of β-*tubulin* are encoded by genes with highly conserved size and number of exons. Both isoforms α and β *tubulins* from plants animals, fungi and protists show more than 80% similarity in amino acid sequences (Dutcher 2001; Morejohn and Fosket, 1991). When phylogenetic analysis of the β-*tubulin* gene was done, it has also shown the conserved nature of the gene (Aman et al., 2013). β-*tubulin* expression was studies in soybean (*Glycine max*) along with other commonly used reference genes (Libault et al., 2005). In the same species its expression was studied in different developmental stages, temperature ranges and photoperiods (Jian et al., 2008). Its constant expression was also observed under different developmental stages and abiotic stresses in tobacco plants (Schmidt and Delaney, 2010). Twenty different β-*tubulin* genes were identified and their expression showed association with cellulose microfibril deposition in secondary wall formation in Populus plant (Oakley et al., 2007). Its expression has also been studied in different developmental stages of trees (Lin and Lai, 2010), five different tissues with two different developmental stages of *Fraxinus* species (Loren et al., 2012), pollens and developmental processes of *O. sativa* (Yoshikawa et al., 2003), leaf development in *P. Jacquemontiana* (Basa et al., 2009) and *Arabidopsis* infested with *F. oxysporum* (Andersen et al., 2004). β-*tubulin* is also validated as the most stable gene in sunflower leaf senescence (Paula et al., 2011). It is also used as a reference gene for normalization in the molecular characterization of *C. album* Cp-sHSPs and its expression in response to different abiotic stresses (Haq et al., 2013) and for the dual role for the Cp-sHSPs of *C. album* including protection from heat and metal stress (Haq et al., 2012).
1.3.4 Ubiquitin

Many proteins like transcription regulators and cyclins are involved in cell cycle. Proteins never last forever and have to be degraded as soon as proteins perform their functions and are finally replaced by new proteins. It helps to degrade proteins as soon as they are tagged for degradations. Similarly, ubiquitin proteosome system is also required for chloroplast development in *A. thaliana* (Ling et al., 2012). Ubiquitin also controls photoperiodic flowering in plants, by degradation of some proteins and ensuring the stability of the others (Pineiro and Jarillo, 2013). Ubiquitin modifications are also important at various levels of plant defense responses (Zeng et al., 2006). Ubiquitin mediated degradation of ethylene receptor; ETR1 has been shown recently in *A. thaliana* (Chen et al., 2010).

Constant expression of *ubiquitin* gene is necessary for validations under given environmental conditions. *Ubiquitin* was also identified as the best internal control gene along with *beta tubulins* and *GAPDH* under herbicide stress in grasses (Petit et al., 2012), in the vegetative tissue samples of root, leaf and stem taken from young tissues and flowers and fruit at different developmental stages in peach (Tong et al., 2009) and in late maturing photoperiod sensitive cultivar of soybean (Jian et al., 2008; Ruibo et al., 2009). It has been shown as a good internal control gene to study biotic and abiotic stresses *O. sativa* (Jain et al., 2006). Similarly, validity of *Ubiquitin-10* as a reference was analyzed in sea grass (*Zostera marina*) under the heat stress (Ransbotyn and Reusch, 2006), in the qRT-PCR profiling of over 1400 *A. thaliana* transcription factors (Czechowski et al., 2005), in six different genotypes and tissues of root, stem, flower and fruit of citrus (Jiawen et al., 2012) and also in cucumber under abiotic stress and growth regulators (Migocka and Papierniak, 2011). Its stable expression is also reported in bamboo (*Phyllostachys edulis*) in a set of six tissue samples of root, stem, mature stem, leaf, flower, and leaf sheath and at two different developmental stages, i.e., before and after flowering (Fan et al., 2013).
Its expression is constant under different environmental conditions (Kawalleck et al., 1993). Because of its stable expression under different biotic and abiotic stresses in plants, it can be used as internal control gene after careful validations under given conditions.

1.3.5 Glyceraldehyde-3-Phosphate Dehydrogenase

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene encodes a classical glycolytic enzyme of about 37 kDa. It catalyzes Glyceraldehyde-3-phosphate conversion to 1, 3-bisphosphiglycerate in the sixth step of glycolysis. Every cell harvests energy as its own demands. Glycolysis is the first phase of both aerobic and anaerobic respiration. In plants, it is an integral enzyme of Calvin cycle. GAPDH is a marker enzyme of chloroplasts (Cerff and Chambers, 1979). Being a vital enzyme of glucose metabolism, GAPDH is functionally required by every prokaryotic and eukaryotic cell. Apart from being important enzyme of energy metabolism, GAPDH also performs some other non-metabolic roles in the cell. It is found to be involved in activation of transcription, apoptosis initiation and transport of vesicles from endoplasmic reticulum to golgi apparatus.

In gene expression data, GAPDH is very commonly used as an endogenous reference gene for relative quantification by real-time PCR. GAPDH was validated as an internal control gene with constant expression in chickpea (Cicer arietinum). It also showed constant expression in all developmental stages of Brassica juncea (Chandna et al., 2012). GAPDH gene has shown the most stable expression in all samples differentiated on the basis of varieties, tissues and developmental stages of L. chinensis plant (Zhang et al., 2011). It was also studied as reference gene in roots and leaf tissues of C. intybus (Maroufi et al., 2010). Similarly, Qi et al., (2010) found GAPDH and ubiquitin as stable genes in Brassica rapa under drought and biotic stress of downy mildew infection.
1.3.6 Elongation Factor 1 Alpha

Elongation factor alpha 1 (EFα1) is a vital enzyme for protein synthesis in the cell. It binds to aminoacyl tRNA complex of ribosome and is a GTP dependent event (Song et al., 1989). Eukaryotic elongation factor alpha has four subunits. Its most abundant subunit is eukaryotic elongation factor 1A (EFα1A) which is a globular protein and constitutes 3 to 10% of the total proteins in the cytoplasm (Merrick, 1992; Hershey et al., 1996). It is very important protein and is a part of proteins involved in translation process at ribosome levels. This gene has multiple copies on different chromosomes. At least one actively coding gene of transcription factor alpha is required for the cell viability (Cottrelle et al., 1985). In many cell types, it is the most abundant protein of translational machinery.

High expression of EFα1 is studied in endosperm of Z.mays and wheat germplasm (Habben et al., 1995). The activity of EFα1 promoter of A. thaliana was also investigated in transgenic Arabidopsis plants quantitatively and histochemically in transgenic plants in comparisons to control plants (Curie et al., 1993). EFα1 gene was also studied as a reference gene in Brachypodium under various stressed conditions and hormonal regimes (Hong et al., 2010).

1.3.7 RNA polymerase II genes

RNA polymerases play central roles in gene regulation to direct multidimensional cellular processes of living organisms. RNA polymerase II is an enzyme of eukaryotic cells. It is a very complex protein of 12 subunits and mass of 550 kDa. Genes for this enzyme are scattered through the genome. Gene coding for the second largest subunit of this enzyme is highly conserved and showed more than 85% amino acid similarity among similar gene of fungi, plants and animals (James et al., 1991). UV-induced pyrimidine dimmers, block the progression of both DNA and RNA polymerases. RNA polymerase II expression analysis showed stable expression in all the cells under pyrimidine dimers repair stress (Fidantsef and Britt, 2011).
1.3.8 Cyclophilins

*Cyclophilins* are found in archeobacteria, eubacteria and eukaryotes (Galat, 1999) and play an important role in folding of newly synthesized proteins. It is also involved in peptide bond isomerization at proline residues (Brandts *et al.*, 1975). *Cyclophilins* are characterized from tomato (*L. esculentum*), maize (*Z. mays*), *A. thaliana*, rice (*O. sativa*) and many other plant species (Kumari *et al.*, 2009; Romano *et al.*, 2004). *Cyclophilins* are also present in DNA containing organelles, mitochondria and chloroplasts (Romano *et al.*, 2004).

Due to stable expression of *cyclophilin* under biotic and abiotic stresses, it is commonly used as reference gene for qRT-PCR analysis. Stable expression of *cyclophilins* was also observed in potato (*Solanum tuberosum*) infected with *F. solani* (Godoy *et al.*, 2000) and tomato roots infected with virus and viroids (Mascia *et al.*, 2010). *Cyclophilins* are also reported as the best reference gene along with elongation factor 4B in *Eucalyptus* under abiotic stress of acibenzolar-s-methyl (Boava *et al.*, 2010). Validations of *cyclophilin* genes have also been studied as reference gene in bamboo (*Phyllostachys edulis*) (Fan *et al.*, 2013).

1.4 Validation of the internal control genes

It is important to validate the candidate genes selected for the gene expression studies because variations in the expression can introduce or mask the exact expression pattern of the target genes rendering the meaningless or misleading results (Bustin, 2002; Dheda *et al.*, 2004; Huggett *et al.*, 2005; Tricarico *et al.*, 2002). Compared to the animals, relatively few reports are available for the validations of reference genes in model plants with sequenced genome (Gutierrez *et al.*, 2008). Validations of reference genes for qRT-PCR normalization has been done in tomato under abiotic stresses like nitrogen, cold and light (Lovdal and Lillo, 2009) and also for the quantitative gene expression analyses in roots, stems, leaves, flowers and fruits of *B. chinense* (Dong *et al.*, 2013).
2011). *Actin, tubulin* and *ubiquitin* have shown highly variable expression among various plant species under different treatments (Kim *et al.*, 2003; Wan *et al.*, 2009; Czechowski *et al.*, 2005; Nicot *et al.*, 2005; Jian *et al.*, 2008; Martin *et al.*, 2008) but still researchers are widely using these genes as internal control genes for expression studies under abiotic stresses in sunflower (Paula *et al.*, 2011; Hewezi *et al.*, 2006). Recently, there are several evidences of the evaluations and validations of candidate internal control genes in plants such as rice (Jain and Khurana, 2009; Kim *et al.*, 2003; Narsai *et al.*, 2010; Jain *et al.*, 2006), potato (Nicot *et al.*, 2005), perennial ryegrass (Martin *et al.*, 2008), *A. thaliana* (Czechowski *et al.*, 2005), soybean (Jian *et al.*, 2008; Brechenmacher *et al.*, 2008), Darnel ryegrass (Dombrowski and Martin, 2009), *Brachiaria brizantha* (Silveira *et al.*, 2009), poplar (Brunner *et al.*, 2004), grape berries (Reid *et al.*, 2006), tomato (Exposito-Rodriguez *et al.*, 2008; Lovdal and Lillo, 2009), *Orobanche ramose* (broomrape, a parasitic plant) (Gonzalez *et al.*, 2008), wheat (Paolacci *et al.*, 2009), coffee (Barsalobres *et al.*, 2009), maize (Scholdberg *et al.*, 2009), peach (Tong *et al.*, 2009), chicory (Maroufi *et al.*, 2010), longan tree (Lin and Lai, 2010), garden Petunia (Mallona *et al.*, 2010), faba bean (Díaz *et al.*, 2010), tobacco (Schmidt and Delaney, 2010), *Salvia miltiorrhiza* (Yang *et al.*, 2010), cotton (Artico *et al.*, 2010), flax (Huis *et al.*, 2010), cucumber (Wan *et al.*, 2009) and water lily (Luo *et al.*, 2010). None of the internal control genes have ever been validated for wild or medicinal plants. These plants are major part of our flora and have several medicinal and sometimes economical values.

### 1.5 Validation methods

No single gene is adequate for the gene expression studies in organism under given conditions. There are several methods available for the selection of internal control genes for the normalization with two or more stable internal control genes of plants or animals by normalization to the geometric means of the selected internal control genes (Vandesompele *et al.*, 2002) e.g., $\Delta C_t$
approach (Silver et al., 2006), geNorm (Vandesompele et al., 2002), Stability index (Brunner et al., 2004), NormFinder (Andersen et al., 2004) and BestKeeper (Pfaffl et al., 2004). These algorithms were developed in order to evaluate the suitability of reference gene(s) for the normalization for real-time PCR data of given biological sample under given conditions (Vandesompele et al., 2002; Pfaffl et al., 2004; Andersen et al., 2004).

Studies have shown that a number reference genes can be used to identify the best suited internal control gene for the normalization of real-time PCR data under specific conditions, this evaluation has been done in various organisms including animals (Chen et al., 2008; Dong et al., 2011), microbes (Hacquard et al., 2012), Arabidopsis (Czechowski et al., 2005), tomato (Lovdal and Lillo, 2009), Petunia hybrida (Mallona et al., 2010), rice (Jain et al., 2006), soybean (Jian et al., 2008), coffee (Cruz et al., 2009), citrus (Jiawen et al., 2012), B.chinense (Dong et al., 2011), wheat (Long et al., 2010) and S.miltiorrhiza (Yang et al., 2010).

1.5.1 geNorm

Single internal control gene that is highly stable in its expression can be considered as the most favorable for the normalization of gene expression data under given conditions but sometimes, the use of two or more genes as internal control for the accurate and reliable data analysis is preferable. geNorm is a method that determines the number of the best internal control genes required for the normalization of gene expression data in any set of samples (Vandesompele et al., 2002). The principle of this approach is to calculate the normalization factor (NF) of two genes that shows the highest expression stabilities and then it adds other internal control genes stepwise in descending order of their gene expression stability values. As a result, pairwise variations (V) of NFn and NFn+1 are calculated that measures the effect of additional internal control genes on NF. It gives a cut-off value below which there is no need of additional internal control gene. The optimal cutoff V value should be around 0.15
(Vandesompele et al., 2002). geNorm is considered one of the best methods to determine the most stable gene in plant research (Gutierrez et al., 2008; Jain et al., 2006). This software is used for the validations of housekeeping genes as internal control for gene expression studies in citrus (Jiawen et al., 2012), rice (Jain et al., 2006), chicory (Maroufi et al., 2010), Bupleurum chinense (Dong et al., 2011), herbicide stress response in grasses (Petit et al., 2012), the expression studies in the sexual and apomictic grass Brachiaria brizantha (Silveira et al., 2009), various tissues of Anoectochilus roxburghii (Zhang et al., 2011), dinoflagellate Prorocentrum minimum (Guo and Jang, 2012). However it has a tendency to assign close rankings to several co-regulated genes, whose expression ratios show less pairwise variations than those of independently regulated genes (Vandesompele et al., 2002).

1.5.2 NormFinder

NormFinder program is a visual basic application of Microsoft Excel and is used frequently for the determination of the expression stabilities of reference genes that ranks all the reference candidate genes based on intra- and inter-group variations, it generates a stability value for each candidate reference gene (Andersen et al., 2004). The Normfinder software (http://www.mdl.dk/publicationsnormfinder.html), another example of a VBA applet and its principle is based on a variance estimation approach which ranks the genes according to their stability under a given set of any experimental conditions. As a result, the more stable the expression of a gene is, the lower is its expression stability value (Andersen et al., 2004). NormFinder has been reported to be used for the validations of internal control genes for gene expression studies in chicory (Maroufi et al., 2010), in response to herbicide stress in grasses (Petit et al., 2012), for the genome wide identification and evaluation of novel internal control genes for qRT-PCR based transcript normalization in wheat (Long et al., 2010).
1.6 Validations of plant internal control genes under abiotic stress

Biotic and abiotic stresses are two main factors affecting the plant growth and production. Living organisms like fungi, bacteria and viruses can also create stress conditions for plants (Dangl and Jones, 2001) and ultimately can lead to the activation of several defense systems of the plants (Collinge and Boller, 2001). Abiotic stress have negative effects on plant growth and any fluctuations in environmental conditions can lead several surveillance mechanisms (Kvaalen and Johnsen, 2008). The affect of cold shock proteins under abiotic stress was studied extensively in *A. thaliana* during growth and seed germination (Park *et al.*, 2009). The studies on the molecular and physiological analysis of drought stress revealed several initial responses that lead to acclimations of *Arabidopsis* (Harb *et al.*, 2010). Similarly, heat stress leads to chloroplast swelling and plastoglobule formation in *Arabidopsis* leaves (Zhang *et al.*, 2010). The effect of cadmium, copper and flavonoids was also analyzed in *Arabidopsis* seedlings (Keilig and Jutta, 2009). Validations of a set of internal control genes for herbicide stress response in grasses have been reported (Petit *et al.*, 2012). Stable internal control genes were identified and validated for heat induced gene expression studies in *Agave americana* (Aman *et al.*, 2012).

We selected *Chenopodium album*, *Ficus carica* and *Mentha spicata* for the validations of suitable internal control genes of these wild plants treated with different types of abiotic conditions. This will be the first report on the validations of internal control genes in these wild plants for their potential use for the normalizations of real-time PCR data. These plants are useful for validation studies as they are wild and no work has been done so far on the gene expression in these plants. This work will also help in understanding the molecular mechanisms of several key genes of different pathways of these plants and other related wild plants by using these internal controls for the normalizations. We used geNorm and NormFinder approaches for suggesting two or more suitable internal control genes for the selected stress conditions.
including cold, drought, heat, metal, salt and growth hormone stresses. Our results will be useful for gene expression studies in other similar plants and will provide insight into the knowledge of gene expression of target genes to understand several metabolic and developmental processes.
1.7 Objectives of the study

The principal objective of this study is to identify stably expressed internal control genes and validations by real-time PCR under multiple environmental stresses in selected plants; *C. album*, *F. carica* and *M. spicata*. The expression of these internal control genes remained consistent under given experimental conditions in one plant or tissue(s). The knowledge of these internal control genes will help to study the key pathways behind the underlying mechanisms of these wild and medicinally important plants. Following are the stepwise objectives of this study.

- To identify and characterize set of selected internal control genes of three wild plants (**C. album**, **F. carica** and **M. spicata**).
- To study the effect of different environmental stresses (heat, metal, cold, drought, growth hormones and salt stress) on internal control gene stabilities of **C. album**, **F. carica** and **M. spicata**.
- To validate the most suitable gene/genes for gene expression studies in the above three selected plants under different conditions.

We presented all the experimental data and related interpretations of above mentioned objectives in form of different chapters of this thesis. First we identified and characterized a set of different internal control genes including ubiquitin (**Ubq**), glyceraldehyde 3-phosphate dehydrogenase (**GAPDH**), actin alpha (**Act α**), actin (**Act**), elongation factor α-1 (**EFα1**), 18S ribosomal RNA (**18SrRNA**), β-actin (**β-act**), beta tubulin 1 (**β-tub1**) and beta-tubulin (**β-tub**) from three different economically or medicinally important plants, **C. album**, **F. carica** and **M. spicata**. And in the last part, we validated the above mentioned genes of **C. album**, **F. carica** and **M. spicata** treated with various abiotic stresses. Conclusively, we suggested set of good reference genes to study the gene expression patterns in these three wild plants in future to explore key players of different underlying pathways.
Material and Methods

2.1 Plant materials and growth conditions

We selected three different plants (Chenopodium album, Ficus carica and Mentha spicata) for the identification, sequencing, characterization and validations of nine internal control genes including three isoforms of actin (actin, β-actin and actin α), two isoforms of tubulin (β-tubulin and β-tubulin-1), GAPDH, EFα1, ubiquitin, and 18S ribosomal RNA gene under different abiotic stresses (heat, drought, cold, salt, growth hormone and metal stress).

1) Chenopodium album

C. album is a dicot C3 weed, commonly known as Fat hen, spread all over the world, that can tolerate extreme environmental conditions like semi-arid, light-saline and other severe abiotic conditions (Yao et al., 2010). The leaves and young shoots of C. album plants are used as a vegetable. Now a days, the leaves of this plant are frequently being used for biosynthesis of silver and gold nanoparticles (Dwivedi and Gopal, 2008). C. album is a wild, edible plant and becoming famous due to its importance in research. There is no report on the validation of any reference gene used for the normalizations of expression data of this plant. Germination of C. album seeds was done in plastic trays containing a mixture of calcined-clay and potting soil (1:1 V/V) in controlled environment growth chambers at either low (26 °C/20 °C day/night) growth temperature regimes with 350 mmol/m2 per second photosynthetic photon flux density (PPFD).

2) Ficus carica

F. carica commonly referred as “Fig plant”, is a dicot and belongs to family Moraceae. It is indigenous to Middle East and mainly to Mediterranean basin (Sadder and Ateyyeh, 2006). The fig tree is famous for its fruit and is cultivated in temperate zone, it has also been investigated for its products like
proteolytic enzymes, organic compounds, and natural rubber (Kang et al., 2000). It has high agricultural and economical value but the physiological and biochemical traits of the fig tree has not been investigated very briefly (Kim et al., 2003). Its fruit is used as food and laxative for children’s cold symptoms. It can also be used as expectorant, emollient and analgesic. Fresh figs can be used for the treatments of boils and small tumors. Its white milky juice extracted from the stems and leaves, is used for the removal of warts (McGovern, 2002; Rubnov et al., 2001; Perez et al., 1999; Canal et al., 2000). It is a wild plant with high medicinal values, no internal control gene have been isolated and validated from this plant so far. We selected this plant for the validations of internal control genes for future gene expression studies under various abiotic stresses. Wild plant samples were collected from different locations of Islamabad.

3) Mentha spicata

M. spicata (Spearmint) is herbaceous, rhizomatous and perennial plant belongs to family Lamiaceae. The leaves of this plant produce an essential oil used for flavor in candies, gums, ice creams and drinks. It is also used commercial preparations of hygienic products like toothpastes and mouth-washes. It is used in many places as alternative treatments due to its antiemetic, antispasmodic, antiseptic, carminative, diuretic, restorative, stimulant, stomachic and tonic properties. We selected this wild plant due to its medicinal value for the validations of internal control genes. We used growth chambers to germinate the seeds of this plant in plastic trays at (26°C/20°C day/night) growth temperature regimes with 350 mmol/m2 per second photosynthetic photon flux density (PPFD).

2.2 Treatments or stress conditions

We treated the above mentioned plants with different types of abiotic stresses including heat, drought, cold, salt and metal stress for the identification, sequencing, characterization and validations of novel internal controls genes specific to these plants for gene expression studies. Almost 8-10 weeks old
plants were treated independently in three biological replicates with six types of abiotic stresses as described below.

**Heat stress** was given to *C. album* plants by incubating the leaves of *C. album* at different temperatures (30 °C, 35 °C and 37 °C) for 4 hours in sterilized incubation buffer (1% sucrose [W/V], 1 mM KPO4 pH-6 [V/V] and 0.02% Tween-20) (Haq *et al.*, 2013). While the leaves of control plants were incubated at 25 °C for 4 hours in the same sterilization incubation buffer (Haq *et al.*, 2013). Cadmium salt (CdCl2.2.5H2O) was used for **heavy metal stress** and *C. album* plants were incubated in different concentrations of cd solutions (10, 15 and 20 mM) for 3 hours with appropriate controls (Haq *et al.*, 2013). Similarly, **drought stress** was given by placing whole plants on dried tissue papers for 1, 3 and 5 hours, while the control plants were kept on wet tissue papers for 5 hours. **Cold stress** was imposed on the *C. album* plants by incubating the potted plants at 4 °C for 3, 6 and 9 hours. While the control plants were kept at 25 °C. Sodium chloride was used for **salt stress** by incubating *C. album* plants in 200, 400 and 600 mM NaCl solution, while control plants were incubated in distilled water (Haq *et al.*, 2013). Leaf samples from all of above treatments and control plants were collected and stored at -80 °C for RNA isolation and expression studies.

*F. carica* plants were independently treated with four different abiotic stresses including, cold, drought, growth hormone and salt stress. **Drought stress** was by placing 12 weeks old plant on dried tissue paper for 1, 3 and 5 hours, while the control plants were kept on wet tissue paper for the same duration. **Cold treatment** was given by placing the potted plants at 4 °C for 3, 5 and 7 hours. While the control plants were kept at 25 °C. Sodium chloride was used for **salt stress** by incubating the plants in 200, 400 and 600 mM NaCl solution with appropriate controls kept in distilled water only for 5 hours. Similarly, growth **hormone stress** was given by incubating the plants in 0.5µM solutions of three plant hormones (IAA, NAA and 2,4 D). While the control
plants were kept in distilled water for four hours and all the samples were stored at -80 °C for further processing.

