Plant Communications
Resource article

CellPress
Partner Journal

CRISPR–Cas9-mediated construction of a cotton *CDPK* mutant library for identification of insect-resistance genes

Fuqiu Wang[1,5], Sijia Liang[2,5], Guanying Wang[1], Tianyu Hu[1], Chunyang Fu[1], Qiongqiong Wang[1], Zhongping Xu[1], Yibo Fan[1], Lianlian Che[1], Ling Min[1], Bo Li[3,*], Lu Long[4,*], Wei Gao[4,*], Xianlong Zhang[1] and Shuangxia Jin[1,*]

[1]Hubei Hongshan Laboratory, National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China

[2]Academy of Industry Innovation and Development, Huanghuai University, Zhumadian, Henan 463000, China

[3]Xinjiang Key Laboratory of Crop Biotechnology, Institute of Nuclear and Biological Technology, Xinjiang Academy of Agricultural Sciences, Urumqi 830091 Xinjiang, China

[4]National Key Laboratory of Cotton Bio-breeding and Integrated Utilization, School of Life Science, Henan University, Henan 475004, China

[5]These authors contributed equally to this article.

***Correspondence: Bo Li (libo@xaas.ac.cn), Lu Long (lulong1826@163.com), Wei Gao (gaowei021@163.com), Shuangxia Jin (jsx@mail.hzau.edu.cn)**

https://doi.org/10.1016/j.xplc.2024.101047

ABSTRACT

Calcium-dependent protein kinases (CDPKs) act as key signal transduction enzymes in plants, especially in response to diverse stresses, including herbivory. In this study, a comprehensive analysis of the *CDPK* gene family in upland cotton revealed that *GhCPKs* are widely expressed in multiple cotton tissues and respond positively to various biotic and abiotic stresses. We developed a strategy for screening insect-resistance genes from a CRISPR–Cas9 mutant library of *GhCPKs*. The library was created using 246 single-guide RNAs targeting the *GhCPK* gene family to generate 518 independent T0 plants. The average target-gene coverage was 86.18%, the genome editing rate was 89.49%, and the editing heritability was 82%. An insect bioassay in the field led to identification of 14 *GhCPK* mutants that are resistant or susceptible to insects. The mutant that showed the clearest insect resistance, *cpk33/74* (in which the homologous genes *GhCPK33* and *GhCPK74* were knocked out), was selected for further study. Oral secretions from *Spodoptera litura* induced a rapid influx of Ca^{2+} in *cpk33/74* leaves, resulting in a significant increase in jasmonic acid content. S-adenosylmethionine synthase is an important protein involved in plant stress response, and protein interaction experiments provided evidence for interactions of GhCPK33 and GhCPK74 with GhSAMS1 and GhSAM2. In addition, virus-induced gene silencing of *GhSAMS1* and *GhSAM2* in cotton impaired defense against *S. litura*. This study demonstrates an effective strategy for constructing a mutant library of a gene family in a polyploid plant species and offers valuable insights into the role of CDPKs in the interaction between plants and herbivorous insects.

Key words: cotton, *CDPKs*, mutant library, CRISPR–Cas9, Ca^{2+} influx, insect resistance

Wang F., Liang S., Wang G., Hu T., Fu C., Wang Q., Xu Z., Fan Y., Che L., Min L., Li B., Long L., Gao W., Zhang X., and Jin S. (2024). CRISPR–Cas9-mediated construction of a cotton *CDPK* mutant library for identification of insect-resistance genes. Plant Comm. **5**, 101047.

INTRODUCTION

Cotton is a vital source of fiber, oil, and feed, making a significant contribution to global development (Chen et al., 2007). Throughout its growth period, cotton is constantly threatened by various herbivorous insects, which seriously affect yield and fiber quality. Insect pests have emerged as factors limiting the development of the cotton industry (Wu and Guo, 2005; Li et al., 2020; Quan and Wu, 2023). The traditional strategy of molecular breeding for insect resistance in cotton involves the introduction of exogenous genes encoding *Bacillus*

Published by the Plant Communications Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

thuringiensis insecticidal protein into the cotton genome to enhance its insect resistance (Wu et al., 2008; Huang et al., 2013; Qiao et al., 2017). Although this method temporarily alleviates the pressure caused by insect pests, issues such as the evolution of target pest resistance and outbreaks of non-target secondary pests have gradually emerged with the long-term cultivation of *B. thuringiensis* cotton (Lu et al., 2010; Heckel, 2012). There is therefore a pressing need to develop new strategies to address insect pest outbreaks.

Over millions of years, the selective pressure exerted by herbivorous insects has led to the evolution of effective defense systems in plants (Erb and Reymond, 2019). Plants respond to herbivorous insect attacks by initiating early signaling events, including depolarization of plasma transmembrane potential (V_m), increases in cytoplasmic Ca^{2+}, generation of reactive oxygen species, and activation of mitogen-activated protein kinases. These pathways play central and conserved roles in enhancing broad-spectrum insect resistance in plants (Howe and Jander, 2008; Erb et al., 2012). Following the influx of Ca^{2+}, Ca^{2+} sensors decode temporal and spatial signal transmission patterns to regulate metabolism and gene expression (Sanders et al., 2002). Typical Ca^{2+} sensors include calmodulin, calmodulin-like proteins, calcineurin B-like proteins, and calcium-dependent protein kinases (CDPKs). Among these, CDPK is the only sensor that combines sensing activity through a EF-hand structure with response activity through a protein kinase structure (Yip Delormel and Boudsocq, 2019). CDPK comprises four domains: the N-terminal variable domain, the Ser/Thr kinase domain, the inhibitory-junction domain, and the calmodulin-like domain (including an EF-hand calcium-binding site) (Batistic and Kudla, 2012). Its distinctive structure enables it to function concurrently as a Ca^{2+} receptor and an effector (Zuo et al., 2013). This specificity has generated particular research interest.

As an important class of Ser/Thr protein kinase genes, *CDPKs* form a large multigene family that is widely distributed in plants, green algae, protozoa, and oomycetes (Valmonte et al., 2014). *CDPKs* can be expressed across multiple tissues, including roots, stems, leaves, flowers, fruits, and seeds (Romeis et al., 2001; Ye et al., 2009). They can also localize to various cellular compartments, including the cytoplasm, nucleus, vacuolar membrane, endoplasmic reticulum, mitochondria, chloroplasts, and peroxisomes (Simeunovic et al., 2016). The wide distribution of CDPKs in plant tissues and multiple organelles indicates that they may have important functional roles in various biological processes of plants. For example, CDPKs play a significant role in the plant signal-transduction network in response to attack by herbivorous insects. In *Arabidopsis*, *AtCPK3/13* regulate the transcript level of *PDF1.2* by phosphorylating HsFB2a, thereby enhancing resistance against *Spodoptera littoralis* (Kanchiswamy et al., 2010). Simultaneous silencing of *NaCPK4* and *NaCPK5* results in significant accumulation of jasmonic acid (JA) in *Nicotiana attenuata*, enhancing its resistance against *Manduca sexta* (Yang et al., 2012). Knockout of *GmCPK38*, a soybean homolog of *NaCPK4/5*, alters numerous resistance-related phosphorylated proteins, genes, and metabolites, thereby enhancing soybean resistance to the common cutworm (Li et al., 2022). In addition, *GmCPK17* has been shown to positively regulate the resistance of soybean hairy roots to the common cutworm (Wang et al., 2022). Various CDPKs from different plant species thus mediate responses to herbivorous insects by regulating gene expression or modulating hormone biosynthesis, exerting positive or negative effects.

The emergence of genome-editing tools has provided a tremendous opportunity for deciphering gene function, and the development of genome sequencing has accelerated the use of gene-editing tools (Pan et al., 2023). Because of its simple design, CRISPR–Cas9 can target the majority of DNA sequence fragments in the genome, enabling the rapid construction of large-scale knockout mutant libraries for forward genetic screening. Therefore, high-throughput loss-of-function screening of mutant libraries constructed using the CRISPR–Cas9 gene-editing system has been rapidly applied to plants (Shalem et al., 2014; Wang et al., 2018; Gaillochet et al., 2021). Lu et al. developed a pooled method for induction of genome-wide gene mutations in rice using a single-guide RNA (sgRNA) library (Lu et al., 2017). They constructed a pooled library of 88 541 sgRNAs targeting 34 234 genes, ultimately producing 91 004 T0 plants. Using CRISPR–Cas9, they performed comprehensive functional screening across the entire rice genome, producing a valuable resource for rice research and breeding. Small-scale functional genomic screening based on CRISPR mutant libraries has also been performed in tomato (*Solanum lycopersicum*), cotton (*Gossypium hirsutum*), soybean (*Glycine max*), and maize (*Zea mays*) through construction of libraries containing 54, 512, 70, and 1368 sgRNAs, respectively (Jacobs et al., 2017; Bai et al., 2020; Liu et al., 2020; Sun et al., 2024). These groundbreaking studies have independently demonstrated the utility of pooled CRISPR screening in plants.

CDPKs play a key role in plant defense against herbivorous insects as early signal-transduction genes (Romeis and Herde, 2014). We therefore developed a new approach to screen insect-resistance genes from the *CDPK* gene family in upland cotton using a CRISPR–Cas9 mutant library. First, through comprehensive analysis of the *GhCPK* gene family, we identified 82 *GhCPKs* in the upland cotton genome and predicted that this gene family was involved in multiple biological processes in cotton, laying the foundation for functional analysis of the *GhCPK* family. Next, we constructed a CRISPR–Cas9 mutant library for the *GhCPK* gene family. Through field phenotyping and insect-resistance experiments, we screened the insect-resistant *cpk33/74* cotton lines from the mutant library, and we performed preliminary investigations of the insect-resistance mechanisms of the screened candidate genes, *GhCPK33* and *GhCPK74*. This study provides valuable insights into the role of CDPKs in plant–insect interactions. In addition, the numerous mutants developed here provide a valuable genetic resource for functional study of the CDPKs and creation of pest-resistant cotton germplasms.

RESULTS

Comprehensive analysis of the *CDPK* gene family in upland cotton

A total of 82 *GhCPK* genes were identified in the upland cotton genome, and genomic information and physicochemical properties of all GhCPK members were summarized, including gene names, chromosomal locations, amino acid residues (which

varied from 331 to 907), relative molecular weights (37.40 to 101.03 kDa), and isoelectric points (4.68 to 9.48) (Supplemental Table 1). N-myristoylation and S-palmitoylation are important post-translational protein modifications; among the 82 predicted *GhCPKs*, 35 had N-terminal myristoylation sites, and 75 had at least one palmitoylation site. Predictions of transmembrane domains, signal peptides, and EF-hand structures revealed that all GhCPK proteins lack transmembrane domains, most lack signal peptides, and 78 have four EF-hands, whereas four have three EF-hands (Supplemental Table 1).

To examine the evolutionary relationships among the *CDPK* genes, we constructed an unrooted phylogenetic tree by aligning the full-length protein sequences of 34 AtCPKs, 31 OsCPKs, and 82 GhCPKs (Supplemental Figure 1). All CDPK proteins from the three species clustered into four distinct subgroups (groups 1–4), implying that these subgroups existed prior to the differentiation of monocots and dicots. We used the MEME website to predict 10 conserved motifs in the GhCPK proteins (Supplemental Figure 2). Motifs 1, 3, and 5 were present in all GhCPK proteins except GhCPK45 and GhCPK50. Motif 4 was present in all proteins except GhCPK50 and GhCPK82. There were also differences in protein motifs among different subgroups; for example, motif 7 was unique to group 3, and all members of group 4 lacked motif 6. GhCPKs from the same subgroup possessed similar numbers, types, and spatial distributions of motifs, supporting their close evolutionary relationships. Differences in motifs among different subgroups imply that CDPKs have diverse functions in upland cotton. An analysis of gene structure is shown in Supplemental Figure 2; despite variation in exon numbers (7 to 16) among family members, *GhCPKs* from the same subgroup exhibited very similar exon–intron structures. Most members of group 1 had seven exons, and all members of group 4 had 12 exons. Conservation of these gene structures among members of each subgroup lends support to their close evolutionary relationships.

Collinearity analysis of the *GhCPKs* was performed with MCScanX and Circos software to investigate their mechanism of amplification (Supplemental Figure 3). A total of 128 collinear gene pairs were identified, each composed of two genes located on distinct chromosomes; because these pairs were not physically adjacent, segmental duplication appears to have been the main amplification mechanism for the *GhCPK* gene family. Collinearity analyses of the *G. hirsutum GhCPKs* with *CPKs of Arabidopsis thaliana*, *Oryza sativa*, and *Gossypium raimondii* revealed 174 pairs of orthologous genes between *G. hirsutum* and *G. raimondii*, 84 pairs between *G. hirsutum and A. thaliana*, and only 13 pairs between *G. hirsutum* and *O. sativa* (Supplemental Figure 4). These results suggest a closer evolutionary relationship of *G. hirsutum* to other dicots. To examine the potential functions and regulatory mechanisms of the *GhCPK* family, we analyzed the *cis*-acting elements in the 2-kb promoter sequences of all *GhCPKs* (Supplemental Figure 5). A total of 29 elements were identified and classified into four categories based on their functions: elements related to plant hormones, plant growth and development, defense response, and light response. These results indicate that CDPKs may function in diverse biological processes, indirectly reflecting the importance of this gene family in upland cotton.

GhCPKs respond to multiple stresses at the transcriptional level

Gene expression is closely related to gene function. To explore the potential roles of the *GhCPKs*, the expression patterns of the 82 *GhCPK* genes were analyzed in the context of their phylogenetic relationships (Figure 1A). Most *GhCPKs* were expressed in multiple tissues, although *GhCPK55* was specifically expressed in petals and *GhCPK64* was specifically expressed in stems. All *GhCPKs* in group 4 showed relatively high expression in all tissues (Figure 1B). Expression of the cotton *GhCPKs* under four abiotic stress conditions (heat, salt, cold, and drought) was also analyzed (Figure 1C). A total of 73% of the *GhCPKs* showed significantly decreased expression after 1 h of high-temperature treatment (38°C) compared with the control group. Most *GhCPKs* demonstrated continuous downregulation in response to low-temperature treatment (4°C), particularly after 12 h. We next analyzed expression of the *GhCPKs* in cotton leaves after exposure to the oral secretions (OS) of *Helicoverpa armigera* and *S. litura* (Figure 1D). At 10 and 30 min after treatment with *H. armigera or S. litura* OS, at least 73% of the *GhCPKs* showed upregulated expression, which peaked at 30 min after treatment, and all *GhCPKs in* group 4 showed marked upregulation at both time points. However, after 240 min of treatment, the induced differential gene expression was no longer evident. We also studied the expression of *GhCPKs* at different time points in two cotton cultivars (exhibiting either strong resistance [HR] or sensitivity [ZS] to whitefly) upon whitefly infection (Figure 1E). The differential expression of *GhCPKs* in cotton leaves was most pronounced after 48 h of whitefly infection. In the strongly resistant cultivar, 54% of *GhCPKs* were significantly induced.

In summary, *GhCPKs* are widely expressed in different tissues and show significant changes in expression under various external stresses. These findings imply that the CDPK gene family is involved in multiple biological processes during cotton's growth, development, and response to adverse environmental conditions, making it an important gene group for upland cotton.

Construction of a mutant library for the upland cotton *CDPK* gene family

To better understand the functional roles of GhCPKs in cotton resistance to insects, we developed an insect-resistance screening method for the *GhCPKs* based on a CRISPR–Cas9 mutant library. First, we designed 246 sgRNAs targeting the *GhCPKs* (Supplemental Table 2); these sgRNAs were classified into two types (Supplemental Figure 6A). The first type was designed to target individual *GhCPKs*; in this case, two different sgRNAs were designed to target an exon in a specific *GhCPK* (Figure 2A). Because of possible functional redundancy among genes of allotetraploid cotton, we also designed a second type of sgRNA to target homologous *GhCPKs* in the A and D subgenomes. In this case, two sgRNAs were designed to target the exon region in a pair of homologous genes, simultaneously mutating the two homologous *GhCPKs* but not other genes (Figure 2B).

To rapidly and efficiently obtain a mutant library of the *GhCPKs*, we used a mutant library construction method based on primer mixing pools (Figure 2C). We designed 246 primers based on the sgRNA sequences and constructed a mixed vector library using six primer

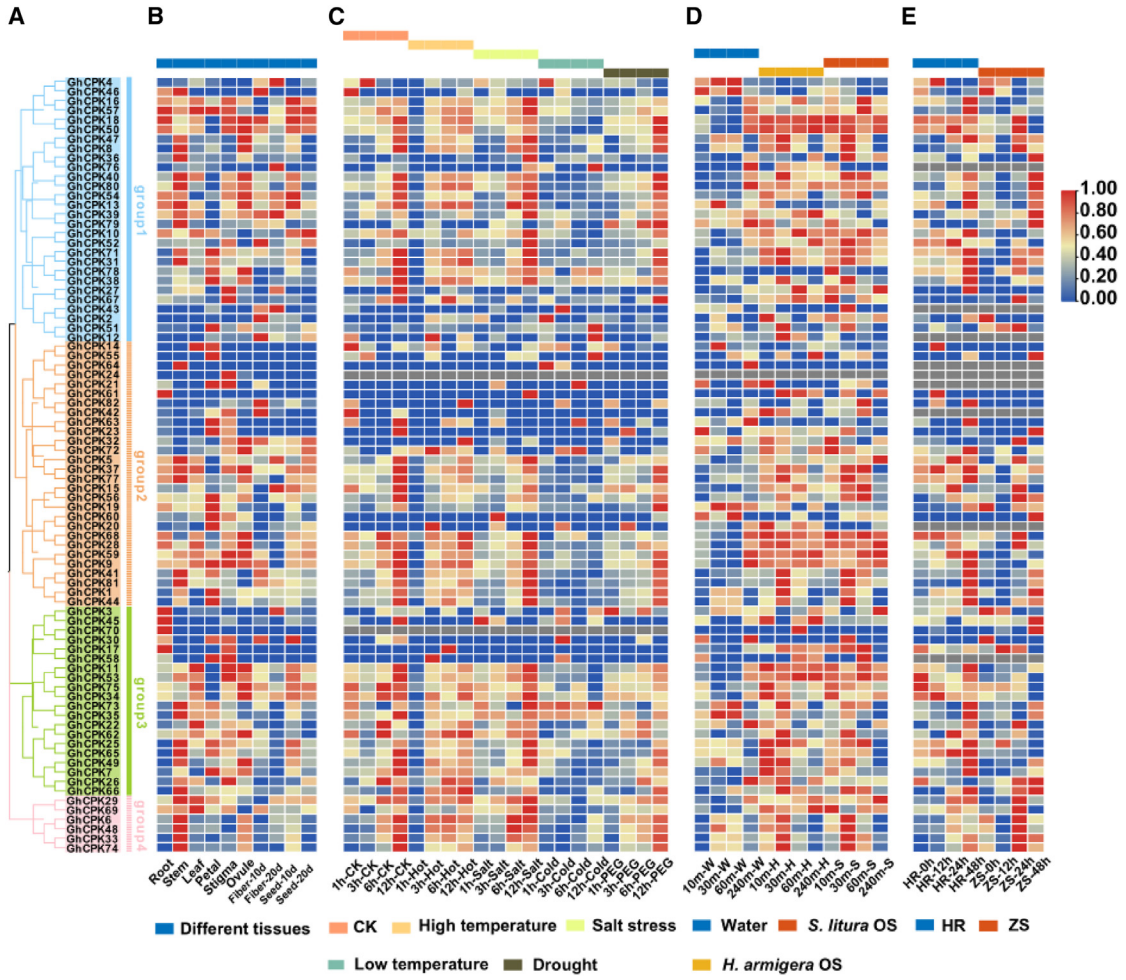

Figure 1. Expression patterns of the *GhCPK* gene family in different tissues and under different stress conditions.

(A) Phylogenetic tree of the *GhCPK* gene family. Different colors represent different subgroups.

(B) Expression profiles of *GhCPKs* in different tissues. From left to right are root, stem, leaf, petal, anther, stigma, ovule, fiber 10 days, fiber 20 days, seed 10 days, and seed 20 days.

(C) Expression profiles of *GhCPKs* induced by four types of abiotic stress: heat, cold, polyethylene glycol (PEG), and salt. The induction times for each type of stress were 1, 3, 6, and 12 h.

(D) Expression patterns of *GhCPK* genes after exposure to OS of *H. armigera* and *S. litura*. Four time points after processing were selected: 10, 30, 60, and 240 min.

(E) Expression patterns of the *GhCPK* gene family at different time points (0, 12, 24, and 48 h after infection) in two cotton cultivars (with strong resistance [HR] or sensitivity [ZS] to whitefly) infected by whitefly. In all the expression profiles mentioned above, red indicates increased expression, and blue indicates decreased expression.

mixing pools (Supplemental Table 3). The mixed-vector library constructed with primer mixing pool 1 was used to validate the feasibility of the method (Supplemental Table 4). By sequencing 100 collected monoclonal clones, we found that 36 sgRNAs were detected, with a coverage rate of 90% (Supplemental Figure 6B). After matching sgRNAs with genes, we found that the 36 sgRNAs could target all 20 genes, achieving a gene coverage of 100% (Supplemental Figure 6C). The majority of sgRNAs were detected two or three times (Supplemental Figure 6D). These results indicate that the method of constructing a mixed-vector library through primer mixing pools is feasible. The plasmids extracted from the mixed-vector library were transferred into *Agrobacterium tumefaciens* for stable transformation of cotton to produce genetically edited plants (Figure 2D), and we obtained 518 regenerated plants in the T0 generation. The method of constructing a CRISPR–Cas9 mutant library based on a primer mixing pool saves significant time and effort, enabling the rapid acquisition of a substantial number of mutant plants.

Molecular detection for the *GhCPK* mutant library

Molecular detection was performed to study the gene-editing status of each plant in the mutant library using the specific detection procedure shown in Figure 2E. Barcode high-throughput sequencing enabled us to quickly match the sgRNA sequences carried by each plant, accurately identifying the edited target gene(s). Among the 518 plants in the mutant library, 16.02% had no detected sgRNAs, 81.08% had one detected sgRNA, and 2.89% had two or three detected sgRNAs. Thus, only one (or a

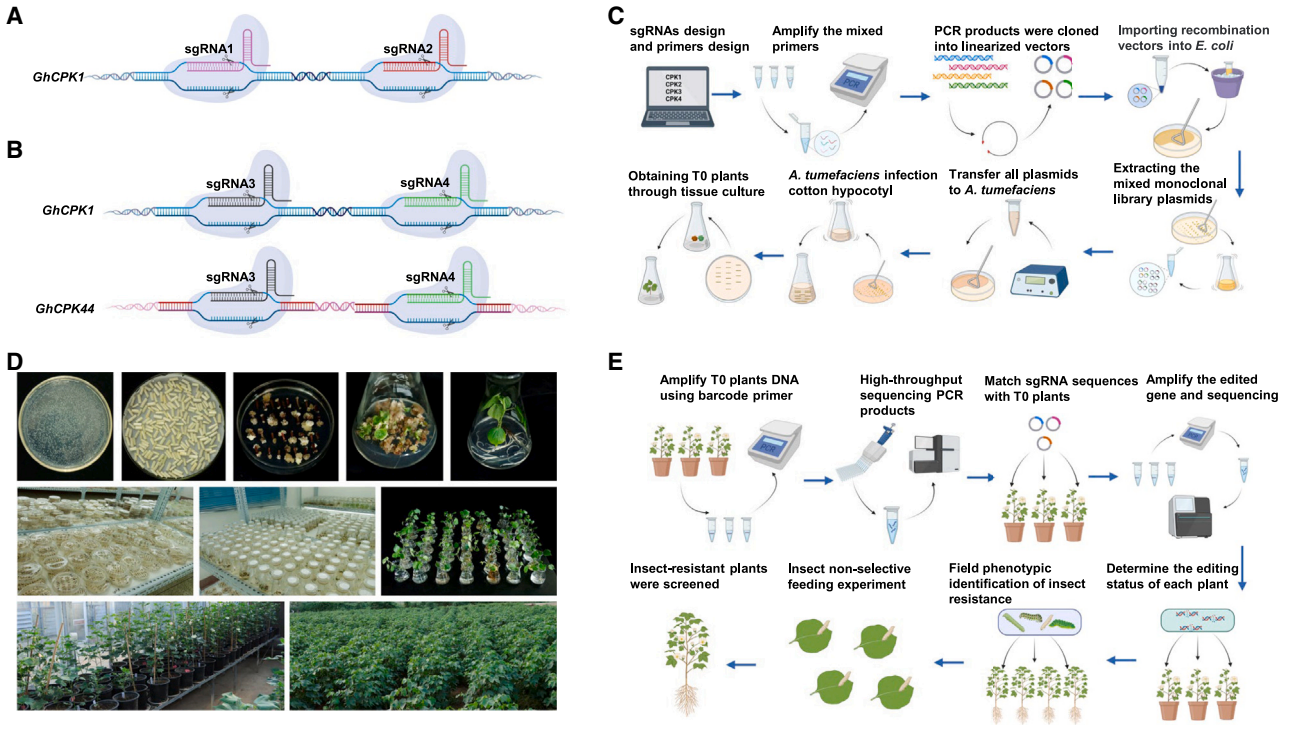

Figure 2. Construction and detection procedures for the mutant library of the *GhCPK* gene family.

