Supplementary Material

# Supplementary Figures and Tables

## Supplementary Figures



**Supplementary Figure 1.** Multiplesequence alignment of ZosmaCHSs with other species. At, *Arabidopsis thaliana*; Ms, *Medicago sativa*; Pp, *Physcomitrium patens*. The catalytic triad of CHSs, CoA-binding residues and CHS family-specific Pro375 were labeled by purple, red and black points, respectively.



**Supplementary Figure 2**. Phylogenetic relationships of the ZosmaCHSs and CHS homologues from other plants using neighbor-joining method. Green, yellow and blue cycles indicate the ZosmaCHSs in clade I, II and III, respectively.



**Supplementary Figure 3**. **(A)** Prokaryotic expression and purification of the empty pET-28a vector in the transformed *E. coli* cells. 1: no IPTG induced; 2: IPTG induced; CL: cell with lysis buffer; FT: flow through buffer; W1: primary washing buffer; W2: secondary washing buffer; E1: primary elution buffer; E2: secondary elution buffer; E3: third time elution buffer. The red arrow represented the corresponding bands of ZosmaCHS (around 43 kDa). **(B)** Prokaryotic expression of the five recombinant ZosmaCHS proteins in the transformed *E. coli* cells. The corresponding protein bands were pointed out by the red arrow at the expected position (around 43 kDa). M: marker; 1: ZosmaCHS01, no IPTG induced; 2: ZosmaCHS01, IPTG induced; 3: ZosmaCHS02, no IPTG induced; 4: ZosmaCHS02, IPTG induced; 5: ZosmaCHS07, no IPTG induced; 6: ZosmaCHS07, IPTG induced; 7: ZosmaCHS08, no IPTG induced; 8: ZosmaCHS08, IPTG induced; 9: ZosmaCHS11, no IPTG induced; 10: ZosmaCHS11, IPTG induced.

**Supplementary Tables**

**Supplementary Table 1:** List of primers used in qRT-PCR.

|  |  |
| --- | --- |
| **Primer Name** | **Sequence (5′--------3′)** |
| *Zosma113g00070-*F | 5' ATACTTGACGAGATGAGGAGGAA 3' |
| *Zosma113g00070-*R | 5' AAGATGGGTTGGCTGTGGA 3' |
| *Zosma263g00020-*F | 5' GTCCAAGGAGGAGGGAAAGG 3' |
| *Zosma263g00020-*R | 5' GCCGTGGAGGACCAGAGTTT 3' |
| *Zosma113g00090-*F | 5' GTCTGACCTTCCATCTCCTCAA 3' |
| *Zosma113g00090-*R | 5' CATCTTATCCTCCATCAACCCT 3' |
| *Zosma11g00330-*F | 5' TTAGGGTTGACGGAGGAGAA 3' |
| *Zosma11g00330-*R | 5' CGTGGAGGACCAGAGTTTCA 3' |
| *Zosma14g00280-*F | 5' GGGTTGACGGAAGATAAGATG 3' |
| *Zosma14g00280-*R | 5' GTTTGTTATGCTGACGCTGTG 3' |
| *Zosma44g00150-*F | 5' CACGATGGACGAGATGAGAAG 3' |
| *Zosma44g00150-*R | 5' CGGAAGGCTACGGAGAACAAT 3' |
| *Zosma52g00210-*F | 5' CCGTACTCGCGGAGTCTTT 3' |
| *Zosma52g00210-*R | 5' TGGTTTCGCATCTCGTTCA 3' |
| *Zosma59g00320-*F | 5' CCGAGGAAGGTGCTGAGTG 3' |
| *Zosma59g00320-*R | 5' CCTGGTCCGAATCCAAAGA 3' |
| *Zosma240g00250-*F | 5' CTCCGACTGGGACTCTTTGTT 3' |
| *Zosma240g00250-*R | 5' TCTGCTGGTCCTCAGTTTGTC 3' |
| *Zosma132g00440-*F | 5' CGAAGACAAACTGAGGGCCACA 3' |
| *Zosma132g00440-*R | 5' CCAAAGAGAACACCCCAATCCA 3' |
| *Zosma699g00010-*F | 5' TTGGAGGACAAGAAGGGGATTT 3' |
| *Zosma699g00010-*R | 5' TTTGGCTTTGGCATAGGAGGTA 3' |
| *Zosma97g00290-*F | 5' GAGAAAGGACATCGGAACGA 3' |
| *Zosma97g00290-*R | 5' CATTAACGGATTGGGAGTGC 3' |

