Genome-wide identification of *GhAAI* genes reveals that *GhAAI66* triggers a phase transition to induce early flowering

Ghulam Qanmber[1,*], Lili Lu[1,*], Zhao Liu[1*], Daoqian Yu[1], Kehai Zhou[1], Peng Huo[1], Fuguang Li[1,2],† and Zuoren Yang[1,2],†

1 State Key Laboratory of Cotton Biology, Key Laboratory of Biological and Genetic Breeding of Cotton, Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang, 455000, Henan, China

2 Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University, Zhengzhou, 4550001, Henan, China;

* These authors contributed equally to this work.

† **Corresponding authors**

Zuoren Yang: yangzuoren4012@163.com

Fuguang Li: aylifug@163.com

Author's information:

Ghulam Qanmber: gqkhan12@gmail.com

Lili Lu: biolll@126.com

Zhao Liu: liuzhaocaas@163.com

Daoqian Yu: yudaoqian88@163.com

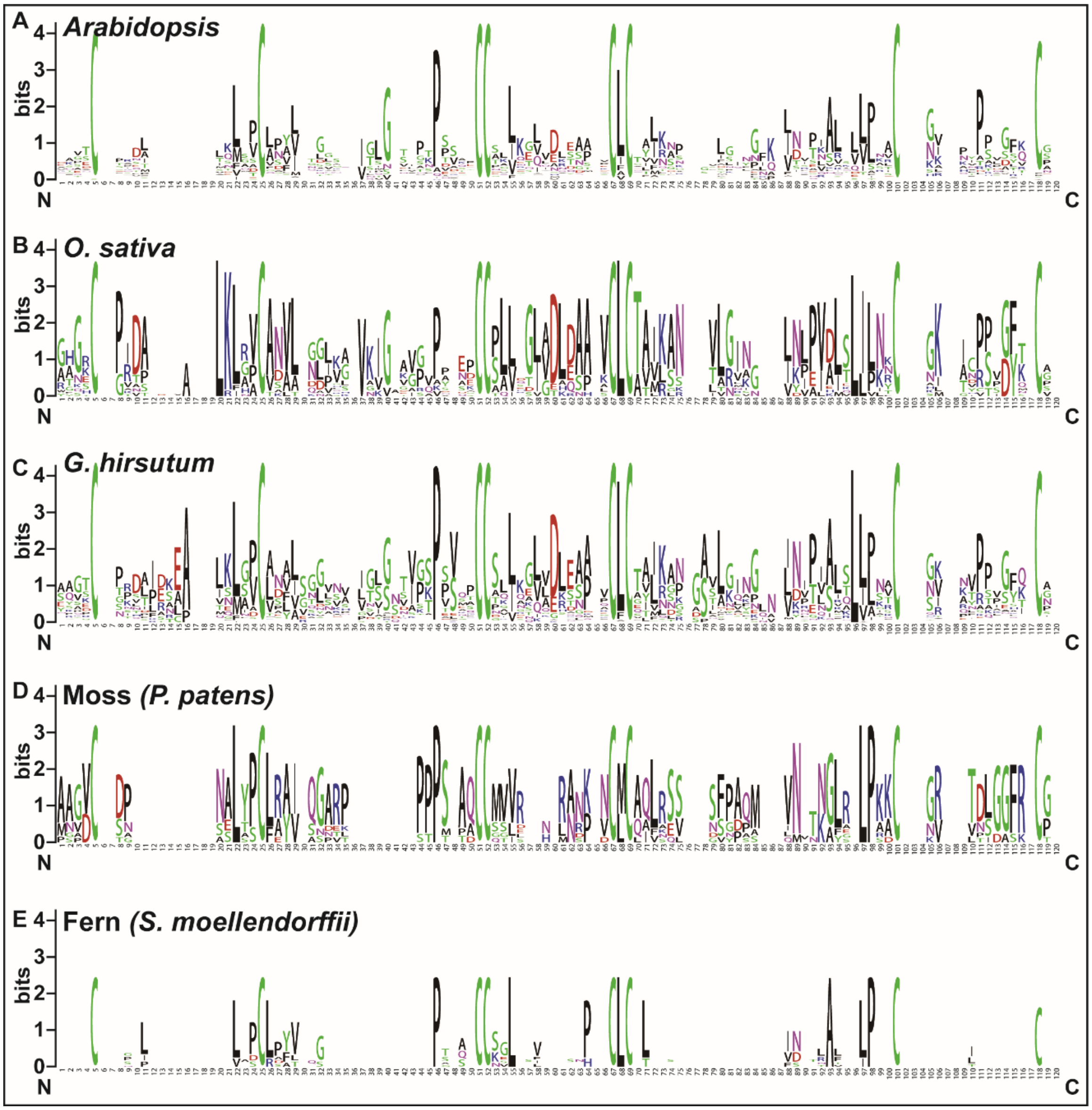
Kehai Zhou: zkhlxy@163.com

Peng Huo: hp20080224@126.com

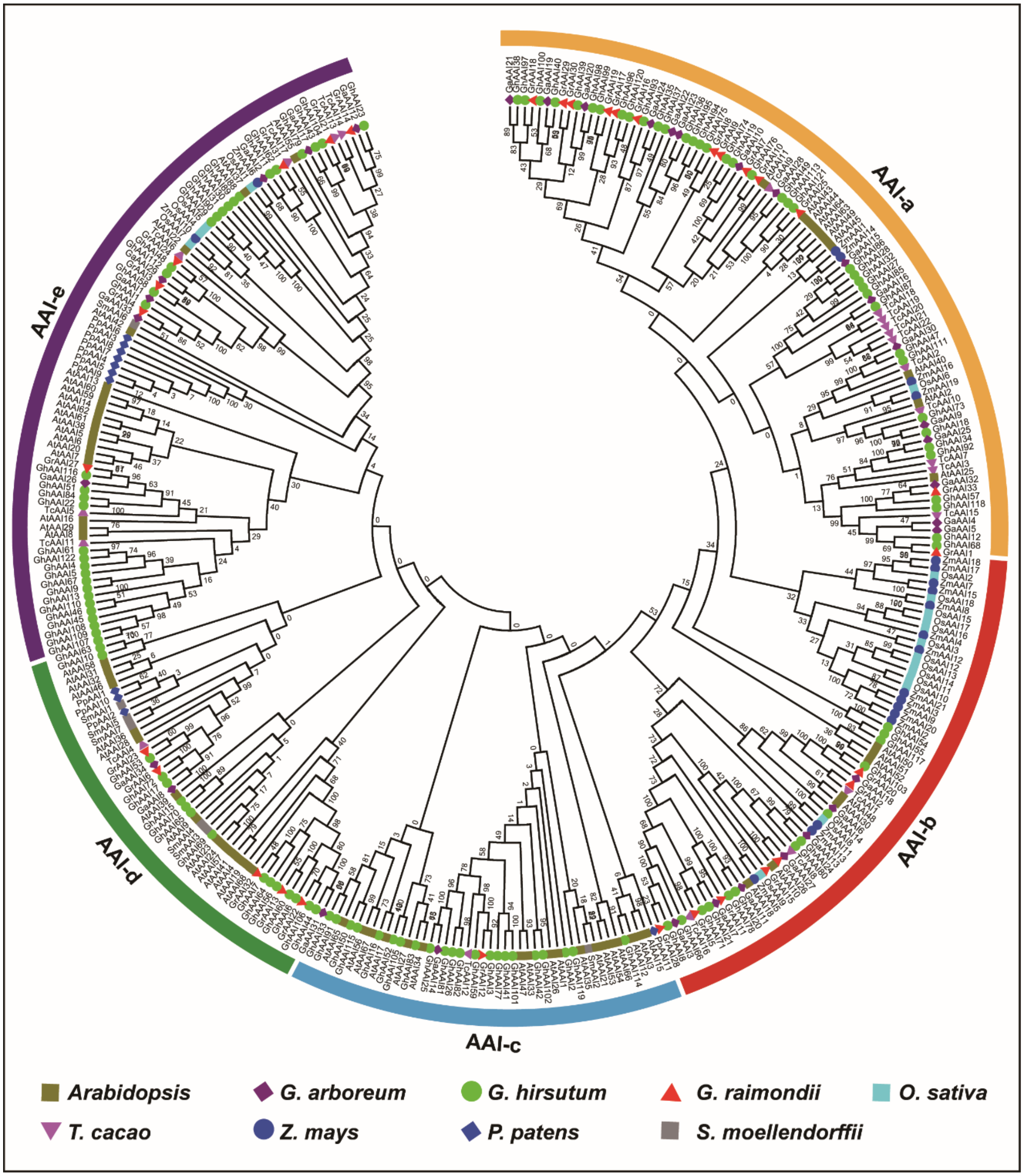
Fuguang Li: aylifug@163.com

Zuoren Yang: yangzuoren4012@163.com

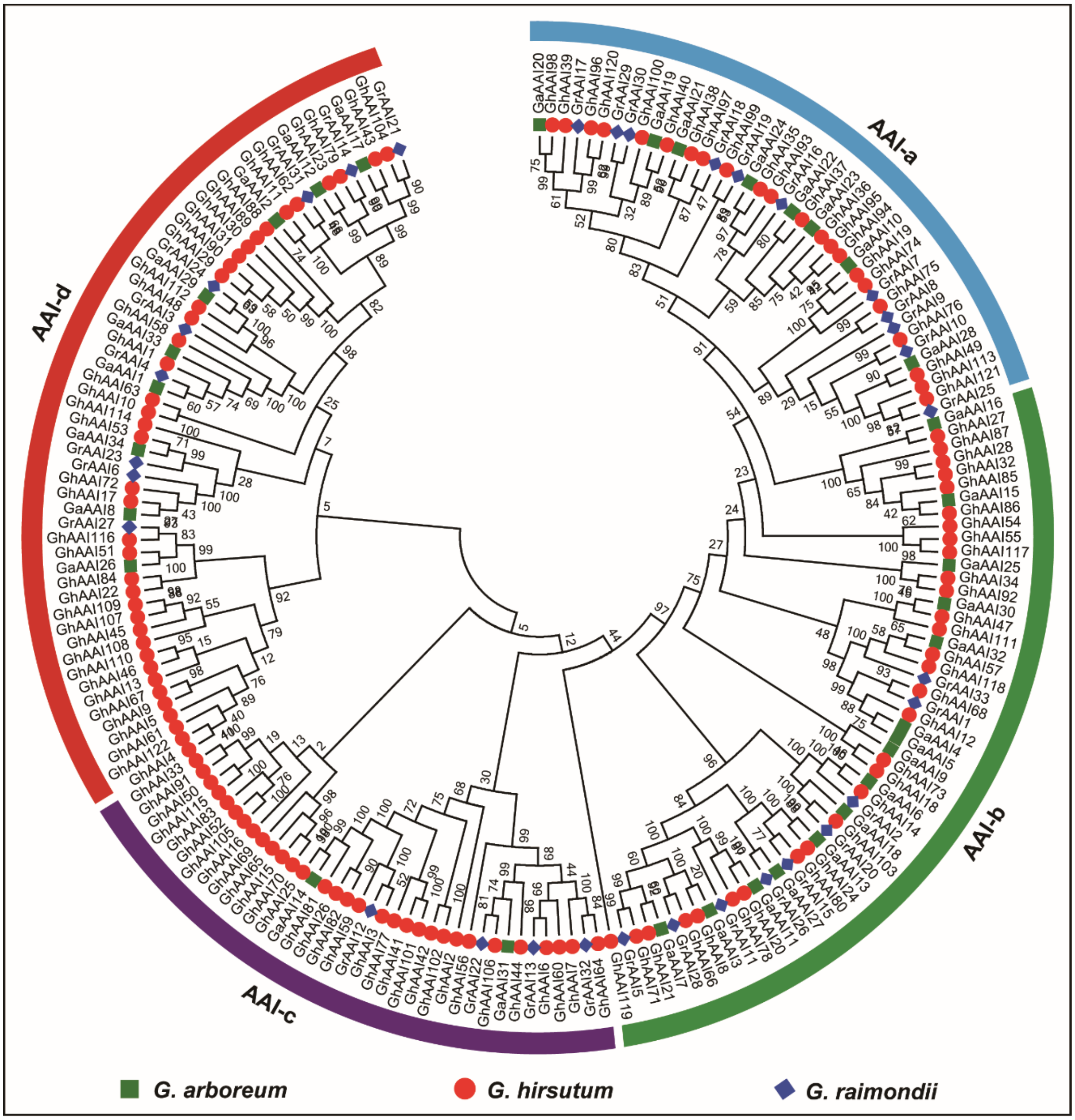
Supplementary Fig. S1. Conserved amino acid residue analysis among (A) Arabidopsis, (B) *O. sativa*, (C) *G. hirsutum*, (D) moss, and (E) fern *AAI* genes.