(A) Schematic diagram of sgRNA design for knockout of a single *GhCPK*. SgRNA1 and sgRNA2 target *GhCPK1* individually and exclusively.

(B) Schematic diagram of sgRNA design for knockout of homologous *GhCPKs*. SgRNA3 and sgRNA4 simultaneously target both *GhCPK1* and *GhCPK44* and exclusively target only *GhCPK1* and *GhCPK44*.

(C) Procedure for construction of a mutant library for the *CDPK* gene family in upland cotton.

(D) The process of cotton genetic transformation.

(E) Molecular detection and insect-resistance screening procedure for the *GhCPK* mutant library.

pair of) *GhCPKs* had been edited in the majority of T0 plants (Supplemental Figure 7B). Only T0 plants carrying a single sgRNA were selected for further investigation. A total of 147 sgRNAs targeting 73 individual *GhCPKs* and 33 pairs of homologous *GhCPKs* were detected. The average target gene coverage was 86.18% (Figure 3A). In general, on the basis of sgRNA detection in the T0-generation plants, we successfully obtained gene-edited materials that met our needs for establishment of a CRISPR–Cas9 mutant library of the *CDPK* gene family in upland cotton. This result implied that construction of a large-scale mutant library using a mixed-vector library is feasible in upland cotton.

Next, primers were designed for each mutant plant based on the position of the sgRNA in the target gene (Supplemental Table 6). By analyzing the amplified sequences, the specific editing status of each gene could be determined. Because upland cotton is an allopolyploid, various types of fragments appear in the amplified sequences after editing. Traditional Sanger sequencing cannot meet the requirements for such complex amplifications. Therefore, high-throughput sequencing was performed. To verify the accuracy of this method, two randomly selected gene-edited plants were subjected to both high-throughput sequencing and Sanger sequencing (Supplemental Figure 7C). The results of the two sequencing methods were consistent, indicating that high-throughput sequencing is suitable for the detection of gene-edited cotton. By analyzing the high-throughput sequencing results of 390 T0-generation plants, we found that 349 plants (89.49%) had undergone effective editing (editing efficiency $\geq$10%) (Supplemental Figure 7D). Among them, 273 plants (78.22%) showed highly efficient editing (editing efficiency $\geq$80%) (Figure 3B). Analysis of editing patterns in T0 plants revealed that the number of editing types varied from one to more than 10 for different plants (Figure 3C). *cpk37-3*, *cpk36/76-1*, and *cpk6/48-1* demonstrated one, two, and more than 10 editing types, respectively (Supplemental Figure 7E). The editing types included deletions, insertions, and substitutions. *cpk45-3* exhibited all three types of mutation simultaneously (Supplemental Figure 7F). Deletions (81.80%) accounted for the largest proportion in all editing types, followed by insertions (12.04%) (Figure 3D). The majority of deletions consisted of one to five missing bases (Figure 3E), whereas most insertion and substitution events consisted of one- or two-base insertions or substitutions (Figure 3F and 3G). These results demonstrate that the CRISPR–Cas9 system enables efficient editing of upland cotton through small-fragment insertions, deletions, and substitutions.

Editing heritability analysis of gene-edited materials in the *GhCPK* mutant library

To determine whether the edited genes could be stably inherited in T1 offspring, 50 gene-edited plants from the T0 generation were randomly selected for genetic analysis. Three plants from

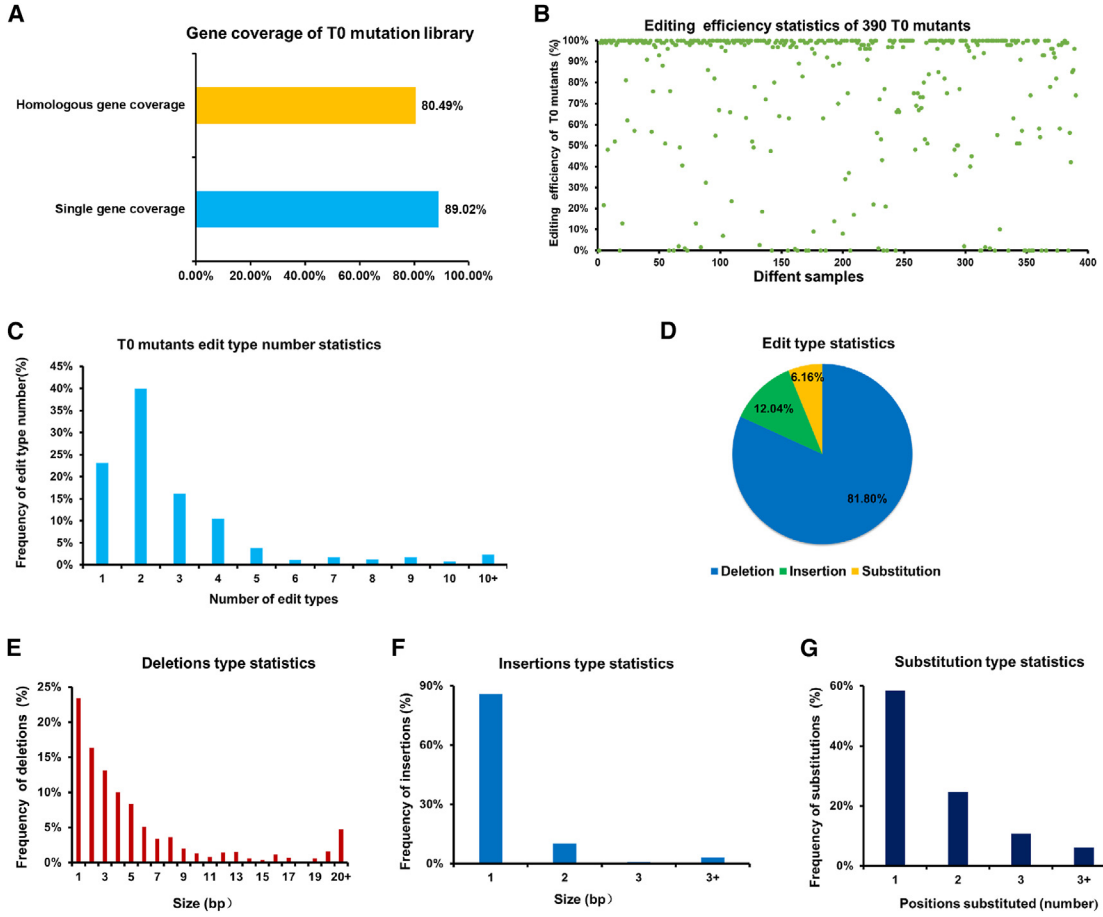

Figure 3. Molecular detection results for T0-generation plants of the *GhCPK* mutant library.

(A) Gene coverage of T0-generation plants from the *GhCPK* mutant library. Yellow represents the coverage of homologous genes, and blue represents the coverage of individual genes.

(B) Editing efficiency of 390 T0-generation plants from the *GhCPK* mutant library. Different samples are represented by green dots.

(C) Types of edits in T0-generation plants of the *GhCPK* mutant library. "Edit type" refers to different types of mutated reads produced after genes have been edited by the Cas9 protein. The number of edit types varies for each plant, ranging from one to more than 10.

(D) Classification of edit types in T0-generation plants of the *GhCPK* mutant library. Blue represents base absence, green represents base insertion, and yellow represents base substitution.

(E) Base deletion statistics. The x axis represents the number of deleted bases, and the y axis represents the frequency.

(F) Base insertion statistics. The x axis represents the number of inserted bases, and the y axis represents the frequency.

(G) Base substitution statistics. The x axis represents the number of substituted bases, and the y axis represents the frequency.

the progenies of each line were randomly chosen for high-throughput sequencing, and editing heritability analysis was performed by comparing their gene-editing types and editing efficiencies with those of the parental plants. If all the gene-editing types found in the T1 plants could be traced back to the parental plants, they were considered to be inherited from the parental generation. Among 150 T1 plants, 123 plants (82%) had gene-editing types that were completely inherited from the parental generation, 24 plants (16%) contained new gene-editing types, and three plants (2%) had not undergone gene editing (Supplemental Figure 8A and 8C). In terms of gene-editing efficiency, 101 plants (67.33%) in the T1 generation inherited 100% editing efficiency from the parental generation. Thirty-one plants (20.67%) had higher gene-editing efficiency than the parental generation, and 18 plants (12%) had lower gene-editing efficiency (Supplemental Figure 8B). These results indicate that most T1 plants inherit the parental editing patterns, but the genetic rules are relatively complicated. For instance, T1-*cpk53-1-2* completely inherited the editing type and efficiency of its T0 parent; *T0-cpk64-2* had an editing efficiency of only 2.91% in the T0 generation, but the offspring showed an inheritance efficiency of 100% (Supplemental Figure 8D).

Screening of insect-resistant materials from the *GhCPK* mutant library

To quickly screen for insect-resistant *GhCPK* mutants, we measured the damage caused by chewing pests on 243 cotton lines (involving editing of 62 independent *GhCPKs* and 28 pairs of homologous *GhCPKs*) grown in the field in two consecutive years and performed no-choice feeding experiments on all cotton lines using *S. littoralis* larvae (Figure 4A and 4B). Based on the damage caused by chewing pests to cotton lines in the field, combined with the results of the no-choice feeding experiment with *S. littoralis*

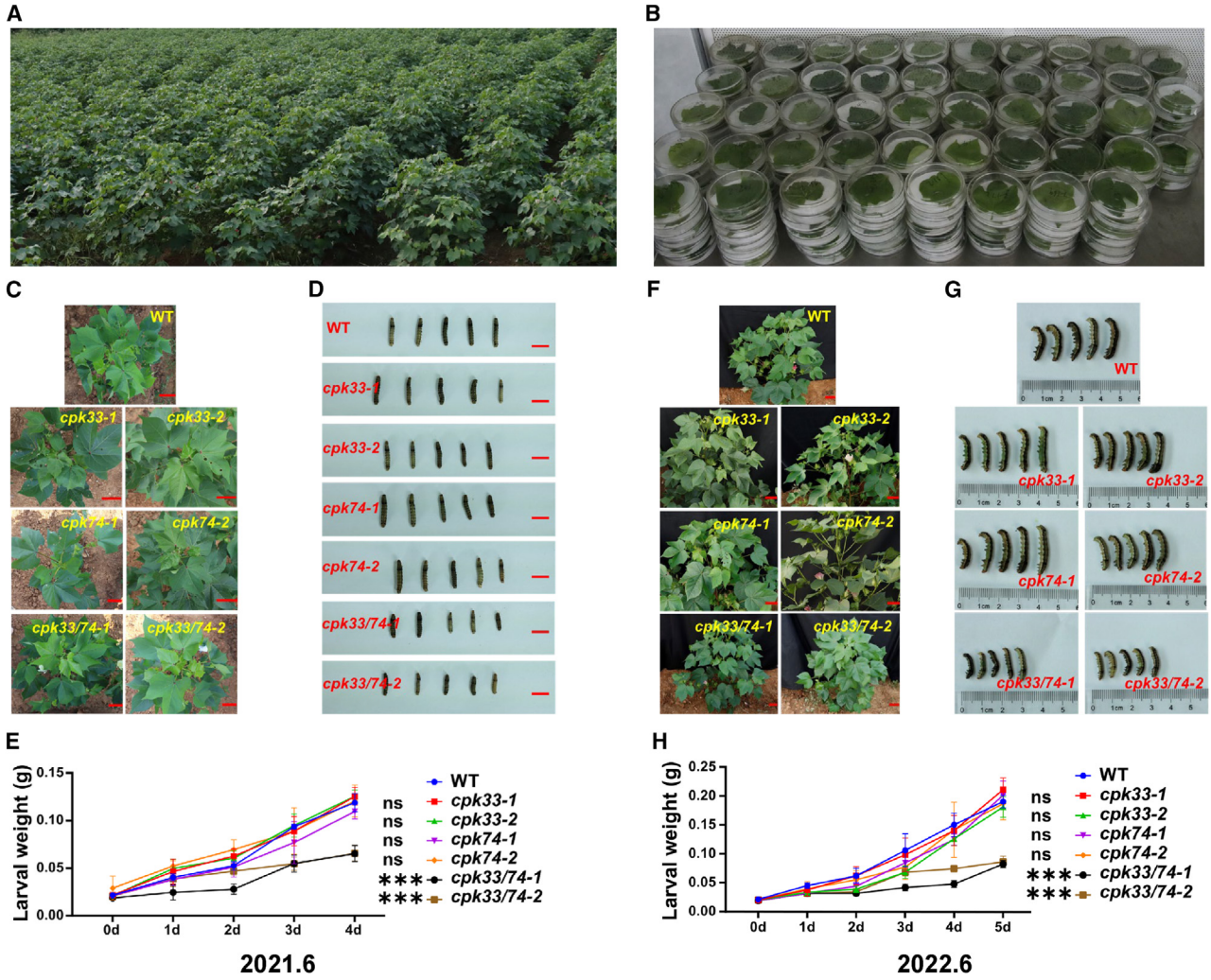

Figure 4. Screening for materials that are resistant or sensitive to chewing pests in the *GhCPK* mutant library.

(A) Field phenotyping for insect resistance in materials from the *GhCPK* mutant library.

(B) Non-selective feeding experiment with *S. litura* larvae on leaves of materials from the *GhCPK* mutant library.

(C) Field phenotyping for insect resistance of T1-generation *cpk33*, *cpk74*, and *cpk33/74* plants. The red line represents 5 cm.

(D) Comparison of body sizes of *S. litura* larvae after 4 days of continuous non-selective feeding on leaves of T1-generation *cpk33*, *cpk74*, and *cpk33/74* plants. The red line represents 1 cm.

(E) Average weight changes in *S. litura* larvae after 4 days of continuous feeding on leaves of T1-generation *cpk33*, *cpk74*, and *cpk33/74* plants.

(F) Field phenotyping for insect resistance of T2-generation *cpk33*, *cpk74*, and *cpk33/74* plants. The red line represents 5 cm.

(G) Comparison of the body sizes of *S. litura* larvae after 5 days of continuous non-selective feeding on leaves of T2-generation *cpk33*, *cpk74*, and *cpk33/74* plants.

(H) Average weight changes in *S. litura* larvae after 5 days of continuous feeding on leaves of T2-generation *cpk33*, *cpk74*, and *cpk33/74* plants. Means ± SE (n = 12). Statistical analyses were performed using Student's t-test. ***$p < 0.001$.

larvae, this serves as a reference standard. Mutant lines whose field pest-resistance phenotypes were consistent with the results of the non-selective feeding experiment were selected for pest resistance evaluation (Supplemental Figure 9A and 9B). We also required that the insect-resistance phenotypes of mutant lines could be replicated across different lines with the same gene knockout. Six stable insect-resistance genes and eight stable susceptibility genes were identified (Table 1). In particular, the *cpk33/74* cotton lines with simultaneous knockout of the homologous genes *GhCPK33* and *GhCPK74* showed the most significant insect-resistance phenotype. Plants of *cpk33/74* showed almost no trace of chewing-insect damage in the field evaluation (Figure 4C and 4F), and larvae that fed on *cpk33/74* had significantly lower weights than controls in a selective feeding assay. However, cotton plants with only *GhCPK33* (*cpk33*) or only *GhCPK74* (*cpk74*) knockout did not exhibit insect resistance (Figure 4D, 4E, 4G, and 4H).

Overexpression of *GhCPK33* and *GhCPK74* reduces cotton defense against *S. litura* larvae

To investigate the role of *GhCPK33* and *GhCPK74* in the interaction between cotton and herbivorous insects, we separately constructed overexpression vectors for *GhCPK33* and *GhCPK74*

Mutant line	Mutated gene ID	Chewing pest resistance	Average editing efficiency
cpk1	Gh_A01G0621	ns	23.76%
cpk6	Gh_A02G1635	+	100.00%
cpk15	Gh_A05G1571	+	100.00%
cpk16	Gh_A05G2355	ns	99.08%
cpk18	Gh_A05G3246	ns	78.81%
cpk24	Gh_A09G1033	− − −	93.88%
cpk26	Gh_A09G1157	ns	100.00%
cpk27	Gh_A09G1248	ns	100.00%
cpk29	Gh_A10G0886	ns	100.00%
cpk33	Gh_A11G1615	ns	100.00%
cpk37	Gh_A12G2686	−	100.00%
cpk39	Gh_A13G0563	−	76.02%
cpk40	Gh_A13G0566	ns	100.00%
cpk45	Gh_D02G0183	− −	70.21%
cpk46	Gh_D02G0663	ns	99.28%
cpk47	Gh_D02G1973	ns	100.00%
cpk48	Gh_D03G0087	ns	100.00%
cpk51	Gh_D04G0895	ns	100.00%
cpk52	Gh_D04G0900	ns	99.69%
cpk53	Gh_D04G1271	ns	100.00%
cpk61	Gh_D07G1198	− − −	100.00%
cpk63	Gh_D08G0142	ns	100.00%
cpk64	Gh_D09G1054	−	100.00%
cpk67	Gh_D09G1249	ns	27.64%
cpk71	Gh_D10G2029	− −	100.00%
cpk74	Gh_D11G1774	ns	100.00%
cpk75	Gh_D11G3329	ns	100.00%
cpk76	Gh_D12G0121	ns	100.00%
cpk77	Gh_D12G2743	ns	100.00%
cpk79	Gh_D13G0560	+	99.78%
cpk80	Gh_D13G0561	ns	96.52%
cpk6/48	Gh_A02G1635 and Gh_D03G0087	+	100.00%
cpk8/47	Gh_A03G1505 and Gh_D02G1973	ns	100.00%
cpk10/52	Gh_A04G0467 and Gh_D04G0900	ns	100.00%
cpk13/54	Gh_A04G1429 and Gh_D04G1486	+++	100.00%
cpk20/60	Gh_A06G1772 and Gh_D06G2206	ns	98.45%
cpk21/61	Gh_A07G1099 and Gh_D07G1198	ns	100.00%
cpk29/69	Gh_A10G0886 and Gh_D10G0863	ns	99.54%
cpk30/70	Gh_A10G1195 and Gh_D10G1303	ns	100.00%
cpk33/74	Gh_A11G1615 and Gh_D11G1774	+++	100.00%
cpk36/76	Gh_A12G0109 and Gh_D12G0121	ns	100.00%
cpk37/77	Gh_A12G2686 and Gh_D12G2743	−	58.89%

Table 1. Evaluation of insect resistance in plants of the *GhCPK* mutant library for defense against chewing pests.
Note: "+" indicates insect resistance, "−" indicates insect susceptibility, and "ns" indicates no difference.

(Supplemental Figure 10A and 10B). Through *Agrobacterium*-mediated transformation, we created independent transgenic plants overexpressing *GhCPK33* (*CPK33-OE*) and *GhCPK74* (*CPK74-OE*). On the basis of *GhCPK33* and *GhCPK74* expression levels, four independent lines (*CPK33-OE1*, *CPK33-OE2*, *CPK74-OE1*, and *CPK74-OE3*) were selected for further study from 22 *CPK33-OE* and 14 *CPK74-OE* lines (Supplemental Figure 10C and 10D). The T2 generation of mutants (*cpk33/74-1* and *cpk33/74-2*) and the T1 generation of overexpression plants (*CPK33-OE1*, *CPK33-OE2*, *CPK74-OE1*, and *CPK74-OE3*), as well as the wild type (WT), were sown simultaneously. After 40 days, selective feeding and non-selective feeding experiments were performed on all materials using *S. litura* larvae. After 24 h of selective feeding by *S. litura* larvae, leaf consumption was significantly higher for the *CPK33-OE* and *CPK74-OE* lines than the WT but significantly lower for the *cpk33/74* lines (Figure 5A and 5B). In the non-selective experiment, after feeding for 5 days, the weight of the *S. litura* larvae fed with *CPK33-OE* and *CPK74-OE* leaves was significantly higher than that of the larvae fed with WT leaves, whereas the weight of larvae fed with *cpk33/74* leaves was significantly lower (Figure 5C and 5D). *CPK33-OE* and *CPK74-OE* thus exhibited significantly reduced defense against *S. litura*, whereas *cpk33/74* plants showed significantly enhanced defense. We hypothesize that *GhCPK33* and *GhCPK74* play a negative regulatory role in cotton defense against *S. litura* larvae.

Evolution and domestication analysis of the *GhCPK33* and *GhCPK74* loci

Resequencing data from 1623 cotton germplasm resources from previous studies was used to investigate the evolution and domestication of the *GhCPK33* and *GhCPK74* loci (Li et al., 2021). The 1623 germplasms included 256 *G. hirsutum* landraces (Ghlandraces), which are naturally insect-resistant varieties; 438 elite *G. hirsutum* cultivars from the USA and other countries (GhImpUSO); and 929 elite *G. hirsutum* cultivars from China (GhImpCHN). The results revealed significant single-nucleotide polymorphism (SNP) variation around the *GhCPK33* and *GhCPK74* loci (20 kb upstream and downstream) in Ghlandraces (Figure 5E). To further investigate the evolution of *GhCPK33* and *GhCPK74* in the Ghlandraces population, we constructed genotype maps for the *GhCPK33* and *GhCPK74* loci (Supplemental Figure 11A and 11B). Haplotype analysis of the *GhCPK33* locus in the Ghlandraces population revealed the presence of four SNPs, two of which were located in introns and two in the 3′ UTR. The major SNP variants were observed at positions 24 527 886 (A to G) and 24 531 788 (G to A), accounting for 24.22% (Supplemental Figure 11A). Haplotype analysis of the *GhCPK74* locus in the Ghlandraces population revealed SNPs at nine positions: three in exons, five in introns, and one in the 5′ UTR. The main SNP variants were observed at positions 20 320 372 (A to G) and 20 319 885 (A to T), accounting for 6.64% (Supplemental Figure 11B). In summary, we speculate that *GhCPK33* and *GhCPK74* may have influenced the insect resistance of upland cotton during evolution and domestication processes.

GhCPK33 and GhCPK74 negatively regulate Ca^{2+} influx induced by *S. litura* OS

Elevation of cytoplasmic Ca^{2+} is an important early signal-transduction step in plants for perceiving attacks from herbivorous insects and initiating defense responses (Howe and Jander, 2008). Here, we used non-invasive micro-test technology (NMT) to investigate dynamic changes in cytoplasmic Ca^{2+} flux in cotton leaf mesophyll cells induced by the OS of *S. litura* and to explore the involvement of *GhCPK33* and *GhCPK74* in mediating post-ingestion Ca^{2+} influx by the insects (Figure 5F). In the untreated state, Ca^{2+} in the leaf mesophyll cells of WT, *cpk33/74*, *CPK33-OE*, and *CPK74-OE* showed an efflux state, with no significant differences in flow rate. The average flow rate within 5 min ranged from 40 to 100 pmol cm^{-2} s^{-1}. After treatment with the OS of *S. litura*, the leaf mesophyll cells of *cpk33/74* exhibited a strong and rapid Ca^{2+} influx, with an average flow rate of 701 pmol cm^{-2} s^{-1} within 5 min. However, this ability was impaired in *CPK33-OE* and *CPK74-OE* compared with the WT. The influx rate of Ca^{2+} was slowed in these lines, with average flow rates of only 52 and 104 pmol cm^{-2} s^{-1} within 5 min, respectively (Supplemental Figure 11C and 11D). These results indicate that *GhCPK33* and *GhCPK74* negatively regulate the Ca^{2+} influx induced by *S. litura*.

GhCPK33 and *GhCPK74* negatively regulate JA synthesis induced by OS of *S. litura*

JA is an important hormone related to plant insect resistance. To investigate the underlying mechanism of *ChCPK33/GhCPK74*-regulated insect resistance, we measured JA content in insect-induced cotton leaves. We cultured T2 generation plants of *cpk33*, *cpk74*, *cpk33/74*, *CPK33-OE*, *CPK74-OE*, and WT in the same environment. After 6 weeks, the leaf surfaces of all materials were scratched using a pattern wheel, and the injured leaf area was then coated with OS of *S. litura* (H + S) to simulate feeding by this insect. Water was applied to the injured area of the leaf (H + W) to simulate mechanical damage, and no treatment was applied to control leaves (N). After a 30-min treatment, JA content was measured in the leaves of all materials. Both H + W and H + S treatments induced JA production in cotton leaves. However, compared with WT leaves, *cpk33/74* leaves showed a more significant increase in JA content after H + S treatment. By contrast, the JA content of *CPK33-OE* and *CPK74-OE* leaves was lower than that of WT leaves after H + S treatment (Figure 5G). These results suggest that *GhCPK33* and *GhCPK74* can negatively regulate the JA synthesis induced by *S. litura* OS. Therefore, the *cpk33/74* material, with simultaneous knockout of *GhCPK33* and *GhCPK74*, shows enhanced resistance against insects. However, JA content in *cpk33* and *cpk74* leaves did not differ significantly from that of WT leaves after H + S treatment. Thus, we hypothesize that *GhCPK33* and *GhCPK74* are functionally redundant in the negative regulation of JA synthesis.