**Supplementary Table 2:** List of primers used inrecombinant protein expression assays.

|  |  |
| --- | --- |
| **Primer Name** | **Sequence (5′--------3′)** |
| *Zosma11g00330-*F | 5' GCGGATCCATGTCTGTGGCGATGAT 3' |
| *Zosma11g00330-*R | 5' GCGAATTCACTGTGAGACGCAGGGC 3' |
| *Zosma14g00280-*F | 5' GCGGATCCATGTCACCGGCGTCGACGATCGAGG 3' |
| *Zosma14g00280-*R | 5' GCGAATTCGTTTGTTATGCTGACGCTGTGGAGA 3' |
| *Zosma113g00070-*F | 5' GCGGATCCATGTCTGTGGCGATGATCGAGGACT 3' |
| *Zosma113g00070-*R | 5' GCGAATTCGGTGGTTAAGATGGGTTGGCTGTGG 3' |
| *Zosma113g00090-*F | 5' GCGGATCCATGTCGGAAGGGATGATCGAGTTGT 3' |
| *Zosma113g00090-*R | 5' GCGAATTCGGTTGTTAAAATGGGCTGACTGTGG 3' |
| *Zosma263g00020-*F | 5' GCGGATCCATGTCTGTGGCGATGAT 3' |
| *Zosma263g00020-*R | 5' GCGAATTCTTCGACGATCGGTTGGC 3' |

**Supplementary Table 3** The number of non-synonymous substitutions per non-synonymous site (*Ka*), the number of synonymous substitutions per synonymous site (*Ks*) and their ratios (*Ka*/*Ks*) between *ZosmaCHS* duplication pairs

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **duplication pairs** | ***K*a** | ***K*s** | ***K*a/*K*s** | ***selective mode*** |
| *ZosmaCHS05*/*ZosmaCHS06* | 0.001101 | -1 | - | - |
| *ZosmaCHS07*/*ZosmaCHS08* | 0.035016 | 0.293130123 | 0.119454959 | negative selection |
| *ZosmaCHS02*/*ZosmaCHS01* | 0.122916 | 1.174929572 | 0.10461582 | negative selection |
| *ZosmaCHS02*/*ZosmaCHS11* | 0.122447 | 1.171402671 | 0.10453011 | negative selection |
| *ZosmaCHS03*/*ZosmaCHS09* | 0.159805 | 1.55046924 | 0.103068976 | negative selection |
| *ZosmaCHS03*/*ZosmaCHS10* | 0.153901 | 1.214635173 | 0.126705705 | negative selection |

1 -, cannot be calculated

**Supplementary Table 4** Detailed information of the 10 conserved motifs in the chalcone synthase (*CHS*) family of *Zostera marina*

|  |  |  |  |
| --- | --- | --- | --- |
| **motif** | **width** | **best possible match** | **E-value** |
| 1 | 50 | EWGQPKSKITHLIFCTTSGVDMPGADYQLTKLLGLRPSVKRLMMYQQGCF | 1.0e-316 |
| 2 | 50 | DYPDYYFRVTKSEHLTDLKKKFQRMCEKSMIRKRHMYLTEEILKENPNVC | 1.0e-291 |
| 3 | 29 | AFTPLGIDDWNSLFWIAHPGGPAILDQVE | 1.9e-181 |
| 4 | 24 | ATRLVLKEYGNMSGARVLFVLDEM | 1.7e-179 |
| 5 | 50 | ERPLFZIISASQTILPDSEGAIDGHLRESGLTFHLLKDVPGLISKNIEKS | 7.1e-211 |
| 6 | 40 | RRKSAEDGKKTTGEGYDWGVLFGFGPGLTVETLVLHSIPI | 1.7e-185 |
| 7 | 21 | IGQAQRADGPAAVJAIGTADP | 5.1e-109 |
| 8 | 29 | YMAPSLNVRQDMVVVEVPKLGKEAATKAI | 3.0e-90 |
| 9 | 15 | AVTFRGPSETHJDSL | 1.6e-47 |
| 10 | 28 | GVTALRLAKDIVENNRGARVLVVTADIM | 6.3e-25 |