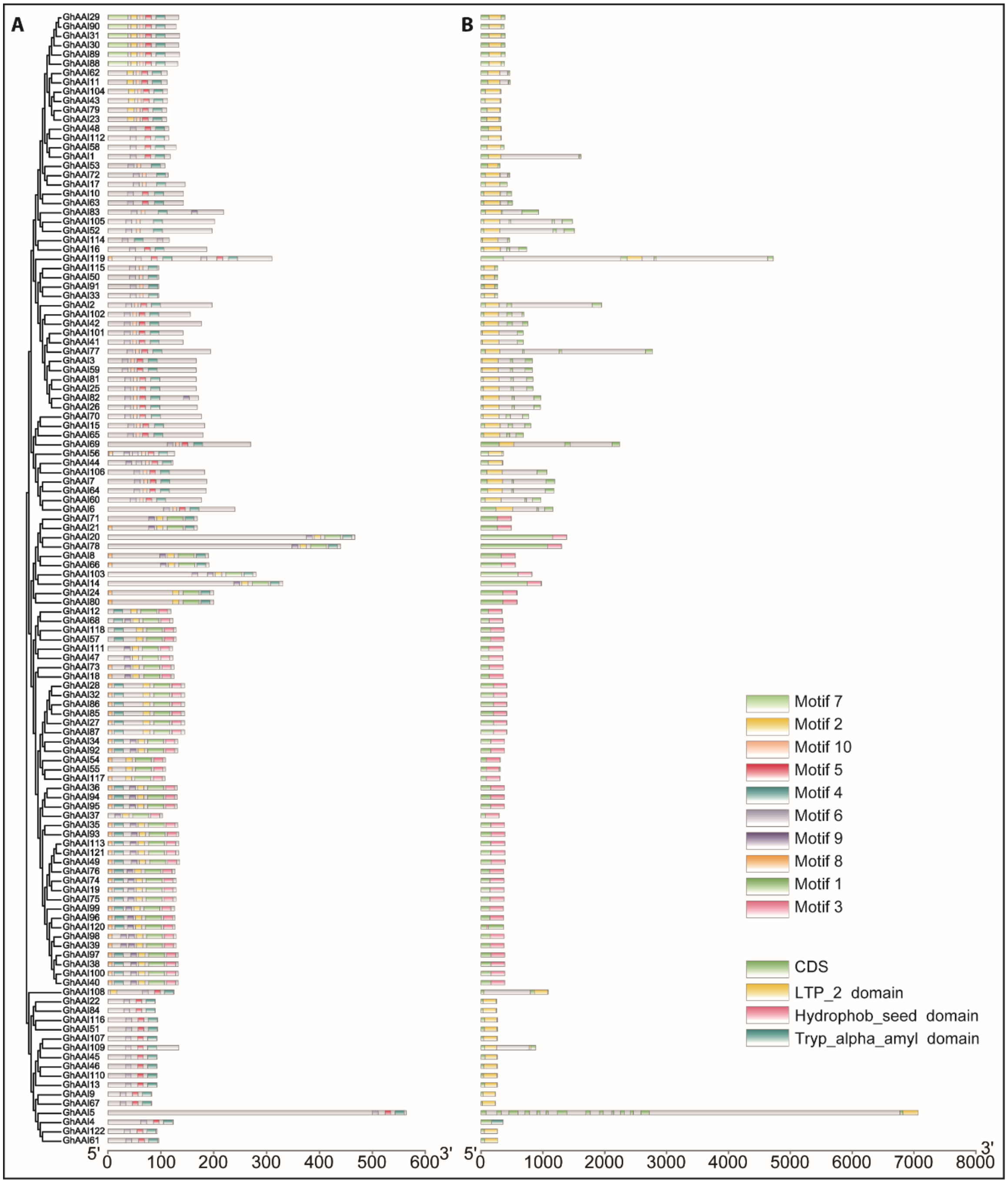
Supplementary Fig. S2. Phylogenetic tree constructed using ME. The phylogenetic tree divided all *AAI* genes from monocots, dicots, moss, and ferns into five major groups from AAI-a to AAI-e. The prefixes At, Ga, Gh, Gr, Tc, Os, Zm, Pp, and Sm were used before the names of *A. thaliana*, *G. arboreum*, *G. hirsutum*, *G. raimondii*, *T. cacao*, *O. sativa*, *Z. mays*, *P. patens* and *S. moellendorffii AAI* genes, respectively. Bootstrap values are noted near nodes of each branch.