GhCPK33 and GhCPK74 interact with GhSAMS1 and GhSAMS2

The subcellular localization of GhCPK33 and GhCPK74 was examined by fusing their full-length coding sequences (CDSs) with N-terminal green fluorescent protein (GFP) and transiently expressing the fusion proteins in tobacco epidermal cells. Fluorescence was observed in both the cell membrane and the nucleus using confocal microscopy (Figure 6A). Yeast two-hybrid (Y2H) assays were then performed to screen interacting proteins of GhCPK33 or GhCPK74. The two S-adenosylmethionine synthases (SAMSs) GhSAMS1 and GhSAMS2 were identified

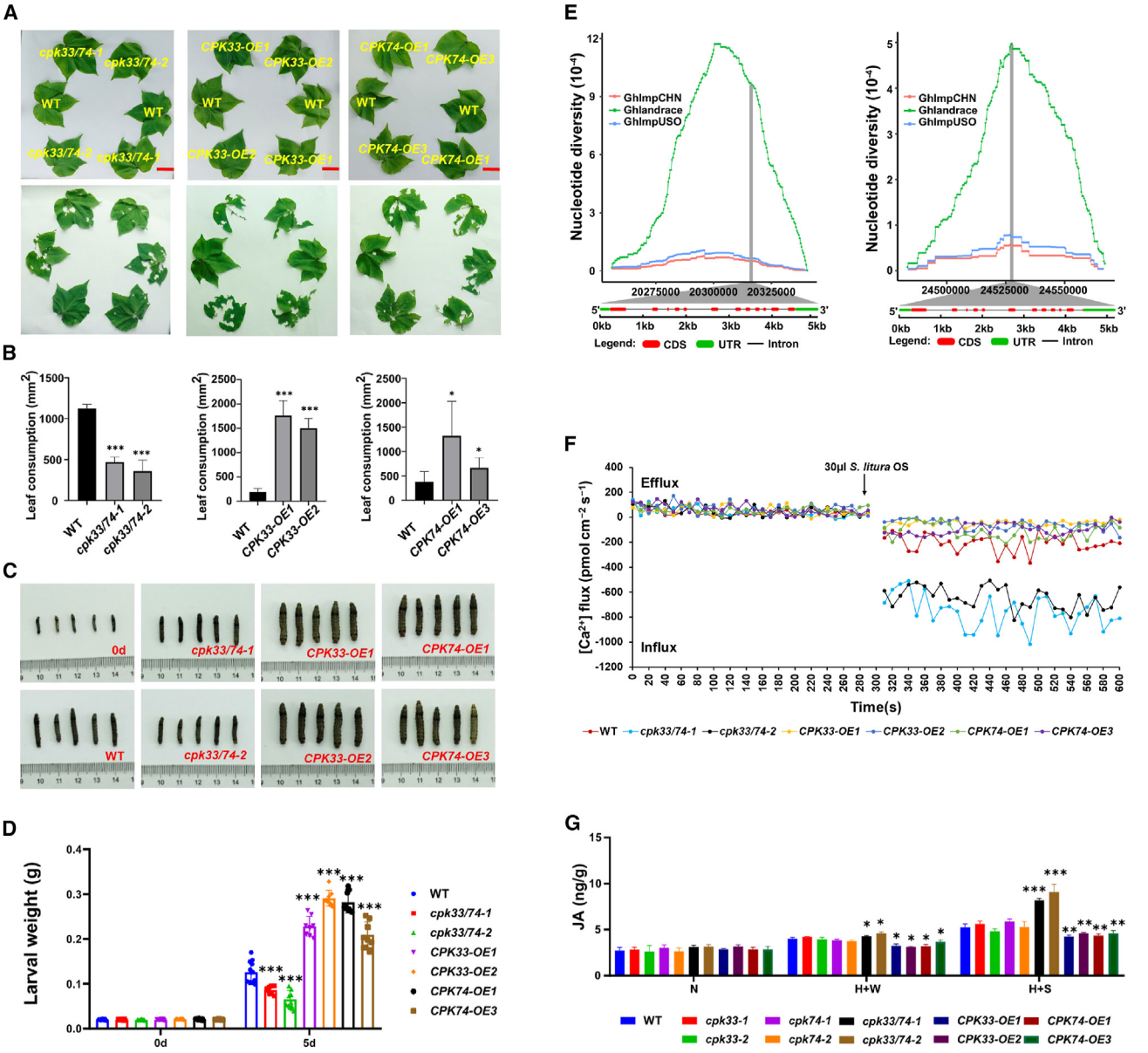

Figure 5. Evidence for the involvement of GhCPK33 and GhCPK74 in pest defense of upland cotton.

(A) Choice feeding assay with *S. litura* larvae and *cpk33/74*, *CPK33-OE*, *CPK74-OE*, and WT plants. The third-instar *S. litura* larvae were pre-starved for 6 h and photographed at 0 and 24 h. The red line represents 3 cm.

(B) Leaf consumption in the choice feeding assay with *S. litura*. Means ± SE (n = 6).

(C) Comparison of the body sizes of *S. litura* larvae after 5 days of continuous non-selective feeding on leaves of *cpk33/74*, *CPK33-OE*, *CPK74-OE*, and WT plants.

(D) Average body weight changes of *S. litura* larvae after 5 days of continuous feeding on leaves of *cpk33/74*, *CPK33-OE*, *CPK74-OE*, and WT plants. Means ± SE (n = 12).

(E) Nucleotide diversity of the Ghlandraces, GhImpUSO, and GhImpCHN populations. The x axis represents the 20 kb upstream and downstream of *GhCPK33* or *GhCPK74*, and the y axis represents nucleotide diversity.

(F) NMT was used to detect dynamic changes in cytoplasmic Ca^{2+} flux in *cpk33/74*, *CPK33-OE*, *CPK74-OE*, and WT leaf mesophyll cells before and after OS induction.

(G) JA content of *cpk33/74*, *cpk33*, *cpk74*, *CPK33-OE*, *CPK74-OE*, and WT leaves. N, not treated; H + W, mechanical damage; H + S, *S. litura* feeding. Means ± SE (n = 3). Statistical analyses were performed using Student's t-test. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

as interacting proteins of GhCPK33 and GhCPK74 (Figure 6B). SAMS is an important protein involved in multiple plant stress responses and is the only enzyme that catalyzes the synthesis of S-adenosyl methionine (Markham et al., 1983; He et al., 2019). Bimolecular fluorescence complementation (BiFC) and luciferase complementation imaging (LCI) assays were performed to verify the protein interactions of the GhSAMS with the GhCPKs. Strong yellow fluorescent protein signals were

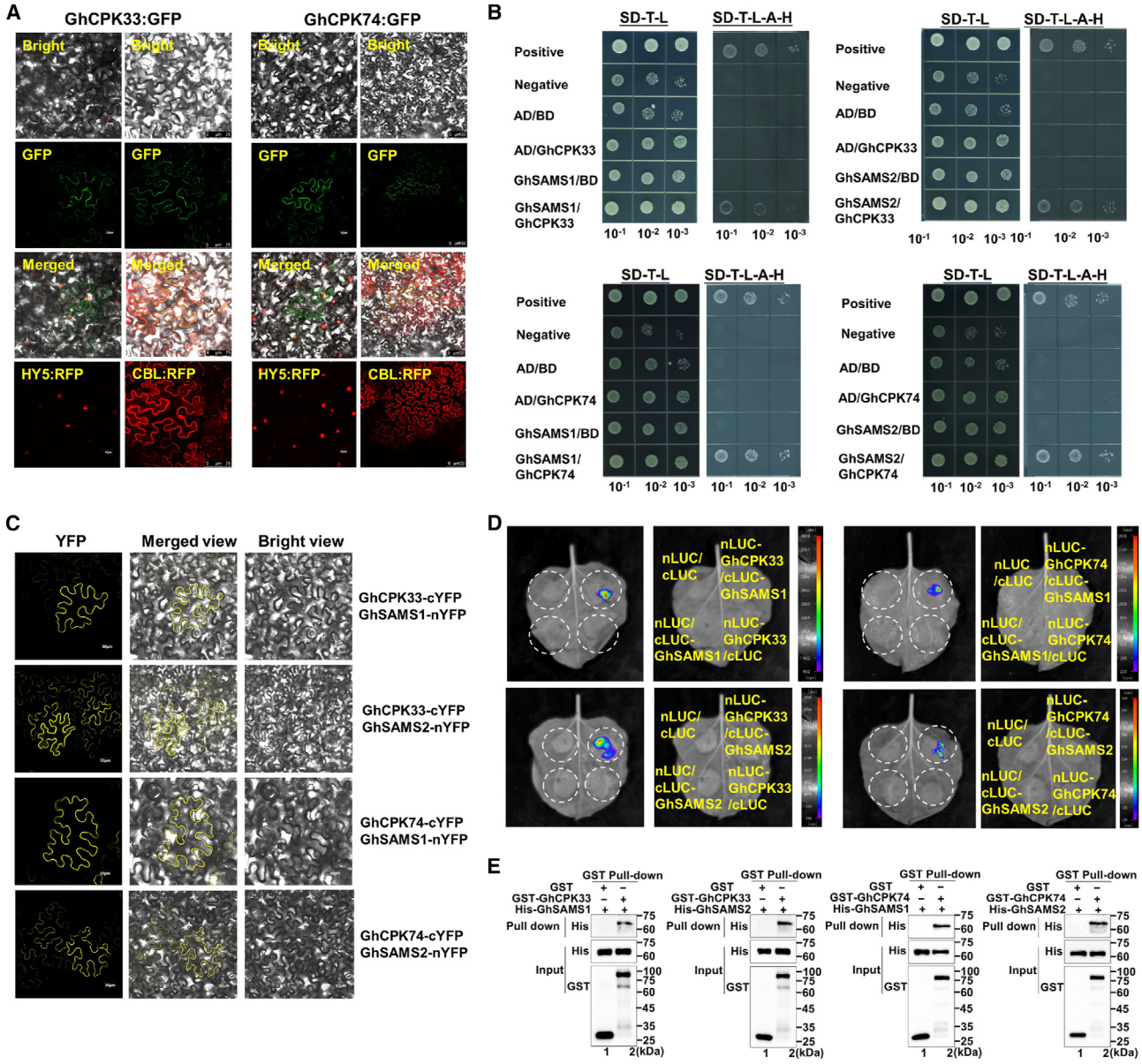

Figure 6. GhCPK33 and GhCPK74 interact directly with GhSAMS1 and GhSAMS2.

(A) Subcellular localization of GhCPK33:GFP protein and GhCPK74:GFP protein. GFP fluorescence was observed after transient expression of GhCPK33:GFP protein and GhCPK74:GFP protein in tobacco epidermal cells. Calcineurin B-Like (CBL) and HY5 were used as plasma membrane and nuclear markers, respectively.

(B) Y2H assay for interactions of GhCPK33 and GhCPK74 with GhSAMS1 and GhSAMS2. SD-T-L medium lacks Trp and Leu, and SD-T-L-H-A medium lacks Trp, Leu, His, and Ade.

(C) BiFC assay of the interaction of GhCPK33-cYFP and GhCPK74-cYFP with GhSAMS1-nYFP and GhSAMS2-nYFP in tobacco epidermal cells. Bars, 30 μm.

(D) LCI analysis of the interaction of GhCPK33-nLUC and GhCPK74-nLUC with GhSAMS1-cLUC and GhSAMS2-cLUC in tobacco leaves.

(E) Pull-down assay verifying the interaction of GhCPK33 and GhCPK74 with GhSAMS1 and GhSAMS2.

observed on the membranes of tobacco cells co-expressing GhSAMS1 and GhCPK33, GhSAMS2 and GhCPK33, GhSAMS1 and GhCPK74, and GhSAMS2 and GhCPK74 (Figure 6C). In the LCI tests, interaction signals between the four pairs of proteins were also detected (Figure 6D). GhSAMS1, GhSAMS2, GhCPK33, and GhCPK74 proteins were expressed and purified in prokaryotic cells, and pull-down experiments were performed to validate their interactions (Figure 6E). The results showed that GhSAMS1 and GhSAMS2 could interact with GhCPK33 and GhCPK74 *in vivo* and *in vitro*.

Silencing *GhSAMS1* and *GhSAMS2* impairs cotton resistance to *S. litura* larvae

To investigate the relationship of GhCPK33 and GhCPK74 with GhSAMS1 and GhSAMS2, expression levels of *GhSAMS1* and

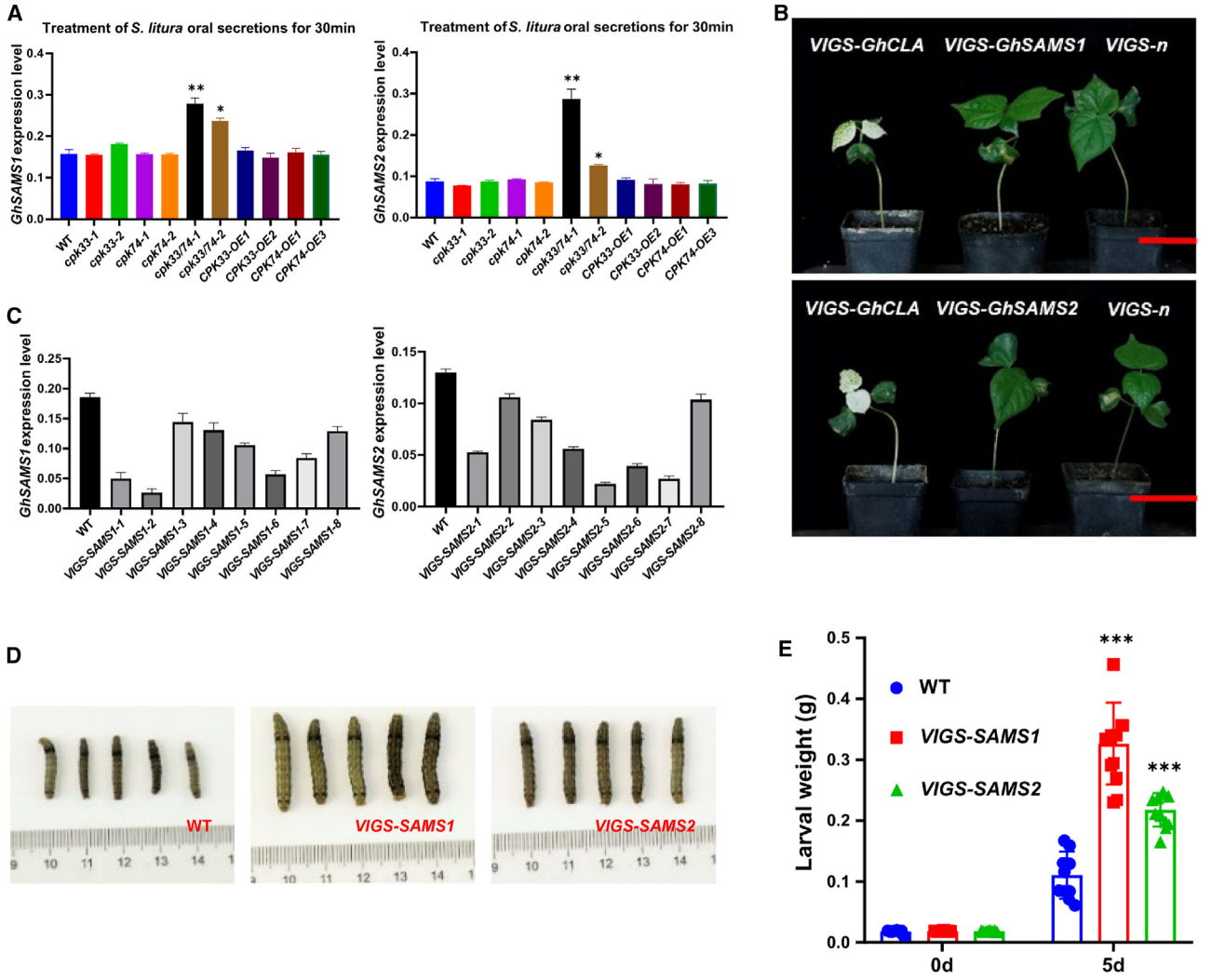

Figure 7. Expression levels of *GhSAMS1* and *GhSAMS2* after H + S treatment and assessment of insect resistance.

(A) qRT-PCR detection of *GhSAMS1* and *GhSAMS2* in *cpk33*, *cpk74*, *cpk33/74*, *CPK33-OE*, *CPK74-OE*, and WT materials after H + S treatment. *GhUB7* was used as the internal control. Means ± SE (n = 3).

(B) VIGS of *GhSAMS1* and *GhSAMS2*. *VIGS-GhCLA* was the positive control, and *VIGS-n* was the negative control. The red line represents 5 cm.

(C) Expression levels of *GhSAMS1* and *GhSAMS2* in leaves of *VIGS-GhSAMS1* and *VIGS-GhSAMS2* plants.

(D) Comparison of the body size of *S. litura* larvae after 5 days of non-selective feeding on WT, *VIGS-GhSAMS1*, and *VIGS-GhSAMS2* plants.

(E) Average weight changes of *S. litura* larvae after 5 days of continuous feeding on WT, *VIGS-GhSAMS1*, and *VIGS-GhSAMS2* leaves. Means ± SE (n = 12). Statistical analyses were performed using Student's t-test. ***$p < 0.001$.

GhSAMS2 were examined after 30 min of H + S treatment in *cpk33*, *cpk74*, *cpk33/74*, *CPK33-OE*, *CPK74-OE*, and WT materials. The expression levels of *GhSAMS1* and *GhSAMS2* were significantly increased in the *cpk33/74* leaves treated with H + S; however, there was no significant difference in *GhSAMS1 and GhSAMS2* expression in *cpk33* and *cpk74* leaves treated with H + S compared with the WT (Figure 7A). To investigate the role of *GhSAMS1* and *GhSAMS2* in defense against insects, we performed virus-induced gene-silencing (VIGS) of *GhSAMS1* and *GhSAMS2*, resulting in *VIGS-GhSAMS1* and *VIGS-GhSAMS2* seedlings (Figure 7B and 7C). Their insect resistance was assessed through a no-choice feeding experiment using *S. litura* larvae. The results showed a significant increase in the weight of *S. litura* larvae fed with *VIGS-GhSAMS1* and *VIGS-GhSAMS2* leaves compared with the control group (Figure 7D and 7E), suggesting that silencing of *GhSAMS1* and *GhSAMS2* can result in decreased resistance of upland cotton to *S. litura* larvae.

DISCUSSION

The *CDPK* gene family has a long evolutionary history, dating back to the earliest land plants, such as ferns and mosses (Hamel et al., 2014). With the development of biotechnology, an increasing number of *CDPKs* have been identified in different species. The *A. thaliana*, *O. sativa*, *Cucumis sativus*, *Populus alba*, and *Cucumis melo* genomes harbor 34, 31, 19, 30, and 18 CDPK genes, respectively (Sanders et al., 2002; Ray et al., 2007; Zuo et al., 2013; Xu et al., 2015; Zhang et al., 2017). The release of upland cotton genome data has provided valuable information

for a comprehensive analysis of upland cotton *CDPKs*. As an allopolyploid, upland cotton originated from a polyploidization event approximately 1–2 million years ago, which combined the *Gossypium arboreum* A subgenome and the *G. raimondii* D subgenome (Wendel, 1989). An assembly of the cotton A and D genomes has been published, and the genome draft of upland cotton was released in 2015, followed by several updated versions, all focusing on TM-1 as the research subject (Wang et al., 2012; Li et al., 2014; Yang et al., 2019). To obtain more comprehensive results, we retrieved *CDPK* family members from two protein databases (NBI and HAU) (Zhang et al., 2015; Wang et al., 2019b). Although there were some differences in alignment results between the two databases, such as the presence of *GhCPK63* in only the NBI database, the majority of the results were consistent. A total of 82 *GhCPKs* were identified. *GhCPKs* were widely expressed in multiple tissues of cotton and were induced by various biotic and abiotic stresses (Figure 1), indicating that they participate in various plant biological processes in cotton. In particular, multiple *GhCPK*s exhibited significant differential expression induced by *H. armigera*, *S. litura*, and whitefly, indirectly confirming the feasibility of searching for insect-resistance genes among the *GhCPKs*. Identification and analysis of the *CDPK* gene family in upland cotton thus provide valuable information for subsequent research.

In this study, we screened for insect-resistance genes using a strategy based on a CRISPR–Cas9 mutant library. The entire study was based on a breakthrough in upland cotton regeneration technology and the development of the CRISPR–Cas9 system for upland cotton (Jin et al., 2006; Wang et al., 2018). In plant species that can be transformed and regenerated, the use of CRISPR for gene mutation to validate gene function is practicable (Bortesi and Fischer, 2015). However, if gene knockout for the 246 sgRNAs were performed by constructing 246 individual vectors for genetic transformation, this would be a labor-intensive and time-consuming process. By creating a mixed pool of sgRNAs, constructing a vector library, and performing *Agrobacterium*-mediated transformation, we demonstrated a new approach for the creation of a large mutant population (Liu et al., 2023). This method is efficient, as it avoids the construction of individual vectors and the performance of individual transformations. It requires only batch transformation of mixed *Agrobacteria* containing the plasmid library, thereby quickly generating the desired gene-edited plants. However, drawbacks of this method are also apparent. Because the sgRNA sequences carried by each transgenic plant are unknown, subsequent molecular detection of the T0-generation materials is relatively complex. First, it is necessary to detect the sgRNA sequences carried by each material, a process that can be simplified by barcode-based high-throughput sequencing. Through our statistical analysis, 147 target sites were detected, targeting 73 individual GhCPKs and 33 pairs of homologous GhCPKs. The average gene coverage was 86.18%. Although our sgRNA coverage was much lower (59.76%) than the nearly 100% sgRNA coverage reported in soybean and tomato (Jacobs et al., 2017; Bai et al., 2020), we designed two sgRNAs for each gene and each homologous gene pair, thus compensating for the low sgRNA coverage and ultimately achieving a gene coverage of at least 85%. Expanding the population size of the transformation groups is another possible approach for enhancing sgRNA coverage. We also failed to rule out the potential influence of knocking out *GhCPKs* related to cotton regeneration, which would affect the total population of the *GhCPK* mutant library.

Through detection of gene editing in T0 plants of the *GhCPK* mutant library, we found that the CRISPR–Cas9 system in upland cotton was remarkably effective, with over 70% of plants exhibiting editing efficiencies >80%. Deletion was the predominant editing type in most plants, accounting for 81.80% of all edits, followed by insertion, which accounted for 12.04%. Through genetic analysis of the T0 and T1 generations, we observed that the heredity of the CRISPR–Cas9 gene edits was complex. Although most T1 plants inherited the editing pattern of their parents, a few produced new editing types owing to the presence of CRISPR–Cas9 expression cassettes. The presence of non-homozygous editing in the T0 generations further contributed to the complexity. When selecting T1 generations, we believe that individuals with a single editing type and without the CRISPR–Cas9 expression cassettes are more suitable for seed preservation.

A small number of T0 plants in the *GhCPK* mutant library exhibited unique phenotypes, such as yellow leaves and dwarf stature (Supplemental Figure 13), and most of these T0 plants were sterile. These plants were valid mutants in which a specific *GhCPK* had been edited. However, not all mutants of the corresponding *GhCPK* showed these phenotypes. We re-transformed CRISPR–Cas9 vectors into cotton to knock out the corresponding *GhCPKs* individually, and the unique phenotypes were not reproducible. We believe that these phenotypes were related to mutations caused by transfer-DNA insertion or the long-term tissue culture process and were not related to *GhCPKs*. To ensure the stability and accuracy of phenotypic identification, it is necessary to select at least two independent knockout lines in the T0 generation of each *GhCPK* for phenotypic screening.

We identified GhSAMS1 and GhSAMS2 as interacting proteins of GhCPK74 and GhCPK33 through BiFC, LCI, and pull-down assays. SAMS is widely involved in the regulation of stress responses and growth/development across plant species. In *Arabidopsis*, *AtSAMS3* is involved in pollen formation (Chen et al., 2016). In cucumber, *CsSAMS* is induced by salt stress and participates in related regulation (Zhu et al., 2021). In tobacco, overexpression of soybean *GmSAMS1* enhances resistance to *S. litura* (Fan et al., 2018). We examined the expression levels of *GhSAMS1* and *GhSAMS2* in cotton leaves induced by OS of *S. litura*. Both *GhSAMS1* and *GhSAMS2* were significantly upregulated in *cpk33/74* materials but showed no obvious changes in *cpk33* and *cpk74* compared with WT plants. We therefore hypothesize that *GhCPK74* and *GhCPK33* act redundantly in the negative regulation of *GhSAMS1* and *GhSAMS2* expression and in cotton resistance to insects. We next performed insect-resistance assays with *VIGS-GhSAMS1* and *VIGS-GhSAMS2* plants and found that silencing of *GhSAMS1* or *GhSAMS2* enhanced the feeding preference of *S. litura* larvae. On the basis of these results, a schematic diagram of the regulation of cotton pest defense by *GhCPK33* and *GhCPK74* is shown in Supplemental Figure 14. We speculate that *GhCPK33* and *GhCPK74*, as upstream signaling factors, may regulate multiple pathways in response to

herbivory stress. In this study, we provide evidence to support the negative regulation of cotton insect resistance by *GhCPK33* and *GhCPK74*. However, the roles of *GhCPK33* and *GhCPK74* in the regulation of JA synthesis are still unclear, and further study is needed to reveal the underlying mechanism of GhCPK-regulated insect resistance.