Four different abiotic stresses including heat, cold, drought and salt stress were given to *M. spicata*. For **heat stress**, the plants were incubated at different temperatures (30 °C, 35 °C and 37 °C) for 4 hours in sterilized incubation buffer as described above, while the control leaves were incubated at 25 °C for 4 hours (Shakeel et al., 2011). **Drought stress** was given by placing the whole plants on dried tissue papers for 1, 3 and 5 hours, while the control plants were kept on wet tissue papers for the same duration (Haq et al., 2012). For **cold stress**, the potted plants were kept at 4 °C for 3, 6 and 9 hours. While the control plants were kept at 25 °C for 3-9 hrs. Different concentrations of sodium chloride (200, 400 and 600 mM NaCl) were used for **salt stress** by incubation in NaCl solution for five hours. Control plants were incubated in distilled water only for 5 hours (Haq et al., 2013). Samples were collected from all treated and control plants and kept at -80 °C.

### 2.3 Primer design and genomic DNA amplification

Gene-specific primers were designed for *Ubiquitin, Actin, Actin α, β-tubulin, GAPDH, EFα1, β-actin, β-tubulin-I and 18S ribosomal RNA* gene from conserved regions of these known housekeeping isoforms downloaded from NCBI database (http://www.ncbi.nlm.nih.gov/) (Aman et al., 2012; Banaras et al., 2012). The genomic DNA was extracted from *C. album, F. carica* and *M. spicata* to amplify by using gene specific primers of selected internal control genes (Table-1) at PCR conditions: denaturation at 95 °C for 5 min, followed by 94 °C for 45 s, 55 °C-62 °C for 45 s, 72 °C for 1 min, for a total of 35 cycles. Finally, the extension of the amplified products was done at 72 °C for 5 min followed by incubation at 15 °C for 15 min. The amplified products were analyzed on agarose gel electrophoresis for subsequent analysis and sequencing.
2.4 Sequeencing and analysis of partial sequence of selected internal control genes

All the amplified internal control genes of *C. album*, *F. carica* and *M. spicata* plants were sequenced and submitted to gene bank (accession numbers KC898957-KC898977) after the purification of PCR products by using rapid PCR Kit (Marligen, USA) as recommended by manufacture. All the PCR purified samples were incubated in water bath at 55°C for 5 min after addition of 8μl of resuspended silica powder. Samples were vortexed and centrifuged at 13000g for 10 sec followed by three washings with 500μl ice cold sequencing-wash buffer. The respective pellets were individually dissolved in 30μl of preheated TE at 65°C. The samples were incubated at room temperature for 3min followed by a short spin at 13000g for 3 min and purified products were confirmed by running on 1.8% agarose gel. Sequencing PCR was done at 95°C for 1min, followed by 95°C for 30 seconds, 55°C-62°C for 30 seconds, 72°C for 4min, for 30 cycles with final elongation step at 72°C for 10 min. Finally the PCR products were purified and resuspended in 20μl Hi-Dye Formamide (HDF) for sequencing in ABI-310 sequencer by Sanger Dideoxy Chain termination method. BLASTn was used to find out the similarity of all the genes with other known sequences downloaded from gene bank as given in the results section. Similarly multiple sequence alignments were done by using the UPGMA option of ClustalW software followed by construction of phylogenetic trees.

2.5 RNA isolation, cDNA synthesis and RT-PCR

Total RNA was isolated from the leaf samples of all the selected plants with and without respective abiotic stresses in three biological and three experimental replicates using TRI Reagent (MRC, TR#118) according to the manufacturer’s instructions. RNA quantification was done by Nanodrop spectrophotometer (ND/-1000) at 230, 260 and 280 nm. A total of 5μg of RNA from each sample was treated with DNase I (RNase-free) (Fermentas Cat.#
EN0521) followed by first strand cDNA synthesis using Revert Aid first strand cDNA synthesis Kit (Fermentas Cat.# K1621). Respective transcript of each internal control gene was amplified in all the three selected plants using real-time PCR.

### 2.6 Real-time PCR

Relative quantification of all the selected internal control genes was done by using Applied Bio-Systems 7500 Fast Real Time PCR system. Real-time RT-PCR was carried out in a 96-well reaction plate using a reaction mixture prepared by adding 12.5μl of the Maxima SYBR Green qPCR Master Mix (2X), forward primer 0.3 μM, reverse primer 0.3 μM, template DNA (<500ng) and nuclease-free water to 25μl. Thermal cycling was performed using a two-step cycling procedure that was repeated three times for three individual biological replicates. Uracil DNA Glycosylase (UDG) pre-treatment at 50 ºC for 2 min, initial denaturation at 95 ºC for 10 mins followed by denaturation at 95 ºC for 15 s, annealing and extension at 55 ºC-60 ºC for 60 s for a total of 40 cycles. Non-treated samples were also used for each run as controls. Real-time PCR was done in three replicates with the same sample. Data analysis was done by using ABI SDS software according to manufacturer’s instructions. All the transcripts were amplified in triplicates using *Ubiquitin, Actin, Actin α, β-tubulin, GAPDH, EFα1, β-actin, β-tubulin 1 and 18S ribosomal RNA* gene specific primers. Melting temperatures of all the primers used ranged from 55-60 ºC for the real-time PCR as discussed in the results Table-1.

### 2.7 Statistical analysis for stability of reference gene

The stability of candidate reference genes was evaluated independently by using geNorm and NormFinder algorithms. Real-time PCR efficiencies were determined by slope of a linear regression model (Pfaffl *et al.*, 2004) for each of the putative internal control gene of selected plants treated with
different types of abiotic stresses as explained above. Ct values were determined for each sample by using ABI SDS software. To analyze the gene expression stabilities, we used geNorm v3.4 software (Vandesompele et al., 2002). All the tested genes were ranked according to their stability and appropriate reference genes were selected in each case. Similarly, NormFinder was used to identify the optimal normalization of a gene among the set of candidate internal control genes. This software generates an descending ordered ranking list of candidate internal genes after normalization according to their stabilities based on their expression patterns in a given set of HKGs under given experimental conditions (Andersen et al., 2004).
Results

3.1 Internal control gene identification and characterization from selected wild plants

For the selection of the best reference/internal control gene for gene expression studies of *C. album*, *F. carica* and *M. spicata*, the respective plants were treated with different types of abiotic stresses (heat, cold, metal, drought and salt stress). We designed gene specific primers by aligning known orthologs or homologues of nine selected candidate reference genes including *ubiquitin* (*Ubq*), *glyceraldehyde 3-phosphate dehydrogenase* (*GAPDH*), *actin alpha* (*Act-α*), *actin* (*Act*), *beta actin* (*β-act*), *elongation Factor α-1* (*EFα1*), *18S ribosomal RNA* (*18S rRNA*), *beta-tubulin* (*β-tub*) and *beta-tubulin-1* (*β-tub-1*). The gene names, accession numbers, plant names, primer sequences with their melting temperatures, approximate product sizes and average PCR efficiencies are given in Table 3.1. These primer pairs were initially used to amplify the respective *C. album*, *F. carica* and *M. spicata* transcripts (control and treated samples) by RT-PCR and qRT-PCR after initial confirmation for the presence of single band and peak in Tm analysis. All the seven amplified products of transcripts of *C. album* specific *Ubq*, *GAPDH*, *Act-α*, *Act*, *EFα1*, *18S rRNA* and *β-tub* genes of approximate sizes of 450, 300, 220, 350, 100, 250 and 250 bp in size were run on agarose gel as shown in (Figure 3.1a) followed by sequencing and genbank submissions after initial analysis. Nine candidate internal control genes including *Ubq*, *GAPDH*, *Act-α*, *Act*, *EFα1*, *18S rRNA*, *β-tub*, *β-act* and *β-tub-1* gene of approximate size of 450, 300, 220, 350, 100, 250, 250, 250 and 110 bp were identified from *F. carica* (Figure 3.1b). Similarly, eight gene fragments including *Act*, *Ubq*, *18S rRNA*, *β-tub*, *Act-α*, *EFα1*, *GAPDH*, *β-tub-1* of 350, 450, 250, 250, 250, 100, 300, 150 bp sizes respectively were identified from *M. spicata* (Figure 3.1c), the sequences were then analyzed by using bioinformatics tools followed by submission to the
genbank (Figure 3.2). We used different bioinformatics tools to analyze these sequences in *C. album*, *F. carica* and *M. spicata* as explained below.

**a. Confirmation of partial sequences of 18S rRNA gene**

We first analyzed newly sequenced partial 18S rRNA gene of *C. album*, *F. carica* and *M. spicata* to find the similarities with other known sequences from databases by using BLAST. Our results have shown ~98% maximum identity (~E-value of 2e-73) with different plants like *Agave americana*, *Ferrocactus glaucescens*, *Arabidopsis thaliana*, *Coffee arabica*, *Pisum sativum*, *Eruca sativa* and *Nicotiana tabacum*. This confirmed that the sequenced products were of 18S rRNA gene of *C. album*, *F. carica* and *M. spicata*.

To further analyze the sequences similarities of this *C. album* gene with a group of known plants, we downloaded several orthologs of 18S rRNA gene from different plants like *F. carica*, *A. americana*, *F. glaucescens*, *C. decidua*, *C. procera*, *C. tetragonaloba*, *E. sativa*, *M. royleana*, *P. juliflora*, *N. tabacum*, *C. frutescens*, *C. arabica*, *O. sativa*, *M. truncatula*, *A. thaliana*, *B. olaracea* and *H. orientalis*. Multiple alignments have shown highly conserved nature of this 18S rRNA gene of *C. album* and *F. carica*, while the partial sequence of 18S rRNA gene of *M. spicata* has shown some variations in the sequence indicating the existence of some changes during the process of evolution (Figure 3.3).

Similarly, phylogenetic analysis has grouped together *C. album* with *C. procera*, *F. glaucescens* and *P. juliflora*, showing similarities of their partial sequences of 18S rRNA gene. *F. carica* was in a separate group along with *C. decidua*, *C. tetragonaloba*, *C. arabica*, *P. hispidum* and *M. truncatula*. *M. spicata* internal control genes have shown similarity with *S. lycopersicum*, *I. hederacea*, *S. indicum*, *B. violifolia*, *M. basjoo*, *T. angustifolia*, *P. sativum* and grouped together with these plants in one of the major sister clade of phylogenetic tree (Figure 3.4).
b. Confirmation of partial sequences of Actin, Actin-α and β-actin (Isoforms) genes

We selected three isoforms of the actin gene (Act, Act-α and β-act) for validations of internal control genes for expression analysis of C. album, F. carica and M. spicata. All these isoforms have shown ~91% maximum identity (~E value of 2e-45) with other known orthologs like Mimosa pudica, Prunus avium, Carica papaya, Litchi chinensis, Helianthus annus, Brassica napus, Glycine max and Nicotiana tabacum. Two isoforms of actin gene (Act and Act-α) were identified and sequenced in C. album, while we could not identify and sequenced β-act from this plant. All the three isoforms of actin gene (Act, Act-α and β-act) were identified and characterized in F. carica and we were able to sequence only two isoforms (Act and Act-α) from M. spicata. The alignment of all of these sequences with G. max, S. lycopersicum, G. hirsutum, L. chinensis, N. tabacum, H. annus, M. truncatula, P. trichocarpa, P. avium, M. pudica and B. napus showed high degree of conservation of actin gene (Act and Act-α). While Act-α isoform have shown less conservation indicating the evolutionary pressure on this part of gene in these plants (Figure 3.5, 3.7 & 3.9).

Phylogenetic analysis based on actin gene has placed C. album with the F. carica, M. spicata and G. max showing that the actin gene in all these plants is highly conserved and they share a common ancestor. M. spicata was grouped with A. thaliana, G. max, N. tabacum and C. album in the tree (Figure 3.6, 3.8 & 3.10).

c. Confirmation of partial sequences of β-tubulin and β-tubulin-1 (isoforms)

Two isoforms of β-tubulin gene (β-tubulin and β-tubulin-1) were identified from three selected plants except C. album; we could not amplify β-tubulin-1 of C. album by using this set of primers. Newly identified genes have shown ~87% maximum identity (~E-value 0.28) with other known orthologs of β-
tubulin genes. For multiple alignments, we downloaded some known plant sequences of β-tubulin genes from genebank. The alignments have shown similarities with β-tubulin genes of Oryza sativa, E. grandis, Arabidopsis thaliana, Brassica napus, S. lycopersicum, Triticum aestivum, Zea mays. We observed few differences between the sequences of A. cepa and F. carica (Figure 3.11 & 3.13). In the phylogenetic analysis C. album was grouped with O. sativa and F. carica was placed in a separate group with A. cepa and M. spicata next to T. aestivum and Z. mays. This indicates the similarity of this newly sequenced M. spicata gene with these monocot plants (Figure 3.12 & 3.14).

d. Confirmation of partial sequences of GAPDH

GAPDH gene was identified and sequenced from the three selected plants; C. album, F. carica and M. spicata. All of these gene sequences showed ~96% identity (~E-value 1.6) with several known GAPDH sequences including Oryza sativa, Lilium longiflorum, Cucurbita maxima, Zinnia elegans, Tragopogon dubius, Nicotiana tabacum, Aralia elata, Solanum tuberosum, Arabidopsis thaliana, Plantago major, Brassica juncea. We downloaded these sequences and aligned together with the newly isolated GAPDH fragments of C. album, F. carica and M. spicata. Alignments showed the conserved nature of this part of GAPDH gene. The newly identified gene sequence of GAPDH showed similarity with T. dubius, N. tabacum, A. elata, S. tuberosum, A. thaliana, P. major, B. juncea, O. sativa, Z. elegans, L. longiflorum, C. maxima (Figure 3.15). In phylogenetic analysis, C. album, F. carica and M. spicata were grouped together one after other indicating the evolutionary relation between the plants (Figure 3.16).

e. Confirmation of partial sequences of Ubiquitin

We identified and characterized Ubiquitin gene in C. album, F. carica and M. spicata too. Our results of BLAST showed ~89% maximum identity (E-value
of 2e-84) with several orthologs of this gene like *Pisum sativum*, *Helianthus annus*, *Arabidopsis thaliana*, *Brassica napus*, *Lotus japonicus*, *Capsicum annuum*, *Populus trichocarpa*, *Oryza sativa*, *Triticum aestivum* and *Brachypodium distachyon*. The alignment of these sequences all the above mentioned sequences showed conserved nature of this gene as expected. Few differences at various nucleotide positions are evidences of evolutionary pressure (Figure 3.17). Phylogenetic analysis showed the grouping of *C. album* next to *P. sativum* and *F. carica* was placed next to *C. album* showing the relatedness of the *ubiquitin* gene in both plants. *C. album*, *F. carica* and *M. spicata* were placed next to each other in the phylogenetic tree indicating the relationship between the *ubiquitin* partial genes of these plants (Figure 3.18).

**f. Confirmation of partial sequences of Elongation factor α 1 (EFα1)**

The newly identified *EFα1* gene sequences of *C. album*, *F. carica* and *M. spicata* showed ~77% similarity with only *EFα1* gene sequence of *Daucus carota*. We could not find any other homologue or orthologues of this gene in the data base. Our results of alignments and phylogentic analysis have shown diverse nature of this gene with less common regions, though this data is too small for such type of analysis. This also shows that the *EFα1* gene has not gain much importance from the plant scientists and it is still unidentified and uncharacterized in several plants (Figure 3.19 & 3.20).
3.2 Validation of *C. album* internal control genes by using real-time PCR

**C. album**

Abiotic stresses

Salt, cold, heat, metal (Cd) and drought

(Total 25 biological controls and stressed samples in 3 replicates each)

**RNA extraction**

**cDNA synthesis**

**Realtime PCR**

**Statistical analysis by using different bioinformatics tools**

1. Tm curve analysis
2. Expression analysis
3. geNorm, NormFinder
There are several rate limiting steps in the selection of the best reference gene or internal control genes for expression studies including RNA quality/quantity, primer designing and testing. Therefore, we first analyzed these parameters on the selected samples amplified with all the selected candidate internal control genes as described below.

Quality of RNA is one of the important factors which can cause the variability in the results of real time PCR. We accessed the quality of isolated RNA initially by A260/A280 ratios and A260/A230 ratios checked by NanoDrop as the presence of contaminants can decrease these ratios. In our samples, the quality of RNA was good enough to proceed based on these ratios. We used RNase-free-DNase I treatment to remove the traces of genomic DNA from each sample of isolated RNA. The second important factor which can increase the variability in qRT-PCR results is the quantification of the initial amount of RNA used for each set of cDNA synthesis. Therefore, we used NanoDrop to find out the concentrations of RNA and equal amounts of all the replicates of treated verses non-treated samples were used for accuracy. Melting curve analysis of all the primers used for qRT-PCR showed the presence of single peak and there were no primer di-mers observed in any case (Figure 3.21). Threshold cycle (Cq) values for all the seven candidate internal control genes were ranged from 18 to 33 (Figure 3.22). Taken together, all the RNA extractions and the primers used for this study were ideally good for relative quantification by qRT-PCR of *C. album*, *F. carica* and *M. spicata*.

### 3.3 Measurement of expression stability of *C. album* internal control genes

We treated *C. album* plants with five different abiotic stresses (heat, cold, drought, metal and salt) and respective transcripts were amplified and analyzed by using seven different internal control genes (*Ubq*, *GAPDH*, *Act-a*, *Act*, *EFα1*, *18S rRNA* and *β-tub*). qRT-PCR was used to evaluate the expression stabilities of above mentioned internal control genes of *C. album*. The amplification curves were generated to determine the cycle threshold (Ct). The
mean C_t value (the average of three biological replicates) in treated samples for each gene was used to measure the expression stabilities by geNorm and NormFinder algorithms as explained in methods (Figure 3.23).

3.3.1 geNorm analysis

The M and V values of all the reference genes were individually calculated for each sample in three replicates. C. album plants treated at 32 °C, 35 °C and 37 °C for 4 hours, while the control samples were kept at 25 °C for same period of time to check the expression stability of different genes. The M values of Ubq and EFa1 were the least; indicative of a good pair of internal controls for heat stress treatment at above conditions. Similarly for cold stress, C. album plants were kept at 4 °C for 3, 6 and 9 hours with appropriate control samples kept at 25 °C. According to our data, the average expression stability value (M) for EFa1 and Act-a was the lowest, showing strong evidence that EFa1 can be used as good C. album internal control in case of cold stress (Figure 3.23). For metal stress, the C. album plants were treated with different concentrations of cadmium chloride (10, 15 and 20 mM) and EFa1 and β-tub were found to be a pair of most stable genes after the geNorm analysis. We also applied different concentrations of NaCl salt solution (200, 400 and 600 mM) for salt stress and we found the same trend like in case of heat stress (lowest values for the pair Ubq and EFa1). Similarly, drought stress was given to C. album plants for 1, 3 and 5 hours and EFa1 and β-tub were selected as the pair of most stable genes under drought stress. We also observed carefully the V values generated by geNorm v3.4 and there is no need of additional internal controls in any case of our treatment to C. album as the V value was less than 0.15 in all treatments and time points (Figure 3.24). A lower cutoff value of 0.15 has been proposed for pairwise variations (V) recently (Vandesompele et al., 2002). Any V value higher than this cutoff value indicates that additional internal controls under given treatments are required for normalizations. Conclusively, Ubq, β- tub, Act-a and EFa1 were selected as good internal
control genes for *C. album* under all the selected abiotic stresses by geNorm in this study. No additional internal control genes are required for normalization of target gene expression under above selected treatments of *C. album*.

### 3.3.2 NormFinder analysis

For further verifications of suitable internal control genes for the *C. album* plant under different abiotic stresses, we used another algorithm, NormFinder, which selects the optimal normalization gene from a set of given genes and rank according to their suitability. The lowest stability value indicates the most stably expressed gene under given condition in given sample. Based on this algorithm, we found almost the same ranking of internal control genes under each abiotic stress like in geNorm analysis as shown in Table 3.2. *EFα1* ranked highest for heat, drought and metal stress while *Act-α* and *Ubq* were found to be the best internal control for salt and cold stress.

Summarized form of this part of my research provided the valuable information regarding identification, sequencing and validation of seven novel housekeeping genes of *C. album* for their use as potential internal controls for normalization of qRT-PCR data of different abiotic stresses. This is the first report on the validations of the internal control genes for gene expression studies of any wild plant. Different reference genes of *C. album* showed variable expression under one or the other type of abiotic stress suggesting that experimental conditions can variably affect the stability of the internal control gene. *Ubq*, *β-tub*, *Actin-α* and *EFα1* were selected as good internal control genes for *C. album* under all abiotic stresses with *EFα1* expressing as the most stable gene under all plant samples. No additional internal control genes are required for normalization of gene expression studies under the above mentioned treatments of these plants. Taken together, our results also suggest that a single reference gene with stable expression under certain condition may not be suitable to normalize target gene expression under other conditions as a careful evaluation of each internal control gene is required for specific experimental conditions (Tong *et al.*, 2009).
Table 3.1: Selected internal control genes used for gene expression studies of *C. album, F. carica* and *M. spicata*

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>Accession number</th>
<th>Plant name</th>
<th>Forward and Reverse Primer Sequence</th>
<th>Primer Efficiency</th>
<th>Approximate product size (bp)</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ubiquitin</td>
<td>Ubq</td>
<td>AK1386753.1</td>
<td>Glycine max</td>
<td>CCTGCCAGCGAGGTTGAGTC TCCCGGAGCTTTCTCCTCAT</td>
<td>1.99</td>
<td>450</td>
<td>62</td>
</tr>
<tr>
<td>Glyceraldehyde</td>
<td>GAPDH</td>
<td>BM666910.1</td>
<td><em>Lycopersicon esculentum</em></td>
<td>GCTTGACCTGCTGTCACCAAAATG GCC AGTACCA ACT GCT</td>
<td>1.99</td>
<td>300</td>
<td>60</td>
</tr>
<tr>
<td>Actin alpha</td>
<td>Act α</td>
<td>EU938070.1</td>
<td><em>Nicotiana tabacum</em></td>
<td>AAGACGGAGGATCG AAC TGG CCA ATG GC  A</td>
<td>1.99</td>
<td>220</td>
<td>55</td>
</tr>
<tr>
<td>Actin</td>
<td>Act</td>
<td>NM_17993.2</td>
<td><em>Arabidopsis thaliana</em></td>
<td>TCCATAATGAGTGATGTGG TA GA CTTGACTC AGT C</td>
<td>1.99</td>
<td>350</td>
<td>55</td>
</tr>
<tr>
<td>Elongation Factor α1</td>
<td>EFα1</td>
<td>NM_125432.3</td>
<td><em>Arabidopsis thaliana</em></td>
<td>CCGACGCTGAACGCGGTAT TAC TCTGAG TC ACT</td>
<td>1.99</td>
<td>100</td>
<td>56</td>
</tr>
<tr>
<td>18S ribosomal RNA</td>
<td>18S rRNA</td>
<td>FR848602.1</td>
<td>Soybean</td>
<td>TCCAGTAAACGAGGAGAC CACC AAG</td>
<td>1.99</td>
<td>250</td>
<td>58</td>
</tr>
<tr>
<td>Beta tubulin b-tub</td>
<td>β-tub</td>
<td>NM_125665.3</td>
<td><em>Arabidopsis thaliana</em></td>
<td>CTAAGCTTTGCTGTAATCT CATC TGG AGTC</td>
<td>1.99</td>
<td>250</td>
<td>58</td>
</tr>
<tr>
<td>Beta actin</td>
<td>β-act</td>
<td>NM_00107227.12</td>
<td><em>Oryza sativa</em></td>
<td>TGCA TGAAGGCTCA AGT TCA ATG GCTG</td>
<td>1.99</td>
<td>250</td>
<td>59</td>
</tr>
<tr>
<td>Beta tubulin b-tub</td>
<td>β-tub-1</td>
<td>NM_106220.2</td>
<td><em>Arabidopsis thaliana</em></td>
<td>AAGAGCTCTGTTGCTGAGGACC A TGC CAGC A</td>
<td>1.99</td>
<td>110</td>
<td>56</td>
</tr>
</tbody>
</table>
Identification and validation of different housekeeping genes for real-time RT-PCR in plants