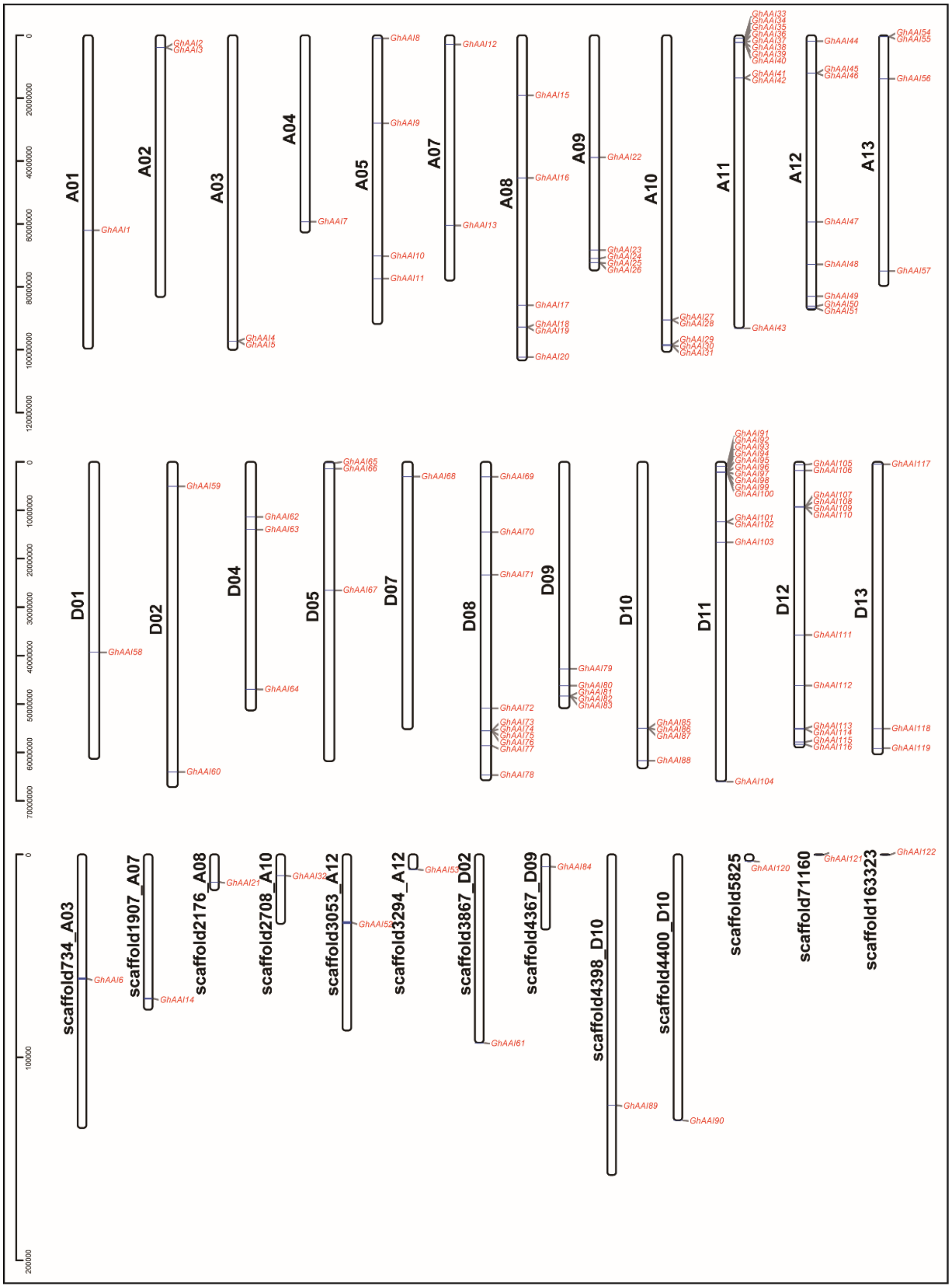
Supplementary Fig. S3. The phylogenetic tree was generated using the NJ method for cotton including *G. arboreum*, *G. hirsutum*, and *G. raimondii AAI* genes in order to estimate common ancestor hypothesis. Tree divided all cotton *AAI* genes into four groups represented by different colors. Bootstrap values are noted near nodes of each branch.



Supplementary Fig. S4. Protein motif distribution and gene structure (exon/intron) analysis of *GhAAI* genes along with phylogenetic tree inferred using the NJ method. (A) Protein motif distribution in *GhAAI* genes. (B) Gene structure (exon/intron) and encoded protein prediction.



Supplementary Fig. S5. Chromosomal location of *GhAAI* genes on different chromosomes. A01 to A13 and D01 to D13 represent At and Dt sub-genomes of *G. hirsutum* chromosomes, respectively. Some *GhAAI* genes' locations were allotted on scaffolds. Chromosomal sizes were calculated from published genome data.



Supplementary Fig. S6. Expression pattern of *GhAAI* genes in different cotton tissues. Heat map of 122 *GhAAI* genes in 22 different tissues. Data was obtained from publicly available transcriptomic data and color bar (down) represents expression level.