METHODS

Genome-wide comprehensive analysis of the *GhCPK* gene family

Genomic data for *G. hirsutum* and *G. raimondii* were obtained from Cottongen (https://www.cottongen.org/). Genomic data for *A. thaliana* and *O. sativa* were acquired from Ensembl Plants (http://plants.ensembl.org/index.html) and the Rice Genome Annotation Project (http://rice.uga.edu/downloads_gad.shtml), respectively. CDPK protein sequences from *A. thaliana*, *O. sativa*, and *G. raimondii* were used as query sequences to search the genomes of *G. hirsutum* (NBI, Zhang et al., 2015; and HAU, Wang et al., 2019b) with the BLASTP program. In addition, a local hidden Markov model–based search (HMMER) was developed based on the CDPK protein sequences to identify *GhCPKs*. The results from BLASTP and HMMER were compared to identify and exclude redundant sequences. The remaining candidate genes were then uploaded to SMART (http://smart.embl.de/) and CDD (https://www.ncbi.nlm.nih.gov/cdd/) to confirm the presence of conserved domains. Amino acid residues, molecular weights, and isoelectric points of the GhCPKs were predicted using the ProtParam tool (https://web.expasy.org/protparam/). DeepLoc-2.0 (https://services.healthtech.dtu.dk/services/DeepLoc-2.0/) was used to predict subcellular localization of the GhCPKs, DeepTMHMM (https://dtu.biolib.com/DeepTMHMM/) to predict transmembrane domains, and SignalP-6 (https://biolib.com/DTU/SignalP-6/) to predict signal peptides. A phylogenetic tree based on the full-length protein sequences of *A. thaliana*, *O. sativa*, and *G. hirsutum* was constructed using the neighbor-joining method in MEGA 7.0 software and visualized with EvolView (http://www.evolgenius.info/evolview/#/). Conserved motifs of the GhCPKs were analyzed using the MEME program and visualized with TBtools (Chen et al., 2020). Gene structures of the *GhCPKs* were analyzed by comparing CDSs with their corresponding genome sequences using GSDS software (http://gsds.gao-lab.org/). *Cis*-acting elements in the *GhCPK* promoters were analyzed using the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). A collinearity analysis between *G. hirsutum* and three other species was performed using MCScanX software. Results of all the aforementioned analyses were visualized with TBtools. Images of *GhCPK* gene locations on chromosomes and genome-wide gene duplication events were produced with Circos software.

Expression profiling of *GhCPKs* in different tissues and stress conditions

Gene expression data for different tissues and abiotic stresses were obtained from TM-1 transcriptome data (Zhang et al., 2015). Transcriptome data for *S. litura* and *H. armigera* OS-induced gene expression at different time points were obtained from the NCBI Sequence Read Archive (NCBI: PRJNA522889) (Si et al., 2020). Transcriptomic data for ZS and HR cultivars at four time points after whitefly infestation were obtained from previous research in our lab (Li et al., 2016). For different RNA sequencing reads, low-quality reads were filtered using Trimmomatic (v.0.39), and clean reads were mapped to the TM-1 reference genome using HISAT2 (v.2.2.1) (Kim et al., 2019). The expression levels (transcripts per million) of genes were calculated using StringTie (v.2.1.4). If a gene had a value of transcripts per million >0, it was considered to be expressed. Differentially expressed genes were identified using the DESeq2 package with false discovery rate <0.05 and $|\log2(\text{fold change})| \geq 1$ (Varet et al., 2016). TBtools was used for heatmap visualization.

sgRNA design

To construct a mutant library of the *GhCPK* gene family, we used CRISPR-P 2.0 software (http://crispr.hzau.edu.cn/cgi-bin/CRISPR2/CRISPR) to design sgRNAs for 82 *GhCPK* genes and 41 pairs of homologous *GhCPK* genes, with the *G. hirsutum* (v1.1) genome as the reference sequence. A total of 246 sgRNAs were designed (Supplemental Table 2). Genome-wide comparison screening was performed, and the principles for sgRNA design were as follows: priority was given to sgRNAs that had more than two mismatches with other non-target genes, sgRNAs with a CG content between 40% and 60% were preferred, and sgRNAs in the key CDS region were prioritized.

Construction of a knockout vector library for the *GhCPK* gene family

We used the efficient CRISPR–Cas9 knockout vector *pRGEB32-GhU6.7*, which was developed specifically for upland cotton (Wang et al., 2018). The two BstBI restriction sites were used for complete enzymatic digestion of the vector, and the linearized plasmid was used as the ligation vector. The 246 selected sgRNA sequences were reverse complemented, and the reverse-complemented sequences were used to synthesize primers with the TTCTAGCTCTAAAAC adapter added to the 5′ end and the TGCACCAGCCGGGAAT adapter added to the 3′ end. To ensure coverage and ease of operation, the 246 synthetic primers were divided into six groups, each containing 40 or 41 primers. The primers in each group were mixed in equal amounts to form a primer pool, and amplification was performed using the *PGTR* plasmid as the template. The PCR products were then cloned into the linearized vector *pRGEB32-GhU6.7* using In-Fusion cloning technology. The cloned plasmid was heat-shock-transformed into the *Escherichia coli* strain Trans5α at 42°C and incubated at 37°C for 15 h. All *E. coli* colonies were collected, and plasmids were extracted using the TIANprep Mini Plasmid Kit (TIANGEN, Beijing, China). The concentration of each group of plasmids was determined using a spectrophotometer, and equal amounts of plasmids from each group were mixed into one plasmid pool for *Agrobacterium*-mediated transformation. Multiple electroporations were performed for *Agrobacterium* GV3101 transformation to ensure a sufficient number of colonies and plasmid coverage. After 2 days of growth, all *Agrobacterium* colonies were scraped and stored at −80°C for cotton genetic transformation. All steps are illustrated in Figure 2C.

Construction of overexpression and VIGS vectors

The full-length CDS sequences of *GhCPK33* and *GhCPK74* were amplified using the cDNA sequence of Jin668 as a template. The amplified sequences were inserted into the *pK2GW7* overexpression vector using Gateway cloning technology (Karimi et al., 2002). The recombinant overexpression plasmid was transformed into *Agrobacterium* by electroporation and reserved for genetic transformation in cotton. VIGS fragments of *GhSAMS1* and *GhSAMS2* were amplified using the cDNA sequence of Jin668 as a template. The PCR product was cloned into the TRV:00 enzyme linear products using In-Fusion cloning technology (Gao et al., 2013). The recombinant VIGS vector plasmid was transformed into *Agrobacterium* by electroporation and used for VIGS experiments.

Barcode high-throughput sequencing detects sgRNAs in T0-generation plants of the *GhCPK* mutant library

To identify the sgRNA sequences carried by each mutant plant, we used a barcode high-throughput sequencing strategy. Because only sgRNA sequences in *pRGEB32-GhU6.7* vectors carried by different plants were different, we designed primer pairs for the common vector sequences upstream and downstream of the sgRNAs according to the location of sgRNAs on the *pRGEB32-GhU6.7* vectors (Supplemental Figure 7A). A 9-nucleotide barcode was then added to the 5′ end of each primer pair. A total of 28 barcode primers were designed, which, depending on the combination of primer pairs, could simultaneously detect 192 samples (Supplemental Table 5). After mixing all amplified sequences,

high-throughput sequencing was performed, and the sgRNA sequences carried by each plant were determined on the basis of their different barcodes.

High-throughput sequencing for detection of gene editing

According to the sgRNA sequence carried by each plant in the *GhCPK* mutant library, the gene ID of the edited gene was determined. Primers were then designed upstream and downstream of the target sequence on the basis of their positions within the full-length sequence of the target gene (Supplemental Table 6). The CRISPR–Cas9 system can introduce a large number of indels at the target editing site, and the upstream and downstream primers should therefore not be designed too close to the target site and should have a minimum spacing of at least 30 bp. The amplified fragments are generally not longer than 280 bp. To maintain consistency in PCR amplification conditions, all primer-annealing temperatures in this experiment were designed to be between 58°C and 60°C. After amplification, all PCR products were mixed for high-throughput sequencing, and the gene-editing status of each plant in the *GhCPK* mutant library was determined.

Identification of insect-resistance phenotypes in plants of the *GhCPK* mutant library

We performed a statistical analysis of the damage caused by chewing pests on T1-generation positive gene-edited materials from the mutant library in our experimental field in Wuhan. All plants were planted on April 18, 2021, and during the following weeks they were infested by chewing pests such as *H. armigera*, *Spodoptera exigua*, and *S. litura* that occurred naturally in the environment. On June 10, 2021, we performed a statistical analysis of the extent of damage caused by chewing insect pests. We repeated the same field trial and statistical analyses in 2022. We also performed no-choice feeding experiments on all materials using *S. litura* larvae. On June 15, 2021, we performed a no-choice feeding experiment with *S. litura larvae* on 243 field plant lines. Twelve *S. litura* larvae with a body weight of about 0.018–0.022 g were selected from each plant line for feeding. Each larva was placed in a glass dish (9 cm in diameter) and fed fresh cotton leaves every day for four consecutive days. The weights of the *S. litura* larvae were recorded daily. The experiment was repeated in 2022.

Evolution and domestication of the *GhCPK33* and *GhCPK74* loci

Resequencing data and RNA sequencing data were obtained from previous studies (Li et al., 2021). VCFtools (v0.1.16) was used to extract SNP variants from the samples, perform quality control, and calculate nucleotide diversity (π) with the specified parameters (–maf 0.01, –hwe 0.01) (Danecek et al., 2011). Plink (version 1.9) was used to convert file formats (Purcell et al., 2007). Gene structures were visualized using the gene structure files through the GSDS website (http://gsds.gao-lab.org/). Haplotype analysis was performed by analyzing SNP variants with R scripts.

Measurement of Ca^{2+} flux in cotton leaf epidermal cells and quantification of JA content in cotton leaves

Seeds of T1-generation *cpk33*, *cpk74*, *cpk33/74*, *CPK33-OE*, *CPK74-OE*, and WT materials were sown at the same time and cultivated under the same conditions (Supplemental Figure 12A). High-throughput sequencing was performed on all T2-generation plants, and all exhibited 100% editing (Supplemental Figure 12B). Expression levels of *GhCPK33 and GhCPK74* were measured in the T2-generation overexpression materials, *CPK33-OE1*, *CPK33-OE2*, *CPK74-OE1*, and *CPK74-OE3* (Supplemental Figure 12C), confirming the sustained high expression of *GhCPK33* and *GhCPK74* in these plants.

Fresh leaves were collected from cotton plants and immediately soaked in test buffer (0.1 mM $CaCl_2$) for 2 h. Ca^{2+} flux was measured using the non-invasive micro-test technique (NMT-YG-100, YoungerUSA, Amherst, MA, USA) as described previously (Wang et al., 2019a). Before treatment of leaves with *S. litura* OS, the steady-state flux of leaf mesophyll cells was recorded continuously for 5 min. Subsequently, 30 μl of *S. litura* OS was slowly added to the measuring buffer, and the transient flux of Ca^{2+} was recorded and sustained for 5 min. At least three leaf mesophyll cells were tested for each line.

To measure the concentration of JA in cotton leaf samples, approximately 100 mg of leaf material was extracted twice with 80% cold methanol (v/v) overnight at 4°C. For each sample, 10 ng (±)-9,10-dihydro-JA (Olchemim) was added as an internal standard. The two extracts were combined, and the mixture was evaporated to the aqueous phase using N_2, dissolved in 0.4 mL of methanol, and filtered using a syringe-facilitated filter (Nylon 66; Jin Teng Experiment Equipment, Tianjin, China). The samples were stored at −80°C before measurement. JA levels were quantified using a high-pressure liquid chromatography–tandem mass spectrometry system (AB SCIEX Triple Quad 5500 LC/MS/MS) with JA (Sigma) as the external standard (Sun et al., 2014).

Subcellular localization

The CDSs of *GhCPK33* and *GhCPK74* were cloned into the C-terminal fusion GFP vector *pGWB405*. The recombinant vectors *pGWB405-GhCPK33* and *pGWB405-GhCPK74* were transiently expressed in tobacco epidermal cells after *Agrobacterium*-mediated transfection. Protein localization was observed using an Olympus FV1200 confocal microscope 2 days after *Agrobacterium* transfection. CBL1:RFP was used as a plasma membrane marker, and HY5:RFP was used as a nuclear marker.

Y2H, LCI, and BiFC assays

The CDSs of *GhCPK33* and *GhCPK74* were cloned into *pGBKT7* to generate a bait vector and transformed into yeast strain Y2H. The CDSs of *GhSAMS1* and *GhSAMS2* were cloned into *pGADT7* to generate prey vectors and transformed into yeast strain Y187. The interactions between bait and prey were assessed by growth on SD-Leu-Trp (SD-2) medium and SD-Leu-Trp-His-Ade (SD-4) medium. For the LCI assays, the CDSs of *GhCPK33* and *GhCPK74* were cloned into the *JW771* vector, and the CDSs of *GhSAMS1* and *GhSAMS2* were cloned into the *JW772* vector. The recombinant vectors were transformed into *A. tumefaciens* GV3101 and transiently expressed in *Nicotiana benthamiana* leaves by injection with needleless syringes. Fluorescence signals of luciferase (LUC) luminescence in LCI were observed with a cryogenically cooled charge-coupled device camera (NightSHADE LB 985). For BiFC assays, the CDSs of *GhCPK33* and *GhCPK74* were cloned into the *pxy104-cYFP* vector, and the CDSs of *GhSAMS1* and *GhSAMS2* were cloned into the *pxy106-nYFP* vector. The vectors were transformed into *A. tumefaciens* strain GV3101 and transiently expressed in *N. benthamiana* leaves by injection with needleless syringes. Fluorescence in the *N. benthamiana* epidermal cells was observed 60 h later using a confocal microscope (Olympus FV1200).

qRT–PCR

For each sample, 3 μg of RNA was reverse transcribed into cDNA using M-MLV reverse transcriptase (Promega). The real-time qRT–PCR reactions were performed using the QuantStudio 6 Flex Real-Time PCR System (Applied Biosystems), and *GhUB7* was used as the internal control gene (Tu et al., 2007).

In vitro pull-down assays

The full-length CDSs of *GhCPK33* and *GhCPK74* were separately cloned into the *pGEX-6P-1* vector (fused with the GST tag) using the Gateway technique. The full-length CDSs of *GhSAMS1* and *GhSAMS2* were separately cloned into the *pet-SUMO* vector (fused with the His tag) using the Gateway technique. In total, four vectors were constructed, *GST-GhCPK33*, *GST-GhCPK74*, *His-GhSAMS1*, and *His-GhSAMS2*. These constructs were transformed into *E. coli* BL21 (DE3) for prokaryotic expression. The GST fusion protein and the target His fusion protein

were co-incubated in pull-down buffer at 4°C for 2 h. GST protein was used as a blank control. After 2 h, 20 μl of the incubation system was taken as the input sample to detect the interaction between GST and His proteins. GST beads were added and incubated at 4°C for 2 h. After four washes with pull-down buffer, samples were prepared for western blot analysis (Chen et al., 2023).

AVAILABILITY OF MATERIALS

The *GhCPK* mutants/transgenic lines generated in this study can be obtained by contacting the corresponding author.

FUNDING

This work was supported by Biological Breeding of Stress Tolerant and High Yield Cotton Varieties (2023ZD04040) to L.M., the National Natural Science Fund of China for Distinguished Young Scholars (32325039) and the National Natural Science Foundation of China (32272128) to S.J., the National Natural Science Foundation of China (32401780) and the Key Scientific and Technological Project of Henan Province (222102110151) to S.L., and the Major Science and Technology Project of Xinjiang Uygur Autonomous Region (2023A02003-2) to B.L.

ACKNOWLEDGMENTS

No conflict of interest is declared.

AUTHOR CONTRIBUTIONS

S.J., X.Z., W.G., L.M., and B.L. provided the experimental design and supervision. F.W., S.L., and L.L. performed the experiments and manuscript preparation. Q.W. and Z.X. provided assistance with data analysis. G.W., T.H., C.F., Y.F., and L.C. provided suggestions and modified the manuscript.

SUPPLEMENTAL INFORMATION

Supplemental information is available at *Plant Communications Online*.

Received: March 2, 2024
Revised: July 10, 2024
Accepted: August 7, 2024
Published: August 12, 2024

REFERENCES

Bai, M., Yuan, J., Kuang, H., Gong, P., Li, S., Zhang, Z., Liu, B., Sun, J., Yang, M., Yang, L., et al. (2020). Generation of a multiplex mutagenesis population via pooled CRISPR-Cas9 in soya bean. Plant Biotechnol. J. **18**:721–731.

Batistic, O., and Kudla, J. (2012). Analysis of calcium signaling pathways in plants. Biochim. Biophys. Acta **1820**:1283–1293.

Bortesi, L., and Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. Biotechnol. Adv. **33**:41–52.

Chen, C., Chen, H., Zhang, Y., Thomas, H.R., Frank, M.H., He, Y., and Xia, R. (2020). TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol. Plant **13**:1194–1202.

Chen, L., Zhang, B., Xia, L., Yue, D., Han, B., Sun, W., Wang, F., Lindsey, K., Zhang, X., and Yang, X. (2023). The GhMAP3K62-GhMKK16-GhMPK32 kinase cascade regulates drought tolerance by activating GhEDT1-mediated ABA accumulation in cotton. J. Adv. Res. **51**:13–25.

Chen, Y., Zou, T., and McCormick, S. (2016). S-Adenosylmethionine Synthetase 3 Is Important for Pollen Tube Growth. Plant Physiol. **172**:244–253.

Chen, Z.J., Scheffler, B.E., Dennis, E., Triplett, B.A., Zhang, T., Guo, W., Chen, X., Stelly, D.M., Rabinowicz, P.D., Town, C.D., et al. (2007). Toward sequencing cotton (Gossypium) genomes. Plant Physiol. **145**:1303–1310.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al.; 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. Bioinformatics **27**:2156–2158.

Erb, M., Meldau, S., and Howe, G.A. (2012). Role of phytohormones in insect-specific plant reactions. Trends Plant Sci. **17**:250–259.

Erb, M., and Reymond, P. (2019). Molecular Interactions Between Plants and Insect Herbivores. Annu. Rev. Plant Biol. **70**:527–557.

Fan, R., Li, X., Wang, S., Yu, D., and Wang, H. (2018). The soybean S-adenosylmethionine synthetase gene GmSAMS1 confers resistance to common cutworm in transgenic tobacco. Soybean Sci. **37**:268–274.

Gaillochet, C., Develtere, W., and Jacobs, T.B. (2021). CRISPR screens in plants: approaches, guidelines, and future prospects. Plant Cell **33**:794–813.

Gao, W., Long, L., Zhu, L.F., Xu, L., Gao, W.H., Sun, L.Q., Liu, L.L., and Zhang, X.L. (2013). Proteomic and virus-induced gene silencing (VIGS) Analyses reveal that gossypol, brassinosteroids, and jasmonic acid contribute to the resistance of cotton to Verticillium dahliae. Mol. Cell. Proteomics **12**:3690–3703.

Hamel, L.P., Sheen, J., and Séguin, A. (2014). Ancient signals: comparative genomics of green plant CDPKs. Trends Plant Sci. **19**:79–89.

He, M.W., Wang, Y., Wu, J.Q., Shu, S., Sun, J., and Guo, S.R. (2019). Isolation and characterization of S-Adenosylmethionine synthase gene from cucumber and responsive to abiotic stress. Plant Physiol. Biochem. **141**:431–445.

Heckel, D.G. (2012). Ecology. Insecticide resistance after Silent spring. Science **337**:1612–1614.

Howe, G.A., and Jander, G. (2008). Plant immunity to insect herbivores. Annu. Rev. Plant Biol. **59**:41–66.

Huang, Y., Wan, P., Zhang, H., Huang, M., Li, Z., and Gould, F. (2013). Diminishing returns from increased percent Bt cotton: the case of pink bollworm. PLoS One **8**:e68573.

Jacobs, T.B., Zhang, N., Patel, D., and Martin, G.B. (2017). Generation of a Collection of Mutant Tomato Lines Using Pooled CRISPR Libraries. Plant Physiol. **174**:2023–2037.

Jin, S., Zhang, X., Nie, Y., Guo, X., Liang, S., and Zhu, H. (2006). Identification of a novel elite genotype for in vitro culture and genetic transformation of cotton. Biol. Plant. (Prague) **50**:519–524.

Kanchiswamy, C.N., Takahashi, H., Quadro, S., Maffei, M.E., Bossi, S., Bertea, C., Zebelo, S.A., Muroi, A., Ishihama, N., Yoshioka, H., et al. (2010). Regulation of Arabidopsis defense responses against Spodoptera littoralis by CPK-mediated calcium signaling. BMC Plant Biol. **10**:97.

Karimi, M., Inzé, D., and Depicker, A. (2002). GATEWAY vectors for Agrobacterium-mediated plant transformation. Trends Plant Sci. **7**:193–195.

Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. **37**:907–915.

Li, F., Fan, G., Wang, K., Sun, F., Yuan, Y., Song, G., Li, Q., Ma, Z., Lu, C., Zou, C., et al. (2014). Genome sequence of the cultivated cotton Gossypium arboreum. Nat. Genet. **46**:567–572.

Li, J., Yuan, D., Wang, P., Wang, Q., Sun, M., Liu, Z., Si, H., Xu, Z., Ma, Y., Zhang, B., et al. (2021). Cotton pan-genome retrieves the lost sequences and genes during domestication and selection. Genome Biol. **22**:119.

Li, J., Zhu, L., Hull, J.J., Liang, S., Daniell, H., Jin, S., and Zhang, X. (2016). Transcriptome analysis reveals a comprehensive insect resistance response mechanism in cotton to infestation by the

phloem feeding insect Bemisia tabaci (whitefly). Plant Biotechnol. J. **14**:1956–1975.

Li, X., Hu, D., Cai, L., Wang, H., Liu, X., Du, H., Yang, Z., Zhang, H., Hu, Z., Huang, F., et al. (2022). CALCIUM-DEPENDENT PROTEIN KINASE38 regulates flowering time and common cutworm resistance in soybean. Plant Physiol. **190**:480–499.

Li, Y., Hallerman, E.M., Wu, K., and Peng, Y. (2020). Insect-Resistant Genetically Engineered Crops in China: Development, Application, and Prospects for Use. Annu. Rev. Entomol. **65**:273–292.

Liu, H.J., Jian, L., Xu, J., Zhang, Q., Zhang, M.L., Jin, M., Peng, Y., Zhang, M., Han, B., Liu, J., et al. (2020). High-Throughput CRISPR/Cas9 Mutagenesis Streamlines Trait Gene Identification in Maize. Plant Cell **32**:1397–1413.

Liu, T.Z., Zhang, X.N., Li, K., Yao, Q., Zhong, D.T., Deng, Q., and Lu, Y.M. (2023). Large-scale Genome Editing in Plants: Approaches, Applications, and Future Perspectives (Curr Opin Biotech), p. 79.

Lu, Y., Wu, K., Jiang, Y., Xia, B., Li, P., Feng, H., Wyckhuys, K.A.G., and Guo, Y. (2010). Mirid bug outbreaks in multiple crops correlated with wide-scale adoption of Bt cotton in China. Science **328**:1151–1154.

Lu, Y., Ye, X., Guo, R., Huang, J., Wang, W., Tang, J., Tan, L., Zhu, J.K., Chu, C., and Qian, Y. (2017). Genome-wide Targeted Mutagenesis in Rice Using the CRISPR/Cas9 System. Mol. Plant **10**:1242–1245.

Markham, G.D., Hafner, E.W., Tabor, C.W., and Tabor, H. (1983). S-Adenosylmethionine Synthetase (Methionine Adenosyltransferase) (Escherichia-Coli). Method Enzymol **94**:219–222.

Pan, C., Li, G., Bandyopadhyay, A., and Qi, Y. (2023). Guide RNA library-based CRISPR screens in plants: opportunities and challenges. Curr. Opin. Biotechnol. **79**:102883.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. **81**:559–575.

Qiao, F., Huang, J., and Wang, X. (2017). Fifteen Years of Bt Cotton in China: Results from Household Surveys. World Dev. **98**:351–359.

Quan, Y., and Wu, K. (2023). Managing Practical Resistance of Lepidopteran Pests to Bt Cotton in China. Insects **14**:179.

Ray, S., Agarwal, P., Arora, R., Kapoor, S., and Tyagi, A.K. (2007). Expression analysis of calcium-dependent protein kinase gene family during reproductive development and abiotic stress conditions in rice (Oryza sativa L. ssp. indica). Mol. Genet. Genomics. **278**:493–505.

Romeis, T., and Herde, M. (2014). From local to global: CDPKs in systemic defense signaling upon microbial and herbivore attack. Curr. Opin. Plant Biol. **20**:1–10.

Romeis, T., Ludwig, A.A., Martin, R., and Jones, J.D. (2001). Calcium-dependent protein kinases play an essential role in a plant defence response. EMBO J. **20**:5556–5567.

Sanders, D., Pelloux, J., Brownlee, C., and Harper, J.F. (2002). Calcium at the crossroads of signaling. Plant Cell **14**:S401–S417.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelson, T., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., et al. (2014). Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. Science **343**:84–87.