Figure 3.1: Genomic DNA amplification of nine candidate internal control genes under different abiotic conditions in (a) *C. album*, (b) *F. carica*, (c) *M. spicata*. Gene specific primers of 18S rRNA, Actin, β-tub, Ubq, GAPDH, Actin-α, β-actin, EFα1 and β-tub-1 genes were used to amplify genomic DNA of *C. album*, *F. carica* and *M. spicata*. All the amplified genes are shown with size markers and lanes with no amplifications are non-template controls (NTC).
Results

Identification and validation of different housekeeping genes for real time RT-PCR in plants

<table>
<thead>
<tr>
<th>LOCUS</th>
<th>Actin</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEFINITION</td>
<td>partial Actin gene.</td>
</tr>
<tr>
<td>ACCESSION</td>
<td></td>
</tr>
<tr>
<td>VERSION</td>
<td></td>
</tr>
<tr>
<td>KEYWORDS</td>
<td></td>
</tr>
<tr>
<td>SOURCE</td>
<td>Chenopodium album</td>
</tr>
<tr>
<td>ORGANISM</td>
<td>Chenopodium album</td>
</tr>
<tr>
<td>Eukaryota;Tracheobionta;Spermatophyta; Magnoliophyta; Magnoliopsida; Dicotyledons; Caryophyllidae; Caryophyllales; Chenopodiaceae; Chenopodium.</td>
<td></td>
</tr>
<tr>
<td>REFERENCE</td>
<td>1 (bases 1 to 355)</td>
</tr>
<tr>
<td>AUTHORS</td>
<td>Aman,S. and Shakci,S.</td>
</tr>
<tr>
<td>TITLE</td>
<td>Characterization of reference genes under various abiotic stresses in Chenopodium album</td>
</tr>
<tr>
<td>JOURNAL</td>
<td>Unpublished</td>
</tr>
<tr>
<td>REFERENCE</td>
<td>2 (bases 1 to 355)</td>
</tr>
<tr>
<td>AUTHORS</td>
<td>Aman,S. and Shakci,S.</td>
</tr>
<tr>
<td>TITLE</td>
<td>Direct Submission</td>
</tr>
<tr>
<td>JOURNAL</td>
<td>Submitted (05-APR-2013) Department of Biochemistry, Quaid-i-Azam University, Islamabad 44000, Pakistan</td>
</tr>
<tr>
<td>COMMENT</td>
<td>##Assembly-Data-START##</td>
</tr>
<tr>
<td>COMMENT</td>
<td>Sequencing Technology :: Sanger dyeode sequencing</td>
</tr>
<tr>
<td>COMMENT</td>
<td>##Assembly-Data-END##</td>
</tr>
<tr>
<td>FEATURES</td>
<td>Location/Qualifiers</td>
</tr>
<tr>
<td>source</td>
<td>1..355</td>
</tr>
<tr>
<td>/organism=&quot;Chenopodium album&quot;</td>
<td></td>
</tr>
<tr>
<td>/mol_type=&quot;genomic DNA&quot;</td>
<td></td>
</tr>
<tr>
<td>CBS</td>
<td>1..355</td>
</tr>
<tr>
<td>/codon_start=1</td>
<td></td>
</tr>
<tr>
<td>/product=&quot;Unknown&quot;</td>
<td></td>
</tr>
</tbody>
</table>
| /translation="- IIVL8GSCLSLSLSLVSSAKRRLLLLPLLPSRRLHHHLHPLHSHNR YCP75LSCCTTIPMHMPHLSRR1TVIQG1TILLYLNIQLCQ5KQYD5SRFNFLEL L"
| BASE COUNT | 88 a 33 c 65 g 111 t |
| ORIGIN | 1 tacattgtgc tccgtgtggg tttaatgatg ttaactgtag tctacttgat cgtgtagt gatcagcag 61 agatttcgg ctctgcctct actacattc agatcaggt tgtgtagcct ctcagagaa 121 gtttaacgtt gttgtactcg ggttttttc ttgcttcct taccacatt caccaggt ac 181 tcctacttc ctctgctcct cttgggact accatgttt atcctgtggt tcttgtagg 241 agaatatca gggatctag ggtgatac ttacatatt tacatattat tttgaagtg 301 tcagttgtc agttgtac ca cttgagccac cttgagccac cctttcttct cttat //
Identification and validation of different housekeeping genes for real time RT-PCR in plants
Identification and validation of different housekeeping genes for real time RT-PCR in plants
### Identification and validation of different housekeeping genes for real time RT-PCR in plants

<table>
<thead>
<tr>
<th>Target Sequence</th>
<th>Actina</th>
</tr>
</thead>
<tbody>
<tr>
<td>GenBank Format</td>
<td></td>
</tr>
<tr>
<td>Mode</td>
<td>Seqin</td>
</tr>
<tr>
<td>Style</td>
<td>Normal</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LOCUS</th>
<th>Actina</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEFINITION</td>
<td>partial Actin-a gene.</td>
</tr>
<tr>
<td>ACCESSION</td>
<td></td>
</tr>
<tr>
<td>VERSION</td>
<td></td>
</tr>
<tr>
<td>KEYWORDS</td>
<td></td>
</tr>
<tr>
<td>ORGANISM</td>
<td>Chenopodium album</td>
</tr>
<tr>
<td>Eukaryota;Tracheobionta;Spermatophyta; Magnoliophyta; Magnoliopsida; Dicotyledons; Caryophyllidae; Caryophyllales; Chenopodiaceae; Chenopodium.</td>
<td></td>
</tr>
<tr>
<td>SOURCE</td>
<td>Chenopodium album</td>
</tr>
</tbody>
</table>

| RELEASE | 1 (bases 1 to 299) |
| AUTHORS | Aman, S. and Shakeel, S. |
| TITLE | Characterization of reference genes under various abiotic stresses in Chenopodium album |
| JOURNAL | Unpublished |
| REFERENCE | 2 (bases 1 to 299) |
| AUTHORS | Aman, S. and Shakeel, S. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (05-AFR-2013) Department of Biochemistry, Quaid-i-Azam University, Islamabad 44000, Pakistan |
| COMMENT | ##Assembly-Data-START## |
| | Sequencing Technology :: Sanger dideoxy sequencing |
| | ##Assembly-Data-END## |
| FEATURES | Location/Qualifiers |
| source | 1..299 |
| /organism="Chenopodium album" |
| /mol_type="genomic DNA" |
| CDS | 1..299 |
| /product="Unknown" |
| /translation="LINTGPELLWEQLQTYEQSRRQLVQIFVGETIMLLQIFKTSYCGFSKULVFIAASIAMGHEVEPQYPIFLSCQFKITFTTINLHLHLIRGERSQIRP" |
| BASE COUNT | 63 a 78 c 47 g 91 t |
| ORIGIN | 1 cctgttacca atactactgg gcggctttta tggsacttcg aagatacac acgaacagtcg 61 cggcgaacc tggctcaaat cctgtttag ggacattta tggggtgct otgcaactt 121 cctggcaaat cctgctctctctct tccaaact tggcttcttt tgcgtctactt 181 ccattqgct atctccac cccctccgct tctctggtc ttcctcaact 241 tacatatta ctatatact acacgtcag ctatctcc gcggagctc aagaatct cc

//
Identification and validation of different housekeeping genes for real time RT-PCR in plants
## Results

Identification and validation of different housekeeping genes for real time RT-PCR in plants

<table>
<thead>
<tr>
<th>LOCUS</th>
<th>GAPDH</th>
<th>288 bp</th>
<th>DNA</th>
<th>linear</th>
<th>05-APR-2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEFINITION</td>
<td>partial GAPDH gene.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORGANISM</td>
<td>Chenopodium album</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REFERENCES</td>
<td>1 (bases 1 to 288)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUTHORS</td>
<td>Aman, S. and Shafeel, S.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TITLE</td>
<td>Characterization of reference genes under various abiotic stresses in Chenopodium album</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JOURNAL</td>
<td>Unpublished</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REFERENCES</td>
<td>2 (bases 1 to 288)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUTHORS</td>
<td>Aman, S. and Shafeel, S.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TITLE</td>
<td>Direct Submission</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JOURNAL</td>
<td>Submitted (05-APR-2013) Department of Biochemistry, Quaid-i-Azam University, Islamabad 44000, Pakistan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BASE COUNT</td>
<td>89 a</td>
<td>70 c</td>
<td>27 g</td>
<td>102 t</td>
<td></td>
</tr>
</tbody>
</table>

//

```plaintext
1     ttctctctat cttcgtgttt gcacagttcc aacatcaaa agcttcctac atctctgttc
61    atactacgtg tgcctcagtt catcaaatat actatttca gatcttcaa cccacatgtt
121   ttcttttatta tatttcttat aggttttctc attcagcga tatccttctc atactgctctc
181   atacttctg tttcctaaa tttttctgctc tttttaatgc tgtgtgtgtg ccccataattt
241   acaacacaaa gatctcaagtt ccgcacgttt aggttttctc ctttccttctctact
```

//
Identification and validation of different housekeeping genes for real time RT-PCR in plants

Results
Identification and validation of different housekeeping genes for real time RT-PCR in plants

Chapter 03

Results

Target Sequence: Actin-alpha

Format: GenBank - Node, Sequin - Style: Normal

LOCUS  Actin-alpha  312 bp  DNA  linear  10-APR-2013
DEFINITION partial Act-alpha gene.
ACCESSION
VERSION
KEYWORDS
SOURCE  Ficus carica
ORGANISM  Ficus carica
          Unclassified.
REFERENCE 1 (bases 1 to 312)
          AUTHORS  Amen, S. and Sheker, I.
          TITLE  Identification and validation of internal control genes under abiotic stresses in Ficus carica
          JOURNAL  Unpublished
REFERENCE 2 (bases 1 to 312)
          AUTHORS  Amen, S. and Sheker, I.
          TITLE  Direct Submission
          JOURNAL  Submitted (10-APR-2013) Department of Biochemistry, Quaid-i-Azam University, Islamabad 44000, Pakistan
COMMENT  #Assembly-Date-START##
          Sequencing Technology :: Sanger dideoxy sequencing
          #Assembly-Date-END##
FEATURES  Location/Qualifiers
          source  1..312
                    /organism="Ficus carica"
                    /mol_type="genomic DNA"
          cds  1..312
                    /protein="Unknown"
                    /translation=""
                    /product="Unknown"
                    /translation=""
                    /product="LQGTYKTVHSISQOHWFYVSVVGIGQMTIPTPKKREKQLQYSLLDGPNWYKRBQTH"
                    VRQYLLKLSMCFNTSFFLLVYVSWYLFNLFCRHSMGFSNVILPP"
BASE COUNT  85 a  75 c  60 g  92 t
ORIGIN  1 aacctacagg gtgctcatta ctccccacatt agccagcatt atggggasag gttcacattc
        61 gtatgctgg tctgactaat taaaaagact atcagacacag aagaggagaa aagggaaagaa
        121 tctacaacg gatctctgg aggccaaasg gatattgaa aagctgttca aaggtatgta
        181 gttctaatc ttacatcaaca tggtcatag gaatggttta ccctcaacttttctttattat
        241 gttctactat gctacttat cacactgttt tggctcatt cataggggtt ttggtggtt
        301 atctcctcc ca
Results

Identification and validation of different housekeeping genes for real time RT-PCR in plants
Results

Identification and validation of different housekeeping genes for real time RT-PCR in plants
Identification and validation of different housekeeping genes for real time RT-PCR in plants
Chapter 03

Results

Identification and validation of different housekeeping genes for real time RT-PCR in plants
Identification and validation of different housekeeping genes for real
time RT-PCR in plants

Chapter 03
Results
Identification and validation of different housekeeping genes for real-time RT-PCR in plants
Identification and validation of different housekeeping genes for real-time RT-PCR in plants
## Results

### Identification and validation of different housekeeping genes for real time RT-PCR in plants

**Target Sequence:** Brubulin-1

**Format:** GenBank

<table>
<thead>
<tr>
<th>Location</th>
<th>Qualifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>source</td>
<td>1..204</td>
</tr>
<tr>
<td></td>
<td>/organism=&quot;Mentha spicata&quot;</td>
</tr>
<tr>
<td></td>
<td>/mol_type=&quot;genomic DNA&quot;</td>
</tr>
<tr>
<td>CDS</td>
<td>1..204</td>
</tr>
<tr>
<td></td>
<td>/codon_start=1</td>
</tr>
<tr>
<td></td>
<td>/product=&quot;Unknown&quot;</td>
</tr>
</tbody>
</table>
|          | /translation="-
|          | STALFLCFPFLVLVSHSLRFFIFKEVRFRLLFKKSYMCAADS8RHGRYLAISAMFRGRK" |

**BASE COUNT:** 51 a 65 c 37 g 51 t

**ORIGIN:**

```
1 cccgcaacag cttacaact ctgctcaott acaccaaco acggttcaact ttttgtaaga
61 tttttatatg tagccacott tcttctcaaa actctttctta tgtgatgagc
121 gctggcaac caagcaacag cgtotactg acggcttca gcatgttcccg gggcaagatg
181 acacaaagc tagttgacg gcag
```
### Results

**Identification and validation of different housekeeping genes for real time RT-PCR in plants**

<table>
<thead>
<tr>
<th>Target Sequence</th>
<th>Actin-alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Format</td>
<td>GenBank</td>
</tr>
<tr>
<td>Mode</td>
<td>Seqin</td>
</tr>
<tr>
<td>Style</td>
<td>Normal</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LOCUS</th>
<th>Actin-alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEFINITION</td>
<td>partial Actin-alpha gene.</td>
</tr>
<tr>
<td>ACCESSION</td>
<td></td>
</tr>
<tr>
<td>VERSION</td>
<td></td>
</tr>
<tr>
<td>KEYWORDS</td>
<td></td>
</tr>
<tr>
<td>SOURCE</td>
<td>Mentha spicata</td>
</tr>
<tr>
<td>ORGANISM</td>
<td>Mentha spicata</td>
</tr>
<tr>
<td>UNCLASSIFIED</td>
<td>Unclassified</td>
</tr>
<tr>
<td>REFERENCE 1</td>
<td>(bases 1 to 267)</td>
</tr>
<tr>
<td>AUTHORS</td>
<td>Aman, S. and Shakeel, S.</td>
</tr>
<tr>
<td>TITLE</td>
<td>Characterization of internal control genes under abiotic stresses in Mentha spicata</td>
</tr>
<tr>
<td>JOURNAL</td>
<td>Unpublished</td>
</tr>
<tr>
<td>REFERENCE 2</td>
<td>(bases 1 to 267)</td>
</tr>
<tr>
<td>AUTHORS</td>
<td>Aman, S. and Shakeel, S.</td>
</tr>
<tr>
<td>TITLE</td>
<td>Direct Submission</td>
</tr>
<tr>
<td>JOURNAL</td>
<td>Submitted (10-APR-2013) Department of Biochemistry, Quaid-i-Azam University, Islamabad 44000, Pakistan</td>
</tr>
<tr>
<td>COMMENT</td>
<td>#Assembly-Data-START#</td>
</tr>
<tr>
<td></td>
<td>Sequencing Technology : Sanger dye dideoxy sequencing</td>
</tr>
<tr>
<td></td>
<td>#Assembly-Data-END#</td>
</tr>
</tbody>
</table>

**FEATURES**

<table>
<thead>
<tr>
<th>source</th>
<th>Location/Qualifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1..267</td>
</tr>
<tr>
<td></td>
<td>/organism=&quot;Mentha spicata&quot;</td>
</tr>
<tr>
<td></td>
<td>/mol_type=&quot;genomic DNA&quot;</td>
</tr>
<tr>
<td>CDS</td>
<td>1..267</td>
</tr>
<tr>
<td></td>
<td>/product=&quot;Unknown&quot;</td>
</tr>
<tr>
<td></td>
<td>/translation=&quot;QNSGSGETYKHI3ISS1FIVRVENL1QFFIQMS3EEGAEAKPELYGYYAPPNRGGGRG LSRYFLTHFICNSREHRSRTIKHDV3IQIL&quot;</td>
</tr>
</tbody>
</table>

**BASE COUNT**

| 75 a | 62 c | 58 g | 72 t |

**ORIGIN**

```
1 caccacctt ttcggtgg gggacacaa aggtagttatat cattcatac taatctcctt attcaaatgt caagcaaga agagacagag
61 gttgcagttg aqaatctta tcattctttt attcaaagt attcaaatgt caagcaaga agagacagag
121 gggatcag aatattaag gggatcag aatattaag gggatcag aatattaag gggatcag aatattaag
181 aggaggttat tccatatac tcctacttctc tccatatac tcctacttctc tccatatac tcctacttctc tccatatac
241 aagcagatgg ttccttatac aatctt
```
> *C. album* Actin partial gene sequence (Genebank Id KC898957)

```
TACATTGTGCTCAGTGGTGTTCACTATGTTCCCTGGTATCGCTGATCGTATCA
GCAAAGAGATTACTGCTCTTTGCCCTACTACCATCAAGATCAAGGTTGTTGCTC
CACCTCAGAGAAGTACAGTGTTGGATACGAGGATCTATCCTTGCATCACCTA
CCACATCTCAACTGCTCTTTCCCTCTTCCCCTTGGTACCACAATTCCTA
ATGCACTTTATGATTCTGAGGAGAAATTATTAGAGTTATTCAAGGTTACTACA
TTACTATATATCCTCAATATTTTGCAGATGTGGATCTCCAAGGGTCAGTATGAC
AAGTCACGTCCAACACCCCTCTCTCCTTCTAT
```

> *C. album* Ubiquitin partial gene sequence (Genebank Id KC898958)

```
TTGCGGAACACTCGAAGGGCGCTCTCCCTGCCGAATACAACATCCAGAAGGAG
TCTACCCTCTCATTGTTGCTCCGCTCCTTGGTGTGGATATGCGATATTCCCTGTTACG
ACTCTTACGGAACGATCATTGAGAGGTTAGAGGCAGCTCCTGCACCACTTTGAC
AATGTAAAGACGAGATTCAGGAGAAGGATATGCCCCACAGCAACAGAGG
TTTCAATTTACGTGCGTGAGCAGTCCTAAGGGATGGCCGTCTTGTGTACACTAAAC
ATCCAGAAGGAAGTGCCACTCTCCTCAATTTGTTGCCTCGGTGGCTAGGGAAGGAAATGCA
AATCCCCGCTAAGACCTGGCAGCAGCAATCGAATCGACACCCAGAGGATTCCCTC
GCCAATCCACATTATTATCTCTACCAAACATTGATATTGTATATACCTCGAG
TATACCACTATGCTCCGCCC
```

> *C. album* β-tubulin partial gene sequence (Genebank Id KC898959)

```
CCCCCGGCCCCCTCCCTCCGCGGGTGGGTTTCGCCCCCTTACCTCTCGGGG
ATCCCAAGCATATGTCCTCCTACTGTCCTCCACAGTAACTCAGCAATATGTTGGGA
TGCCAAGAAACATGATGTCCGCTGCTGATCCCGACATGGAGGATGATCTCAGCT
CTGCCAGTGTCCGTTAGATGAGCAGCAAAAGAGGATTGACGACGAG
```
Chapter 03

Results

>C. album Actin-α partial gene sequence (Genebank Id KC898960)

CCGCTTACCAATACTACTGGGCCTTTTTATGGGAACCTGCAGACATACACACAGAA
CAGTCCGCGACAAACTTGTGTTAATATCCTCTTGGAGGAACTATTATGTTGCCG
CTTCTGCAAATCCCTTTTCCAAACCTCCTTTATTGTATCGGATTTTTCCAAAACCTTTG
GTCCTTTATTGCGTCAATTGCAATGGGACATGGTGCTGTCGATAATACACATC
A

>C. album EFα1 partial gene sequence (Genebank Id KC898961)

CGGTCACAATACCTATATATATTACCCTCCGCGGTTTACGGTACAAATTGT
TCTACGGGGGGGGCTTGGGGGGCCCTTATTGGAGCGAAGTTAGAGCAATTGGT

>C. album GAPDH partial gene sequence (Genebank Id KC898962)

TCTTCCTCATCCTGCTTTGTGTACAAGTCCAACATCAAATTATATCATCAACAATC
CTGTACATACATCAGCTGTGATCCCTGATCATGTTCTGGGGCCGCACGCGCTAACA
CTGATGTATTCAACGAGTCTATAGCTTTGAAACAGACAGGTCCGGGTAATCTTG
AGGGTTTTCATCTCCCTTTACT

>C. album 18S rRNA partial gene sequence (Genebank Id HQ827790.1)

TCCGCTAGACCTTAAAAGGTTCTCCCGCCTTTAGGCCACGGAAGTTTGAGGCA
ATAACAGGTCTGTGTGATGCCCTTAGATGTTCTGGGGCCGCACGCGCTAACA
CTGTATGGATCTCAACGAGTCTATAGCTTTGAAACAGACAGGTCCGGGTAATCTTG
GGAATTTTCATCGTGATT

>F. carica GAPDH partial gene sequence (Genebank Id KC898963)

AAAGTAATCTTCTATTTGACAAACAAATTTTATACATTCTTACCTTCTGTTGTGCAACAA
TTGGCTTTACGGATATGTGTCAGGGCATGTGAACGTGTGCTCCTATACATTGGGCAT
GTGACTTTATATGTGTCACCTCTCCTAAAGATTACCTATTTATCATGCCAACATG
GTGTCGCGTTTCCCAACGGTCGCTCATTCTTTATCAACAGCTGTGCTTCTATA

Identification and validation of different housekeeping genes for real
time RT-PCR in plants 56
Identification and validation of different housekeeping genes for real time RT-PCR in plants

Chapter 03

Results

TCTTCCCTTTGTTGGTCACAGCGAAGTCAGCAAGAATATGGAGTTTTGG
TTACAGCAGGTC

>F. carica Actin-α partial gene sequence (Genebank Id KC898964)
ACCCCTACAGGGGTTCCTTACATCTCCCTCCCATACGGATCATTTGGGAAAGGTT
CATTACGTATCGTCCAGTCAGTGGAAATTCAAAGACTATCCGAACGGAAGT
AAAAGGAAGAAACCTTCAATCAGCAGTCTCTTGGGAGGCCCACCGATATTTG
CGTCAACAAACCGCATGTAGTACATATCCTACACTACATTTGCTTCAATG

>F. carica Actin partial gene sequence (Genebank Id KC898965)
AACGGCCTTTGTATGGAACATCGTGTACATGTGTTGTTGACCATGACATTT
CTTCTCTGCTCTCCACTACCTTCAACAGGTAACGCCATCTCCGCAATCA
AACAGCTGATATCCTTATTTCTGTGTTCTTCTACTATTGCTTTTCATC
ATCATTCCATGGGTTTTCGCCGGTTATCCTCTCCCCCATATATATCAATATT
TGCAGATGATGATCTCTAGTACGGTCTGATGAC

>F. carica β-actin partial gene sequence (Genebank Id KC898966)
ACGCAATACACAAACGGAAGTTCACCCTCAAAGGCTGTGTACCTTCAAACCT
GTCAGAAGAAATCCTTACATCTGACCTCTGGAGGTCTACATCAACCCGACT
GGACAAGGGACAAACTGTGATCTCTTCTACTCACCTATGATCCTTCAATCC
CTTACCCTTTACTCCCTTTCAAAATCTCTTCTATTTCTATTTCTACTTC
GGGGGGGCCTGGGGGGCTCT