**Supplementary Table 5** The putative *cis*-elements in 2.0 kb upstream from the start codon

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *cis*-elmments | group I | | | | group II | | | group III | | | |
| *Zosma*  *CHS05* | *Zosma*  *CHS06* | *Zosma*  *CHS07* | *Zosma*  *CHS09* | *Zosma*  *CHS10* | *Zosma*  *CHS04* | *Zosma*  *CHS02* | *Zosma*  *CHS03* | *Zosma*  *CHS08* | *Zosma*  *CHS11* | *Zosma*  *CHS01* |
| CGTCA-motif | 1 | 1 | 4 | 8 | 2 | 2 | 3 | 2 | 3 | 1 | 2 |
| Box 4 | 3 | 3 | 2 | 1 | 6 | 6 | 5 | 5 | 6 | 4 | 4 |
| TGACG-motif | 1 | 1 | 4 | 8 | 2 | 2 | 3 | 2 | 3 | 1 | 2 |
| MYB | 5 | 5 | 3 | 5 | 0 | 6 | 2 | 3 | 4 | 3 | 5 |
| MYC | 7 | 7 | 5 | 7 | 1 | 6 | 3 | 0 | 1 | 0 | 2 |
| MBS | 1 | 1 | 1 | 2 | 0 | 2 | 1 | 1 | 0 | 4 | 3 |
| GT1-motif | 1 | 1 | 2 | 2 | 0 | 1 | 1 | 1 | 3 | 2 | 0 |
| ABRE | 2 | 2 | 2 | 0 | 0 | 1 | 3 | 0 | 5 | 3 | 5 |
| ARE | 0 | 0 | 2 | 0 | 8 | 1 | 1 | 1 | 3 | 2 | 2 |
| TCT-motif | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 2 | 1 | 1 | 1 |
| GA-motif | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| MRE | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 1 | 1 |
| TCA-element | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| circadian | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| I-box | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| GATA-motif | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 2 |
| GARE-motif | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| TC-rich repeats | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 3 | 1 |
| A-box | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 2 | 4 |
| AE-box | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 2 |
| TGA-element | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 2 | 2 | 0 | 1 |
| AuxRR-core | 0 | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 2 | 0 | 1 |
| G-box | 2 | 2 | 4 | 1 | 0 | 0 | 2 | 0 | 6 | 2 | 4 |
| TCCC-motif | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| GCN4\_motif | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| CAT-box | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| GARE-motif | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| LTR | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| Unnamed\_1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| P-box | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 |
| L-box | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| chs-CMA2a | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 0 |
| ACE | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| ATC-motif | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Sp1 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 |
| Box II | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |

**Supplementary Table 6** The accession numbers of all CHSs used for the tree

|  |  |
| --- | --- |
| **Protein Name** | **Accession Number** |
| Pinus sylvestris CHS | P30079' |
| Medicago sativa CHS | P30075 |
| Ipomoea purpurea CHS E | [O22047](https://www.uniprot.org/uniprot/O22047) |
| Ipomoea purpurea CHS D | [O04934](https://www.uniprot.org/uniprot/O04934) |
| Arabidopsis thaliana CHS | [P13114](https://www.uniprot.org/uniprot/P13114) |
| Oryza sativa CHS | [Q2R3A1](https://www.uniprot.org/uniprot/Q2R3A1) |
| Zea mays CHS01 | P24824 |
| Zea mays CHS02 | P24825 |
| ZosmaCHS02 | A0A0K9PYG7 |
| ZosmaCHS08 | [A0A0K9Q2V7](https://www.uniprot.org/uniprot/A0A0K9Q2V7) |
| ZosmaCHS07 | [A0A0K9Q4S5](https://www.uniprot.org/uniprot/A0A0K9Q4S5) |
| ZosmaCHS01 | [A0A0K9Q3B7](https://www.uniprot.org/uniprot/A0A0K9Q3B7) |
| ZosmaCHS11 | A0A0K9PF20 |
| Hydrangea macrophylla CTAS | [O82145](https://www.uniprot.org/uniprot/O82145)' |
| Rheum palmatum BAS | [Q94FV7](https://www.uniprot.org/uniprot/Q94FV7) |
| ZosmaCHS03 | A0A0K9P0Q7 |
| ZosmaCHS09 | [A0A0K9PZ39](https://www.uniprot.org/uniprot/A0A0K9PZ39) |
| ZosmaCHS10 | [A0A0K9PJD4](https://www.uniprot.org/uniprot/A0A0K9PJD4) |
| Ipomoea purpurea CHS A | [P48397](https://www.uniprot.org/uniprot/P48397) |
| Ipomoea purpurea CHS C | [P48399](https://www.uniprot.org/uniprot/P48398) |
| Ipomoea purpurea CHS B | [P48398](https://www.uniprot.org/uniprot/P48398) |
| ZosmaCHS04 | [A0A0K9NZH3](https://www.uniprot.org/uniprot/A0A0K9NZH3) |
| ZosmaCHS06 | A0A0K9NV10' |
| ZosmaCHS05  Escherichia coli FABH | [A0A0K9NWQ5](https://www.uniprot.org/uniprot/A0A0K9NWQ5)  P0A6R2 |