Si, H., Liu, H., Sun, Y., Xu, Z., Liang, S., Li, B., Ding, X., Li, J., Wang, Q., Sun, L., et al. (2020). Transcriptome and metabolome analysis reveal that oral secretions from Helicoverpa armigera and Spodoptera litura influence wound-induced host response in cotton. Crop J. **8**:929–942.

Simeunovic, A., Mair, A., Wurzinger, B., and Teige, M. (2016). Know where your clients are: subcellular localization and targets of calcium-dependent protein kinases. J. Exp. Bot. **67**:3855–3872.

Sun, L., Alariqi, M., Wang, Y., Wang, Q., Xu, Z., Zafar, M.N., Yang, G., Jia, R., Hussain, A., Chen, Y., et al. (2024). Construction of Host Plant Insect-Resistance Mutant Library by High-Throughput CRISPR/Cas9 System and Identification of A Broad-Spectrum Insect Resistance Gene. Adv. Sci. **11**:e2306157.

Sun, L., Zhu, L., Xu, L., Yuan, D., Min, L., and Zhang, X. (2014). Cotton cytochrome P450 CYP82D regulates systemic cell death by modulating the octadecanoid pathway. Nat. Commun. **5**:5372.

Tu, L., Zhang, X., Liu, D., Jin, S., Cao, J., Zhu, L., Deng, F., Tan, J., and Zhang, C. (2007). Suitable internal control genes for qRT-PCR normalization in cotton fiber development and somatic embryogenesis. Chinese Sci Bull **22**:3110–3117.

Valmonte, G.R., Arthur, K., Higgins, C.M., and MacDiarmid, R.M. (2014). Calcium-dependent protein kinases in plants: evolution, expression and function. Plant Cell Physiol. **55**:551–569.

Varet, H., Brillet-Guéguen, L., Coppée, J.Y., and Dillies, M.A. (2016). SARTools: A DESeq2- and EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data. PLoS One **11**:e0157022.

Wang, H., Li, X., Su, F., Liu, H., Hu, D., Huang, F., Yu, D., and Wang, H. (2022). Soybean CALCIUM-DEPENDENT PROTEIN KINASE17 Positively Regulates Plant Resistance to Common Cutworm (Spodoptera litura Fabricius). Int. J. Mol. Sci. **23**:15696.

Wang, J., Liu, X., Zhang, A., Ren, Y., Wu, F., Wang, G., Xu, Y., Lei, C., Zhu, S., Pan, T., et al. (2019a). A cyclic nucleotide-gated channel mediates cytoplasmic calcium elevation and disease resistance in rice. Cell Res. **29**:820–831.

Wang, K., Wang, Z., Li, F., Ye, W., Wang, J., Song, G., Yue, Z., Cong, L., Shang, H., Zhu, S., et al. (2012). The draft genome of a diploid cotton Gossypium raimondii. Nat. Genet. **44**:1098–1103.

Wang, M., Tu, L., Yuan, D., Zhu, D., Shen, C., Li, J., Liu, F., Pei, L., Wang, P., Zhao, G., et al. (2019b). Reference genome sequences of two cultivated allotetraploid cottons, Gossypium hirsutum and Gossypium barbadense. Nat. Genet. **51**:224–229.

Wang, P., Zhang, J., Sun, L., Ma, Y., Xu, J., Liang, S., Deng, J., Tan, J., Zhang, Q., Tu, L., et al. (2018). High efficient multisites genome editing in allotetraploid cotton (Gossypium hirsutum) using CRISPR/Cas9 system. Plant Biotechnol. J. **16**:137–150.

Wendel, J.F. (1989). New World Tetraploid Cottons Contain Old-World Cytoplasm. P Natl Acad Sci USA **86**:4132–4136.

Wu, K.M., and Guo, Y.Y. (2005). The evolution of cotton pest management practices in China. Annu. Rev. Entomol. **50**:31–52.

Wu, K.M., Lu, Y.H., Feng, H.Q., Jiang, Y.Y., and Zhao, J.Z. (2008). Suppression of cotton bollworm in multiple crops in China in areas with Bt toxin-containing cotton. Science **321**:1676–1678.

Xu, X., Liu, M., Lu, L., He, M., Qu, W., Xu, Q., Qi, X., and Chen, X. (2015). Genome-wide analysis and expression of the calcium-dependent protein kinase gene family in cucumber. Mol. Genet. Genomics. **290**:1403–1414.

Yang, D.H., Hettenhausen, C., Baldwin, I.T., and Wu, J. (2012). Silencing Nicotiana attenuata calcium-dependent protein kinases, CDPK4 and CDPK5, strongly up-regulates wound- and herbivory-induced jasmonic acid accumulations. Plant Physiol. **159**:1591–1607.

Yang, Z., Ge, X., Yang, Z., Qin, W., Sun, G., Wang, Z., Li, Z., Liu, J., Wu, J., Wang, Y., et al. (2019). Extensive intraspecific gene order and gene structural variations in upland cotton cultivars. Nat. Commun. **10**:2989.

Ye, S., Wang, L., Xie, W., Wan, B., Li, X., and Lin, Y. (2009). Expression profile of calcium-dependent protein kinase (CDPKs) genes during the whole lifespan and under phytohormone treatment conditions in rice (Oryza sativa L. ssp. indica). Plant Mol. Biol. **70**:311–325.

Yip Delormel, T., and Boudsocq, M. (2019). Properties and functions of calcium-dependent protein kinases and their relatives in Arabidopsis thaliana. New Phytol. **224**:585–604.

Zhang, H., Wei, C., Yang, X., Chen, H., Yang, Y., Mo, Y., Li, H., Zhang, Y., Ma, J., Yang, J., et al. (2017). Genome-wide identification and

expression analysis of calcium-dependent protein kinase and its related kinase gene families in melon (Cucumis melo L.). PLoS One **12**:e0176352.

Zhang, T., Hu, Y., Jiang, W., Fang, L., Guan, X., Chen, J., Zhang, J., Saski, C.A., Scheffler, B.E., Stelly, D.M., et al. (2015). Sequencing of allotetraploid cotton (Gossypium hirsutum L. acc. TM-1) provides a resource for fiber improvement. Nat. Biotechnol. **33**:531–537.

Zhu, H., He, M., Jahan, M.S., Wu, J., Gu, Q., Shu, S., Sun, J., and Guo, S. (2021). CsCDPK6, a CsSAMS1-Interacting Protein, Affects Polyamine/Ethylene Biosynthesis in Cucumber and Enhances Salt Tolerance by Overexpression in Tobacco. Int. J. Mol. Sci. **22**:11133.

Zuo, R., Hu, R., Chai, G., Xu, M., Qi, G., Kong, Y., and Zhou, G. (2013). Genome-wide identification, classification, and expression analysis of CDPK and its closely related gene families in poplar (Populus trichocarpa). Mol. Biol. Rep. **40**:2645–2662.

Plant Communications, Volume *5*

Supplemental information

CRISPR–Cas9-mediated construction of a cotton *CDPK* mutant library for identification of insect-resistance genes

Fuqiu Wang, Sijia Liang, Guanying Wang, Tianyu Hu, Chunyang Fu, Qiongqiong Wang, Zhongping Xu, Yibo Fan, Lianlian Che, Ling Min, Bo Li, Lu Long, Wei Gao, Xianlong Zhang, and Shuangxia Jin

Figure S1

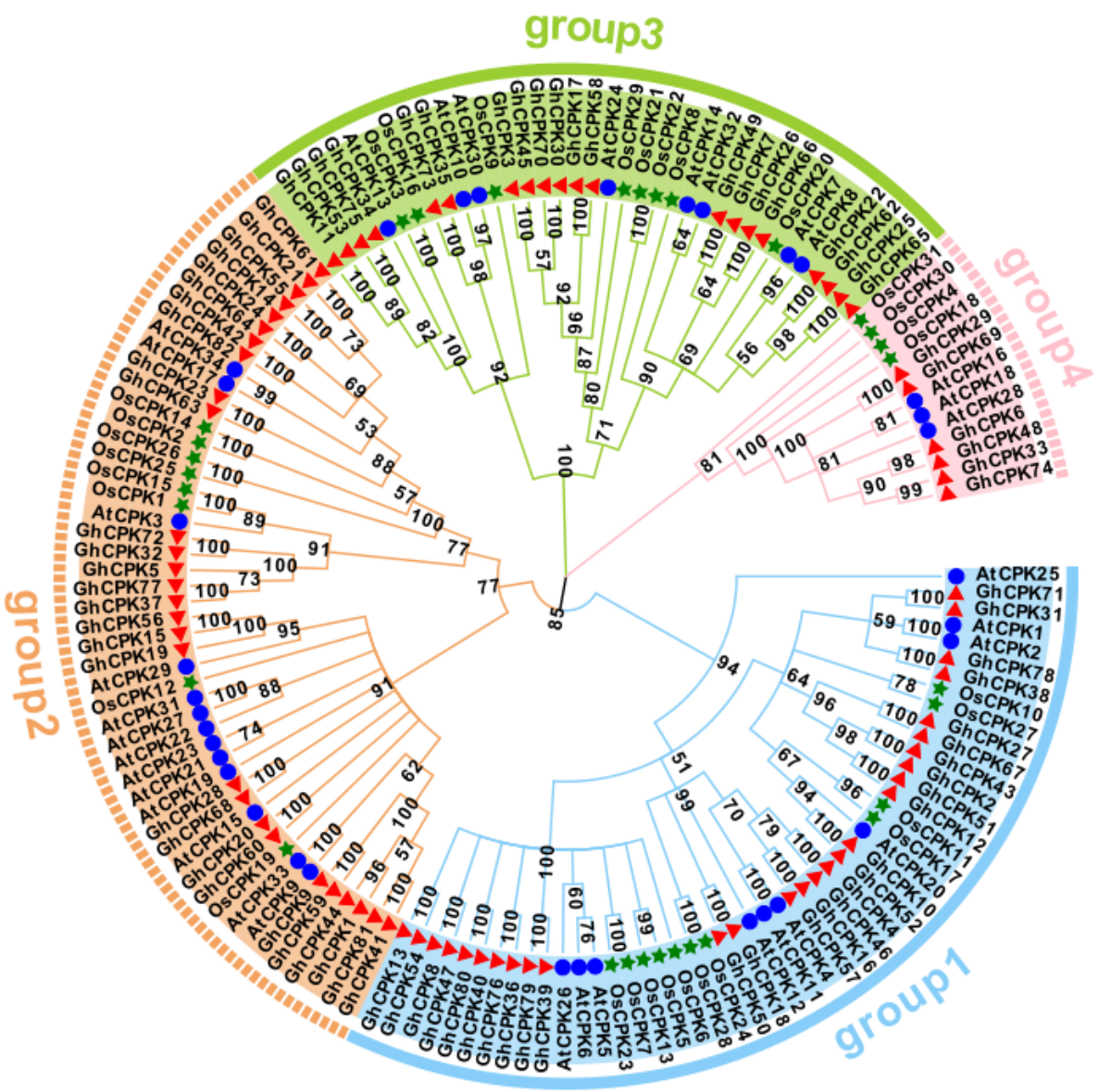

Figure S1 Phylogenetic analysis of *CDPK* gene families in *G. hirsutum*, *A. thaliana* and *O. sativa*. Groups 1, 2, 3, and 4 indicate different gene clusters. Different colors of symbols represent different species. Red triangles represent *G. hirsutum*, blue circles represent *A. thaliana* and green pentagrams represent *O. sativa*.

Figure S2

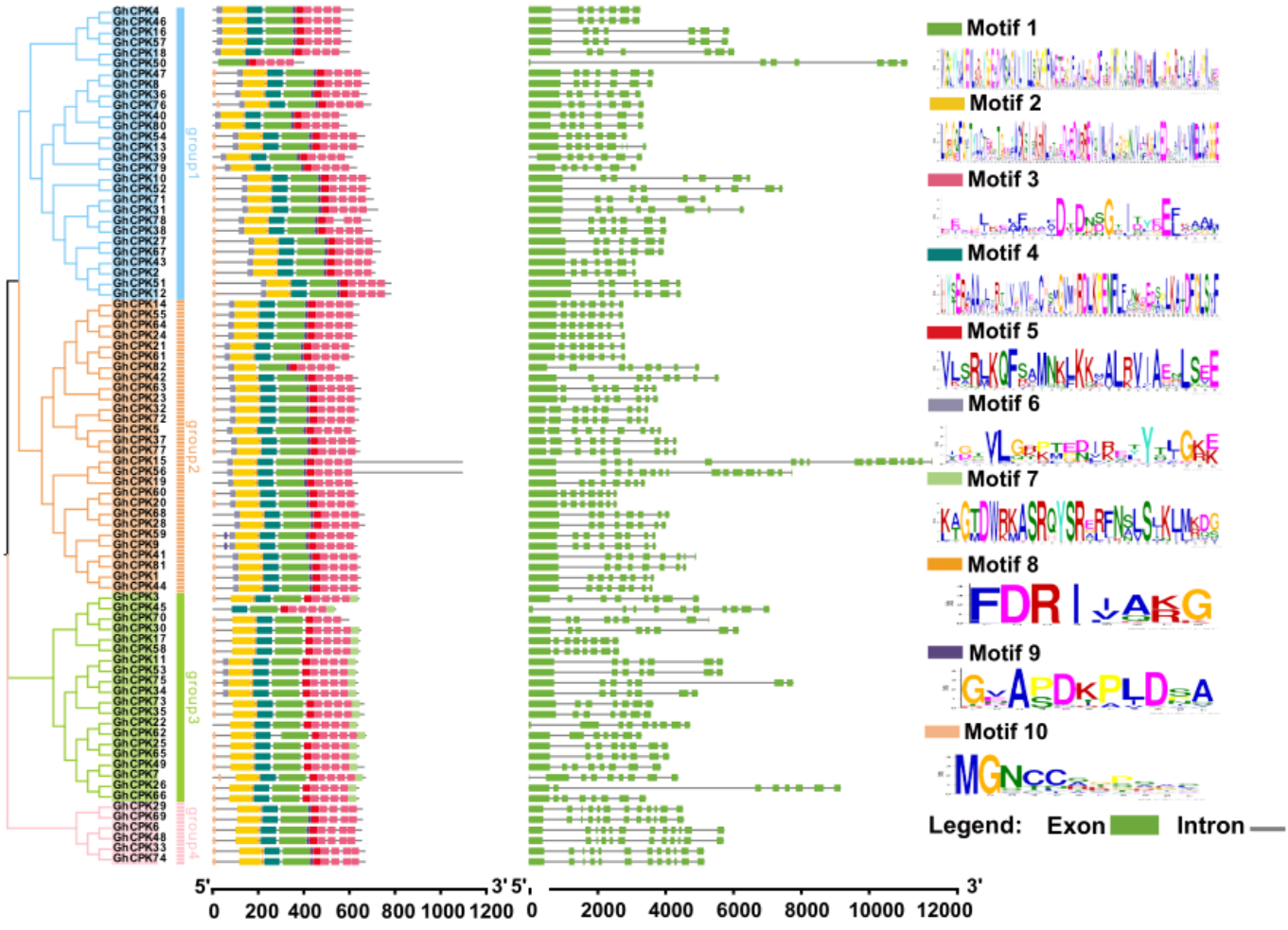

Figure S2 Phylogenetic tree, conserved protein motifs and gene structures of *GhCPKs* gene family. From left to right, it shows the phylogenetic tree, conserved protein motifs and gene structure of the *G. hirsutum CDPK* gene family. In the phylogenetic tree, different colors represent different subgroups. Within the protein conserved motifs, a total of 10 conserved motifs have been identified, distributed in different colored boxes. On the far right are the sequence logos of these protein motifs. In the exon-intron structure diagram, the green box represents exons and the black line represents introns.

Figure S3

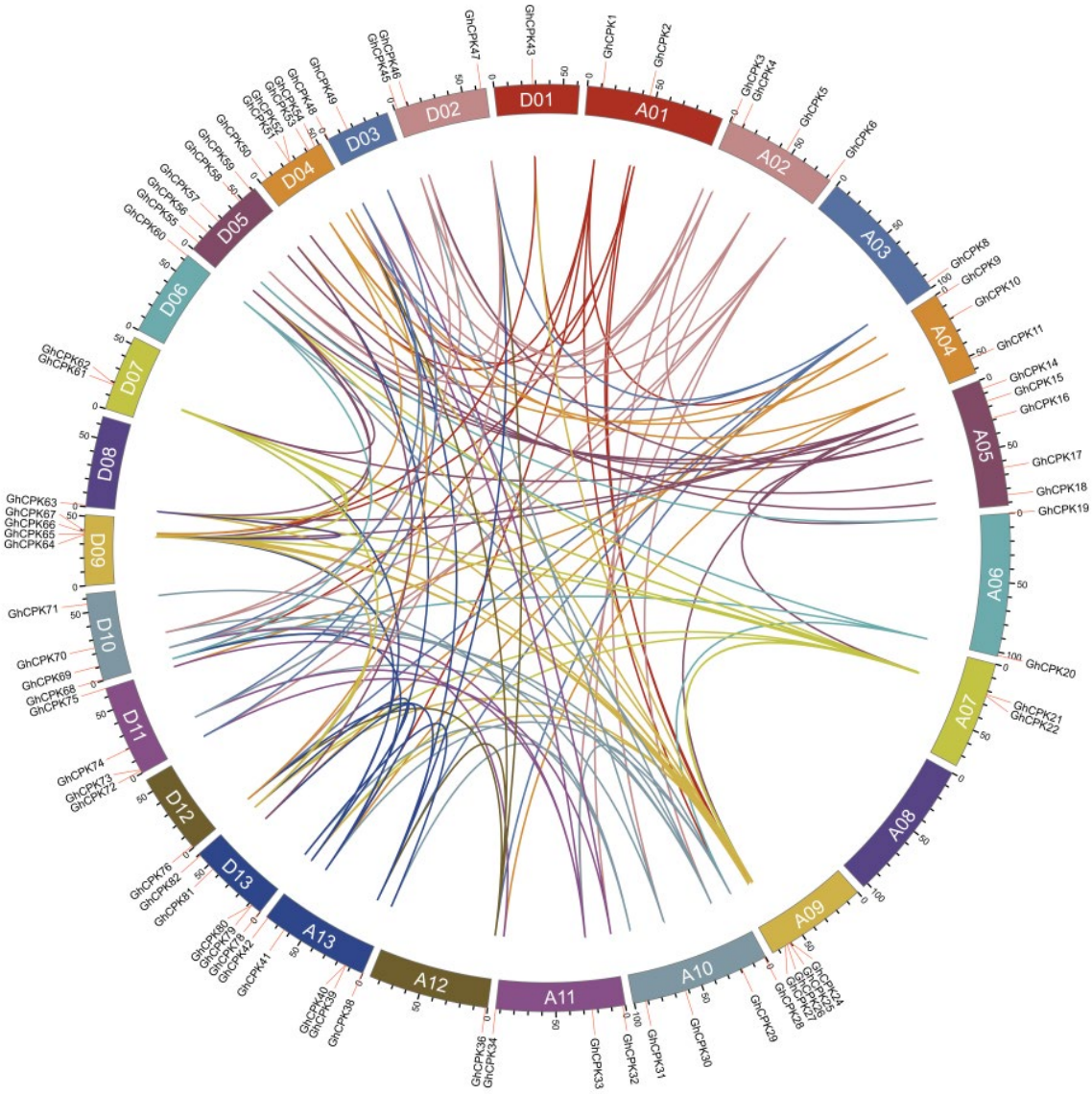

Figure S3 Collinearity analysis of the *GhCPKs* gene family. Chromosome of *G. hirsutum* was distinguished by different color. Each *GhCPK* gene is marked with a short red line on the chromosome and collinear gene pairs are represented by a color curve.

Figure S4

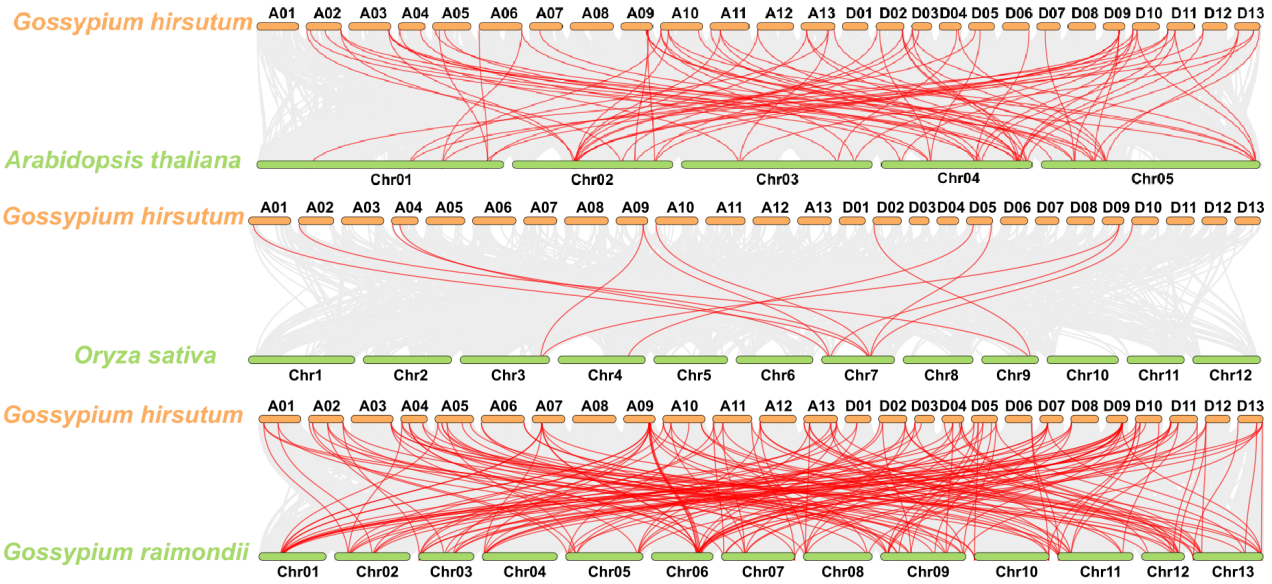

Figure S4 Collinearity analysis of *CDPK* genes between *G. hirsutum* with *A. thaliana*, *O. sativa* and *G. raimondii*. Grey lines indicate collinear blocks within the *G. hirsutum* genome and other plant genomes, and the red curve indicates *CDPK* genes with collinearity.

Figure S5

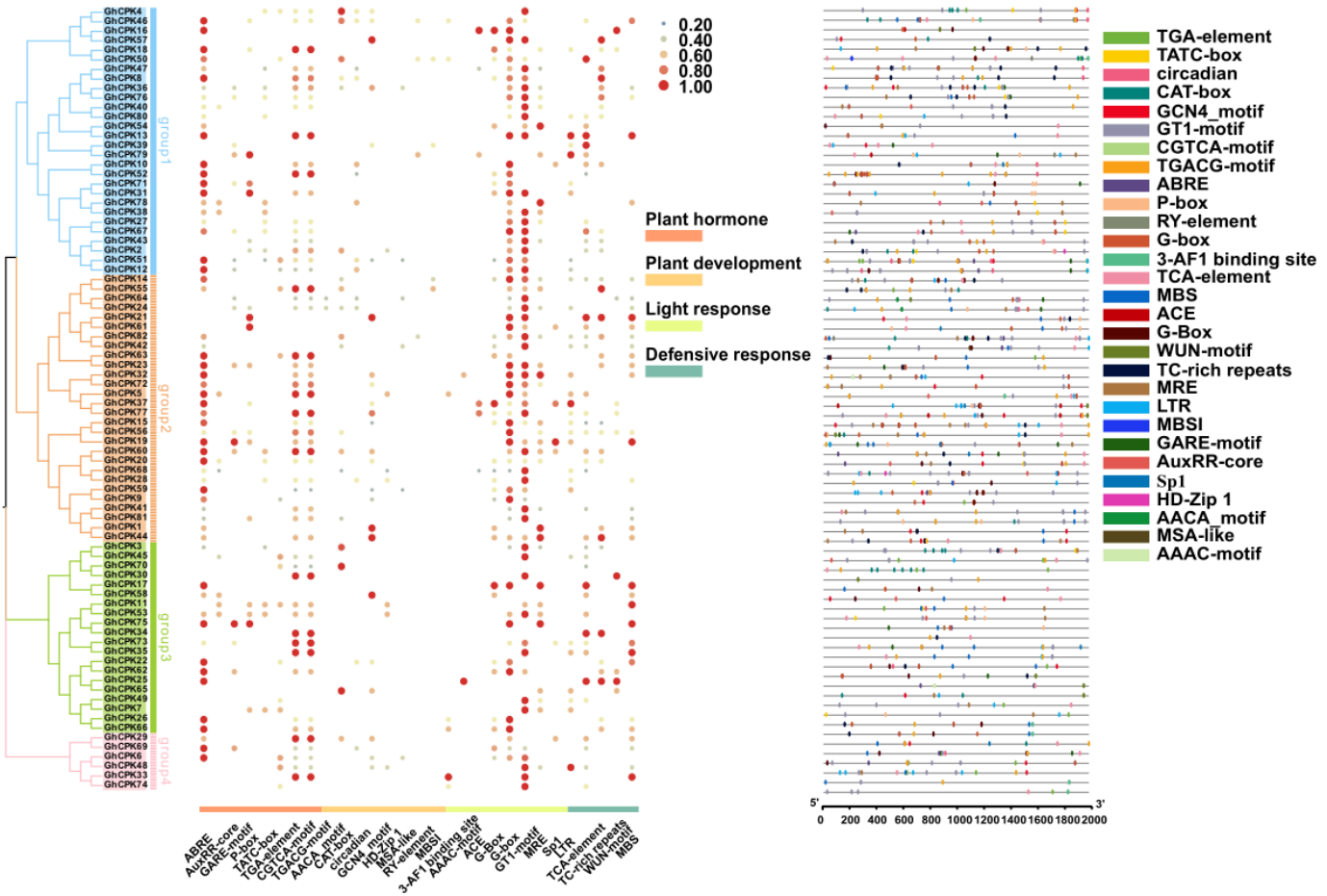

Figure S5 Cis-acting elements in the promoter regions of *GhCPKs*. On the left is the phylogenetic tree of the *GhCPKs* gene family. In the middle is a heat map of the number of all cis-acting elements. The larger and redder the circle, the greater the quantity represented. The diagram on the right shows the positional distribution of promoter cis-acting elements of the *GhCPK* genes. The black line indicates the promoter length of the *GhCPK* genes. The different colored boxes on the right represent cis-acting elements with different functions.