>F. carica β-tubulin-1 partial gene sequence (Genebank Id KC898967)
TACCCCTACTGCTACTACACATTGTTACACGAGGAAGATTAGACAAACAAG
AGTGGAAAGGGAAGGGCCACAGTGAGCGCCCGCCGCGGTGCTCAGCGAC
CCGTTCCACG
Chapter 03

Results

> *F. carica* Ubiquitin partial gene sequence (Genebank Id KC898967)

GGTCAAGGATACCAGAACAACCTCCATCCCGCCACCAGTTCCATTGCAGGACCGTCAA
ATGCGGCTGGGCTCCTTTCCGACGAGATGCTGCTTGGGGCTTATTTGACTACGAAA
TGCTTTTACTGTTTTCAATAGGAACCTCTCTACTATCGAGTGGAGAGATGTAACCCGA
TTACATCGCCGCGTGGGAAGGACTGCGATCGTACCGTGCCGGACCAACGCAATTCGA
ACCAAATCTGGTCGAAAGGGCTCAAGGAATCCACCGGACAAGGGAGGGGCATC
CCTCCGACCAGCGAGGGCTTATCTTCCGCTGGCAAGCAGCTCGAGGATGGAAGG
ACTCCTGCGGCGGAGTTACCCGGAGGCTACCGCAGTCGACTCCGCTGCTCCTCTATG
TTATCCTACCCCTCTCACCACACACTACGTTAGTATATATATACATGATACAG
CCTATCTCCGGCC

> *F. carica* β-tubulin partial gene sequence (Genebank Id KC898959)

ACAAGACAAGATTATACGATCACTACTCGAACTCGTACACCTCAGACTTCAACC
CAGACAAGATGTTGGGATGCAAGAAGAAACATCATGTGCGACTGCTCGACCCTCGT
GCATCGGTCCGTGTACCTCAACTGCTACTCCATCATTTGCTACTAGGAGGAAAGATT
AGACAACAAAAGGGAAGTTGGAACGAGG

> *F. carica* EFα1 partial gene sequence (Genebank Id KC898961)

CCTCTTTTCAAAAGGCAATCGGCCCTGTGGTTACTGCTGCGGACTGATACCTGAT
CTACCGGCGGATTTGCGCTATTACGATATTTGCTTTTCTCATCGGTGGTCTACAT
CCTTGGAACATTT

> *F. carica* 18S rRNA partial gene sequence (Genebank Id JX444507)

TCCCGTTGCGCCAGGCTTTCTTAGGAGAACCTATGGCCGGCTTAGGGCAAGGAAG
TTTGAGGCAATACAGGTCTGTGTATGGCCTTCAAGATGTGTGTGGGCGACCGCGG
CTACACTGATGTATTTCAACGAGTCTATAGCCTTGCCGACAGCCCGGAAGTTAC
TTGGAATAATT

> *M. spicata* actin partial gene sequence (Genebank Id KC898969)

AAGGATTTTGTATGGGAAACATCGTTACTGCTGCTGATGACCTATGTTCCCTGGA
ATTGCCAGATACTGAGTCTACCGGACATACCTTTCCTGGCTCTGCTGAGTACCATGG
ATTACGGTGTGCAGCGCCCTGCAAGAAGTGACGAGTACATGTTGAGTACGAGGATCA
ATTCTTTGATCATCTCAGTTACTTTCCAAACAGGTACGATACATACATCTTCTTAT

Identification and validation of different housekeeping genes for real
time RT-PCR in plants
Identification and validation of different housekeeping genes for real time RT-PCR in plants

Chapter 03

Results

CCTCTTTCTCTTCAAAAATTGT CATCACTATTTTCACTTTGCTCTAGTCAAAGGG
GAAAGCTCACTGCAAAATACATGCCTTACATACCATTATGCTTGGTCTTCTTGTG
CATGTATATATGCT TACGCAAATGCGG CATGCTGGATTTTTC

>M. spicata ubiquitin partial gene sequence (Genebank Id KC898970)

CTCTCTTCTCTTAAATTGCATCACACTATTTTTATTGCGCTACTGCTAAGGGTCAAAGGG
GAAAGCTCAGCTGCAAAATACATGCCCTTACATACCATTATGCTTGGTCTTCTTGTG
CATGTATATATGCT TACGCAAATGCGG CATGCTGGATTTTTC

>M. spicata 18S rRNA partial gene sequence (Genebank Id JX444508)

CTCTCTTCTCTTAAATTGCATCACACTATTTTTATTGCGCTACTGCTAAGGGTCAAAGGG
GAAAGCTCAGCTGCAAAATACATGCCCTTACATACCATTATGCTTGGTCTTCTTGTG
CATGTATATATGCT TACGCAAATGCGG CATGCTGGATTTTTC

>M. spicata β-tubulin partial gene sequence (Genebank Id KC898971)

CTCTCTTCTCTTAAATTGCATCACACTATTTTTATTGCGCTACTGCTAAGGGTCAAAGGG
GAAAGCTCAGCTGCAAAATACATGCCCTTACATACCATTATGCTTGGTCTTCTTGTG
CATGTATATATGCT TACGCAAATGCGG CATGCTGGATTTTTC

CTCTCTTCTCTTAAATTGCATCACACTATTTTTATTGCGCTACTGCTAAGGGTCAAAGGG
GAAAGCTCAGCTGCAAAATACATGCCCTTACATACCATTATGCTTGGTCTTCTTGTG
CATGTATATATGCT TACGCAAATGCGG CATGCTGGATTTTTC

CTCTCTTCTCTTAAATTGCATCACACTATTTTTATTGCGCTACTGCTAAGGGTCAAAGGG
GAAAGCTCAGCTGCAAAATACATGCCCTTACATACCATTATGCTTGGTCTTCTTGTG
CATGTATATATGCT TACGCAAATGCGG CATGCTGGATTTTTC

CTCTCTTCTCTTAAATTGCATCACACTATTTTTATTGCGCTACTGCTAAGGGTCAAAGGG
GAAAGCTCAGCTGCAAAATACATGCCCTTACATACCATTATGCTTGGTCTTCTTGTG

Identifica 59
Results

Identification and validation of different housekeeping genes for real time RT-PCR in plants

>>M. spicata Actin-α partial gene sequence (Genebank Id KC898973)

CACCAAAATTCCTCCGGCGGGAGACATACAAGTTTATAATTTCCTTCGATCAGTTTATCGTTCGAGTTGAGAATCTTACTCAATTCTTTATTCAAAATGTCAGCGGAAGAAGGACAGAGGGCAAAAACAGAATTATACGGGGGTATTACGCCCCACCAAATCGGGGGGTCGGGGGTTGAGCCGTTA

>>M. spicata EFα1 partial gene sequence (Genebank Id KC898961)

GCGGTCACAAATACTATATATATATTATATTATTTGTATCTCAGCTATATCCTTCTTTTTGTCTACATCTCCTTTTACAAATCCC

>>M. spicata GAPDH partial gene sequence (Genebank Id KC898974)

ACTCTTAAGGGAGAACCTAACATTTCTTACTTTTAGATCTTTTTGTTTGTAAATTC

>>M. spicata β-tubulin-1 partial gene sequence (Genebank Id KC898972)

CCCAGCAGCAGCGCTTACCTCGCCACCTTACCAACACACTGGTCTCACATC

Figure 3.2: Sequencing of C. album, F. carica and M. spicata internal control genes (GAPDH, Actin-α, Actin, β-actin, β-tubulin-1, Ubiquitin, β-tubulin, EFα1, 18S rRNA) of almost 300, 250, 350, 250, 150, 450, 250, 100, 250bp length respectively.
Figure 3.3: Sequence alignments of newly identified partial sequence of 18S rRNA gene of C. album, F. carica and M. spicata done by BioEdit software.

Identification and validation of different housekeeping genes for real time RT-PCR in plants
Figure 3.4: Phylogenetic relationships of 18S rRNA gene of C. album, F. carica and M. spicata constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI genebank.
### Results

<table>
<thead>
<tr>
<th>Gene Family</th>
<th>1270</th>
<th>1280</th>
<th>1290</th>
<th>1300</th>
<th>1310</th>
<th>1320</th>
<th>1330</th>
</tr>
</thead>
<tbody>
<tr>
<td>F. carica</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>C. album</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>M. spicata</td>
<td>CATC</td>
<td>CATC</td>
<td>CATC</td>
<td>CATC</td>
<td>CATC</td>
<td>CATC</td>
<td>CATC</td>
</tr>
<tr>
<td>G. max</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>S. lycopersicum</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>G. hirsutum</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>L. chinensis</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>N. tabacum</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>H. annuus</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>C. papaya</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>M. truncatula</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>F. trichocarpa</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>P. avium</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>M. pudica</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>B. napus</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene Family</th>
<th>1340</th>
<th>1350</th>
<th>1360</th>
<th>1370</th>
<th>1380</th>
<th>1390</th>
<th>1400</th>
</tr>
</thead>
<tbody>
<tr>
<td>F. carica</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>C. album</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>M. spicata</td>
<td>CATC</td>
<td>CATC</td>
<td>CATC</td>
<td>CATC</td>
<td>CATC</td>
<td>CATC</td>
<td>CATC</td>
</tr>
<tr>
<td>G. max</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>S. lycopersicum</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>G. hirsutum</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>L. chinensis</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>N. tabacum</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>H. annuus</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>C. papaya</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>M. truncatula</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>F. trichocarpa</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>P. avium</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>M. pudica</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>B. napus</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene Family</th>
<th>1410</th>
<th>1420</th>
<th>1430</th>
<th>1440</th>
<th>1450</th>
<th>1460</th>
<th>1470</th>
</tr>
</thead>
<tbody>
<tr>
<td>F. carica</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>C. album</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>M. spicata</td>
<td>CATC</td>
<td>CATC</td>
<td>CATC</td>
<td>CATC</td>
<td>CATC</td>
<td>CATC</td>
<td>CATC</td>
</tr>
<tr>
<td>G. max</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>S. lycopersicum</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>G. hirsutum</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>L. chinensis</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>N. tabacum</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>H. annuus</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>C. papaya</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>M. truncatula</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>F. trichocarpa</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>P. avium</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>M. pudica</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
<tr>
<td>B. napus</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
<td>CACC</td>
</tr>
</tbody>
</table>

Continued.......

Identification and validation of different housekeeping genes for real time RT-PCR in plants
Figure 3.5: Sequence alignments of newly identified and sequenced partial Actin gene of C. album, F. carica and M. spicata done by the BioEdit software.
Figure 3.6: Phylogenetic tree of partial Actin gene of *C. album*, *F. carica* and *M. spicata* constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI genebank.
Figure 3.7: Sequence alignment of newly identified partial Actin-a gene of C. album, F. carica and M. spicata done by BioEdit software.

Identification and validation of different housekeeping genes for real time RT-PCR in plants.
Identifcation and validation of different housekeeping genes for real time RT-PCR in plants

Figure 3.8: Phylogenetic tree of newly identified partial Actin-a gene of C. album, F. carica and M. spicata constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI genebank.
Results

Identification and validation of different housekeeping genes for real time RT-PCR in plants
Figure 3.9: Sequence alignments of newly identified β-actin gene of *F. carica* done by BioEdit software.
Figure 3.10: Phylogenetic tree of β-actin of *F.* *carica* constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI database.
Figure 3.11: Sequence alignments of newly identified $\beta$-tubulin gene of C. album, F. carica and M. spicata done by the BioEdit software.
Figure 3.12: Phylogenetic tree of *C. album*, *F. carica* and *M. spicata* partial β-tubulin gene constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI database.
Figure 3.13: Sequence alignments of newly identified and sequenced partial β-tubulin-1 gene of *F. carica* and *M. spicata* done by the BioEdit software.
Figure 3.14: Phylogenetic tree of *F. carica* and *M. spicata* partial β-tubulin-1 gene constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI database.
Figure 3.15: Sequence alignments of newly identified and sequenced, partial \textit{GAPDH} genes of \textit{C. album}, \textit{F. carica} and \textit{M. spicata} by using BioEdit software.
Identification and validation of different housekeeping genes for real time RT-PCR in plants.

Figure 3.16: Phylogenetic relationships C. album, F. carica and M. spicata partial GAPDH genes constructed by ClustalW multiple alignments with other known gene sequences downloaded from genebank.
Identification and validation of different housekeeping genes for real
time RT-PCR in plants
Figure 3.17: Sequence alignments of newly identified *Ubiquitin* gene of *C. album*, *F. carica* and *M. spicata* done by BioEdit software.
Chapter 03

Results

Identification and validation of different housekeeping genes for real time RT-PCR in plants

Figure 3.18: Phylogenetic tree of Ubiquitin gene of C. album, F. carica and M. spicata constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI.
Chapter 03

Results

Identification and validation of different housekeeping genes for real time RT-PCR in plants

Figure 3.19: Sequence alignments of newly identified \( EF\alpha 1 \) gene of \( C. \) album, \( F. \) carica and \( M. \) spicata done by BioEdit software.

\[
\begin{align*}
F. \text{carica} & \quad \vdots \\
M. \text{spicata} & \quad \vdots \\
C. \text{album} & \quad \vdots \\
D. \text{carota} & \quad \vdots
\end{align*}
\]

Figure 3.20: Phylogenetic tree of \( EF\alpha 1 \) gene of \( C. \) album, \( F. \) carica and \( M. \) spicata constructed by ClustalW multiple alignments with other known gene sequences downloaded from NCBI genebank.

\[
\begin{align*}
M. \text{spicata} & \\
D. \text{carota} & \\
C. \text{album} & \\
F. \text{carica} &
\end{align*}
\]
Identification and validation of different housekeeping genes for real-time RT-PCR in plants
Chapter 03

Results

Identification and validation of different housekeeping genes for real-time RT-PCR in plants

Figure 3.21: Dissociation curves of different putative internal control gene primers used for real time PCR in *C. album*, *F. carica* and *M. spicata*. The gene specific primers of *Ubiquitin*, *Actin*, β-actin and Actin α (isotypes), β-tubulin and β-tubulin-1 (isotypes), GAPDH, EFα1 and 18S ribosomal RNA used to amplify the *C. album*, *F. carica* and *M. spicata* transcript by real-time PCR.
Figure 3.22: Expression levels of seven candidate reference genes of *C. album*. The cycle threshold numbers (Ct values) are given as a mean of the three biological replicates. These values represent the points at which a particular gene shows the maximum expression.
Identification and validation of different housekeeping genes for real-time RT-PCR in plants
Identification and validation of different housekeeping genes for real time RT-PCR in plants
Figure 3.23: Gene expression stabilities of the seven candidate genes of *C. album* as given by geNorm. Mean expression stabilities (M) follows the exclusion of the least stable gene stepwise across all the stressed samples within an experiment, (a) cold stress, (b) drought stress, (c) heat stress, (d) metal stress and (e) salt stress. The least stable genes are on the left of the graph and the most stable pair of the genes are on the right as calculated by the geNorm.
Figure 3.24: Determining the optimal number of reference genes of *C. album*. Pairwise variations were calculated by geNorm to determine the minimum number of reference genes for accurate normalization under different abiotic stress samples (a) salt stress, (b) cold stress, (c) heat stress, (d) metal stress and (e) drought stress. Each bar represents the change in normalization accuracy when stepwise adding more endogenous reference genes according to ranking in figure 3.23.
Table 3.2: Candidate internal control gene ranking of *C. album* under different types of abiotic stresses according to NormFinder

<table>
<thead>
<tr>
<th>Rank</th>
<th>Heat</th>
<th>Drought</th>
<th>Cold</th>
<th>Metal</th>
<th>Salt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em>EFa1</em></td>
<td>0.005</td>
<td><em>EFa1</em></td>
<td>0.213</td>
<td><em>Actin</em></td>
</tr>
<tr>
<td>2</td>
<td><em>Actin</em></td>
<td>0.124</td>
<td><em>Actin</em></td>
<td>0.330</td>
<td><em>EFa1</em></td>
</tr>
<tr>
<td>3</td>
<td><em>GAPDH</em></td>
<td>0.135</td>
<td><em>Ubiquitin</em></td>
<td>0.340</td>
<td><em>Actin</em></td>
</tr>
<tr>
<td>4</td>
<td><em>Ubiquitin</em></td>
<td>0.196</td>
<td><em>Actin</em></td>
<td>0.369</td>
<td><em>Ubiquitin</em></td>
</tr>
<tr>
<td>5</td>
<td><em>Actin</em></td>
<td>0.234</td>
<td><em>Ubiquitin</em></td>
<td>0.419</td>
<td><em>GAPDH</em></td>
</tr>
<tr>
<td>6</td>
<td><em>Ubiquitin</em></td>
<td>0.258</td>
<td><em>18S rRNA</em></td>
<td>0.570</td>
<td><em>Ubiquitin</em></td>
</tr>
<tr>
<td>7</td>
<td><em>18S rRNA</em></td>
<td>0.515</td>
<td><em>GAPDH</em></td>
<td>0.580</td>
<td><em>18S rRNA</em></td>
</tr>
</tbody>
</table>
3.4 Validations of *F. carica* internal controls by using real time PCR

**F. carica**

↓

Abiotic stresses

Salt, cold, drought and growth hormone (NAA, 2,4-D, ABA)

(Total 25 biological controls and stressed samples with 3 replicate each)

↓

RNA extraction

↓

cDNA synthesis

↓

Real-time PCR

↓

Statistical analysis by using different bioinformatics tools

1. Tm curve analysis
2. Expression analysis
3. geNorm, NormFinder
Measurements of gene expression stability of *F. carica* internal control genes

qRT-PCR was used to evaluate the expression stability of nine selected internal control genes like *Ubq*, *GAPDH*, *Act-a*, *Act*, *EFα1*, *18S rRNA*, *β-tub*, *β-act* and *β-tub-1* of *F. carica* treated with four different abiotic stresses (cold, drought, salt and growth hormone). The amplification curves were generated to determine the cycle threshold (Ct). The mean Ct values (average of three biological replicates) in a treatment sample for each gene was used to measure the expression stability by geNorm and NormFinder algorithms as explained in methods (Figure 3.25).

3.4.1 geNorm analysis

The M and V values of the nine internal control genes were individually calculated in three replicates by exposing the plant to the cold stress at 4 °C for 3, 5 and 7 hours, while control samples were kept at 25 °C. The average expression stability value (M) for *18S rRNA* and *β-actin* was the lowest, showing a strong evidence that *β-actin* and *18S rRNA* can be used as a pair of good internal controls in case of cold stress. For hormonal stress, *F. carica* plants were treated with 0.5uM concentration of three different growth hormones (IAA, NAA, 2,4 D) solutions, while the control samples were kept in distilled water for four hours. *18S rRNA* and *β-tub-1* were found to be the most stable pair of genes under growth hormone stress. Three different concentrations of NaCl salt solutions (200, 400 and 600 mM) were used to treat the samples under salt stress. The samples with salt stress showed the lowest value for *β-tub* and *Act-α*, hence ranked as the most stable pair of genes.

Similarly, drought stress was given to the plant for 1, 3 and 5 hours and control plants were kept on the wet tissue paper. *Actin* and *Actin-α* were considered as the most stable pair under drought stress. No additional
internal controls are required in the case of drought stress in *F. carica* because the highest V value (0.01) of all the internal control genes tested was less than 0.15. V value under growth hormone stress is 0.008 with no need of additional internal controls under growth hormone stress. The V value (0.015) of the candidate internal controls shows that no additional controls are necessary in case of cold stress. The V value of all the internal control genes used for salt stress was also less than the cutoff value of 0.15 and could support the data without any additional internal control genes. Conclusively, 18S rRNA, β-tub and Act were selected as good internal control genes for *F. carica* under all stressed conditions. The potential use of 18S rRNA was shown by studying the expression profiles of chloroplast specific small HSP gene of Agave (Aman et al., 2012). No additional internal control genes are required for normalization of gene expression under the given treatments, geNorm analysis results are shown in Figure 3.26 & 3.27.

3.4.2 NormFinder analysis

For further confirmations of the suitable internal control genes for the *F. carica* plant under different types of abiotic stresses, we used another algorithm NormFinder, which selects the optimal normalization gene from a set of given genes and rank according to their suitability, but in this case the lowest stability value indicates the most stably expressed gene. We found almost the same ranking pattern of internal control genes under each abiotic stress generated by geNorm analysis with slight variations as shown in Table 3.3 in case of individual treatments. While, 18S rRNA gene was ranked as the most stable internal control gene for the growth hormones related, Act-a for drought stress related, β-tub-1 for cold stress related and β-tub for salt stress related experiments.

In summary, we identified and validated nine novel housekeeping genes of *F. carica* to use as internal controls for normalization of qRT-PCR data of
different abiotic stresses. So far there is no report on validations of internal controls for gene expression studies of this plant. Different reference genes of *F. carica* have shown variable expression under one or the other type of abiotic stress indicating that experimental conditions can variably affect the stability of the internal control gene. Conclusively, 18S rRNA, β-tub and *Act* were selected as good internal control genes for *F. carica* under all stressed samples. The potential use of 18S rRNA was shown by studying the expression profiles of chloroplast specific small HSP gene of *A. americana* (Aman et al., 2012). β-tub can be used as a stable internal control gene because of its high conservation among various plants (Aman et al., 2013). No additional internal control genes are required for normalization of gene expression under the given treatments.
Figure 3.25: Expression levels of nine candidate internal control genes in *F. carica*. The values in the graph are given as cycle threshold numbers (Ct values) of mean of the three biological replicates. These values represent the points at which a particular gene shows the maximum expression.
Identification and validation of different housekeeping genes for real time RT-PCR in plants

Chapter 03

Results

Average expression stability values of remaining control genes

Salt stress

Average expression stability values of remaining control genes

Growth hormone stress
Chapter 03

Results

Identification and validation of different housekeeping genes for real time RT-PCR in plants

Figure 3.26: Gene expression stabilities of the nine candidate genes of *F. carica* as predicted by geNorm. Mean expression stabilities (M) following stepwise exclusion of the least stable gene across all the stressed samples within an experimental set (a) salt stress, (b) drought stress, (c) growth hormone stress and (d) cold stress. The least stable genes are on the left of the graph and the most stable pair of the genes are on the right as calculated by the geNorm.
Results

Identification and validation of different housekeeping genes for real-time RT-PCR in plants

Figure 3.27: Determination of the optimal number of reference genes of *F. carica*. Pairwise variations were calculated by geNorm to determine the minimum number of reference genes for accurate normalization under different abiotic stress samples (a) salt stress, (b) drought stress, (c) growth hormone stress and (d) cold stress. Each bar represents the change in normalization accuracy when stepwise adding more endogenous reference genes according to ranking in figure 3.26.
Table 3.3: Candidate internal control gene ranking of *F. carica* with different abiotic stresses according to NormFinder

<table>
<thead>
<tr>
<th>Rank</th>
<th>Gene</th>
<th>Stability</th>
<th>Gene</th>
<th>Stability</th>
<th>Gene</th>
<th>Stability</th>
<th>Gene</th>
<th>Stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18S rRNA</td>
<td>0.184</td>
<td>Actin α</td>
<td>0.025</td>
<td>β-tub 1</td>
<td>0.078</td>
<td>β-tubulin</td>
<td>0.062</td>
</tr>
<tr>
<td>2</td>
<td>β-tub 1</td>
<td>0.392</td>
<td>GAPDH</td>
<td>0.651</td>
<td>18S rRNA</td>
<td>0.172</td>
<td>Actin α</td>
<td>0.096</td>
</tr>
<tr>
<td>3</td>
<td>GAPDH</td>
<td>0.501</td>
<td>Actin</td>
<td>0.138</td>
<td>β-actin</td>
<td>0.196</td>
<td>GAPDH</td>
<td>0.119</td>
</tr>
<tr>
<td>4</td>
<td>β-tubulin</td>
<td>0.502</td>
<td>18S rRNA</td>
<td>0.290</td>
<td>β-tubulin</td>
<td>0.330</td>
<td>β-actin</td>
<td>0.132</td>
</tr>
<tr>
<td>5</td>
<td>Actin</td>
<td>0.588</td>
<td>Ubiquitin</td>
<td>0.308</td>
<td>EFa1</td>
<td>0.433</td>
<td>Ubiquitin</td>
<td>0.154</td>
</tr>
<tr>
<td>6</td>
<td>Ubiquitin</td>
<td>0.591</td>
<td>EFa1</td>
<td>0.403</td>
<td>Ubiquitin</td>
<td>0.530</td>
<td>Actin</td>
<td>0.209</td>
</tr>
<tr>
<td>7</td>
<td>β-actin</td>
<td>0.619</td>
<td>β-actin</td>
<td>0.497</td>
<td>Actin</td>
<td>0.546</td>
<td>β-tub 1</td>
<td>0.326</td>
</tr>
<tr>
<td>8</td>
<td>EFa1</td>
<td>0.704</td>
<td>β-tubulin</td>
<td>0.511</td>
<td>Actin α</td>
<td>0.640</td>
<td>18S rRNA</td>
<td>0.476</td>
</tr>
<tr>
<td>9</td>
<td>Actin α</td>
<td>0.819</td>
<td>β-tub 1</td>
<td>0.643</td>
<td>GAPDH</td>
<td>0.670</td>
<td>EFa1</td>
<td>0.962</td>
</tr>
</tbody>
</table>
3.5 Validations of *M. spicata* internal controls by using qRT-PCR

*Abiotic stresses*
- Cold, drought, heat and salt
  
  (Total 25 biological controls and stressed samples with 3 replicates each)

- RNA extraction
- cDNA synthesis
- Realtime PCR

Statistical analysis by using different bioinformatics tools

1. Tm curve analysis
2. Expression analysis
3. geNorm, NormFinder
Measurement of expression stabilities of *M. spicata* internal control genes

*M. spicata* plants were treated with four different abiotic stresses (drought, cold, heat and salt) and eight different internal control genes (*Act*, *Ubq*, *18S rRNA*, *β-tub*, *Act-a*, *EFa1*, *GAPDH*, *β-tub-1*) were used to amplify the transcripts. qRT-PCR was used to evaluate the expression stabilities of the above mentioned eight internal control genes of *M. spicata*. The amplification curves were generated to determine the cycle threshold (C<sub>T</sub>). The mean C<sub>T</sub> values (average of three biological replicates) in treated samples for each gene was used to measure the expression stabilities by using the geNorm and NormFinder algorithms (Figure 3.28).

