In-silico genome wide analysis of Mitogen Activated Protein Kinase Kinase Kinase gene family in *C. sinensis*

Abhirup Paul^{1†}, Anurag P. Srivastava^{2†}, Shreya Subrahmanya³, Guoxin Shen^{4†*}, Neelam Mishra^{3*}

¹Department of Biochemistry REVA University Bangalore, Karnataka, India

²Department of life Sciences Garden City University Bangalore, Karnataka, India

³Department of Botany St. Joseph's College autonomous Bangalore, Karnataka, India

⁴Sericultural Research Institute, Zhejiang Academy of Agricultural Sciences Hangzhou 310021, China

[†]These authors contributed equally to this work.

*Corresponding authors:

Guoxin Shen, Ph.D., Professor, Tel: +86-571-86404298; Fax: +86-571-86404298

Email address: guoxin.shen@ttu.edu

Neelam Mishra, Ph.D., Assistant professor

Email address: neelamiitkgp@gmail.com; neelammishra@sjc.ac.in

Orcid id:

- 1. Abhirup Paul: 0000-0003-2143-7511
- 2. Neelam Mishra: 0000-0001-6191-5392







S1 Fig. Transmembrane helices of MAPKKKs of MEKK subfamily in *C. sinensis.* TMHMM Server, v.2.0 (http://www.cbs.dtu.dk/services/TMHMM/), was used to predict the presence of transmembrane helices. The plot shows the posterior probabilities of inside/outside/TM helix. The plot is obtained by calculating the total probability that a residue sits in helix, inside, or outside summed over all possible paths through the model. The red peaks represent the possible transmembrane helices present. If the red peak crosses the transmembrane line (represented in pink colour), then its calculated as 1 transmembrane helix.











S2 Fig. Transmembrane helices of MAPKKKs of Raf subfamily in *C. sinensis*. TMHMM Server, v.2.0 (http://www.cbs.dtu.dk/services/TMHMM/), was used to predict the presence of transmembrane helices. The plot shows the posterior probabilities of inside/outside/TM helix. The plot is obtained by calculating the total probability that a residue sits in helix, inside, or outside summed over all possible paths through the model. The red peaks represent the possible transmembrane helices present. If the red peak crosses the transmembrane line (represented in pink colour), then its calculated as 1 transmembrane helix.





S3 Fig. Transmembrane helices of MAPKKKs of ZIK subfamily in *C. sinensis***.** TMHMM Server, v.2.0 (http://www.cbs.dtu.dk/services/TMHMM/), was used to predict the presence of transmembrane helices. The plot shows the posterior probabilities of inside/outside/TM helix. The plot is obtained by calculating the total probability that a residue sits in helix, inside, or outside summed over all possible paths through the model. The red peaks represent the possible transmembrane helices present. If the red peak crosses the transmembrane line (represented in pink colour), then its calculated as 1 transmembrane helix.



S4 Fig. Motif logos of the 10 identified motifs in 59 MAPKKKs of *C. sinensis***.** The motif logos were generated by MEME suite.



S5 Fig. The ds/dn cumulative graph of MAPKKKs of MEKK subfamily in *C. sinensis*. SNAP server (https://www.hiv.lanl.gov/content/sequence/SNAP/SNAP.html) has been used to generate the graph.



S6 Fig. The ds/dn cumulative graph of MAPKKKs of Raf subfamily in *C. sinensis*. SNAP server (https://www.hiv.lanl.gov/content/sequence/SNAP/SNAP.html) has been used to generate the graph.



S7 Fig. The ds/dn cumulative graph of MAPKKKs of ZIK subfamily in *C. sinensis*. SNAP server (https://www.hiv.lanl.gov/content/sequence/SNAP/SNAP.html) has been used to generate the graph.



S8 Fig. GO analysis of all the 59 MAPKKKs in *C. sinensis***.** The results have been grouped into three main categories: Biological Process, Cellular Component and Molecular function. The y-axis represents the frequency of genes while the x-axis represents the putative functions.



S9 Fig. Heat maps for tissue-specific expression patterns of A) MEKK; B) Raf and; C) ZIK genes in *C. sinensis.* The relative expression of these genes were analysed in different tissues, by using GraphPad Prism 8 software. The level of expression was in transcript per million (TPM). The 8 different tissues are represented on the right and the tea genes are marked below.



S10 Fig. Heat maps for cold stress expression patterns of A) MEKK; B) Raf and; C) ZIK genes in *C. sinensis*. The relative expression of these genes were analysed in different stages, by using GraphPad Prism 8 software. The level of expression was in transcript per million (TPM). The different stages are represented on the right and the tea genes are marked below.



S11 Fig. Heat maps for drought stress expression patterns of A) MEKK; B) Raf and; C) ZIK genes in *C. sinensis.* The relative expression of these genes were analysed in different stages, by using GraphPad Prism 8 software. The level of expression was in transcript per million (TPM). The different stages are represented on the right and the tea genes are marked below.



S12 Fig. Heat maps for salt stress expression patterns of (A) MEKK (B) Raf and (C) ZIK genes in *C. sinensis*. The relative expression of these genes were analysed in different stages, by using GraphPad Prism 8 software. The level of expression was in transcript per million (TPM). The different stages are represented on the right and the tea genes are marked below.



S13 Fig. Heat maps for MeJA treatment expression patterns of (A) MEKK (B) Raf and (C) ZIK genes in *C. sinensis*. The relative expression of these genes were analysed in different stages, by using GraphPad Prism 8 software. The level of expression was in transcript per million (TPM). The different stages are represented on the right and the tea genes are marked below.