Figure S6

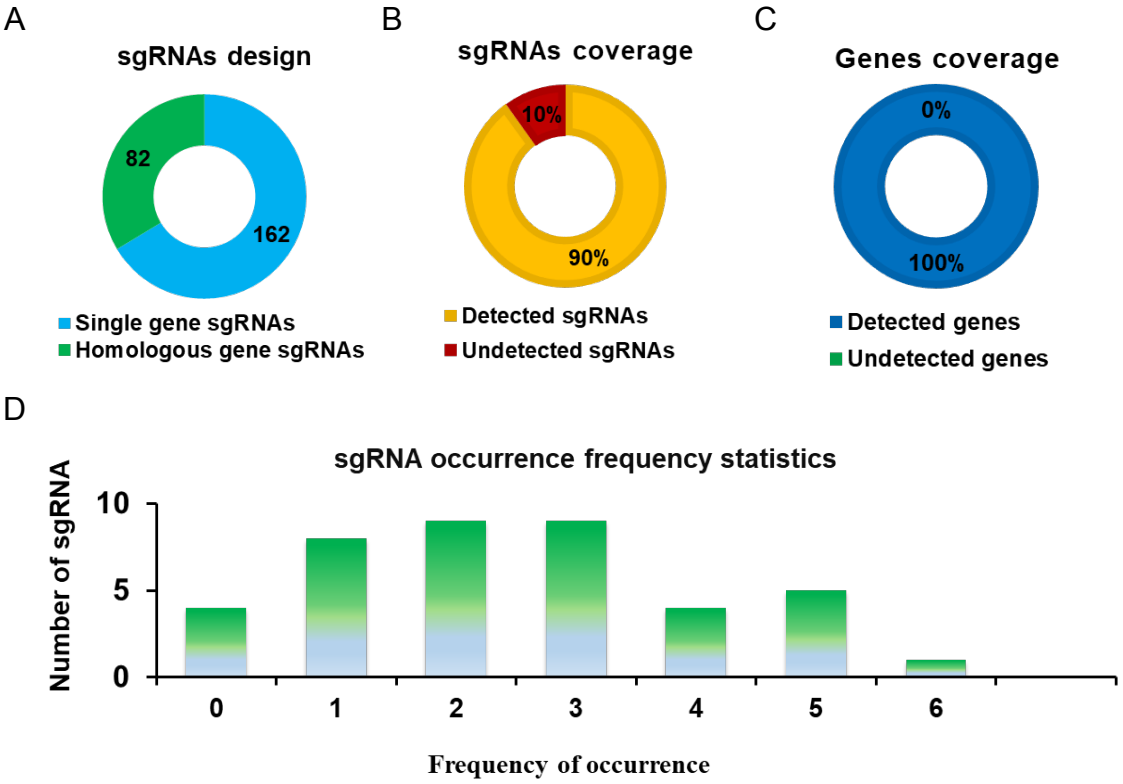

Figure S6 The number of sgRNAs in the *GhCPKs* gene family mutant library and the preliminary experimental results of mixed pool 1. (A) The quantity of different types of sgRNAs. Blue represents the quantity of sgRNAs targeting individual *GhCPKs*, green represents the quantity of sgRNAs targeting homologous *GhCPKs*. (B) Coverage of sgRNA in the preliminary experiment. Yellow color represents detected sgRNAs, red color represents undetected sgRNAs. (C) Coverage of gene in the preliminary experiment. Deep blue color represents detected genes, green color represents undetected genes.(D) Frequency statistics of sgRNA. The X-axis represents the frequency of occurrence of target sgRNAs, and the Y-axis represents the number of sgRNAs.

Figure S7

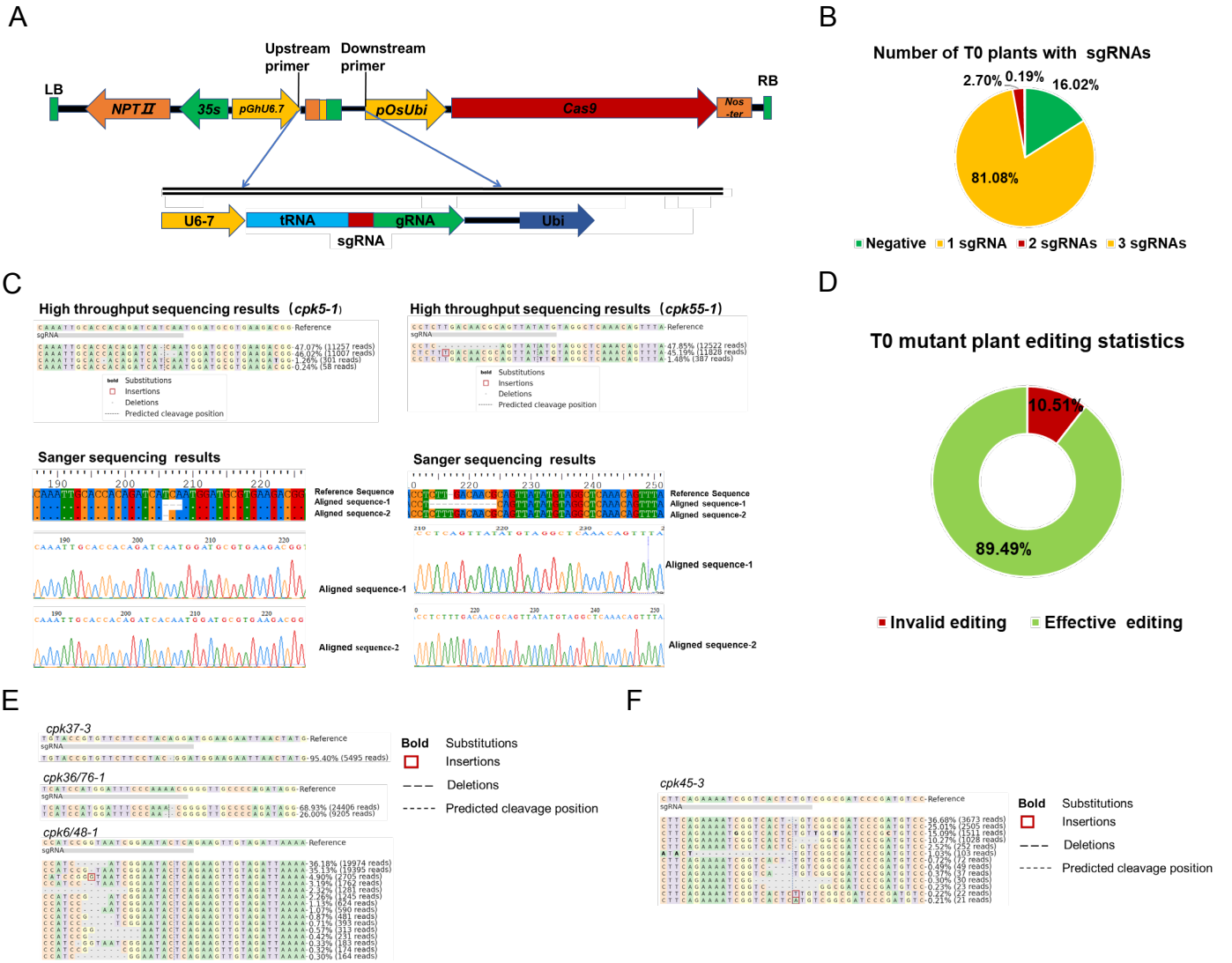

Figure S7 Molecular detection of T0 generation *GhCPKs* mutant library plants. (A) PCR amplification sgRNA vector sequence diagram. The different colored boxes represent different vector components. (B) Statistical analysis of the number of sgRNAs present in the T0 generation *GhCPKs* mutant library plants. Yellow means only one sgRNA, red means two sgRNA, and green means three sgRNA. (C) Comparison of high-throughput sequencing and Sanger sequencing results for gene editing plants. (D) Editing statistics of T0 generation *GhCPKs* mutant library plants. Green means that gene editing is detected, red means that gene editing is not detected. (E) Examples of one, two and multiple types of editing. Bold represents base substitutions, red boxes represent base insertions, dashes represent base deletions, and dashed lines represent predicted cleavage positions. (F) The most common editing types in the mutant library. Bold represents base substitutions, red boxes represent base insertions, dashes represent base deletions, and dashed lines represent predicted cleavage positions.

Figure S8

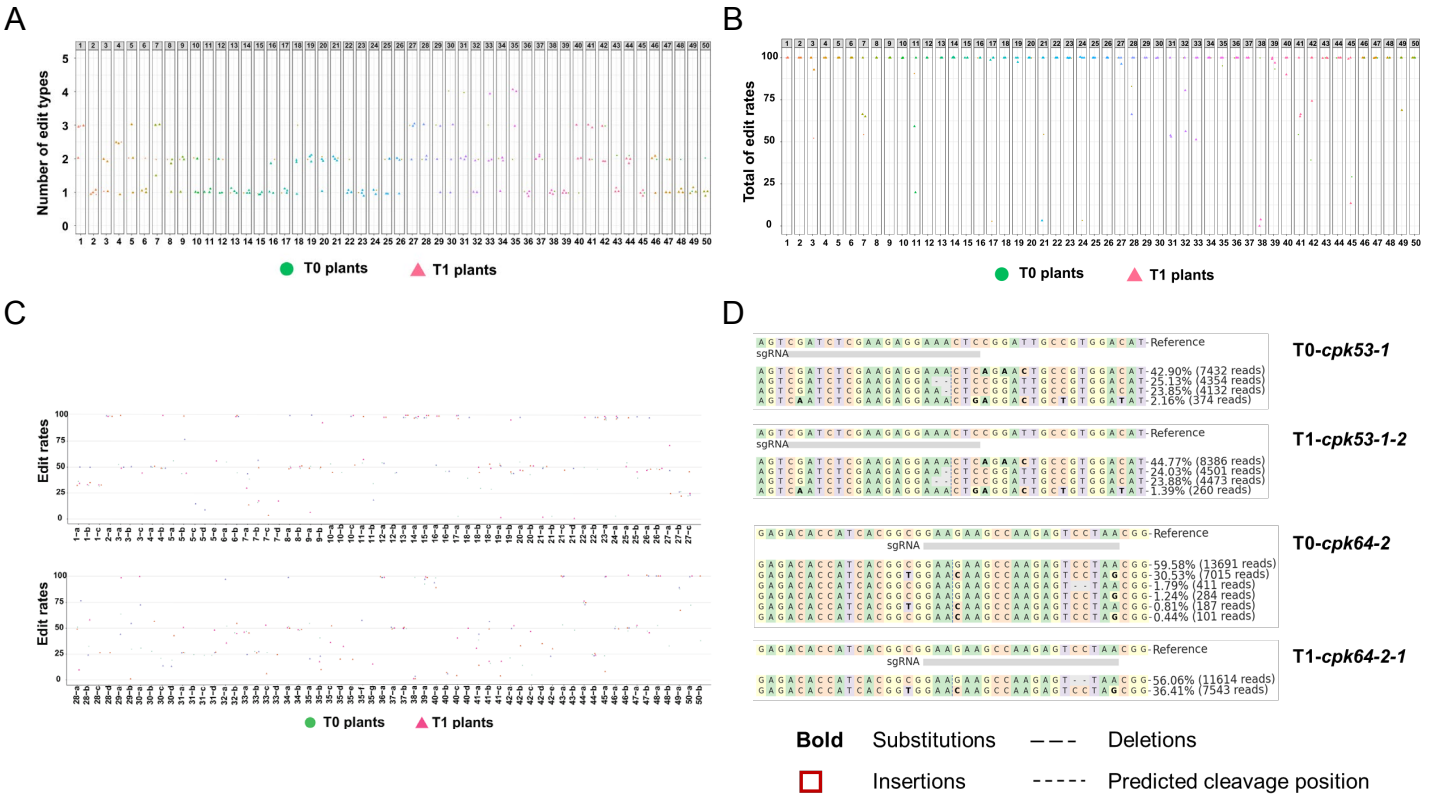

Figure S8 Genetic analysis of 50 *GhCPKs* gene-edited plants and their 150 progeny. (A) Editing type statistics of 50 *GhCPKs* gene-edited plants and their 150 progeny. The circle represents T0 generation plants, the triangle represents T1 generation plants. (B) Editing efficiency statistics of 50 *GhCPKs* gene-edited plants and their 150 progeny. (C) Statistical analysis of editing efficiency for each editing type in 50 *GhCPKs* gene-edited plants and their 150 progeny. (D)) Examples of genetic states in different progenitor and progeny gene-edited plants. Bold represents base substitutions, red boxes represent base insertions, dashes represent base deletions, and dashed lines represent predicted cleavage positions.

Figure S9

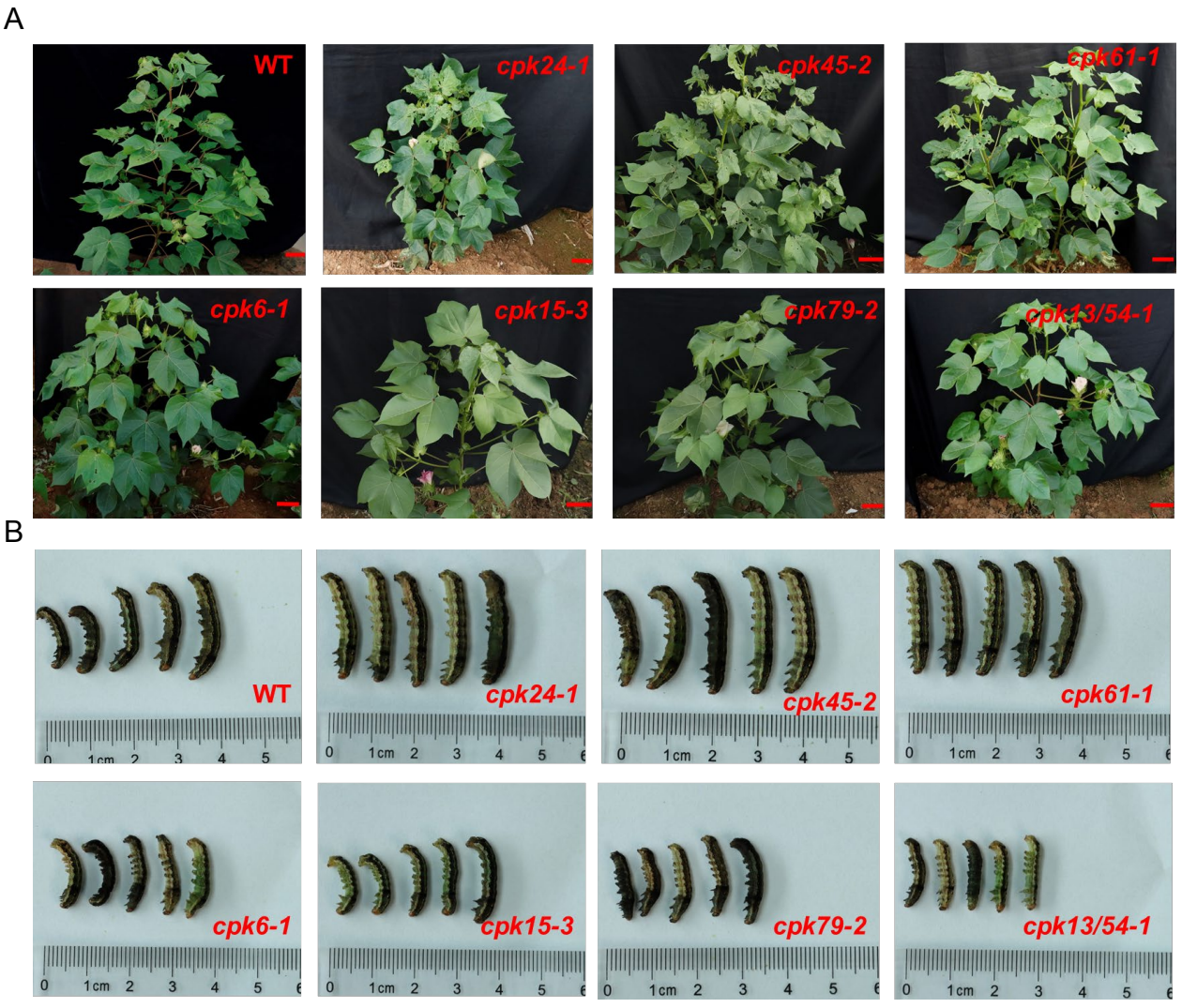

Figure S9 Example of insect resistance phenotypic screening in *GhCPKs* mutant library plants. (A) Field insect resistance phenotype of corresponding numbered plants. (B) Phenotype of *S. litura* larvae feeding on the corresponding numbered plants five days later.

Figure S10

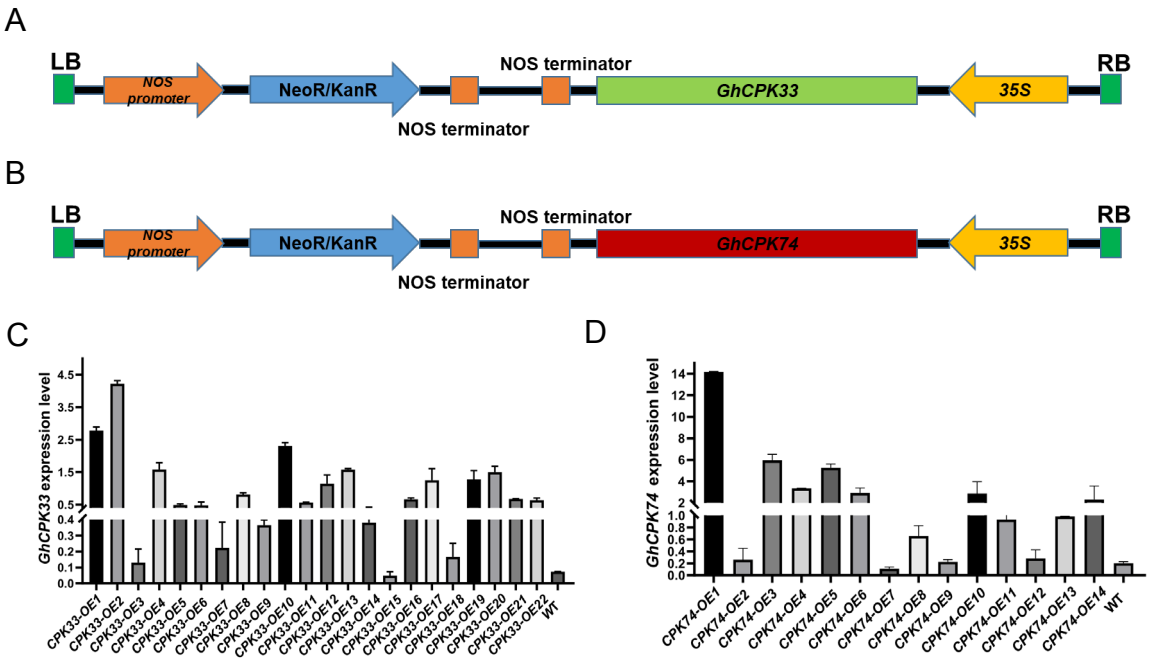

Figure S10 Construction of expression vectors and detection of expression levels for *GhCPK33* and *GhCPK74* overexpression materials. (A) Construction of *GhCPK33* overexpression vectors. (B) Construction of *GhCPK74* overexpression vectors. Different color boxes represent different carrier elements. (C) Detection of *GhCPK33* expression levels in T0-generation *CPK33-OE* overexpression materials. (D) Detection of *GhCPK74* expression levels in T0-generation *CPK74-OE* overexpression materials. The expression of *GhUB7* was used as internal control. Means ± SE (n = 3).

Figure S11

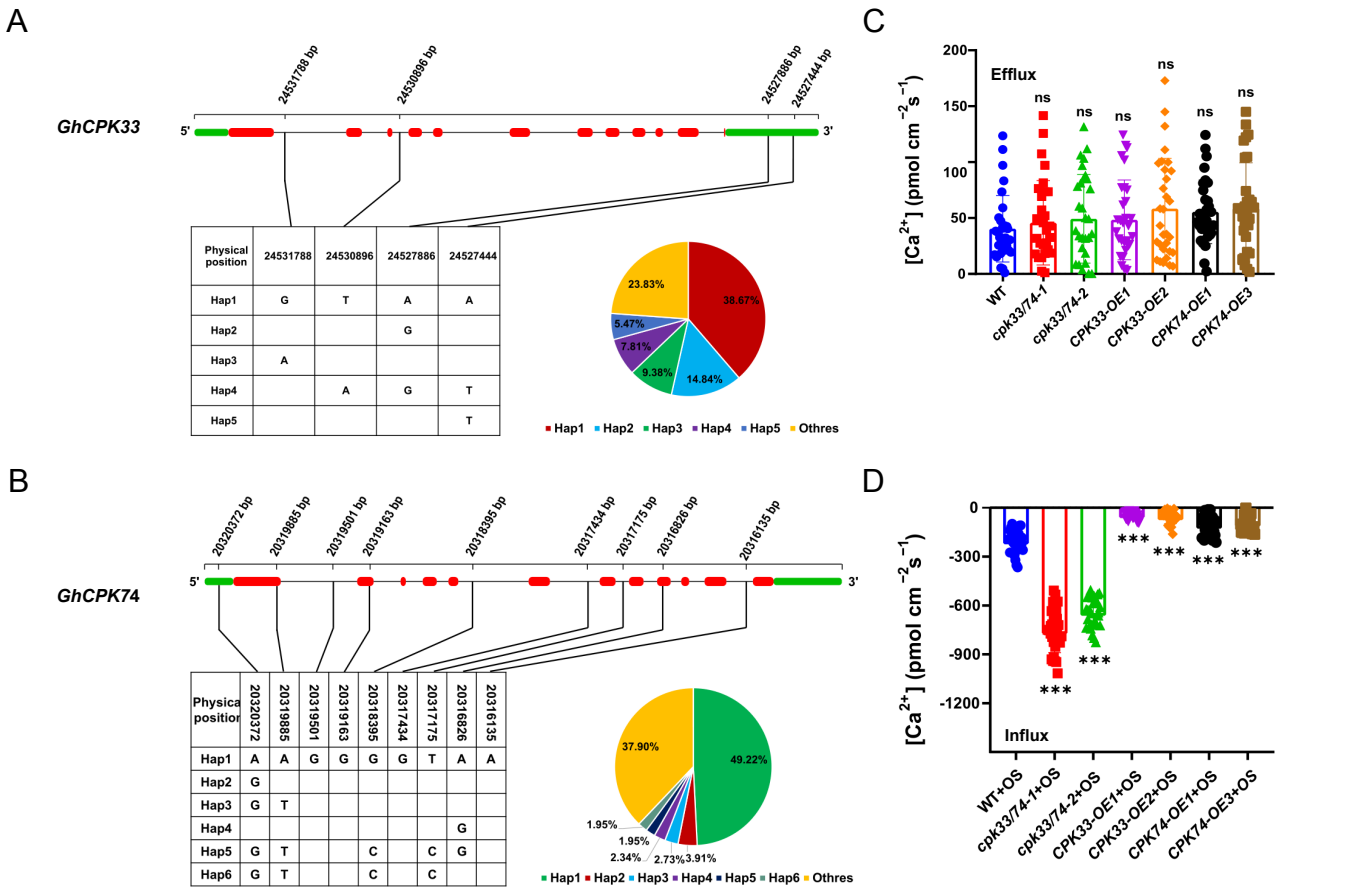

Figure S11 Haplotype analysis of *GhCPK33* and *GhCPK74* and the detection of average Ca^{2+} flow rate in leaf mesophyll cells. (A) The gene *GhCPK33* structure and single-nucleotide polymorphisms (SNPs) variants contained in the Ghlandraces population showed the five haplotypes with highest proportions, and the blank boxes were the same nucleotides as the reference genome (Hap1). The pie charts show the proportion of the five haplotypes. (B) The gene *GhCPK74* structure and single-nucleotide polymorphisms (SNPs) variants contained in the Ghlandraces population showed the six haplotypes with highest proportions, and the blank boxes were the same nucleotides as the reference genome (Hap1). The pie charts show the proportion of the six haplotypes. (C) In the untreated state, the average Ca^{2+} flow rate in leaf mesophyll cells of *cpk33/74*, *CPK33-OE*, *CPK74-OE* and WT over a 5-minute period. (D) After treatment with the OS of *S. litura*, the average Ca^{2+} flow rate in leaf mesophyll cells of *cpk33/74*, *CPK33-OE*, *CPK74-OE* and WT over a 5-minute period.