### 3.5.1 geNorm analysis

M and V values of the eight reference genes were individually calculated by treating the *M. spicata* plants for 4 hours at 32 °C, 35 °C and 37 °C for heat stress treatment, while control samples were kept at 25 °C to check the expression stability of different genes. The M value of *GAPDH* and *β-tub-1* was the lowest indicating these two genes as a pair of good internal controls. Similarly for cold stress, the *M. spicata* plants were kept at 4 °C for 3-9 hours, while control samples were kept at 25 °C. The average expression stability value (M) for *GAPDH* and *EFa1* was the lowest, showing a strong evidence that *GAPDH* can be used as good internal controls in case of cold stress. Furthermore different concentrations of NaCl salt solution (200, 400 and 600 mM) were used to treat the samples under salt stress; the most stable genes with the lowest values were *GAPDH* and *β-tub*. Similarly, drought stress was given by placing a whole plant on dried tissue paper for 1-5 hours, while the control plants were kept on wet tissue paper for the same duration; which showed the same trends of the lowest value for *GAPDH* and *β-tub-1* as shown under heat stress as shown in Figure 3.29. No additional internal controls
are required in the case of heat stress in *M. spicata* because the highest V value (0.012) of all the internal control genes tested was less than 0.15. V value under drought stress is 0.012 with no need of additional internal controls under drought stress. The V value (0.015) of the candidate internal controls shows that no additional controls are necessary in case of cold stress. The V value (0.015) of all the internal control genes used for salt stress, was also less than the cutoff value of 0.15 and could support the data without any additional reference genes as shown in Figure 3.30.

Conclusively, *GAPDH*, *β-tub*, *β-tub1* and *EFα1* were selected as good internal control genes for *M. spicata* under all abiotic stresses with *GAPDH* expressing as the most stable gene under almost all stressed plant samples. *GAPDH* have also shown a stable expression during stress in grasses (Jarosova and Kundu, 2010; Hong et al., 2010). No additional internal control genes are required for normalization of gene expression under the given treatments.

### 3.5.2 NormFinder analysis

For further confirmations of suitable internal control genes for the *M. spicata* plant under different types of abiotic stresses, we used another algorithm NormFinder, which selects the optimal normalization gene from a set of given genes and rank according to their suitability, but in this case the lowest stability value indicates the most stably expressed gene. We found almost the same ranking patterns of internal control genes under each abiotic stress generated by geNorm analysis as shown in Table 3.4 in case of individual treatments. *B-tub-1* ranked highest for heat and drought while *EFα1* was found to be the best internal control for salt and cold stress. *GAPDH* was the second stable internal control gene according to NormFinder.

**Conclusively**, we identified and validated eight novel housekeeping genes of *M. spicata* to use as internal controls for normalization of qRT-PCR.
data of different abiotic stresses. This is the first report on validations of internal controls for gene expression studies of this plant. Different reference genes of *M. spicata* showed somewhat variable expression under one or the other type of abiotic stress suggesting that experimental conditions can variably affect the stability of the internal control gene. *GAPDH*, *β-tub*, *β-tub-1* and *EFα1* were selected as good internal control genes for *M. spicata* under all abiotic stresses with *GAPDH* expressing as the most stable gene under almost all stressed plant samples. *GAPDH* have also shown a stable expression during stress in grasses (Jarosova and Kundu, 2010; Hong *et al*., 2010). No other internal control genes are required for normalization of gene expression under the given treatments.
Figure 3.28: Expression levels of eight candidate internal control genes in *M. spicata*. The values in the graph are given as cycle threshold numbers (Ct values) of mean of the three biological replicates. These values represent the points at which a particular gene shows the maximum expression.
Chapter 03

Results

Identification and validation of different housekeeping genes for real

Average expression stability values of remaining control genes

Cold stress

Average expression stability values of remaining control genes

Drought stress

Identification and validation of different housekeeping genes for real
time RT-PCR in plants 103
Identification and validation of different housekeeping genes for real time RT-PCR in plants

Figure 3.29: Gene expression stability index of the eight candidate genes of *M. spicata* as predicted by geNorm. Mean expression stability (M) following stepwise exclusion of the least stable gene across all the stressed samples within an experimental set (a) drought stress, (b) cold stress, (c) heat stress and (d) salt stress. The least stable genes are on the left of the graph and the most stable pair of the genes are on the right as calculated by the geNorm.
Figure 3.30: Determination of the optimal number of reference genes. Pairwise variations were calculated by geNorm to determine the minimum number of reference genes for accurate normalization under different abiotic stress samples (a) salt stress, (b) heat stress, (c) cold stress and (d) drought stress. Each bar represents the change in normalization accuracy when stepwise adding more endogenous reference genes according to ranking in figure 3.29.
Table 3.4: Candidate internal control gene ranking of *M. spicata* with different abiotic stresses according to NormFinder

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>β-actin</td>
<td>0.057</td>
<td>β-actin</td>
<td>0.057</td>
<td>EFα1</td>
<td>0.055</td>
<td>EFα1</td>
<td>0.113</td>
</tr>
<tr>
<td>2</td>
<td>Ubiq</td>
<td>0.244</td>
<td>Ubiq</td>
<td>0.244</td>
<td>GAPDH</td>
<td>0.148</td>
<td>GAPDH</td>
<td>0.249</td>
</tr>
<tr>
<td>3</td>
<td>β-tubulin</td>
<td>0.466</td>
<td>β-tubulin</td>
<td>0.466</td>
<td>Ubiq</td>
<td>0.173</td>
<td>Ubiq</td>
<td>0.304</td>
</tr>
<tr>
<td>4</td>
<td>EFα1</td>
<td>0.499</td>
<td>EFα1</td>
<td>0.499</td>
<td>β-actin</td>
<td>0.217</td>
<td>β-actin</td>
<td>0.375</td>
</tr>
<tr>
<td>5</td>
<td>Actin</td>
<td>0.503</td>
<td>Actin</td>
<td>0.503</td>
<td>18SrRNA</td>
<td>0.327</td>
<td>β-actin</td>
<td>0.450</td>
</tr>
<tr>
<td>6</td>
<td>Actin α</td>
<td>0.579</td>
<td>Actin α</td>
<td>0.579</td>
<td>β-tubulin</td>
<td>0.437</td>
<td>Actin α</td>
<td>0.626</td>
</tr>
<tr>
<td>7</td>
<td>GAPDH</td>
<td>0.590</td>
<td>GAPDH</td>
<td>0.590</td>
<td>Actin α</td>
<td>0.562</td>
<td>Actin α</td>
<td>0.689</td>
</tr>
<tr>
<td>8</td>
<td>18SrRNA</td>
<td>0.986</td>
<td>18SrRNA</td>
<td>0.986</td>
<td>Actin</td>
<td>0.771</td>
<td>18SrRNA</td>
<td>0.804</td>
</tr>
</tbody>
</table>
Discussion

Abiotic stresses are debilitating to plant growth and metabolism. The regulation and expression patterns of a number of genes can be affected by plant exposure to any single type of stress or combinations of stresses for even a short period of time. Gene expression studies can help us to explore many key components and their roles in different pathways of a plant or to find their cross-species variations. qRT-PCR is a robust and specific method for studying gene expression before and after the given stress conditions. This technique is very sensitive and can detect small amount of changes in the transcript abundance in a quantitative manner, provided the data is normalized with a suitable housekeeping gene. Accurate and reliable gene expression studies with qRT-PCR depend upon the selection of good internal control genes for normalizations (Czechowski et al., 2005; Jain et al., 2006). There are several reports on the comparisons of the stability of different internal control genes that can be used for gene expression studies under given conditions (Exposito-Rodriguez et al., 2008; Cortleven et al., 2009; Artico et al., 2010).

Two or more reference genes can also be used for the accuracy of qRT-PCR data. An ideal reference gene should not be affected by any experimental conditions (Butte et al., 2001). There are several reports which show that the expression of these reference genes or internal controls varies with the experimental conditions (Thellin et al., 1999; Suzuki et al., 2000; Lee et al., 2002; Czechowski et al., 2005). Therefore it is necessary to validate different reference genes carefully before their use in gene expression analysis as an internal control under specific experimental conditions in given organism or sample. This objective can be easily achieved for model plants or important crops due to availability of sequencing data, but there is limited genomic information available for the wild or some medicinal plants such as C.album, F.carica and
M. spicata. We identified, characterized and validated seven different housekeeping genes of *C. album*, nine of *F. carica* and eight in the case of *M. spicata*, to use as internal controls for the normalization of gene expression analysis under different abiotic (heat, cold, metal, drought, growth hormone and salt) stresses. We used gene specific primers of *Ubq*, *GAPDH*, *Act-α*, *Act*, *EFα1*, 18S rRNA, β-tub, β-act and β-tub-1 designed from other plant species to sequence *C. album*, *F. carica* and *M. spicata* specific internal control genes. Initial analysis of the partial sequences was done by BLAST which showed similarities with previously known homologs of different plants including *Ubq* (89%, E-value 2e-84), *GAPDH* (96%, E-value 1.6), *Act-α*, *Act*, *β-act* (91%, E-value 2e-45), *EFα1* (77%, E-value 2e-10), 18S rRNA (98%, E-value of 2e-73), β-tub, β-tub-1 (87%, E-value 0.28). *Ubq* gene sequences of *C. album*, *F. carica* and *M. spicata* showed similarity with *P. sativum*, *H. annus*, *A. thaliana*, *B. napus*, *L. japonicus*, *C. annum*, *P. trichocarpa*, *O. sativa*, *T. aestivum* and *B. distachyon*. While, partial sequences of *GAPDH* gene isolated from *C. album*, *F. carica* and *M. spicata* have shown high percentage of identity with *O. sativa*, *L. longiflorum*, *C. maxima*, *Z. elegans*, *T. dubius*, *N. tabacum*, *A. elata*, *S. tuberosum*, *A. thaliana*, *P. major*, *B. juncea*. Similarly, *Act-α*, *Act* and *β-act* (isoforms of actin gene) isolated from *C. album*, *F. carica* and *M. spicata* have shown similarity with *G. max*, *S. lycopersicum*, *G. hirsutum*, *L. chinensis*, *N. tabacum*, *H. annus*, *M. truncatula*, *P. trichocarpa*, *P. avium*, *M. pudica* and *B. napus*. And isoforms of β-tubulin gene (β-tub and β-tub-1) were similar to *O. sativa*, *E. grandis*, *A. thaliana*, *B. napus*, *S. lycopersicum*, *T. aestivum*, *Z. mays*, *F. carica* and *A. cepa*.

18S rRNA partial genes isolated from *C. album*, *F. carica* and *M. spicata* have shown similarity with *A. americana*, *F. glauescens*, *C. decidua*, *C. procera*, *C. tetragonaloba*, *E. sativa*, *M. royleana*, *P. juliflora*, *N. tabacum*, *C. frutescens*, *C. arabica*, *O. sativa*, *M. truncatula*, *A. thaliana*, *B. olearacea* and *H. orientalis* (Aman *et al.*, 2012). Partial gene sequences of
EFα1 isolated from above mentioned three plants were similar to *D. carota* only, as we could not find any other homologue or orthologue of this gene in the database. Overall, our all the data including multiple alignments and phylogenetic analysis have shown high degree of conservation of these internal control partial gene sequences among wide range of similar sequences of different plants including, *O. sativa, E. grandis, A. thaliana, B. napus, S. lycopersicum, T. aestivum, Z. mays, F. carica* but *A. cepa* have shown some differences in the alignment pattern indicating the changes. In case of EFα1 gene the results of alignments and phylogenetic analysis have shown that the gene is diverse in nature containing less common regions, though this data is too small for such type of analysis. This also indicates that the EFα1 gene has not gain much importance from the plant scientists and it is still unidentified and uncharacterized in several plants.

We used two well known statistical algorithms, geNorm (Vandesompele *et al.*, 2002) and NormFinder (Andersen *et al.*, 2004) for the validations of the above mentioned candidate reference genes isolated from *C. album, F. carica* and *M. spicata*. One or combinations of approaches can be used for the validation of two or more stable internal control genes of plants or animals (Vandesompele *et al.*, 2002) e.g., Stability index (Brunner *et al.*, 2004), geNorm (Vandesompele *et al.*, 2002), ΔCt approach (Silver *et al.*, 2006), NormFinder (Andersen *et al.*, 2004) and BestKeeper (Pfaffl *et al.*, 2004). Most of these methods are Ct based, so the quality, quantity of RNA and initial input of cDNA was critical. Keeping this in mind, we tried to use more than one method to check the quality and quality of extracted RNA as discussed in results (Aman *et al.*, 2012; Shakeel *et al.*, 2011; Haq *et al.*, 2013). Several studies have use similar methods for the extraction of total RNA from *B. chinense* tissues for qRT-PCR of citrus (Dong *et al.*, 2011; Jiawen *et al.*, 2012).

The gene expression stabilities of different candidate genes of *C. album, F. carica* and *M. spicata* were analyzed and compared for the
Identification and validation of different housekeeping genes for real
time RT-PCR in plants

Discussion

selection of best HKG in case of different abiotic stresses. According to
our results of gene expression stabilities of \textit{C.album} specific internal
controls under abiotic stress by geNorm, \textit{EFaI} was ranked the most stable
gene in heat, metal, salt, cold and drought stressed samples, whereas its
pairing as combinations of genes varied for its potential use under these
conditions. \textit{EFaI} gene was also validated as a stable reference control in
\textit{Brachypodium} which was either grown under various stressed conditions
or treated with different growth hormones (Hong \textit{et al.}, 2010). Similarly,
high expression of \textit{EFaI} is studied in endosperm of \textit{Z.mays} and wheat
germplasm (Habben \textit{et al.}, 1995). In our study, \textit{EFaI} and \textit{Act-a} were
considered as a good pair of internal control genes in cold stressed
\textit{C.album} plants. Similarly, \textit{β-tub} was ranked as the most stable gene with
combinations of \textit{EFaI} under drought stress followed by \textit{Act-a}.
Phylogenetic analysis of the \textit{β-tub} gene has also shown the conserved
nature of these genes (Aman \textit{et al.}, 2013). \textit{β-tub} expression was also
studied in soybean (\textit{G.max}) along with other commonly used reference
genes (Libault \textit{et al.}, 2005). In the same species, \textit{β-tub} expression was
studied under different developmental stages, temperature ranges and
photoperiods (Jian \textit{et al.}, 2008). Its constant expression was also observed
under different developmental stages and abiotic stresses in tobacco
(\textit{N.tabacum}) plants (Schmidt and Delaney, 2010). Twenty different \textit{β-tub}
gen genes were identified showing their expression associated with cellulose
microfibril deposition in secondary wall formation in Populus plant
(Oakley \textit{et al.}, 2007).

In case of heat stressed \textit{C.album}, the validation of internal control
genes suggested us the use of two genes (\textit{Ubq} and \textit{EFaI}) as a best pair for
expression analysis. Our data showed, \textit{EFaI} paired with \textit{β-tub} can be used
as the most stable genes under heat stress. \textit{Ubq} in combinations with \textit{EFaI}
can be used as stable genes for drought stress in \textit{C.album} \textit{Ubq} was already
identified as the best internal control gene for herbicide stress in grasses
(Petit et al., 2012), in vegetative tissues (root, leaf and stem), flowers and fruits in different developmental stages in peach (*P. persica*) (Tong et al., 2009) and also in late maturing photoperiod sensitive cultivars of soybean (*G. max*) (Jian et al., 2008; Ruibo et al., 2009). It has also been shown as a good internal control gene to study biotic and abiotic stresses in *O. sativa* (Jain et al., 2006). Validity of *Ubq-10* as a reference gene was carefully analyzed in seagrass (*Z. marina*) subjected to heat stress (Ransbotyn and Reusch, 2006) by qRT-PCR profiling for more than 1400 transcription factors of *A. thaliana* (Czechowski et al., 2005), six different genotypes and four different tissues including root, stem, flower and fruit of citrus (Jiawen et al., 2012) and cucumber (*C. sativus*) plants subjected to abiotic stresses and different growth regulators (Migocka and Papierniak, 2011). Similarly, its stable expression is also been reported in bamboo (*P. edulis*) in a set of six tissues (root, stem, mature stem, leaf, flower, and leaf sheath) and at two different developmental stages (before and after flowering) (Fan et al., 2013). Our data provide evidence that no additional internal control genes are required in the case of heat, cold, metal, drought and salt stresses in *C. album*, because the highest $V$ value of all the internal control genes tested was less than 0.15. Similar studies on twenty short term cultured normal fibroblasts samples, from different individuals generated the cut-off value of 0.15 by geNorm and was considered as the value below the cut off for the inclusions of an additional internal control gene. This $V$ value calculated by geNorm is the systematic variation also known as the pairwise variation, for repeated qRT-PCR experiments on the same gene reflects the enzymatic, inherent machine, and pipet variation (Vandesompele et al., 2002). Collectively, our study of *C. album* gene internal control gene validation by geNorm proved *Ubq, β-tub, Act-α* and *EFα1* as good internal controls under heat, metal, cold, drought and salt stresses. Heat and metal stress was also applied to study differential regulation of novel Cp-sHsp family members in *C. album* by using β-tub.
internal control (Haq et al., 2012) and role of Cp-sHSPs under salt, drought and cold stresses was evaluated (Haq et al., 2013).

To further verify the HKGs suitability for *C. album* under abiotic stress in this study, we used another statistical algorithm, NormFinder. This approach is based on the application of a mathematical model that describes the expression values measured by qRT-PCR, sample subgroups can be analyzed separately by this software, intra- and the intergroup expression variation both can be calculated, and finally the stability value of the candidate gene can be calculated. In order to develop a model-based approach several choices were made regarding the statistical methods used and all the other choices were made that gave the similar results. This helped us to select the best gene by the optimal normalization method from a set of genes followed by a gene ranking. The validity of this approach depends upon the number of samples and candidates that are to be analyzed, *i.e.*, the more the number of samples and candidates, the better are the result estimations. The sample set should contain minimum 8 samples/group, and the candidate number should be at least 3 for this software to work, but usually 5–10 are recommended. Further more, it requires the candidates to be chosen from a set of genes that does not show difference in expression levels between the groups. It means that the average expression level is approximately the same in the different groups. Thus, instead of assuming the individual candidate genes to show no systematic intergroup variation, the average of the candidate genes to show no systematic variation can be assumed (Andersen et al., 2004). In our study, NormFinder has identified the same ranking pattern for selection of top one or top two HKGs under given stress to *C. album*. *EFα1* was ranked at the highest place for heat, drought and metal stress. Similarly, *Act-α* and *Ubq* for cold and salt stressed *C. album* plants. Our results were in accordance with the previous studies in *Arabidopsis*, where *Ubq-10* was recommended as a good internal control gene (Czechowski et al., 2005).
The expression of *Ubg5* and *EFα1* genes was also observed as the most stable among all the studied tissues of rice (*Oryza sativa*) under growth hormone, salt and drought stress (Jain *et al*., 2006). Three or four genes can be selected as the most suitable reference genes for all the sample pools as stated by (Tong *et al*., 2009). *EFα1* was also considered as the most stably expressed in potato plants grown under biotic and abiotic stress conditions (Nicot *et al*., 2005). *Act2* and *EF1β* were used in the combinations for the characterization of reference genes in *A.roxburghii* (Zhang *et al*., 2011).

In this study, the expression of nine internal control genes was also evaluated under four different abiotic stresses (growth hormone, salt, drought and cold) in *F.carica*. When the gene expression stability under salt stress was analyzed *β-tub* and *Act-α* were considered as a stable pair of internal control genes in *F.carica*. While under drought stress *Act* and *Act-α* were the most stable genes in *F.carica*. Stable expression of *β-tub-1* and *18S rRNA* was observed under the effect of different concentrations of growth hormones in case of *F.carica*. For cold stress, *18S rRNA* and *β-act* were considered as stable gene pair in the same plant. No additional internal controls were required in case of all selected treatments in *F.carica*. Conclusively, *β-tub*, *β-tub-1*, *Act-α* and *18S rRNA* were selected as the good internal control genes for *F.carica* under cold, drought, salt and growth hormone stresses respectively.

While, *18S rRNA* was ranked highest in stability index by using Normfinder under hormone stress, *Act-α* under drought stress, *β-tub-1* gene under cold stress and *β-tub* under salt stressed conditions of *F.carica*. *Act2* in combination of other genes has already been evaluated as candidate reference gene for gene expression normalization in *Brassica juncea* vegetative stages (Chandna *et al*., 2012). Similarly, *18S rRNA* has been identified and validated as a stable internal control for heat induced gene expression of *A.americana* (Aman *et al*., 2012). It has also been
identified and comparatively analysis of wide range of medicinal plants (Banaras et al., 2012). β-tub was also identified in various plants and have shown the conserved nature and can be use as a good and stable internal control (Aman et al., 2013). In soybean, β-tub expression was studied in different developmental stages, temperature ranges and photoperiods (Jian et al., 2008). Its constant expression was also observed under different developmental stages and abiotic stresses in tobacco (N.tabacum) plants (Schmidt and Delaney, 2010). Twenty different β-tub genes were identified and their stable expression associated with cellulose microfibril deposition in secondary wall formation was showed in P.trichocarpa (Oakley et al., 2007).

We also provided evidences In this study for the characterization of different internal control genes followed by validations of their expression stabilities pattern for eight different housekeeping genes of M.spicata by using geNorm. M.spicata plants were treated with four types of abiotic stresses including drought, cold, heat and salt stress. β-tub-1 and GAPDH were evaluated as stable internal control genes under drought and heat stress in M.spicata. GAPDH and EFα1 were ranked as stable gene pair under cold stress and the gene pair (GAPDH and β-tub) was referred as stable internal control genes under salt stress in M.spicata. We showed experimentally that no additional internal control genes are required for M.spicata in case of above mentioned stresses. This all data was counterchecked by NormFinder analysis, which showed that β-tub-1 is the most stable gene in case of heat and drought stress in this plant. The constant and stable expression of β-tub was also observed under different developmental stages and abiotic stresses in tobacco (N.tabacum) plants (Schmidt and Delaney, 2010). We proved that EFα1 can be selected as the best internal control gene for the analysis of cold and salt stressed M.spicata. EFα1 gene has already been extensively studied as a reference control in Brachypodium either grown under various stressed conditions or
treated with different growth hormones (Hong et al., 2010). Conclusively, our data showed that β-tub, β-tub1, EFα1 and GAPDH are good internal control genes for *M. spicata* under drought, cold, heat and salt stresses respectively.

Five different families are included in the tubulin superfamily and only present in kinetoplastid protozoa are the alpha-, beta-, gamma-, delta-, and epsilon-tubulins and a sixth family (zeta-tubulin) (Kumar, Singh et al. 2012). The paralogs of the alpha/beta-tubulin superfamily genes have been applied in *V. vinifera* (Reid et al., 2006), *A. thaliana* (Gutierrez et al., 2008), *P. persica* (Tong et al., 2009) and *G. max* (Reid et al., 2006). In addition GAPDH have shown the most stable expression for the identification of drought-response genes during sucrose accumulation and water deficit in different plant tissues of sugarcane (Iskandar et al., 2004). EFα1 and GAPDH also exhibited consistent expression in all the tissue samples of chickpea and it was considered ideal to include Ubq10 and EFα1 as reference genes to normalize gene expression data in these plant tissues (Garg et al., 2011). The paralogous genes of *Act* have also been reported in some studies for target gene normalization in plants (Jian et al., 2008; Gutierrez et al., 2008; Caldana et al., 2007; Ruibo et al., 2009; Wan et al., 2009) and specially in *O. sativa* using eight paralogs of *Act* (Zhang et al., 2009). The assumption was the balanced expression of different amplified genes of the same family, but this is not true in all the cases. For example, Ubq5 was found to be one of the most favorable reference genes in a set of different tissue samples in rice, while the expression of Ubq10 parologue was unstable (Jain et al., 2006). Similarly, paralogues of *Act* gene family was used to study the gene expression in samples collected at different development stages of soybean, where Act2/7 was the most stably expressed gene while the expression of Act11 was variable (Jian et al., 2008). To find out the expression stability in the panicles of *O. sativa* under drought stress, eight paralogous genes of *Act* were used and all of them
have shown variable expression stability (Zhang et al., 2009). Our results also support the use of two or three family members of same HKGs because their stability and expression might vary in a given experimental conditions, for example the expression patterns of Act, Act-a and β-act (isoforms of actin gene) and β-tub, β-tub-1 (isoforms of β-tubulin gene) were quite different in the C.album, F.carica and M.spicata plants used for this study. These results indicate that the paralogous sequences as control genes in qRT-PCR can show variable expression, even within the same organism.
Conclusions

We identified and validated seven novel housekeeping genes of *C. album*, nine novel housekeeping genes of *F. carica* and eight genes of *M. spicata* to use as internal controls for normalization of qRT-PCR data of different abiotic stresses. This is the first report on validations of such internal control genes for gene expression studies from wild plants with medicinal values. Our data showed highly variable expression of all of these genes under one or the other type of abiotic stress in the selected plants. This suggests that experimental conditions variably affect the stability of these internal control genes in these plants. More importantly our data showed that in *C. album* β-tub, EFα1, Ubq and Act-α were the most stable genes under cold, drought, heat, metal and salt stress as suggested by geNorm and NormFinder algorithms. In *F. carica*, 18S rRNA, β-tub and Act were the three selected stable genes under cold, drought, salt and growth hormone stress. Similarly, GAPDH, β-tub, β-tub-1 and EFα1, were selected as stable internal control genes for drought, cold, heat and salt stress in *M. spicata* respectively. These results suggested that a reference gene exhibiting stable expression under certain condition can also express stably to normalize the gene expression under another condition, which shows high specificity to the experimental conditions and requires careful evaluation of internal control gene for every experimental setup (Tong et al., 2009). The gene stability index of these HKGs can also facilitate the selection of more than one internal control for the verifications of gene expression studies of these or any other related plant species and will provide insight into the knowledge of gene expression of target genes to understand several metabolic and developmental processes.
References


Chapter 05

References


<table>
<thead>
<tr>
<th>Rank</th>
<th>Source</th>
<th>Title</th>
<th>Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Garg, R.</td>
<td>&quot;Validation of internal control genes for quantitative gene expression s..&quot;</td>
<td>1%</td>
</tr>
<tr>
<td>2</td>
<td><a href="http://www.cropj.com">www.cropj.com</a></td>
<td></td>
<td>1%</td>
</tr>
<tr>
<td>3</td>
<td>Paula Fernandez.</td>
<td>&quot;Comparison of predictive methods and biological validation ...&quot;</td>
<td>1%</td>
</tr>
<tr>
<td>4</td>
<td>Zhaoguo Tong.</td>
<td>&quot;Selection of reliable reference genes for gene expression stud..&quot;</td>
<td>1%</td>
</tr>
<tr>
<td>5</td>
<td><a href="http://www.biomedcentral.com">www.biomedcentral.com</a></td>
<td></td>
<td>1%</td>
</tr>
<tr>
<td>6</td>
<td><a href="http://www.plosone.org">www.plosone.org</a></td>
<td></td>
<td>1%</td>
</tr>
<tr>
<td>7</td>
<td>Sellars, M.J.</td>
<td>&quot;Real-time RT-PCR quantification of Kuruma shrimp transcripts: A...&quot;</td>
<td>1%</td>
</tr>
<tr>
<td>8</td>
<td>Shin-Young Hong.</td>
<td>&quot;Exploring valid reference genes for gene expression stud...&quot;</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>9</td>
<td>Xiang-Yu Long.</td>
<td>&quot;Genome-wide identification and evaluation of novel internal c...&quot;</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>10</td>
<td>Fan, Chunjie, Jinmin Ma, Qirong Guo, Xiaotie Li, Hui Wang, and Mengzhu Lu...</td>
<td>&lt;1%</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td><a href="http://www.gene-quantification.de">www.gene-quantification.de</a></td>
<td></td>
<td>&lt;1%</td>
</tr>
<tr>
<td>12</td>
<td>Jain, M.</td>
<td>&quot;Validation of housekeeping genes as internal control for studying ge...&quot;</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>13</td>
<td>de Oliveira, L. A., M. C. Breton, F. M. Bastolla, S. d. S. Camargo, R. Margis, J...</td>
<td>&lt;1%</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Rongying TANG.</td>
<td>&quot;Validation of Zebrafish (Danio rerio) Reference Genes for Q...&quot;</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>15</td>
<td>Jin, Xuehua, Jianxin Fu, Silan Dai, Yi Sun, and Yan Hong.</td>
<td>&quot;Reference gene se...&quot;</td>
<td>&lt;1%</td>
</tr>
</tbody>
</table>
Lei Chen. "Validation of reference genes for RT-qPCR studies of gene expression in plant development." Publication

Du, Yishuai, Linlin Zhang, Fei Xu, Baoyu Huang, Guofan Zhang, and Li Li. "Validation of reference genes for studying gene expression in plant development." Publication

M. Migocka. "Identification of suitable reference genes for studying gene expression in plants." Publication

Peters, I.R. "Development and application of multiple internal reference (housekeeping) genes for gene expression studies in plants." Publication

Aline Borges. "Validation of reference genes for RT-qPCR normalization in crop plants." Publication


Kumar, K. "Differential regulation of rice mitogen activated protein kinase kinases in response to abiotic stresses." Publication


Swapna Priya Rajarapu. "Validation of reference genes for gene expression studies in rice." Publication

Carolina V Morgante. "Reference genes for quantitative reverse transcription-polymerase chain reaction (qRT-PCR) in plant tissue culture." Publication


www.jimmunol.org "Validation of reference genes for gene expression studies in mice." Internet Source

David Hoogewijs. "Selection and validation of a set of reliable reference genes for quantitative RT-PCR in Arabidopsis thaliana." Publication


Siang Ling Chua. "UBC and YWHAZ as suitable reference genes for accurate quantitative RT-PCR in rice." Publication

Zivny, J.H. "Reduced erythroid cell and erythropoietin production in response to cold stress." Publication
<table>
<thead>
<tr>
<th>Number</th>
<th>Title</th>
<th>Author/Authors</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>Reference Gene Selection for Real-Time Quantitative Polymerase Ch.</td>
<td>Jiani Qi</td>
<td>Publication</td>
</tr>
<tr>
<td>35</td>
<td>Reference gene selection for qPCR analysis during somatic embry...</td>
<td>Lin, Y.L.</td>
<td>Publication</td>
</tr>
<tr>
<td>36</td>
<td>Validation of reference genes for quantitative RT-PCR studies of...</td>
<td>Cassan-Wang, H., M. Soler, H. Yu, E. L. O. Camargo, V. Carocha, N. Ladou...</td>
<td>Publication</td>
</tr>
<tr>
<td>37</td>
<td>Evaluation of candidate reference genes in Clostridium difficile for</td>
<td>Metcalf, D.</td>
<td>Publication</td>
</tr>
<tr>
<td>38</td>
<td>Evaluation of internal control genes for qRT-PCR normalization in tis...</td>
<td>Su, J.</td>
<td>Publication</td>
</tr>
<tr>
<td>39</td>
<td>Effects of inhibiting CDK9 on the expression of primary response</td>
<td>Keskin, H.</td>
<td>Publication</td>
</tr>
<tr>
<td>40</td>
<td>Isolation of Stress-Related Genes of Rubber Particles and Latex in...</td>
<td>J. S. Kim</td>
<td>Publication</td>
</tr>
<tr>
<td>41</td>
<td><a href="http://www.gene-quantification.org">www.gene-quantification.org</a></td>
<td></td>
<td>Internet Source</td>
</tr>
<tr>
<td>42</td>
<td>Cao, Xi-mei, Xu-guang Luo, Jun-hong Liang, Chao Zhang, Xiao-ping Meng, a...</td>
<td></td>
<td>Publication</td>
</tr>
<tr>
<td>43</td>
<td>Validation of reference genes for gene expression studies in p...</td>
<td>Xiaoyuan Chi</td>
<td>Publication</td>
</tr>
<tr>
<td>44</td>
<td>Zhu, Jianfeng, Lifeng Zhang, Wanfeng Li, Suying Han, Wenhua Yang, and Liw...</td>
<td></td>
<td>Publication</td>
</tr>
<tr>
<td>45</td>
<td>Gopaulchan, David, Adrian M. Lennon, and Pathmanathan Umaharan.</td>
<td></td>
<td>Publication</td>
</tr>
<tr>
<td>46</td>
<td>xms.sicau.edu.cn</td>
<td></td>
<td>Internet Source</td>
</tr>
<tr>
<td>47</td>
<td>Klattig, Jürgen Thomas(Banahmad, Aria, Englert, Christoph and Schartl, Manfr...</td>
<td></td>
<td>Publication</td>
</tr>
<tr>
<td>48</td>
<td>Tony Remans. Normalisation of real-time RT-PCR gene expression measure...</td>
<td></td>
<td>Publication</td>
</tr>
<tr>
<td>49</td>
<td>G. J. Olsen. Sequence of the 16S rRNA gene from the thermoacidophilic arc...</td>
<td></td>
<td>Publication</td>
</tr>
<tr>
<td>ID</td>
<td>Title</td>
<td>Author(s)</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>Evaluation of the in vitro cell-material interactions and in vivo o...</td>
<td>Sophie Verrier</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>Small, B.C.. &quot;Stability of reference genes for real-time PCR analyses in channe...</td>
<td>Small, B.C.</td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>Laura Beekman. &quot;Evaluation of suitable reference genes for gene expression...</td>
<td>Laura Beekman</td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>Smith, Jennifer . &quot;Characterization of the Mechanism of Action of Eribulin and A...</td>
<td>Smith, Jennifer</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>Gen-Hong Wang. &quot;Reference genes identified in the silkworm Bombyx mori du...</td>
<td>Gen-Hong Wang</td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>Zampieri, M.. &quot;Validation of suitable internal control genes for expression studi...</td>
<td>Zampieri, M.</td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>Huang, Chien-Hsun, Mu-Tzu Chang, and Lina Huang. &quot;Species identification ...</td>
<td>Huang, Chien-Hsun</td>
<td></td>
</tr>
<tr>
<td>78</td>
<td><a href="http://www.agrf.org.au">www.agrf.org.au</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>Fijnvandraat, A.C.. &quot;Cardiomyocytes purified from differentiated embryonic ste...</td>
<td>Fijnvandraat, A.C.</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>Ruan, W.. &quot;Actin, a reliable marker of internal control?&quot;, Clinica Chimica Acta, 2...</td>
<td>Ruan, W.</td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>Lise AA Sorby. &quot;Evaluation of suitable reference genes for normalization of re...</td>
<td>Lise AA Sorby</td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>Vito R Cicinnati. &quot;Validation of putative reference genes for gene expression s...</td>
<td>Vito R Cicinnati</td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>McDonald, Mary Kathleen. &quot;Characterizing neuroadaptation processes in chro...</td>
<td>McDonald, Mary Kathleen</td>
<td></td>
</tr>
<tr>
<td>84</td>
<td><a href="http://www.fermentas.com">www.fermentas.com</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>Asad Maroufi. &quot;Validation of reference genes for gene expression analysis in ...</td>
<td>Asad Maroufi</td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>Tiago R. Jacob. &quot;Selection and validation of reference genes for gene expres...</td>
<td>Tiago R. Jacob</td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Source</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effects of hypoxia on mRNA expression of housekeeping genes</td>
<td>Publication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional identification and mapping of a gene</td>
<td>Publication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normalizing genes for real-time polymerase chain reaction in epithel</td>
<td>Publication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validation of internal control genes for expression studies</td>
<td>Publication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First molecular evidence of diatom effects in the copepod Call...</td>
<td>Publication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stem cell-based adipose tissue engineering</td>
<td>Publication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selection of appropriate reference genes for gene expression studies</td>
<td>Publication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Development of Quantitative Real-time Polymerase Chain Reaction...</td>
<td>Publication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A new mouse model of Peyronie's disease: An increased expr...</td>
<td>Publication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improved RNA isolation from Laminaria japonica Aresch (Lamin...</td>
<td>Publication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selection of reference genes for quantitative real-time PC.</td>
<td>Publication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genetics of polyphenol oxidase (PPO) activity in wheat (Triticum...</td>
<td>Publication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selection of suitable reference genes for real-time PCR s...</td>
<td>Publication</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

www.sbi.uni-rostock.de
Internet Source

physiolgenomics.physiology.org
Internet Source

Fernandes, J.M.O.. "Selection of suitable reference genes for real-time PCR s... | Publication
IDENTIFICATION AND VALIDATION OF STABLE INTERNAL CONTROL FOR HEAT INDUCED GENE EXPRESSION OF AGAVE AMERICANA

SITWAT AMAN1, NOOR UL HAQ1 AND SAMINA N. SHAKEEL1*

1Department of Biochemistry, Quaid-i-Azam University, Islamabad, 45320 Pakistan.
*Corresponding author’s e-mail: snq@qau.edu.pk

Abstract

18S ribosomal RNA (18S rRNA) has been used as housekeeping gene for normalization in gene expression studies of plants. Recently, the effect of experimental conditions and nature of samples have been shown on the stability of internal control gene. Agave americana is a monocot heat tolerant plant adapted to arid conditions with Crassulacean acid metabolism (CAM). Few reports have shown the gene expression studies of this or other CAM plants due to lack of suitable reference gene. Here, we partially sequenced 18S rRNA gene of agave and evaluated its potential use as reference gene under heat stress conditions.

Gene specific primers were designed from highly conserved regions of known 18S rRNA genes and amplified by using genomic DNA and transcript of Agave followed by sequencing (submitted to gene bank with accession # HM991824). To validate the potential use of Agave 18S rRNA gene for real-time PCR data normalization, we evaluated the expression stabilities of this gene in different tissues and various heat stress conditions. The plants were treated with different temperatures viz., 35°C, 40°C, 45°C, 50°C and 60°C. The relative abundance of a heat regulated gene, Cp-sHSP (chloroplast small heat shock protein) was examined by real-time PCR. Varied levels of Cp-sHSP gene expression under different heat treatments showed the heat regulated expression. Maximum Cp-sHSP gene expression was observed in the leaves of Agave after heat stress for four hours at 45°C. No significant difference in 18S rRNA expression was observed among control and heat treated samples. Conclusively, this 18S rRNA gene could be used as a stable internal control for normalization of real-time PCR data of A. americana. This work will help to explore many key players in the heat stress related pathways of CAM plants.

Introduction

Gene expression analysis is becoming an important step in analyzing the biological processes in any living organism. Study of the underneath mechanisms and gene expression of complex mechanisms and their networking can lead to the identification of genes necessary for that particular biological process (Teixeira et al., 2009; Zeller et al., 2009; Zuniga et al., 2009; Chou & Huang, 2010; Di Matteo et al., 2010; Miyazaki et al., 2008; Wang & Xu, 2010; Wang et al., 2010). Reverse transcription real-time quantitative polymerase chain reaction (RTqPCR) is supposed to be an accurate and sensitive method of quantifying mRNA transcripts (Bustin et al., 2005; Duquenne et al., 2010; Regier & Frey, 2010; Vaudano et al., 2010). The detection of amplicon accumulation can be observed from beginnings in this method by use of fluorogenic probes or the dyes which can intercalate such as SYBR Green-I, rather. qPCR is a highly sensitive and specific method with excellent reproducibility and less post amplification procedures (Bustin et al., 2005; Valasek & Repa, 2005; Francino et al., 2006; Imène et al., 2011). That’s why qPCR has become one of the favorite method for validation of microarray data or a smaller set of genes, transgenic gene expression, molecular diagnostics andiotic or abiotic stress (Arikawa et al., 2008; Kant et al., 2008; Cortleven et al., 2009; Chang et al., 2010; Di Matteo et al., 2010). It is extremely powerful technique to study the expression patterns of any gene in given conditions but a careful normalization of the data is required. Several experimental strategies have been proposed for normalization of transcript data of different organisms (Bonefeld et al., 2008; Olbrich et al., 2008; Axtner & Sommer, 2009; Artico et al., 2010; Borges et al., 2010; De Santis et al., 2010).

Ideally a good reference gene used for normalization of real-time PCR data should express in a stable way in selected tissues and given experimental conditions. In short the selection of a stable reference gene is very important for accurate gene expression quantification and data analysis (Axtner & Sommer, 2009; Gubern et al., 2009; Bagnall & Kotze, 2010; Boava et al., 2010; Chervoneva et al., 2010). Due to the importance of selection of a stable reference gene in normalization of expression data, different housekeeping genes have been identified and evaluated for stable expression in a given set of conditions in various organisms. In case of plants, many housekeeping genes such as ubiquitin, 18S rRNA, glyceraldehyde-3-phosphate dehydrogenase and actin, b-tubulin were proved to be as good internal controls for expression studies because of their uniform expression in various samples and experimental conditions (Exposito-Rodriguez et al., 2008; Barsalobres-Cavallari et al., 2009; Cortleven et al., 2009; Hu et al., 2009; Artico et al., 2010; Boava et al., 2010). It has been shown that most of the well-known and frequently used reference genes cannot be used as a good internal control for normalization in qPCR because the levels transcript vary considerably under given experimental conditions (Thorrez et al., 2008; Teste et al., 2009). The choice of an appropriate internal control is very important for an accurate quantitative data analysis (Teste et al., 2009; Bentsink et al., 2010; Boava et al., 2010; De Santis et al., 2010).

So far only few reports have shown the identification and validations of appropriate housekeeping gene in plants (Wan et al., 2010; Exposito-Rodriguez et al., 2008; Cortleven et al., 2009; Tong et al., 2009; Artico et al., 2010; Jarosova & Kundu, 2010). Therefore quantitative transcript analysis of many important genes involved in novel pathways of non-model and wild plants are limited due to lack of suitable reference genes. Comparative sequence analysis is a powerful approach toward understanding the structure and translational mechanisms. Different internal control genes of cross species plants have been employed in several studies (Artico et al., 2010; Huis et al., 2010).
18S rRNA gene can be used as an internal control for relative quantification (Tong et al., 2009; Bagnall & Kotze, 2010; Jarosova & Kundu, 2010) due to highly conserved regions at the 3’ region and greater tendency to form hairpin loops. It is highly conserved throughout many plant species and several recent studies have led to the conclusion that the ribosomal RNA was the best choice of internal control in a variety of cell systems (Goidin et al., 2001; Nicot et al., 2005). One of the reasons, why some researchers avoid using ribosomal RNA for normalization of real-time PCR data is its high abundance as compared to target gene expression (Schmittgen & Zakrajsek, 2000), although this ratio varies based on the samples and treatments. There are reports of use of 18S and 25S rRNA as good internal control for highly expressed transcript quantifications (Kim et al., 2003). 18S ribosomal RNA show less fluctuation under the conditions that affect the expression of mRNAs, due to high abundance i.e., more than 80% and different polymerases are involved for transcription of mRNAs and rRNAs (Goidin et al., 2001). Due to their constant basal level expression, no dependence on cell cycle, and non-responsiveness to external treatments or developing stages, these genes can be used as internal controls.

Production of a specific set of proteins, heat shock proteins (HSPs) has been reported in almost all organisms in response to elevated temperatures and many other types of stresses (Vierling, 1991). This is a universal stress response and the mechanism of plant protection under elevated temperatures or other types of stresses is well conserved. Heat shock proteins production is the key for plant survival under heat stress (Wang et al., 2004). HSPs range from 15 to 110 KD. Small heat shock proteins (sHSPs), range in size from approximately 15 to 30kD, are more abundant and diverse in plants than other organisms (Vierling, 1991). Five classes of sHSPs in plants have been reported with different subcellular localizations. Chloroplast small heat shock proteins (CP-sHSPs) are produced in the cytoplasm and then targeted to the chloroplast with the help of a transit peptide that can be cleaved off later on (Chen & and Vierling, 1991; Wang & Luthe, 2003) where they can localized to thylakoids or stroma (Osteryoung & and Vierling, 1994; Heckathorn et al., 2002).

In the present work, 18S rRNA gene was identified and sequenced from Agave americana (A. americana) as a house keeping gene, which can be used as an internal control to check the relative transcript expression under particular treatments of this plant. As this gene is highly conserved, we can also use it in cross-species plants like, A. americana, a heat tolerant CAM plant. Chloroplast specific small heat shock proteins have been chosen for study the difference in the gene expression in this plant due to its heat regulated differential expression in many plants under stress.

Materials and Methods

Plant material and treatments: Maguey or Agave americana is originally adapted to Mexico, although it is now cultivated in many parts of the world. This is very important plant with many uses like, produces fibers for clothing, ropes, bags and to make many tools. In past its thorns were used as an important tool for perforators in bloodletting rituals. A mildly beverage (alcoholic) called pulque, can also be obtained by the fermentation of aguamiel, the sweet, milky juice extracted from the leaves of the plant.

Experiments were performed using A. americana, collected from surroundings of Islamabad and were grown at 30°C/20°C (day/night) in growth chamber. Approximately 100mg plant tissues was used for heat stress in sterilized incubation buffer (1% sucrose, 1mM Potassium phosphate pH 6 and Tween-20) at particular temperature starting from 28°C to 55°C and samples were collected each at 0 (control), 1-4hrs respectively in three replicates. Similarly different tissues of agave plant including leaf tips, leaf base and roots were also treated with elevated temperatures at 45°C for four hours. All the samples were frozen immediately in liquid nitrogen and stored at -70°C until used.