Figure S12

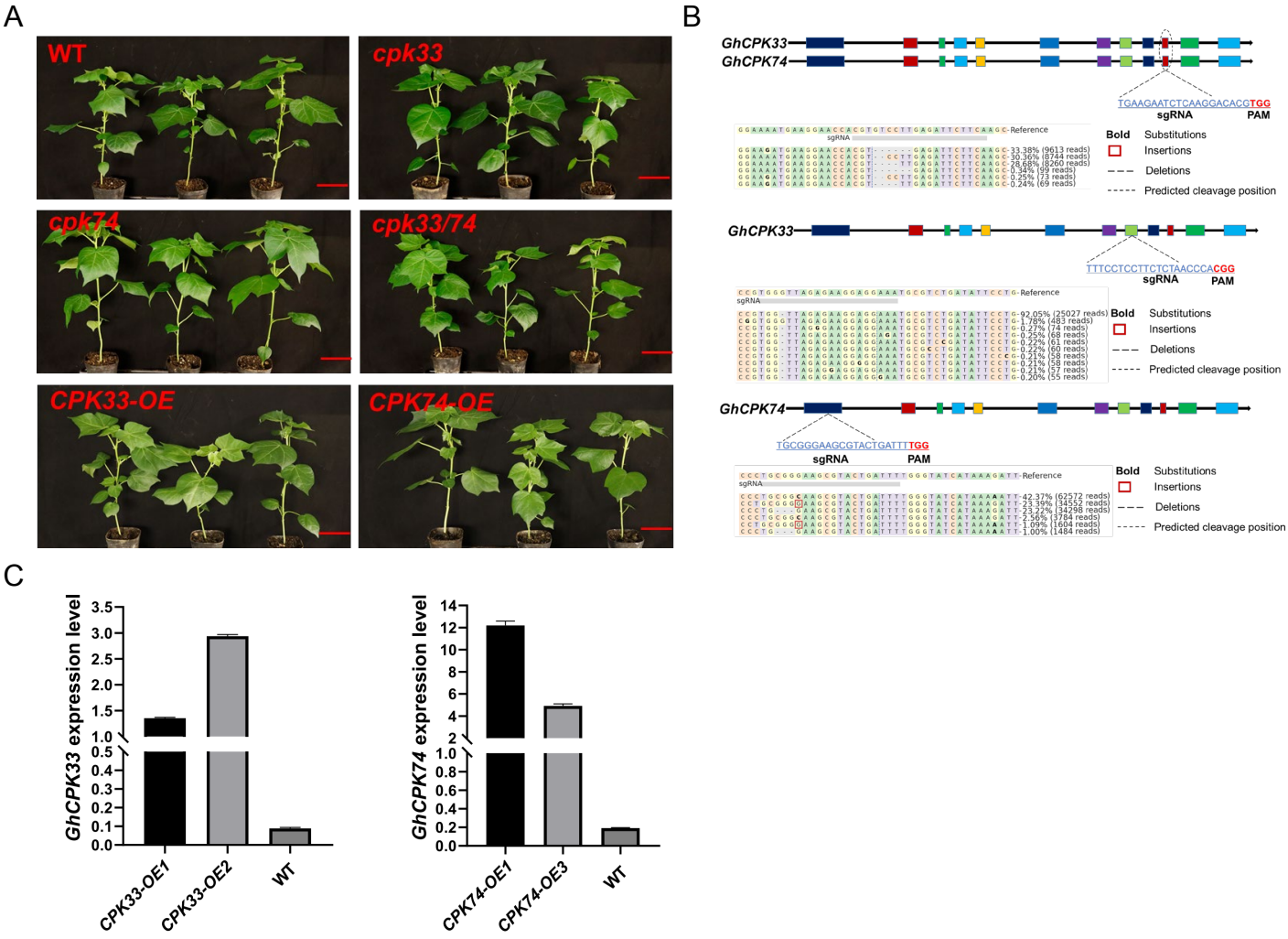

Figure S12 Cultivation, gene editing and expression level detection of T2 generation *cpk33/74*, *cpk33*, *cpk74*, *CPK33-OE*, *CPK74-OE* and WT plants. (A) Cultivation of T2 generation *cpk33/74*, *cpk33*, *cpk74*, *CPK33-OE*, *CPK74-OE* and WT plants. (B) Detection of editing in T2 generation *cpk33, cpk74* and *cpk74/33* plants. (C) Detection of *GhCPK33* and *GhCPK74* expression levels in T2 generation *CPK33-OE* and *CPK74-OE* overexpression materials. The expression of *GhUB7* was used as internal control. Means ± SE (n = 3).

Figure S13

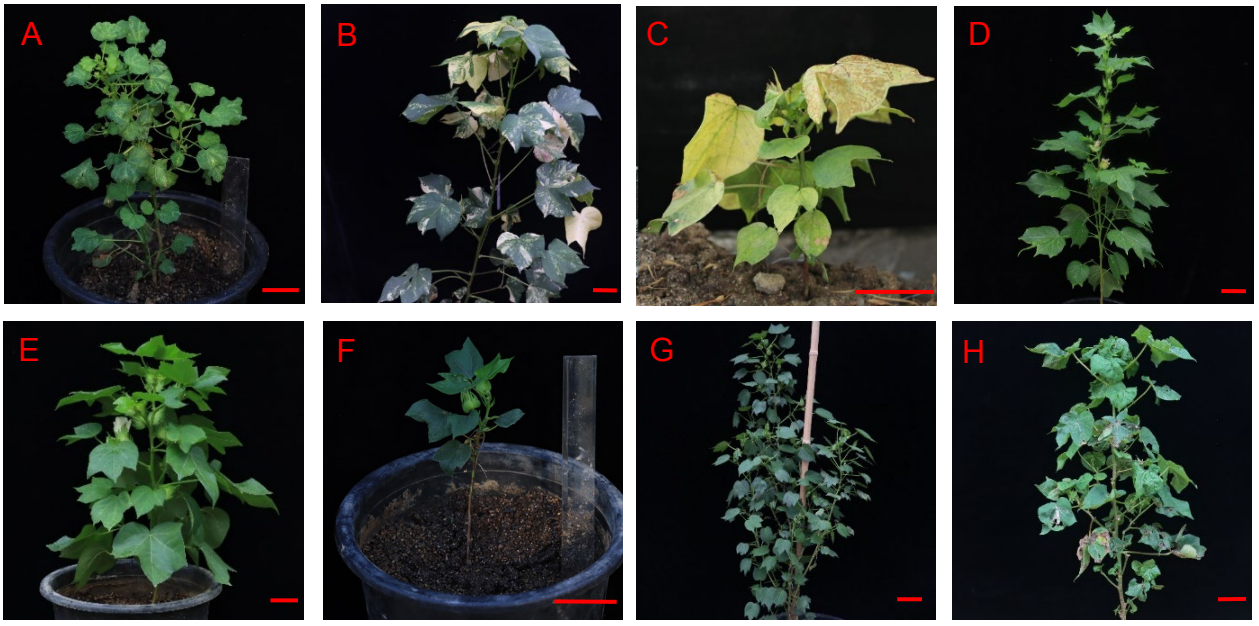

Figure S13 Identification of unique phenotypes in the T0 generation *GhCPKs* mutant library plants. (A) The leaf is round. (B) The leaves showed chimeric yellowing. (C) The plant type is dwarfed and the leaves turn yellow. (D) The fruit branches become shorter. (E) Dwarfing and compact plants. (F) Plant dwarfing and early maturing. (G) The leaves become smaller. (H) Necrotic spots appeared on the leaves. The red line represents 5 cm.

Figure S14

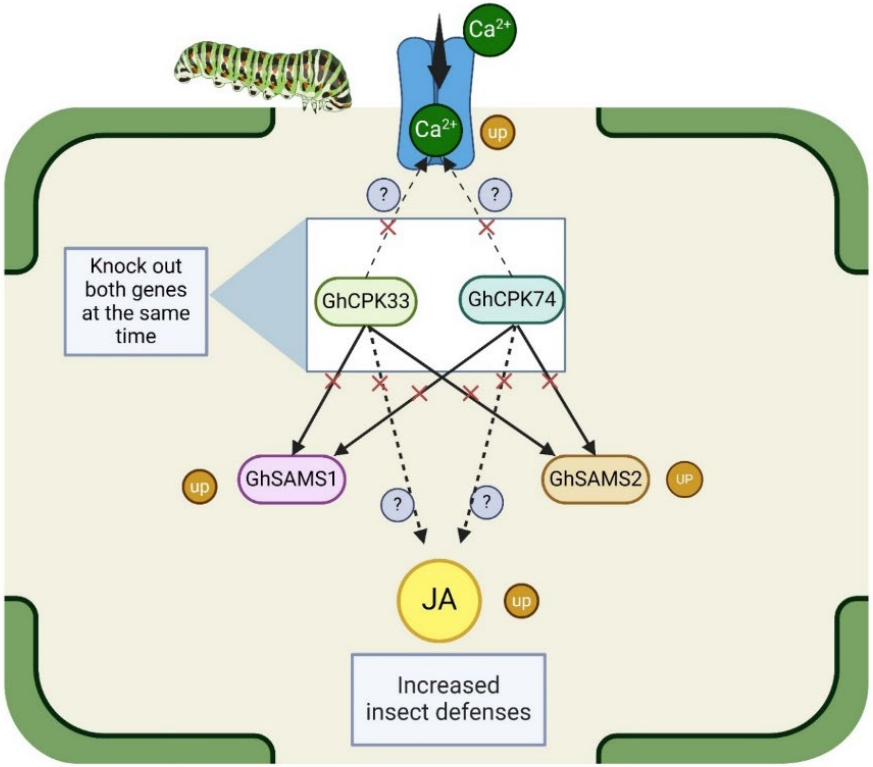

Figure S14 Models of *GhCPK33* and *GhCPK74* regulating pest defense in upland cotton. Under the induction of OS from the *S. litura*, simultaneously knocking out *GhCPK33* and *GhCPK74* can accelerate the influx of Ca^{2+} in cotton leaf mesophyll cells, enhance the synthesis of JA, and increase the expression of *GhSAMS1* and *GhSAMS2*, resulting in an increased defense capability of cotton against herbivorous insects.

Table S1 Physical and chemical aspects of *CDPK* genes in the *Gossypium hirsutum*

Gene name	NAU gene ID	HAU gene ID	Chromosomal Location				Putative Proteins			N-terminal	N-Myristoylation	S-Palmitoylation	No. of EF hands	Group	Subcellular location	Transmembrane Domain	Signal Peptide (Sec/SPI)
			Chr	Chr_start	Chr_end	Direction	length(aa)	MW(Da)	pI								
GhCPK1	Gh_A01G0621	Ghir_A01G007100.1	A01	11154745	11157809	+	537	60535.5	6.47	MGTCLTKS	Y	Y	4	II	M	0	N
GhCPK2	Gh_A01G1119	Ghir_A01G012430.1	A01	45532122	45534741	+	590	66237.2	5.24	MGNSCAKS	Y	Y	4	I	C|M	0	N
GhCPK3	Gh_A02G0144	Ghir_A02G001700.1	A02	1508323	1512491	-	532	60688.1	5.59	MGSCVARP	Y	Y	4	III	M	0	N
GhCPK4	Gh_A02G0609	Ghir_A02G006230.1	A02	9500456	9503183	+	511	57142.6	5.51	MNNQSSSI	N	Y	4	I	C	0	N
GhCPK5	Gh_A02G1029	Ghir_A02G011170.1	A02	45630623	45633866	+	519	58599.3	5.77	MGNCNGLP	N	Y	4	II	C|M	0	N
GhCPK6	Gh_A02G1635	Ghir_A02G018670.1	A02	82802954	82807746	+	541	61333.5	9.26	MGACLSAT	Y	Y	4	IV	M	0	N
GhCPK7	Gh_A02G1796	Ghir_A02G012330.1	scaffold430_A02	50467	54125	+	555	63114.6	6.93	MVSNNKKN	N	N	4	III	M	0	Y
GhCPK8	Gh_A03G1505	Ghir_A03G018780.1	A03	95507304	95510335	-	568	63601.5	6.02	MGNTCRGS	N	Y	4	I	M	0	N
GhCPK9	Gh_A04G0148	Ghir_A04G001660.1	A04	2396366	2399478	-	524	58823.5	6.36	MGGCLTKT	Y	Y	4	II	M	0	N
GhCPK10	Gh_A04G0467	Ghir_A04G005600.1	A04	19716614	19722043	-	572	63553.5	5.24	MGNACAGP	N	Y	4	I	C|M	0	N
GhCPK11	Gh_A04G0780	Ghir_A04G009910.1	A04	52438907	52443668	-	527	59393.3	6.39	MGNCCRSP	N	Y	4	III	M	0	N
GhCPK12	Gh_A04G1372	Ghir_A04G005530.1	scaffold882_A04	35941	39666	+	648	71746.9	5.19	MGNVCATS	N	Y	4	I	C|M	0	Y
GhCPK13	Gh_A04G1429	Ghir_A04G012050.1	scaffold1014_A04	44428	47300	+	547	61320.2	6.1	MGNTCRGP	N	Y	4	I	M	0	N
GhCPK14	Gh_A05G0617	Ghir_A05G007640.1	A05	6512015	6514333	+	532	59518.7	5.31	MGNLCSRS	Y	Y	4	II	M	0	N
GhCPK15	Gh_A05G1571	Ghir_A05G018710.1	A05	16092204	16102107	-	907	100917	7.08	MGICQSLC	N	Y	4	II	M	0	N
GhCPK16	Gh_A05G2355	Ghir_A05G028110.1	A05	28825850	28830758	+	502	56466.1	5.53	MNKKIAGS	N	N	4	I	C	0	N
GhCPK17	Gh_A05G2859	Ghir_A05G033000.1	A05	65056518	65058712	-	534	60698.6	7.1	MGSCISAP	Y	Y	4	III	M	0	N
GhCPK18	Gh_A05G3246	Ghir_A05G038740.1	A05	84817227	84822268	-	497	55945.2	5.2	MSRTSSGT	N	N	4	I	C	0	N
GhCPK19	Gh_A06G0013	Ghir_A06G000080.1	A06	49515	52357	-	526	59240.9	6.34	MGLCQSLG	Y	Y	4	II	M	0	N
GhCPK20	Gh_A06G1772	Ghir_A06G022050.1	A06	102737441	102739595	-	527	59315	5.08	MGCCSSKN	Y	Y	4	II	M	0	N
GhCPK21	Gh_A07G1099	Ghir_A07G012870.1	A07	22889554	22891905	-	513	57260.3	5.14	MGNCCTRG	Y	Y	4	II	M	0	N
GhCPK22	Gh_A07G1123	Ghir_A07G013130.1	A07	23758317	23762270	-	527	59502.1	7.07	MIFDNYFR	N	N	3	III	C	0	N
GhCPK23	Gh_A08G2530	Ghir_A08G001480.1	scaffold2271_A08	65298	68459	-	538	60507.2	5.24	MGNCCSRG	Y	Y	4	II	M	0	N
GhCPK24	Gh_A09G1033	Ghir_A09G011860.1	A09	60455043	60457397	-	523	58551	5.9	MGNCCSCG	Y	Y	4	II	C|M	0	N
GhCPK25	Gh_A09G1067	Ghir_A09G012230.1	A09	61271009	61274418	-	531	59856.1	6.58	MGNCCATP	N	Y	4	III	M	0	N
GhCPK26	Gh_A09G1157	Ghir_A09G013220.1	A09	62989834	62997485	-	531	60213.3	6.76	MGNCCVTS	N	Y	4	III	M	0	N
GhCPK27	Gh_A09G1248	Ghir_A09G014160.1	A09	64587018	64590334	-	610	68241.9	4.85	MGNNCFKT	N	Y	4	I	C|M	0	N
GhCPK28	Gh_A10G0124	Ghir_A10G000830.1	A10	1022813	1026173	-	552	62205.6	6.35	MGCFSSKH	Y	Y	4	II	M	0	N
GhCPK29	Gh_A10G0886	Ghir_A10G009700.1	A10	19193297	19197076	+	544	61921.2	8.63	MGICLSTT	Y	Y	4	IV	M	0	N
GhCPK30	Gh_A10G1195	Ghir_A10G013610.1	A10	62477215	62482365	+	537	61297.6	5.91	MGSCISTQ	Y	Y	4	III	M	0	N
GhCPK31	Gh_A10G1756	Ghir_A10G020050.1	A10	91965125	91970397	-	600	66967.4	5	MGNTCVGP	N	Y	4	I	M	0	N
GhCPK32	Gh_A11G0213	Ghir_A11G002410.1	A11	2006974	2009900	-	530	60394.3	6.87	MGNCNRPP	N	Y	4	II	M	0	N
GhCPK33	Gh_A11G1615	Ghir_A11G018500.1	A11	23975683	23979976	-	554	62700	9.48	MGACLSTT	Y	Y	4	IV	M	0	N
GhCPK34	Gh_A11G2941	Ghir_A11G035540.1	A11	93261063	93265206	+	529	59527.5	6.79	MGNCCRSP	N	Y	4	III	M	0	N
GhCPK35	Gh_A11G3011	Ghir_A11G003380.1	scaffold2728_A11	87030	90022	-	550	62325.9	6.53	MGNCNACV	N	Y	4	III	C|M	0	N
GhCPK36	Gh_A12G0109	Ghir_A12G001310.1	A12	1542851	1545587	-	560	62563.4	5.81	MGNTCRGS	N	Y	4	I	M	0	N
GhCPK37	Gh_A12G2686	Ghir_A12G024600.1	scaffold3405_A12	154335	157956	-	534	60239.7	6.28	MGNCNSQP	Y	Y	4	II	C|M	0	N
GhCPK38	Gh_A13G0017	Ghir_A13G000100.1	A13	149337	152705	-	579	64644.2	6.05	MGNTCVGP	N	Y	4	I	M	0	N
GhCPK39	Gh_A13G0563	Ghir_A13G006450.1	A13	13153568	13156341	-	508	56850.9	6.72	MGYNTTTG	N	Y	4	I	C	0	N
GhCPK40	Gh_A13G0566	Ghir_A13G006470.1	A13	13367845	13370645	-	487	54624	4.86	MRRAIDHQ	N	N	4	I	C	0	N
GhCPK41	Gh_A13G1164	Ghir_A13G014410.1	A13	64133647	64137746	-	536	60366.4	6.32	MGGCLTKN	Y	Y	4	II	M	0	N
GhCPK42	Gh_A13G1891	Ghir_A13G023110.1	A13	78586052	78590707	+	527	58760.4	6.18	MGNCCTRG	N	Y	4	II	M	0	N
GhCPK43	Gh_D01G1194	Ghir_D01G013420.1	D01	27904260	27906874	+	591	66336.2	5.18	MGNSCAKS	Y	Y	4	I	C|M	0	N
GhCPK44	Gh_D01G2360	Ghir_D01G007400.1	scaffold3726_D01	13436	16465	-	537	60459.3	6.56	MGSCLTKS	Y	Y	4	II	M	0	N
GhCPK45	Gh_D02G0183	Ghir_D02G001790.1	D02	2005202	2011106	-	447	51027.1	5.08	MHAFYILR	N	N	4	III	C	0	N
GhCPK46	Gh_D02G0663	Ghir_D02G006620.1	D02	9261983	9264700	+	508	56841.3	5.5	MNNQSSSI	N	Y	4	I	C	0	N
GhCPK47	Gh_D02G1973	Ghir_D02G020140.1	D02	63684341	63687392	-	568	63571.4	6.02	MGNTCRGS	N	Y	4	I	M	0	N
GhCPK48	Gh_D03G0087	Ghir_D03G000930.1	D03	613069	617844	-	541	61285.4	9.36	MGACLSAT	Y	Y	4	IV	M	0	N
GhCPK49	Gh_D03G0609	Ghir_D03G007610.1	D03	15802257	15805491	+	550	62289.8	7.07	MGNCCATP	Y	Y	4	III	C|M	0	N
GhCPK50	Gh_D04G0366	Ghir_D04G004340.1	D04	5618562	5627848	+	331	37021.8	4.68	MEGLGGMC	N	Y	4	I	C	0	N
GhCPK51	Gh_D04G0895	Ghir_D04G010070.1	D04	25443458	25447176	+	648	71828	5.13	MGNVCATL	N	Y	4	I	C|M	0	Y
GhCPK52	Gh_D04G0900	Ghir_D04G010120.1	D04	25701269	25707494	-	572	63599.6	5.17	MGNACAGP	N	Y	4	I	C|M	0	N
GhCPK53	Gh_D04G1271	Ghir_D04G014130.1	D04	41779274	41784032	-	527	59442.4	6.39	MGNCCRSP	N	Y	4	III	M	0	N
GhCPK54	Gh_D04G1486	Ghir_D04G016440.1	D04	46703911	46706311	-	552	61943.4	5.13	MGNTCRGP	N	Y	4	I	M	0	N
GhCPK55	Gh_D05G0748	Ghir_D05G007720.1	D05	6137190	6139492	+	532	59505.8	5.31	MGNLCSRS	Y	Y	4	II	M	0	N
GhCPK56	Gh_D05G1749	Ghir_D05G018740.1	D05	15794186	15800643	-	907	101032	6.92	MGICQSLC	N	Y	4	II	M	0	N
GhCPK57	Gh_D05G2622	Ghir_D05G028130.1	D05	27103112	27107994	+	502	56400	5.42	MNKKIAGS	N	Y	4	I	C	0	N
GhCPK58	Gh_D05G3156	Ghir_D05G033690.1	D05	47939538	47941747	-	534	60716.6	7.24	MGSCISAP	Y	Y	4	III	M	0	N
GhCPK59	Gh_D05G3567	Ghir_D05G037820.1	D05	59026817	59029919	+	524	58700.3	6.47	MGGCLTKT	Y	Y	4	II	M	0	N
GhCPK60	Gh_D06G2206	Ghir_D06G023100.1	D06	63778915	63781070	+	527	59318	5.23	MGCCSSKN	Y	Y	4	II	M	0	N
GhCPK61	Gh_D07G1198	Ghir_D07G013040.1	D07	18231735	18234096	-	513	57311.2	5.14	MGNCCSRG	Y	Y	4	II	M	0	N
GhCPK62	Gh_D07G1228	Ghir_D07G013310.1	D07	18817674	18820435	-	557	62469.5	7.58	MGNCCATT	N	Y	3	III	M	0	N
GhCPK63	Gh_D08G0142		D08	1197380	1200506	-	538	60511.2	5.1	MGNCCSRG	Y	Y	4	II	M	0	N
GhCPK64	Gh_D09G1054	Ghir_D09G011400.1	D09	36618092	36620409	-	524	58565.9	5.41	MGNCCSCG	Y	Y	4	II	C|M	0	N
GhCPK65	Gh_D09G1074	Ghir_D09G011730.1	D09	36952932	36956372	-	531	59847.1	6.52	MGNCCATP	N	Y	4	III	M	0	N

GhCPK66	Gh_D09G1163	Ghir_D09G012710.1	D09	38278169	38281037	-	531	60227.3	6.76	MGNCCVTS	N	Y	4	III	M	0	N
GhCPK67	Gh_D09G1249	Ghir_D09G013620.1	D09	39552322	39555623	-	610	68213.9	4.96	MGNNCFKT	N	Y	4	I	C|M	0	N
GhCPK68	Gh_D10G0130	Ghir_D10G001600.1	D10	1030951	1034394	-	551	62014.3	6.17	MGCFSSKH	Y	Y	4	II	M	0	N
GhCPK69	Gh_D10G0863	Ghir_D10G009490.1	D10	10909734	10913532	-	544	61950.2	8.81	MGICLSTT	Y	Y	4	IV	M	0	N
GhCPK70	Gh_D10G1303	Ghir_D10G013990.1	D10	24177574	24181997	-	495	56317.6	5.41	MGSCISTQ	Y	Y	3	III	M	0	N
GhCPK71	Gh_D10G2029	Ghir_D10G021750.1	D10	56010875	56015207	-	584	64984.3	5.14	MGNTCVGP	N	Y	4	I	M	0	N
GhCPK72	Gh_D11G0227	Ghir_D11G002390.1	D11	1905499	1908419	-	530	60355.2	6.56	MGNCNRPP	N	Y	4	II	M	0	N
GhCPK73	Gh_D11G0314	Ghir_D11G003360.1	D11	2722348	2725395	+	550	62311.9	6.53	MGNCNACV	N	Y	4	III	C|M	0	N
GhCPK74	Gh_D11G1774	Ghir_D11G018640.1	D11	19767240	19771544	-	554	62788.1	9.48	MGACLSTT	Y	Y	4	IV	M	0	N
GhCPK75	Gh_D11G3329	Ghir_D11G036410.1	D11	66059614	66066098	+	527	59284.3	6.39	MGNCCRSP	N	Y	4	III	M	0	N
GhCPK76	Gh_D12G0121	Ghir_D12G001340.1	D12	1591928	1594742	-	575	64573.8	6.07	MRLHYCMR	N	Y	4	I	N	0	Y
GhCPK77	Gh_D12G2743	Ghir_D12G024580.1	scaffold4587_D12	20270	23886	-	534	60242.7	6.39	MGNCNSQP	Y	Y	4	II	C|M	0	N
GhCPK78	Gh_D13G0033	Ghir_D13G000300.1	D13	246690	250046	-	573	64084.6	6.22	MGNTCVGP	N	Y	3	I	M	0	N
GhCPK79	Gh_D13G0560	Ghir_D13G005790.1	D13	7586741	7589365	-	523	58400.4	6.41	MGNTCLGS	N	Y	4	I	M	0	N
GhCPK80	Gh_D13G0561	Ghir_D13G005800.1	D13	7614829	7617628	-	487	54631	4.86	MRRAIDHQ	N	N	4	I	C	0	N
GhCPK81	Gh_D13G1455	Ghir_D13G015090.1	D13	45574159	45578015	+	536	60415.3	6.37	MGGCLTKN	Y	Y	4	II	M	0	N
GhCPK82	Gh_D13G2278	Ghir_D13G023730.1	D13	58814168	58818344	+	461	51476	6.4	MGNCCTRG	N	Y	4	II	M	0	N