Genomic DNA extraction and PCR: Genomic DNA of A. americana was isolated by using a modified CTAB method (Zhang & Stewart., 2000). The samples were resuspended in 30µl 10mM Tris HCl. 5µl sample was loaded on 1% agarose gel and run at 80volts for 90mins to check the quality and quantity of DNA.

Sequences of known 18S rRNA of different plant species were downloaded from NCBI (http://www.ncbi.nlm.nih.gov/) database and were aligned by using Bioedit program. The conserved regions were selected for primer design with Primer3 software. Similarly, Chloroplast small heat shock protein (Cp-sHSP) gene (unpublished data) has been used for Cp-sHSP primers.

Sequencing of 18S rRNA: Amplified 18S rRNA gene of Agave was sequenced as follows and submitted to gene bank with accession number HM991824. PCR products were purified using the rapid PCR Kit (Marligen, USA) as recommended by manufacture, by adding 100µl conversion buffer and 900ul of binding solution (H1= Conc. Guanidine HCl, EDTA, Tris HCl and Isopropanol). The samples were incubated in water bath at 55°C for 5min after adding 8µl of resuspended silica powder. Samples were vortexed and centrifuged at 13000g for 10sec followed by three times washings with 500µl ice cold sequencing wash buffer. The pellet was dissolved in 30µl of preheated TE at 65°C. The samples were incubated at room temperature for 3min followed by spin at 13000g for 3min. Purified products were confirmed by running on 1.8% agarose gel. Sequencing PCR was done at 95°C for 1min, followed by 95°C for 30 seconds, 54°C for 30 seconds, 72°C for 4min, a total of 30 cycles were repeated followed by a final elongation at 72°C for 10min. Finally PCR products were purified and resuspended in 20µl Hi Dye Formamid (HDF) for sequencing in ABI-310 sequencer by Sanger Dideoxy Chain termination method.

Sequence comparisons and phylogenetic tree: BLASTn was used to find out the similarity of agave specific 18S rRNA gene with other known 18S rRNA sequences downloaded from gene bank. Similarly multiple sequence
alignment was also done by using ClustalW software followed by construction of Phylogenetic tree for further comparisons.

**RNA extraction and Reverse transcriptase PCR:** Total RNA was isolated by using TRI Reagent (MRC, TR# 118) according to the manufacturer’s instructions from 100mg tissue of control and heat stressed plant samples in three replicates. Samples were homogenized in TRI reagent and 200µl chloroform was added, vortexed vigorously for 30 sec, followed by spin at 11000g for 15min at 4°C. Supernatant (0.5ml) was transferred into a fresh tube. Precipitation was done by Isopropanol at room temperature and precipitated at 5000g for 8min at 4°C. Pellet was washed twice with 75% ethanol with centrifugation at 4000g for 5min at 4°C. Pellet was air dried and dissolved in DEPC water. To check the quality and quantity of RNA, 7µl of each sample was used for 1% agarose gel electrophoresis. RNA was quantified both by gel electrophoresis and Nanodrop spectrophotometer (ND-1000 V3.7.1) at 230, 260 and 280nm. A total of 5µg of RNA of each sample was treated with DNase I RNase free (Fermentas Cat. # EN0521).

Synthesis of cDNA was done by using Revert Aid first strand cDNA synthesis Kit (Fermentas Cat. # K1621) as described by manufacturer by using oligo dT primers. After cDNA synthesis, PCR was performed to confirm the quality and integrity of samples. A total of 5µg RNA was used to make cDNA. Then cDNA of control and treated samples were amplified with candidate internal control (gene) primers: (forward primer 5’-TCCTGAGTAAAGGAAGAGACC-3’, reverse primer 5’-CAGATGAAATTCCCAAGAT-3’) and target gene (Cp-HSPs) specific primers (forward primer, 5’-CAGGATGTTTGAAGATGGC-3’ and reverse primer, 5’-ATGACCGAGCCATTTGCTTCCA-3’) by using following conditions: denaturation at 95°C, then at 94°C for 45sec, annealing at 54°C for 45sec, extension at 72°C for 1min for 30 cycles followed by final extension at 72°C for 10min. 10µl of each amplified transcript was checked on 1.2% agarose gel.

**Real-time PCR:** Relative quantification of the target gene was done by using 18S rRNA as reference gene and Cp-sHSP as target gene in ABI 7500 real time PCR system (Applied Biosystems). Real-time PCR was carried out in a 96-well reaction plate using a reaction mixture of 12.5µl of the containing Maxima SYBR Green qPCR Master Mix (2X), forward primer 0.3µM, reverse primer 0.3µM, template DNA (~50ng) and nuclease-free water to make up the final volume of 25µl. cDNA of control and treated samples were amplified with candidate internal control (gene) primers and target gene (Cp-HSPs) specific primers as described above. DNA or RNA contamination was ruled out by addition of negative controls. Two types of controls were used including non-DNA control and non-treated controls for each run. Thermal cycling was performed by using a two-step cycling procedure in the three replicates. Uracil DNA Glycosylase (UDG) pre-treatment at 50°C for 2min and 1 cycle, initial denaturation at 95°C for 10 min followed by denaturation at 95°C for 15sec, annealing and extension at 54°C for 60sec for a total of 40 cycles.

Relative quantification of transcript was done with ABI 7500 systems SDS software V1.4 (Applied BioSystems, USA). All analysis was based on Ct values of PCR products. Three biological replicates of each sample were used for real-time PCR. Average Ct values were obtained for each sample and data was normalized with this newly sequenced 18S rRNA of Agave as reference gene. Non treated (control) sample was used as a calibrator. Relative quantities were determined for each sample. Error bars indicate the standard deviations of individual replicates.

**Results and Discussion**

**RNA quality/quantity assessment:** Quality of RNA is one of the important factors which can cause the variability in the results of real time PCR. We accessed the quality of RNA initially by A260/A280 ratio and A260/A230 ratio checked by NanoDrop, because presence of contaminants can decrease these ratios. Quality of RNA isolated was found to be good and pure based on these ratios (1.8 and 1.9 respectfully) in each sample. To remove the traces of genomic DNA in isolated RNA, we used RNase-free-DNase I treatment as described in methods. The second important factor which can increase the variability of real-time PCR results is the initial amount of RNA used for cDNA synthesis. We used NanoDrop to find out the concentration of RNA and equal amounts were used for further accuracy. These results demonstrated that RNA used for this study was good enough for relative quantification by real-time PCR.

**Identification and characterization of 18S rRNA gene of agave:** 3’ region of 18S rRNA gene is highly conserved in eukaryotes (Hagenbüchle et al., 1978; Azad & Deacon, 1980). Therefore, to identify and sequence of 18S rRNA gene of A. americana, we designed primers from the conserved regions of known 18S rRNA sequences of different plant species, downloaded from NCBI database. Genomic DNA and cDNA was used to amplify approximately 190bp partial sequence of 18S rRNA gene of agave as shown in Fig. 1a. These products were sequenced individually and submitted to gene bank after confirmation by its homology with other known 18S rRNA genes downloaded from gene bank. The highly conserved nature of the Agave 18S rRNA genes was confirmed by alignment with other known 18S rRNA sequences of Typha angustifolia, Dioscorea sylvara, Burmannia congesta, Chimonanthus campanulatus, Hortonia floribunda, Doryphora sasafiras, Restio tetraphyllus, Puya raimondii, Peumus boldus, Laurelia novae-zelandiae, Juncus effuses, Gomortega keule, Galbulimima belgraveana, Daphnandra micrantha, Atherosperma moschatum, Glomero pitcairnia penduliflora, Calycanthus floridus, Acorus gramineus, Calamus caesius, Thalictrum simplex, Aconitum carmichaelii, Clematis gratopsis, Clematis gratopsis, Aquilegia vulgaris, Actaea cimicifuga, Berberis bealei, Fuchsia tinctoria, Coscinium fenestratum, Arcangelsia flava, Nelumbo nucifera, Gymnosiphon bekensis, Gladiolus buckerveldii, Triticum aestivum, Zea mays and Solanum tuberosum. The 18S rRNA sequence of agave
has 98-99% similarity with other known 18S rRNA genes of diverse plants as shown in Fig. 1b. Similarly, phylogenetic tree was constructed to check the evolutionary relationship among 18S rRNA of A. americana and that of from other plants as shown in the Fig. 1c. Analysis of this tree has revealed the conserved nature of 18s RNA of CAM plants and interestingly it grouped together with another monocot, Zea mays which is C4 grass and has different metabolism. These similarities with other monocots have shown the evidence of some changes in the past during their evolution from a common ancestor.

Expression patterns of candidate reference gene and target gene by relative quantification: There are many reports on the stable and constant expression of different reference genes to use for gene expression studies (Wan et al., 2010; Exposito-Rodriguez et al., 2008; Cortleven et al., 2009; Tong et al., 2009; Artico et al., 2010; Jarosova & Kundu, 2010). For an ideal reference gene, the expression should remain nearly constant under any experimental conditions (Butte et al., 2001). In order to check the effect of experimental conditions on the expression of agave 18S rRNA, we used agave specific Cp-sHSP gene as a target gene for relative quantification by real-time PCR under varied levels of heat treatments. We used agave specific Cp-sHSP primers to amplify one of the heat induced chloroplast small heat shock protein by using 18S rRNA gene a reference gene for relative quantification in real time PCR. All primers were first tested with standard RT-PCR and melting curve analysis by using real time PCR to optimize the conditions and to exclude the possibility of presence of primer di-mers. The single products of each gene were verified by gel electrophoresis (Fig. 1a). Melting curve analysis of the primers used has revealed the presence of unique melting peak (Tm) showing specific nature of primers with no mismatch or false priming to the selected genes.
CONTROL FOR HEAT INDUCED GENE EXPRESSION OF AGAVE AMERICANA

Fig. 2. Sequence alignment and phylogenetic relationship of 18S rRNA gene of A. americana with other known genes downloaded from public database.

Time course expression studies of target gene (Cp-sHSP) and reference gene (18S rRNA) was done by giving heat stress to agave plant at 45°C for 1–4hrs. Control plants were kept at 28°C for the given period of time. Total RNA was isolated from all samples. Maximum expression of Cp-sHSP was seen after heat treatment of 4hrs as compared to controls by using 18S rRNA as reference gene (Fig. 2a). While no significant difference in the expression of 18S rRNA gene was seen under heat stress.

Similarly the effect of dose response was seen after heat stress of four hrs at 30°C, 35°C, 40°C, 42°C, 45°C, 50°C and 55°C. Whereas all the control plants were kept at 28°C room temperature and expression levels of target and candidate reference gene was compared. The expression of agave chloroplast small heat shock protein was induced by increasing the heat stress, and maximum expression was seen at 45°C (Fig. 2b). This was followed by a decrease in transcript levels as temperature increases to 55°C, showing the degradation of transcript at higher temperatures. While the expression of 18S rRNA gene remained constant throughout the heat treatments as shown in Fig. 2. Almost no variability in absolute Ct values of 18S rRNA of each sample generated with SDS software were examined individually and plotted as shown in the Fig. 2.

To check the expression levels of 18S rRNA in different tissues or parts of agave plant, we examined the relative expression of target gene and reference gene in Leaf tip, leaf base and roots of the agave plant with and without heat stress. Leaf samples accumulated maximum Cp-sHSP as compared to other tissues after heat stress at 45°C for four hours, while no apparent effect of heat stress was seen on 18S rRNA gene expression Fig. 3. Similarly amplification plot of target and reference gene generated by SDS software of real-time PCR has shown less difference in the amplification verses cycles of PCR, which is indicative of variations in the initial levels of transcripts in the given samples (Fig. 3b). Melting curve analysis of primers used for real-time PCR was also done as given in the Fig. 3c; presence of single peak excludes the possibility of presence of primer di-mers. Agave 18S rRNA can be used as a good internal control for heat stress related studies of high expression genes like sHSPs, because in our hands it had stable expression and no effect of experimental conditions and tissue used was observed. These results are consistent with use of 18S and 25S rRNA as good internal control for quantification of genes with high expression levels in rice (Kim et al., 2003).
**Fig. 3.** Time, dose dependence and tissue specific response of target gene and candidate reference gene under heat stress. To study the time course (A), plant were grown at 45 ºC for 1-4 hours, for dose response (B) the plants were treated at 28ºC, 30ºC, 35ºC, 40ºC, 42ºC, 45ºC, 50ºC and 55ºC for 4 hours and to check the tissue specific response (C) different tissues (leaf tip, leaf base and roots) with and without heat stress (45ºC for 1-4 hours) were used. Total RNA was extracted from three biological replicates of each treatment and respective controls (without heat treatment). Message levels for \(Cp\)-sHSP \((mRNA)\) were determined by real-time PCR by using gene specific primers of \(Cp\)-sHSP genes and \(18S\) rRNA gene as a reference gene to normalize the data. 6µl of amplified product of \(18S\) rRNA was run at 1.5% agarose gel for visualizations. Similarly graphical representation of absolute ct values of \(18S\) rRNA gene of \(A. americana\) generated by SDS software is given which was used for normalization of target gene.

**Conclusion**

A lot of transcript information’s and data is available now days for the analysis of major model plant species to facilitate the identification of appropriate reference genes for gene expression normalization. However, very less is known for most of the other plant species, especially wild plants. In this study, we have identified a suitable reference gene for studying heat regulated gene expression in \(A. americana\). The potential use of \(18S\) rRNA as a reference gene has been shown by studying the expression profiles of chloroplast specific small HSP gene of Agave. As less is known about the CAM plants and CAM metabolism, other researchers can use this gene as an internal control for several gene expression studies of agave and related plants.

**Acknowledgements**

This work was supported by Quaid-i-azam University, Research Grant (2010-11) for basic research.

**References**


(Received for publication 5 September 2010)
MOLECULAR AND COMPARATIVE ANALYSIS OF NEWLY ISOLATED BETA-TUBULIN PARTIAL GENE SEQUENCES FROM SELECTED MEDICINAL PLANTS

SITWAT AMAN1, MUNEEZA IQBAL1, SAFDAR ABBAS1, SADIA BANARAS2, MUHAMMAD AWAI51, IMTIAZ AHMAD1, ZABTA KHAN SHINWARI1 AND SAMINA N SHAKEEL1*

1Department of Biochemistry, Faculty of Biological Sciences, Quaid-i-Azam University 45320, Islamabad, Pakistan
2Department of Biotechnology, Faculty of Biological Sciences, Quaid-i-Azam University 45320, Islamabad, Pakistan
*Corresponding author’s e-mail: snq28@yahoo.com

Abstract

Dynamic nature of beta tubulin (β-tubulin) gene is unleashed by recent studies reporting that apart from being a reliable reference gene serving for normalization purposes in gene expression analysis, β-tubulin encodes for structural proteins playing important role in cell cytoskeleton, microtubules and regulation of cell networking. This study is focused on the identification, isolation and characterization of 6 novel β-tubulin genes isolated from diverse range of 6 medicinal plants including Ficus carica, Pisum sativum, Capsicum annum, Capparis decidua, Maytenus royleana and Eruca sativa. The genomic sequences of newly isolated β-tubulin genes were analyzed and confirmed by using different bioinformatics tools followed by submission to genebank. We used ClustalW for pairwise alignment of these novel sequences with other known 18S rRNA sequences to find out their phylogenetic relationships. Our results have shown conserved nature of β-tubulin with some variable regions might be landmarks of some historical signals. Being housekeeping gene, β-tubulin can be used as good marker and internal control for several types of molecular analysis followed by validations of their consistent expression in the several plant species in future. This study will provide a platform for the molecular biologist interested in studying novel mechanisms of different medicinal plants.

Introduction

An essential pillar in maintaining dynamic microtubule network and wider aspects disclosed in recent researches are the factors making β-tubulin a gene of interest for future studies. Tight regulation of the dynamic behavior and function of the microtubule cytoskeleton is essential for axonal transport, organelle positioning, formation and assembly of cilia and flagella (Dutcher, 2001), cell motility, transport and maintenance of cell shape (Berrieman et al., 2004) as well as for the development and survival of neurons (Tischfield & Engle, 2012) due to their involvement in various cellular processes (Jaglin & Chelly, 2009; Jaglin et al., 2009). Alpha and β-tubulin heterodimers are the major components of the microtubules (Maccioni & Cambiazo, 1995; Dutcher, 2001).

β-tubulin is a GTP binding protein having 445 amino acids residues in sequence (Luduen, 1998; Wood et al., 2001), encoded by four exons of β-tubulin gene (Diaz & Andreu, 1993). After their formation alpha and β-tubulin heterodimers polymerize in a head to tail array to form protofilaments, which assemble forming microtubules (Hesse et al., 1987) containing different α- and β-tubulin isoforms each encoded by distinct genes (Lopata & Cleveland, 1987). Microtubule targeting agents bind to the β-tubulin subunit of the α/β-tubulin heterodimers forming microtubules (Stengel et al., 2010). Microtubules needs to polymerize and depolymerize for performing its functions properly during cell division (Jordan & Wilson, 1998). Further researches on β-tubulin revealed new aspects of this gene. Recently many studies have shown differences in the β-tubulin gene expression is linked with drug resistance in cancer patients (Correnti et al., 1995; Kavallaris et al., 1997; Hasegawa et al., 2003; Mozzetti et al., 2005; Seve et al., 2005; Urano et al., 2006; Tommasi et al., 2007). It has distinct multiple isoatypes having dynamic properties (Banerjee et al., 1992; Panda et al., 1994) and all of these isoatypes are conserved among different species (Berrieman et al., 2004). However small differences in the properties of β-tubulin isoatypes can influence the structure or assembly of microtubules, e.g., βIII-depleted tubulin polymerizes and assembles into microtubules at a faster rate than the unfractionated tubulins (Banerjee et al., 1990; Ranganathan et al., 1998; McKeen et al., 2001). Future studies that examine the role of each isotype in specific organism will greatly impact our overall understanding of microtubule function and behavior, and may provide avenues for future therapeutic intervention (Tischfield & Engle, 2012).

Moreover β-tubulin is also reported as a significant housekeeping gene to be used as internal control for gene expression analysis, because several studies account it as a very reliable reference gene for data normalization (Brunner et al., 2004; Liu & Xu, 2006; deAlmeida et al., 2010; Fernandez et al., 2011). Those genes which encode the transcripts involved in basic cellular processes and cell survival, shows consistent expression in altering experimental conditions, so they are preferred for normalization of data (Czechowsk et al., 2005). β-tubulin is an example of such gene as it encodes for proteins playing role in structure of cytoskeleton, thus it is recommendable for normalization purposes. These wider angles of β-tubulin gene make us to carry out its identification and analysis in diverse group of medicinal plants.
This paper attempts to identify sequence and characterize novel isoforms of β-tubulin genes from six plants including *Ficus carica*, *Pisum sativum*, *Capsicum annuum*, *Capparis decidua*, *Maytenus royleana* and *Eruca sativa*. These species are reported to be medicinally important (Duke, 1981; Konyalouglu et al., 2005; Lynn et al., 2006; Mahla et al., 2010; Rafatullah et al., 2012) this is a follow up of our earlier studies on *18s Ribosomal RNA* of these species (Banaras et al., 2012). This wide range of β-tubulin gene sequences from non-model plants for their potential use as an internal control gene for future studies related to above mentioned plants.

**Material and Methods**

**Plant material:** Six diverse plants i.e., *F. carica*, *P. sativum*, *C. annuum*, *C. decidua*, *M. royleana* and *E. sativa* were collected from different parts of Pakistan for identification and sequencing of different homologues of β-tubulin gene. Their brief description is as follows;

*F. carica* commonly known as 'fig' belongs to family Moraceae and has many medicinal benefits. Its leaves and fruits are very famous as having laxative, stimulant, antitussive and emollient properties. They are also very effective against various throat diseases (Konyalouglu et al., 2005). Its latex also possesses potent pharmacological activities, the most important being the anticarcinogenic and antioxidant properties (Oliveria et al., 2010). Moreover its methanolic extracts can lower the serum level of alanine aminotransferase, aspartate aminotransferase, bilirubin and malondialdehyde equivalents as an index of lipid peroxidation (Mohan et al., 2007).

*P. sativum* belongs to family Fabaceae and is commonly called as 'pea'. It is cool season vegetable crop commonly used for culinary purposes. Its seeds contain trypsin and chymotrypsin that can be used as ecbolic, contraceptive, fungistatic and spermicide (Duke, 1981).

*C. decidua* belongs to family Capparaceae and is important drought resistant plant. It grows with dense, tufty and xerophytic shrub with significant medicinal value besides many socioeconomic and ecological benefits. It possesses many pharmacological properties like hypercholesterolemic, anti-inflammatory and analgesic, antidiabetic, antimicrobial, antiplaque, antihypertensive, antihelminthic & purgative activities (Mahla et al., 2010; Singh et al., 2011). Its intake also results in reduction in plasma triglycerides, total lipids and phospholipid concentrations (Goyal & Grewal, 2003).

*M. royleana* belongs to family Celastraceae and grows in foot-hill zones. It is drought tolerant plant and can grow in arid or semi-arid areas. Its bark and leaves are involved in medicinal uses mainly for the treatment of bone fractures (Rauf et al., 2012).

*E. sativa* belongs to Brassicaceae rocket species that are commonly used as salads vegetable and spice throughout the world (Lamy et al., 2008). It possesses various therapeutic and medicinal properties like inhibition of tumorigenesis, anti-ulcer, hepatoprotective, stimulant, aphrodisiac, diuretic and in treatment of stomach diseases (Lynn et al., 2006; Alqasoumi et al., 2008; Rafatullah et al., 2008) as well as owns dynamic anti bacterial activity (Gulfrazi et al., 2011).

The fresh leaves of these plants were taken for the extraction of genomic DNA.

**Genomic DNA extraction:** Genomic DNA was extracted by CTAB (Cetyl Trimethyl Ammonium Bromide) method. Leaves tissues were washed with distilled water and 70% ethanol. One gram of washed leaves tissues from each plant was ground into fine powder in presence of liquid nitrogen in triplicates. The finely ground tissues were homogenized completely with 3ml of extraction buffer (E.B) in falcon tube and kept at room temperature. The composition of E.B was 100mM Tris HCl with pH 8, 20mM EDTA, 1M NaCl, 2% PVP-40, 0.002% CTAB, 0.02% phenanthroline and 0.2% β-mercaptoethanol. The mixture was kept at 65°C in incubator for one hour with regular shaking and then was kept at the room temperature (25°C). Chloroform and isoamyl alcohol (1:24 ratio) was added in phenol (1:2 ratios) and 3 ml of this mixture was added in the falcon tube. All the samples were centrifuged at 12,000 rpm for 15mins to separate out the debris. The supernatant was taken after careful washing and poured into new falcon tubes. An equal volume of ice cold iso-propanol was added. The mixture was kept at -20°C overnight. Next day, the mixtures were centrifuged at 8000 rpm for 12mins and the pellets were washed twice with 15mM ammonium acetate in 80% ethanol first and then with 100% ethanol. The pellets were air dried and dissolved in 30µl 10mM Tris + EDTA (TE Buffer). Then at last, in order to check DNA quality, 5µl of each sample was loaded on 1% agarose gel and run at 90volts for 70mins. The gel was then checked on gel doc system at proper resolution and image was saved.

DNA quantification was done by NanoDrop-1000 spectrophotometer (ND-1000 V3.7.1, ThermoScientific). The instrument was calibrated to remove zero error with the help of T.E buffer in which DNA was dissolved and was taken as blank. 1µl from each aliquot was used in nanodrop for quantification and concentration was recorded in ng/µl units. Then different dilutions of stock genomic DNA was prepared in order to use required concentrations of DNA. 200ng/µl concentration of DNA was used for polymerase chain reaction (PCR).

**Polymerase chain reaction (PCR):** Genomic DNA of all the six plants was amplified by using gene specific primers under following conditions. First denaturation was done at 95°C for 5 min, followed by 35 cycles denaturation for 45 sec at 94°C, annealing at 57°C for 1 min followed by extension for 1 min at 72°C. Final extension was done for 10 min at 72°C. PCR products were checked on 1% agarose gel and UV trans-illuminator was used to scan the gel.

**Sequencing of β-tubulin gene:** Sequencing PCR products were purified by using Axygen prep kit (Tischfield et al., 2010) according to the manufacturer’s
instructions. Sequencing was performed by using Beckman CEQ 8800 sequencer. Sequencing PCR reaction mixture was made by adding RRv3.1 master mix as recommended by suppliers. Sequencing PCR was done by denaturation at 95°C for 1min followed by 30 cycles of denaturation at 95°C, annealing at 57°C (β-tubulin gene) for 30 seconds each and extension at 72°C for 4min followed by final extension at 72°C for 10min.

**Analysis of sequence:** The β-Tubulin gene sequence from genomic DNA amplification obtained after complete analysis by Beckman CEQ-8800 sequencer was then analyzed in Bioedit software to remove N’s from the sequence and BLASTn software was used for alignment. This alignment information was then used for deducing phylogenetic information among these plants by constructing dendrogram in clustalW multiple alignment application.

**Results and Discussion**

We selected 6 different plants (Ficus carica, Pisum sativum, Capsicum annum, Capparis decidua, Maytenus royleana, Eruca sativa) having significant medical importance in order to disclose β-tubulin gene identification, sequencing and evolutionary analysis. Good quality and quantity of DNA was extracted by CTAB method which was then quantified by Nano Drop. After DNA extraction, β-tubulin gene was amplified using forward and reverse primers which were designed from the conserved regions of known β-tubulin genes from NCBI database. The amplified product of ~250 bp of β-tubulin gene from respective plants is shown in the Fig. 1. We sequenced each product individually by Sanger dideoxy chain termination method in Beckman CEQ-8800 sequencer and β-tubulin partial gene sequences of F. carica, P. sativum, C. annum, C. decidua, M. royleana, E. sativa were submitted to genebank (Genebank group-grp 3974929). Initial analysis of these sequences by BLAST showed high degree of similarities with previously known homologues of β-tubulin genes of a number of other dicot plants, hence confirmed that the identified and sequenced gene is β-tubulin gene of selected medicinal plants. The dicot plants showing similarity in sequences, with isolated and sequenced β-tubulin gene were B.napus, A.thaliana, G.hirsutum, C.maxima, S.tuberosum, P.trichocarpa, R.Communis, L.albus, P.pyrifolia, P.salicina, A.pyhllitidis, E.grandis, D.carota, T.tetragonioides, M.truncatula. Sequences showing similarity were aligned with the new isolated sequences of respective medicinal plants, by using Bio Edit software as shown in Fig. 2. Alignment results indicate that β-tubulin gene sequence is highly conserved. These results are consistent with reports substantiating the conserved nature of β-tubulin gene (Burns & Surridge, 1990; Guenette et al., 1991; Lai et al., 1994).

Our data suggests that this partial fragment of ~250bp of β-tubulin gene is highly conserved among the selected plant species in this study. Most of these sequences have less variability as compared to other sequences.

Phylogenetic tree was constructed based on DNA alignment generated by ClustalW. Result of phylogenetic tree showed evolutionary relationships of isolated β-tubulin gene with β-tubulin gene sequences of other known dicot plants as shown in Fig. 3. It clearly indicates that β-tubulin gene in these dicot plants remained conserved during evolution and is placed with the other dicots in the tree with slight changes in the sequence. Analysis of these newly isolated partial β-tubulin gene patterns and evolution in angiosperms fully or partially supports the clades of phylogenetic analysis of several previous studies (Banaras et al., 2012).

This work can be further extended to study of expression profiles of β-tubulin gene as done previously with other genes (Aman et al., 2012). β-tubulin being a reliable reference gene (Brunner et al., 2004; Liu & Xu, 2006; deAlmeida et al., 2010; Fernandez et al., 2011) can serve for normalization purposes so this work can be proceeded to its validation in respective plants Ficus carica, Pisum sativum, Capsicum annum, Capparis decidua, Maytenus royleana and Eruca sativa.

To participate in exploring dynamic β-tubulin gene, we identified and characterized it in 6 different medicinal plants, we hope study of β-tubulin to flourish and further unleash its potentials in future. Newly revealed dynamic nature of β-tubulin proposes to examine the role of its each isotype, which will greatly impact our understanding of microtubule function and behavior, in relation to other structural proteins of the cell. For all of the future studies to explore different pathways and for gene analysis in the selected set of medicinal plants as internal control after appropriate validations for a given set of experimental conditions. This may provide avenues for future therapeutic intervention and growth of scientific knowledge of major/ minor mechanisms of plant life and physiology.
Fig. 2. Alignment of newly sequenced β-tubulin gene of selected plants including, *F. carica*, *P. sativum*, *C. annum*, *C. decidua*, *M. royleana* and *E. sativa* with related sequences downloaded from NCBI database. Amino acids conserved throughout all sequences are marked by a black background, which shows sequences of plants under study to be highly conserved.
Fig. 3. Phylogenetic tree generated by using newly isolated sequences of β-tubulin gene of respective plants including, F. carica, P. sativum, C. annuum, C. decidua, M. royleana and E. sativa with already known sequences of β-tubulin gene downloaded from NCBI.

Acknowledgments

This research was supported by a QAU Research Grant for 2010 to SNS at Quaid-e-azam University, Pakistan.

References


(Rceived for publication 10 December 2011)
MOLECULAR IDENTIFICATION AND COMPARATIVE ANALYSIS OF
NOVEL 18S RIBOSOMAL RNA GENOMIC SEQUENCES OF
A WIDE RANGE OF MEDICINAL PLANTS

SADIA BANARAS1, SITWAT AMAN2, MARYAM ZAFAR2, MUDASSAR KHAN2, SAFDAR ABBAS2,
ZABTA KHAN SHINWARI1 AND SAMINA N SHAKEEL2

1Department of Biotechnology, Faculty of Biological Sciences, Quaid-i-Azam University 45320, Islamabad, Pakistan
2Department of Biochemistry, Faculty of Biological Sciences, Quaid-i-Azam University 45320, Islamabad, Pakistan

Abstract

Ribosomal RNAs (rRNAs) are universally distributed and known for their functional equivalence among all the known organisms. Analysis of small-subunit rRNAs (16-18S rRNAs) can permit the accurate statistical measurement of a broad range of phylogenetic relationships due to highly conserved sequences. Therefore, we identified and partially sequenced novel isoforms of 18S rRNA gene from 7 wild, medicinal plants (Ferocactus glaucescens, Capparis decidua, Calatropis procera, Maytenus royleana, Prosopis Juliflora, Ficus carica and Mentha spicata) and three cultivated plants (Cyamopsis tetragonoloba, Eruca sativa and Solanum lycopersicum). The genomic sequences of 18S rRNA from all these diverse plants were analyzed and confirmed by using bioinformatics tools and submitted to genebank. We used ClustalW for pairwise alignment of these novel sequences with other known 18S rRNA sequences to find out their phylogenetic relationships. Our results have shown highly conserved nature of 18S rRNA with variable regions might be indications of some historical signals. Secondary structure constrains of rRNA can affect their phylogenetic interpretations rarely. These novel 18S rRNA sequences can also be used as internal controls for several types of molecular analysis after accurate validations of their consistent expression in the given plant species in future studies, as less is known about these housekeeping genes of wild plants.

Introduction

Recently several analyses together with molecular evidences have surprisingly improved our understanding of plant phylogenies (Shinwari & Shinwari, 2010). However, several evolutionary relationships within many major groups of land plants still remained unclear and can be further explored by phylogenetic analysis of several interesting housekeeping genes (Shinwari et al., 2011). Housekeeping genes are integral part of cellular metabolism. We need to find elsewhere for additional and novel characteristics to elaborate the existing picture of plant phylogeny. Ribosomal RNA (rRNA) has frequently been used for reconstruction of deep branches of plant evolutionary history. Small-subunits (16S, 18S) rRNA sequences were used in several attempts to infer the life history (Woese & George, 1977; Woese, 1987; Olsen & Woese, 1996; Woese, 1998). In past, the 18S rRNA sequence analysis was used to predict early eukaryotic diversifications (Bhattacharya & Medlin, 1995) and based on those predictions, the fungi were placed in a sister group to animals (Wainright et al., 1993). Similarly, 18S rRNA sequences have been successfully used in reconstruction of eukaryotic phylogeny into many groups of plants including algae, bryophytes, gymnosperms and angiosperms (Buchheim & Chapman, 1991; Chaw et al., 1993; Chaw et al., 1995; Hedderston et al., 1996; Chaw et al., 1997; Solits et al., 1997; Chapman et al., 1998; Hedderston et al., 1998). Similarly several other studies showed morphological and physiological characterization of plants (Muntaz et al., 2011). Molecular polymorphism and phylogenetic relationships has also been extensively studied (Akbar et al., 2011).

18S rRNA sequences have been used in several studies with main focus on the origin of land plants with their placement into distinct phylogenetic groups (Mishler et al., 1994; Hedderston et al., 1996; Hedderston et al., 1998). All of these studies actually portray the unique patterns of land plant’s relationships in many clades and usually cannot provide an accurate general outline of few plant phylogenies due to limited samplings of taxons (Hedderson et al., 1996), which can leads to some disagreements on the relationships of major lineages. The larger sampling can improve the inferences of plant phylogenies based on 18S rRNA or related sequences. This seems a challenging job due to the limited number of known genome/gene sequences covering wide range of plants. Most of the available genomic information is unfortunately still limited to the model plants and few crops and not much is known about wild or some cultivated plants with more or less economic importance and medicinal values (Shinwari & Qaisar, 2011). The grouping of such sequences into functional groups based on their expression levels is very beneficial as it can provide a basic framework to direct further research for defining their particular roles in form of gene products in evolution.

18S rRNA are housekeeping genes (HKGs) and are ubiquitously expressed in all tissues and cell types for the maintenance of the basic cellular functions in living cells. Furthermore, the expression of such genes is assumed to be comparatively constant or nearly constant during all the environmental or experimental conditions. Most commonly used plant housekeeping genes are β-actin (ACT), α-tubulin (TUA), ubiquitin (UBQ), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), 18S or 6S ribosomal RNA and elongation factors (EF) etc., (Nicot et al., 2005; Hu et al., 2009; Garg et al., 2010; Maroufi et al., 2010).

Normalization of a target gene expression to the HKGs in several molecular expression analyses is required to minimize the variations in target gene quantifications irrespective of experimental conditions. But an accurate
selection of housekeeping genes with stable expression under given conditions is a prerequisite to achieve the above objective (Jain et al., 2006; Majerowicz et al., 2011; Borges et al., 2012). Alternatively 2 or more housekeeping genes can also be used as internal controls for data normalization to minimize the experimental errors (Thellin et al., 1999; Vandesompele et al., 2002).

Again, most of the internal control gene identifications and validation studies are still limited to model plants or crops. Recent studies have shown that different experimental conditions can affect the stable expression of some internal controls due to their involvement in multiple cell signalling/defense pathways (Vandesompele et al., 2002; Sitwat et al., 2012). It’s really important to identify the novel homologues or orthologues of known housekeeping genes of non-model plant species to study their metabolic pathways and related variations according to their habitats or experimental conditions.

This paper attempts to identify, sequence and characterize novel homologues and orthologues of 18S rRNA genes from a diverse group of plants including seven wild, medicinal plants (Fereroxoccus glucuescens, Capparis decidua, Calatropis procera, Maytenus royleana, Prospis juliflora, Ficus carica and Mentha spicata) and three cultivated plants (Cyanopsis tetragonoloba, Eraca sativa and Solanum lycopersicum) to describe their phylogenetic relationships based 18S rRNA sequences. As we need to explore wide range of housekeeping genes from non-model plants for their potential use as an internal control gene (reference gene). These novel, partial gene sequences of 18S rRNA, isolated from above mentioned plants can also be used as internal controls for normalizations of several types of gene expression studies of these selected plants after accurate validations of their consistent expression in particular plant species under given experimental conditions in future.

Materials and Methods

Plant materials: Here, we used ten diverse plants for identifications and sequencing of different homologues of 18S rRNA gene as given below with their economical and medicinal values.

*Cyanopsis tetragonoloba* is commonly known as Guar or cluster beans, belongs to family leguminase. It is a drought-tolerant crop, commonly grown in arid and semi-arid regions with annual rainfall of 200-600mm. Guar beans can be used as vegetables for human consumption, also grown for cattle feed and for green manure. It contains large endosperm comprises of significant amounts of galactomannan gum that forms a viscous gel in cold water. The gum obtained is the primary marketable product of the plant. Highly refined guar gum is used as a stabilizer for cheeses, stiffener in ice cream, and is a meat binder whereas the lower grade of guar gum is used in cloth and paper manufacturing industries (Undersander et al., 1991).

*Eraca sativa* commonly called taramira belongs to the family Brassicaceae, and is used as spice and vegetable for human consumption. It is important for preparations of some traditional medicines and remedies (Flanders & Karim, 1985). It is also well known for the drought resistance and salt tolerance (Shannon & Grieve, 1999).

*Calotropis procera* belongs to family Asclepiadaceae and is important for its medicinal properties. Its different parts have been reported to exhibit antioxidant, analgesic, and anti-inflammatory properties. It has bactericidal and vermicidal effects and can be used to treat leprosy and elephantiasis (Sing et al., 2002). Its latex has been reported as a useful remedy for the coetaneous infections, leprosy, inflammation, eczema, and malarial and low grade fevers (Kumar & Basu, 1994). It is a drought-resistant and salt tolerant plant.

*Capparis decidua* is a member of family Capparaceae. Fruits (green berries) can be used as vegetables and have anti-diabetic action. The bark has been reported for the treatments of cough, inflammations and asthma. Roots are useful to cure fever and buds are good to cure boils. Leaves can be used as appetizers, and are helpful in cardiac problems. Shoots are commonly used for antifertility tonic. Root bark act as anthelmintic and purgative and wood coal is effective for muscular injuries. It is highly tolerant to persistent drought conditions and is known for its adaptations to the arid conditions.

*Prospis juliflora* belongs to family Mimosaceae. It is a leguminous, perennial phreatophyte. It grows in very hot and dry areas with high temperatures like 48°C with annual precipitation of 150-750mm (Darke, 1993; Geilfur, 1994). Its pods are one of the earliest known foods of ancient man. Pods are fermented to make wine. Leaves can be used as forage. Wood is used for floors, furniture, and many items. Toasted seeds can be added to coffee. The gum is used as an emulsifying agent. Gum is used in confectionary. Roots also contain 6–7% of tannin, which discourage the Rhizobia. It is used as a folk remedy for colds, catarrh, diarrhea, dysentery, eyes, inflammations, itching, measles, stomachaches, sore throats, and wounds (Duke & Wain, 1981). Aqueous and alcoholic extractions are remarkably antibacterial.

*Maytenus royleana* belongs to family Celastraceae. It is highly drought tolerant plant and can survive in arid/semi-arid regions. Bark or leaves in powdered form is used for herbal therapy for the treatment of bone fractures (Rauf et al., 2012).

*Ficus carica* is dicot and belongs to family Moraceae. It is a monoeocious and deciduous tree or a large shrub. Common Fig plant is used as a laxative, expectorant, emollient and analgesic. It is usually used in preparations of laxative syrups in combinations with Senna and carminatives. The fruit can be used treatments of colds. Fresh figs can be used for treatment of boils and very small tumors. Its white milky juice extracted from the stems and leaves is used for removal of warts.

*Mentha spicata* (Spearmint) is a herbaceous, rhizomatous and perennial plant, belongs to family Lamiaeae. Its leaves produce an essential oil used for flavor in candies, gums, ice creams, drinks. It is also used commercially for preparations of hygiene products (toothpaste, mouth-washs, etc). It has been used in many continents as an alternative treatments due to its anitmic, antispasmodic, antiseptic, carminative, diuretic, restorative, stimulant, stomachic and tonic. The medicinal herb tea made from the leaves is used in the treatment of fevers, bronchitis, chills, cramps, chronic gastritis, common cold,
diuretic, morning sickness, nasal congestions, halitosis, nausea, painful menstruation, and many minor problems.

**Genomic DNA extraction:** Genomic DNA was extracted from leaves of all the selected plants by CTAB (Cetyl Trimethyl Ammonium Bromide) method (Richards, 1997). Plant leaves (~0.3 g) were harvested, washed with 70% ethanol and homogenized in preheated (65°C) 2X CTAB buffer followed by incubation at 65°C for 45 minutes and centrifugation at 10,000 rpm for 10 minutes. The supernatant was then collected and transferred to new tubes. Equal volume of chloroform-isooamylic alcohol (24:1) was added and mixed with the supernatant followed by centrifugation at 10,000 rpm for 10 minutes. Equal volume of chilled isopropanol and 1 M sodium acetate was added to the supernatant. The mixture was kept at -20°C for 30 minutes for DNA precipitation. Finally centrifuged at 12,000 rpm for 10 minutes and subsequent washings were done to remove impurities followed by air drying. The pellet was resuspended in 40µl of Tris EDTA buffer containing 10 µg/µl of RNase. The DNA samples were incubated at 37 °C for 30 minutes to remove RNA impurities and purified samples were stored at -20°C for further use.

DNA samples were quantified by using NanoDrop-1000 spectrophotometer (ND-1000 V3.7.1, Thermoscientific) and the DNA samples were diluted to a final concentration of 200ng/µl for further molecular analysis. Similarly the DNA samples were also analyzed by loading samples on 1% agarose gel stained with ethidium bromide for gel electrophoresis.

**Polymerase chain reaction (PCR):** PCR was performed to amplify of 18s rRNA gene from all of the above selected plants by using gene specific primers (Haq et al., 2010) and Promega’s master mix (Cat. # M7502) according to manufacturer’s instructions at following PCR conditions for amplification. First denaturation was done at 95°C for 5 min, followed by 35 cycles denaturation for 45 sec at 94°C, annealing at 55°C for 1 min followed by extension for 1 min at 72°C. Final extension was done for 10 min at 72°C. PCR products were checked on 1% agarose gel.

**Sequencing of partial 18s rRNA gene:** Sequencing PCR products were purified by using Axygen prep kit (Catalog No. AP-PCR-250) according to the manufacturer’s instructions. Sequencing was performed by using Beckman CEQ 8800 sequencer. Sequencing PCR reaction mixture was made by adding RRv3.1 master mix as recommended by suppliers. Sequencing PCR was done by denaturing the template at 95°C for 1 min, followed by 30 cycles of denaturation at 95°C, annealing at 55°C (18s rRNA) for 30 seconds each, and extension at 72°C for 4 min, followed by final extension at 72°C for 10 min.

**Analysis of sequences:** The sequences were initially analyzed by using BioEdit software. To confirm the identified partial 18s rRNA gene sequences, we first used BLAST with “somewhat similar sequences (blastn)” options to find out the similarities of these genes with other known plant genes. Then several other plant genes with high similarity were downloaded from NCBI with their accession numbers and were used for alignment by using BioEdit software followed by construction of heuristic parsimony phylogenetic trees for evolutionary analysis.

**Results and Discussion**

Ten novel partial sequences of 18s rRNA gene were isolated and characterized from a group of seven diverse wild plants (Ferocactus glaucescens, Capparis decidua, Calatropis procera, Maytenus royleana, Prosopis Juliflora, Ficus carica and Mentha spicata) and three cultivated plants (Cyanopsis tetragonoloba, Eruca sativa and Solanum lycopersicum). For identification and characterizations of 18s rRNA genes we isolated the genomic DNA from the leaves of all of the above plants by using CTAB method (Richards, 1997). Good quality genomic DNA is one of the prerequisite for PCR and other PCR based technologies. The quality and quantity of extracted DNA was further analyzed by NanoDrop (ND/-1000 V3.7.1) and agarose gel electrophoresis, which showed the presence of high molecular weight DNA with minimum degradations in each case. These genomic DNAs of all the selected plants were used as template for polymerase chain reaction (PCR) to amplify 18SrRNA gene individually by using gene specific primers. Approximately 200 to 290bp products were amplified from each selected plant species as shown in Fig. 1 and absence of amplified product of 18SrRNA gene in case of non-template control was indicative of gene specific amplifications in PCR in each case as expected. We sequenced this entire product individually and the 18s rRNA partial gene sequences of F. glaucescens, S. lycopersicum, C. decidua, C. procera, C. tetragonoloba, E. sativa, M. royleana, P. juliflora, F. carica and M. spicata were submitted to genebank (Genebank accession numbers JX444499-JX444508 respectively) after initial analysis by BLAST, which showed high degree of similarities with previously known homologues of plant 18s rRNA genes downloaded from NCBI database. Our results confirmed that all the sequenced products are unique and novel sequences of 18s rRNA gene isolated from non-model plant species.

![Fig. 1. Amplification of partial sequences of 18s rRNA gene from a diverse group of plants including Ferocactus glaucescens, Capparis decidua, Calatropis procera, Maytenus royleana, Prosopis Juliflora, Ficus carica, Mentha spicata, Cyanopsis tetragonoloba, Eruca sativa and Solanum lycopersicum. Gene specific primers of 18s rRNA were used to amplify genomic DNA extracted from all of the above selected plants (Lane 2 to 11 respectively). PCR products were run on 1% agarose gel along with 100bp DNA ladder. No amplification was seen in the non-template controls (Lanes 1 and 12).](image-url)
To find out the similarities and conserved patterns of newly isolated sequences of 18S rRNA genes from above mentioned plants, we aligned these novel partial sequences together with already known 18S rRNA genes of other plant species as shown in Fig. 2. Our data suggests that this particular segment of ~200-290 bp of 18S rRNA is highly conserved among these selected plant species. Most of these sequences had less variable sequences as compared to few sequences with high degree of variability. Interestingly, we could not see any major differences or unique conserved regions characteristic of monocots or dicots.

Fig. 2. Alignment of novel genomic DNA sequences of partial 18S rRNA gene isolated from Ferocactus glaucescens, Capparis decidua, Calatropis procera, Maytenus royaleana, Prosopis Juliflora, Ficus carica, Mentha spicata, Cyamopsis tetragonoloba, Erucas sativa and Solanum lycopersicum with genomic DNA sequences of 18S rRNA genes of previously known plant 18S rRNA genes.
To find out the similarities and conserved patterns among newly isolated partial sequences of 18S rRNA genes, we aligned these novel partial sequences together with already known 18SrRNA genes of other plant species as shown in Fig. 2. Our data suggests that this particular segment of ~200-290 bp of 18S rRNA is highly conserved among these selected plant species. Most of these sequences have less variability as compared to other sequences. Interestingly, we could not see any major differences or unique conserved regions characteristic of monocots or dicots.

Similarly, phylogenetic trees were also constructed by using all of above 18S rRNA partial sequences as shown in Fig. 3. Our results showed high degree of conservation among this part of 18SrRNA gene of all selected plants. Analyses of 18S rRNA patterns and evolution in angiosperms fully or partially supported the clades showing transitions and transversions in the previous studies. Our data showed that these selected plants fall into three major clades. One major clade comprises of 14 plant species including E. sativa, B. oleracea, A. thaliana, O. sativa, C. decidua, H. orientalis etc. as shown in Fig. 3.

Fig. 3. Phylogenetic relationships of partial 18S rRNA gene isolated from diverse group of plants including, Ferocactus glaucescens, Capparis decidua, Calatropis prodera, Maytenus royleana, Prosopis juliflora, Ficus carica, Mentha spicata, Cyamopsis tetragonoloba, Eruca sativa and Solanum lycopersicum, with already known sequences of 18S rRNA genes by using Paup. The evolutionary history was inferred using Neighbor-Joining method. The percentage of replicate trees in which the associated taxa were clustered together in the bootstrap analysis of 1000 replicates are shown above the branches. The molecular systematics of the green algae. In: The molecular systematics of plants II (Eds.): P.S. Soltis, D.E. Soltis, and J.J. Doyle, Kluwer Academic Publishers, Norwell, Massachusetts, USA. pp. 508-540.

References