MW molecular weight, pI isoelectric points, M Cell membrane,C Cytoplasm, N nucleus
The myristoylation sites were predicted by the Myristoylator program (http://web.expasy.org/myristoylator/).
The palmitoylation sites were predicted by CPS-Palm (http://lipid.biocuckoo.org/index.php)
Number of EF-hands were predicted by SMART database (http://smart.embl-heidelberg.de/)
The subcellular localizatio were predicted by DeepLoc-2.0 (https://services.healthtech.dtu.dk/services/DeepLoc-2.0/)
The transmembrane domains were predicted by DeepTMHMM (https://dtu.biolib.com/DeepTMHMM/)
The signal peptides were predicted by SignalP-6 (https://biolib.com/DTU/SignalP-6/)

Table S2 sgRNA of *CDPK* gene family mutant library in the *Gossypium hirsutum*

sgRNA sequence	Targeted gene ID	Targeted gene name	sgRNA sequence	Targeted gene ID	Targeted gene name
TCCTCTTCTCCGTTTCTGTG TGGTGGAATTGGGGTTGAAC	Gh_A01G0621	GhCPK1	ATCAAAGCTCACAGAGGCAG CTCTCTGAAATAGAAGGCCA	Gh_A01G0621 and Gh_D01G2360	GhCPK1 and GhCPK44
AATGAAGCAGACCATCTAGT ACAAGGAAGGGTCCAAGACA	Gh_A01G1119	GhCPK2	CTTTGCATCGAGAAAGGTAC TGCTCGACAGAGATACGAAG	Gh_A01G1119 and Gh_D01G1194	GhCPK2 and GhCPK43
GCAAGTTGTCGGCAACCACC CCTTCCCTAAGTGTTTGAAT	Gh_A02G0144	GhCPK3	CGACAAGTATCCATTCTCGT AGAGACGGACTTCAGAAAAT	Gh_A02G0144 and Gh_D02G0183	GhCPK3 and GhCPK45
GACAGCAGCAACCAAAGGAT CGACAGCAGCAACCAAAGGA	Gh_A02G0609	GhCPK4	CTTCAAGCGTGATAGCACGG TGGAACATCAGGTTGTGCCT	Gh_A02G0609 and Gh_D02G0663	GhCPK4 and GhCPK46
CAGTGGAACAATTAGCTTCG ATTGCACCACAGATCATCAA	Gh_A02G1029	GhCPK5	TGCAATGGTCTTCCGTCCAC GGAAAGCAGCTAACTTGTGT	Gh_A02G1029 and Gh_D03G0701	GhCPK5 and Gh_D03G0701
TCTTAGGGTAAAAGCACCAA AGTGCACCGTGACATGAAAC	Gh_A02G1635	GhCPK6	TGAGTATTCCGATTACCGGA AACTCCGGAGGAACTACGAA	Gh_A02G1635 and Gh_D03G0087	GhCPK6 and GhCPK48
GCAAGCTTAGGTAGAACTTA TCACACCCAAGTTCATATCT	Gh_A02G1796	GhCPK7	GTTCTAAATGATCCAACTGG AACAGCTTCTCATGAGAAGA	Gh_A02G1796 and Gh_D03G0609	GhCPK7 and GhCPK49
GCAAATATGGAGAGATCCGT CAGTCCATATTACGTTGCTC	Gh_A03G1505	GhCPK8	ATAAGCATGAGAGATGCCCC AGGCCAAGTTTTTACCGACG	Gh_A03G1505 and Gh_D02G1973	GhCPK8 and GhCPK47
GTTGGACACAAGTATGTGAA ATTCACAATTGCCCTGCACA	Gh_A04G0148	GhCPK9	TGAGAAGTGGAACCCAACAA TGTCCGAAAGATGTTAACGC	Gh_A04G0148 and Gh_D05G3567	GhCPK9 and GhCPK59
GAACAATGCCAATAAGGCTT GGTTGGAAGTCCATATTATG	Gh_A04G0467	GhCPK10	AAAGGCAGCTGCGCTTGCTA GTAGCTATGATGCAAGACAC	Gh_A04G0467 and Gh_D04G0900	GhCPK10 and GhCPK52
GTAGTAATGAACCAACCCGT GAGGAAACTCAGAACTGCCG	Gh_A04G0780	GhCPK11	TGGCTCCTGAGGTTCTCAAA TGGAACCAGACCCAAAGCTC	Gh_A04G0780 and Gh_D04G1271	GhCPK11 and GhCPK53
GGTCCTAAACCAGCCGAAGA TTGCAGAAAACTGCTTCATG	Gh_A04G1372	GhCPK12	GATCTTGGTAAGAAGACCCT GTTCCTTTCAGGCCATCCTT	Gh_A04G1372 and Gh_D04G0895	GhCPK12 and GhCPK51
GTCGATATCGAAGACGAAGT GCTAAGAAGAAGATGGTTGG	Gh_A04G1429	GhCPK13	CTTTCTGCCTTATTTCAGGA GACGATTAAAGGCACGTACG	Gh_A04G1429 and Gh_D04G1486	GhCPK13 and GhCPK54
CGCTGCTTCTTTGCTTCGCA GTCTAATATTGTCGAGCTCA	Gh_A05G0617	GhCPK14	CTTACTTAGAACCTGCACTG CTGAAGCACATACCAAGAAG	Gh_A05G0617 and Gh_D05G0748	GhCPK14 and GhCPK55
CTAACCACTGACAAGGACGT GTTATTTGAATCACGGCCAA	Gh_A05G1571	GhCPK15	GCTGATGCCTTGAATTCGAG AGAGGAAAACATGAGCCGCG	Gh_A05G1571 and Gh_D05G1749	GhCPK15 and GhCPK56
TTCAGCATAGCCGATGCCTT TGTTGGGTCTTTCATTCCAA	Gh_A05G2355	GhCPK16	CATAGCAGAGAGACTCTCTG GATAATACAGCGGAATCCAG	Gh_A05G2355 and Gh_D05G2622	GhCPK16 and GhCPK57
TAACATAAGACAACGTACCG ATCCAAAGCCAAGCTGCGGA	Gh_A05G2859	GhCPK17	TCGTTGGAGGAGCTTAAAGA TTTGTCAGGATGGTAATACG	Gh_A05G2859 and Gh_D05G3156	GhCPK17 and GhCPK58
GGGTTCTGAACTTATGGAAT GTTCGGAGTGACGGCAAATG	Gh_A05G3246	GhCPK18	CAGTTTATTCAAGTGCACCG GAAAGGGAAGAGAATCTGGT	Gh_A05G3246 and Gh_D04G0366	GhCPK18 and GhCPK50
GATTTGTTGGACAGGATGGG ACAGAAGCAAGCCCGATTGA	Gh_A06G0013	GhCPK19	CATCAAGCAAATAATCTCCG ACATAGCACCCGAAGTTCTT	Gh_A06G1772 and Gh_D06G2206	GhCPK20 and GhCPK60

AGTGAGGAAGATGTTGACAA CTTGTTTTGTTGGCTAATGG	Gh_A06G1772	GhCPK20	GATGCGCCAAAGCAAGCAGA GATTACTAACCTGCCCAGAA	Gh_A07G1099 and Gh_D07G1198	GhCPK21 and GhCPK61
GCAGAAGGTTCACCACCAGC GATGAACTCATCGTAGTCTA	Gh_A07G1099	GhCPK21	TTGTGTACTGATCCCAACTC GAAGGAAGCTTCCCCCTTGA	Gh_A07G1123 and Gh_D07G1228	GhCPK22 and GhCPK62
GCTAACATTGGCCAGACGAG GCAAACAAGAACATCCACAA	Gh_A07G1123	GhCPK22	ACCATAGATTCCGATAACAG GGTCATTGCAGGATGTTTAT	Gh_A08G2530 and Gh_D08G0142	GhCPK23 and GhCPK63
CACTTCGACAAAGACAATAG GTATGGCACGTTCATGATGT	Gh_A08G2530	GhCPK23	GAAAAGGAAACCCAAAGAGG TGAGAAAAGGAAACCCAAAG	Gh_A09G1033 and Gh_D09G1054	GhCPK24 and GhCPK64
CTTCACCTGTAAGAGATCCA GGGGAGGCTGTATTATTCCT	Gh_A09G1033	GhCPK24	GCTGAAGAAGAGAGCACTAA TTGTCAGTTGAGGAAGTAGC	Gh_A09G1067 and Gh_D09G1074	GhCPK25 and GhCPK65
CAAGGGGACACTACACTGAA TGGAGATGTTTTGGTTCAAG	Gh_A09G1067	GhCPK25	AGTGGATGTTTGGAGTGCTG TGATTGTTGAACGTGCAGGA	Gh_A09G1157 and Gh_D09G1163	GhCPK26 and GhCPK66
TGAGGAACTAAGAGATGAGT AGATGAGTTGGCTGATGAGC	Gh_A09G1157	GhCPK26	CATCAGCCTGAACCCAACTG CCTAGGACTAAAGAATCAAG	Gh_A09G1248 and Gh_D09G1249	GhCPK27 and GhCPK67
TGGTGAACCTGACTTCACAT CATCAGCCTGAACCCAACTG	Gh_A09G1248	GhCPK27	AGACAAGCCAATAGGTAGCG GGACGGGAGAATTAACTACG	Gh_A10G0124 and Gh_D10G0130	GhCPK28 and GhCPK68
GGTGGATGAAGCCAGCATCA TAGCAGCAAGCATAAATTGC	Gh_A10G0124	GhCPK28	GCAGCAGTCAGTCTAGCCCG CTGTCACTAATTGTTGGCCA	Gh_A10G0886 and Gh_D10G0863	GhCPK29 and GhCPK69
CCATTCATGACGTGAACCAT TTAGAGGCCCTGAAAGGCCA	Gh_A10G0886	GhCPK29	GCTCTACTACAAGACGACCC ATGATAGATCCGGTACCAAA	Gh_A10G1195 and Gh_D10G1303	GhCPK30 and GhCPK70
TCGAAGATATTCTTCCCCGA AAGTGGAACACTAAGCTACG	Gh_A10G1195	GhCPK30	AGGAGCTTATGAGGATGCAG GGAAGCTTGCCATTCTTTGG	Gh_A10G1756 and Gh_D10G2029	GhCPK31 and GhCPK71
TAAACCAGAGACCAAACAAG CGGTGCGTATGTGCATTAGC	Gh_A10G1756	GhCPK31	GATGCTACGACAAGATCCCA TACGACAAGATCCCAAGGAA	Gh_A11G0213 and Gh_D11G0227	GhCPK32 and GhCPK72
TTCCGACCAGAATCGGACAC TATCTTTCAGGTCTGCAGGA	Gh_A11G0213	GhCPK32	TGAAGAATCTCAAGGACACG TTGCAGTTCGTAGAAGTCGG	Gh_A11G1615 and Gh_D11G1774	GhCPK33 and GhCPK74
TTTCCTCCTTCTCTAACCCA AGCTTCTCCAGCACACCCGT	Gh_A11G1615	GhCPK33	TTTCACGTCTTCTCTTGCAA TTACACTGAAAGAGCTGCCG	Gh_A11G2941 and Gh_D11G3329	GhCPK34 and GhCPK75
GTAAACATCCTGTCAGCGTC TACCTCGTGGATCGCAAGCT	Gh_A11G2941	GhCPK34	ACATTATAGCAACCTCGCGT TGTTCGGCAGATGTTAGAAC	Gh_A11G3011 and Gh_D11G0314	GhCPK35 and GhCPK73
AGGTAATGCCAAACTCACCT GAGTCACCGAACACGAATAG	Gh_A11G3011	GhCPK35	TCCATGGATTTCCCAAAACG GAATCGGAAGACGAACCATG	Gh_A12G0109 amd Gh_D12G0121	GhCPK36 amd GhCPK76
CAAGCATTACGGACCAGAAG ATAGTGGGAGCATCGATTAT	Gh_A12G0109	GhCPK36	ATGTAGCTCCAGAAGTATTG TTTGTGCCGGCAGATCGTGA	Gh_A12G2686 and Gh_D12G2743	GhCPK37 and GhCPK77
ACCGTGTTCTTCCTACAGGA TCCATCCTGTAGGAAGAACA	Gh_A12G2686	GhCPK37	CTAAGCCTCCTCATGTGAAG ATATTCGACCAAGTCTTGCA	Gh_A13G0017 and Gh_D13G0033	GhCPK38 and GhCPK78
TAGCGGCACAATTGATTATG TCTTGCCTAGTCTCCGATTT	Gh_A13G0017	GhCPK38	GAGCTGTCAACCGTTCTGAA AGGAACTGTAAACAGCGGGA	Gh_A13G0563 and Gh_D13G0560	GhCPK39 and GhCPK79
TGTGCATTCACCCTTCAGAA TTGACTTCAGCTCTTACCCA	Gh_A13G0563	GhCPK39	CCTCAATAAACTAGAACGCG TGGCTTGCGAGAAATGTTTA	Gh_A13G0566 and Gh_D13G0561	GhCPK40 and GhCPK80

CAGCCAAATGGTGCATAATT TATGTACCGAGATTTCGACT	Gh_A13G0566	GhCPK40	CAGCCTGCAAAATAAAACCG CTATCTGATAACAAGCCCGG	Gh_A13G1164 and Gh_D13G1455	GhCPK41 and GhCPK81
GTGTTAACATCCTGCAGACT AACATCGATACTGACAACAG	Gh_A13G1164	GhCPK41	AAGGAGGTTAACAGCTGCCC ACACTTGAAGAACTCAGGCA	Gh_A13G1891 and Gh_D13G2278	GhCPK42 and GhCPK82
ATAGGGCAAGCAATGAACAG TTGGTGATGGAATTGTGCGG	Gh_A13G1891	GhCPK42			
CTTGCAAGTCGATCGCGAAA CGTGCTCTAATGCAGGCCGT	Gh_D01G1194	GhCPK43			
AGATCAACTACGAGGAGTTC CCTGATTAAGATGGAAGGTA	Gh_D01G2360	GhCPK44			
CAGAAAATCGGTCACTCTGT CAGAAAATCGGTCACTCTGT	Gh_D02G0183	GhCPK45			
TCAACCGTAAAATGAGCGGA CCGACGACAACAGCAACCAA	Gh_D02G0663	GhCPK46			
AGCTGAAAGTCTTTCCGAGG ACTTCGGGAGCAACGTAATA	Gh_D02G1973	GhCPK47			
CACTTACGTTGCAATAGACA TAGCAGTTCGTAAAAGCCGT	Gh_D03G0087	GhCPK48			
GGTCAACATCACCATGGCAA GGATGACAATGCAGTCCATT	Gh_D03G0609	GhCPK49			
GCGTTCGGGGTGAAGTCAAA TCAATTTGGGAGACGCGTTC	Gh_D04G0366	GhCPK50			
TGGCTTAGGGTCTCTTCTCC CCTGTCTGTGCAATAAGCAA	Gh_D04G0895	GhCPK51			
GGAAGGGAACAATGCCAAAA AAGGGCAGCTGCACTTGCTA	Gh_D04G0900	GhCPK52			
ACAAAAACCATGTCATTGGG CGATCTCGAAGAGGAAACTC	Gh_D04G1271	GhCPK53			
TGTGTACAGAGCTTTCCACG GTCGATATCGAAGAGGAAGT	Gh_D04G1486	GhCPK54			
CATATAACTGCGTTGTCAAG CAGAGCTTGCTCTAGTTCTT	Gh_D05G0748	GhCPK55			
GGTTTCTTCAAGTTGAACAC GATAAACCTGAATGAAAGCG	Gh_D05G1749	GhCPK56			
GTCTTTCATTCCAAACGCAT TGGTTGTTGCTGAATTGTGG	Gh_D05G2622	GhCPK57			
TCCCAATCCTCAATCCCGGA TGAGGTTCAAATGTTGATGC	Gh_D05G3156	GhCPK58			
TGAAAGAGCAGCTGCTTCCA GAAAGAGCAGCTGCTTCCAT	Gh_D05G3567	GhCPK59			

TTAGTTTCGAAAACGACCCA TGACGATGGACCCGAATGAG	Gh_D06G2206	GhCPK60
GTGTTGTGTAGGTGATTGCG ATCACCAACCCACTGAATAG	Gh_D07G1198	GhCPK61
TGCTCTTTGTAACAGGTGAA GCACTTCACCCATATTATGT	Gh_D07G1228	GhCPK62
AATGTTCTTCCCGATCCATT TGATGTAATGTTCGACAATT	Gh_D08G0142	GhCPK63
GAAGAAGCCAAGAGTCCTAA TAGTGCTTGCTCTAGTTCTT	Gh_D09G1054	GhCPK64
TATGCAAGAGAATGAACGAG AGGTTGATGTCTGGAGTGCT	Gh_D09G1074	GhCPK65
GTCACATTACTCAGAACCAG AGATGAGTTGGCTGATGAGC	Gh_D09G1163	GhCPK66
ACAGACTGCTATAAAGAACG AATCCTGAAGTTGGTAAGAG	Gh_D09G1249	GhCPK67
TAGCAGCAAACATAAACCGT GCTATGAAGGAGTATGGAAT	Gh_D10G0130	GhCPK68
GTTGCCATAGGCAATGAAAA GGATACACCTACGTTGCCAT	Gh_D10G0863	GhCPK69
TTTCCCCCAGATTAACGTTG CTACGACGAGTTCACAACAA	Gh_D10G1303	GhCPK70
GAAGTCGGACTGAAAAGAGT CTTGCAAATCGATTGCAAAG	Gh_D10G2029	GhCPK71
TCATTATGATGGGTACCGCC CATTATGATGGGTACCGCCT	Gh_D11G0227	GhCPK72
AAAGAGTCCCACAGAAACAG AGCTCGGCCGAGGTGAATTT	Gh_D11G0314	GhCPK73
AAATCAGTACGCTTCCCGCA CGTGGTTCCTTCATCTTCCA	Gh_D11G1774	GhCPK74
GACGTGAAATCGAATTTCTC CGTGCTTGCAGGAGTTCCAA	Gh_D11G3329	GhCPK75
CAAGCATTACGGACCGGAGG TAGCGGGAGCATCGATTACG	Gh_D12G0121	GhCPK76
TTGATTTCGCATCAGATCCA TAAGTGTTCTTCCTACAGGA	Gh_D12G2743	GhCPK77
AATCTGTTTCAGCCGTAATG ATAGCGGCACAATGATTATG	Gh_D13G0033	GhCPK78
GGCGAATGCACAACATCTTT GGATGTGGAGGATGCTAGGA	Gh_D13G0560	GhCPK79

GGCCTAAAGTGTAAAGCTCA TATGTACCGAGATTTTGACC	Gh_D13G0561	GhCPK80
GTGCCAAGGATCTAGTCTGC ATTGTTGGCCAGAAACATGC	Gh_D13G1455	GhCPK81
ATCGGAAAAGAACTAGGTCG AACTAGGTCGTGGTCAATTT	Gh_D13G2278	GhCPK82

Table S3 Primers for constructing the vector mixing pool

Downstream primer	Primer name	Upstream primer	Primer name
ttctagctctaaaacCACAGAAACGGAGAAGAGGAtgcaccagccgggaat	GhCPK1	AAGCATCAGATGggcaAACAAAGCACCAGTGGTCTAG	pRGEB32-7S
ttctagctctaaaacGTTCAACCCCAATTCCACCAtgcaccagccgggaat			
ttctagctctaaaacACTAGATGGTCTGCTTCATTtgcaccagccgggaat	GhCPK2		
ttctagctctaaaacTGTCTTGGACCCTTCCTTGTtgcaccagccgggaat			
ttctagctctaaaacGGTGGTTGCCGACAACTTGCtgcaccagccgggaat	GhCPK3		
ttctagctctaaaacATTCAAACACTTAGGGAAGGtgcaccagccgggaat			
ttctagctctaaaacATCCTTTGGTTGCTGCTGTCtgcaccagccgggaat	GhCPK4		
ttctagctctaaaacTCCTTTGGTTGCTGCTGTCGtgcaccagccgggaat			
ttctagctctaaaacCGAAGCTAATTGTTCCACTGtgcaccagccgggaat	GhCPK5		
ttctagctctaaaacTTGATGATCTGTGGTGCAATtgcaccagccgggaat			
ttctagctctaaaacTTGGTGCTTTTACCCTAAGAtgcaccagccgggaat	GhCPK6		
ttctagctctaaaacGTTTCATGTCACGGTGCACTtgcaccagccgggaat			
ttctagctctaaaacTAAGTTCTACCTAAGCTTGCtgcaccagccgggaat	GhCPK7		
ttctagctctaaaacAGATATGAACTTGGGTGTGAtgcaccagccgggaat			
ttctagctctaaaacACGGATCTCTCCATATTTGCtgcaccagccgggaat	GhCPK8		
ttctagctctaaaacGAGCAACGTAATATGGACTGtgcaccagccgggaat			
ttctagctctaaaacTTCACATACTTGTGTCCAACtgcaccagccgggaat	GhCPK9		
ttctagctctaaaacTGTGCAGGGCAATTGTGAATtgcaccagccgggaat			
ttctagctctaaaacATAATCGATGCTCCCACTATtgcaccagccgggaat	GhCPK10		
ttctagctctaaaacCATAATATGGACTTCCAACCtgcaccagccgggaat			
ttctagctctaaaacACGGGTTGGTTCATTACTACtgcaccagccgggaat	GhCPK11		
ttctagctctaaaacCGGCAGTTCTGAGTTTCCTCtgcaccagccgggaat			
ttctagctctaaaacTCTTCGGCTGGTTTAGGACCtgcaccagccgggaat	GhCPK12		
ttctagctctaaaacCATGAAGCAGTTTTCTGCAAtgcaccagccgggaat			
ttctagctctaaaacACTTCGTCTTCGATATCGACtgcaccagccgggaat	GhCPK13		
ttctagctctaaaacCCAACCATCTTCTTCTTAGCtgcaccagccgggaat			
ttctagctctaaaacTGCGAAGCAAAGAAGCAGCGtgcaccagccgggaat	GhCPK14		
ttctagctctaaaacTGAGCTCGACAATATTAGACtgcaccagccgggaat			
ttctagctctaaaacACGTCCTTGTCAGTGGTTAGtgcaccagccgggaat	GhCPK15		
ttctagctctaaaacTTGGCCGTGATTCAAATAACtgcaccagccgggaat			
ttctagctctaaaacAAGGCATCGGCTATGCTGAAtgcaccagccgggaat	GhCPK16		
ttctagctctaaaacTTGGAATGAAAGACCCAACAtgcaccagccgggaat			
ttctagctctaaaacCGGTACGTTGTCTTATGTTAtgcaccagccgggaat	GhCPK17		
ttctagctctaaaacTCCGCAGCTTGGCTTTGGATtgcaccagccgggaat			
ttctagctctaaaacATTCCATAAGTTCAGAACCCtgcaccagccgggaat	GhCPK18		
ttctagctctaaaacCATTTGCCGTCACTCCGAACtgcaccagccgggaat			
ttctagctctaaaacCCCATCCTGTCCAACAAATCtgcaccagccgggaat	GhCPK19		

ttctagctctaaaacTCAATCGGGCTTGCTTCTGTtgcaccagccgggaat	GhCPK19
ttctagctctaaaacTTGTCAACATCTTCCTCACTtgcaccagccgggaat	GhCPK20
ttctagctctaaaacCCATTAGCCAACAAAACAAGtgcaccagccgggaat	GhCPK21
ttctagctctaaaacGCTGGTGGTGAACCTTCTGCtgcaccagccgggaat	GhCPK22
ttctagctctaaaacTAGACTACGATGAGTTCATCtgcaccagccgggaat	GhCPK23
ttctagctctaaaacCTCGTCTGGCCAATGTTAGCtgcaccagccgggaat	GhCPK24
ttctagctctaaaacTTGTGGATGTTCTTGTTTGCtgcaccagccgggaat	GhCPK25
ttctagctctaaaacCTATTGTCTTTGTCGAAGTGtgcaccagccgggaat	GhCPK26
ttctagctctaaaacACATCATGAACGTGCCATACtgcaccagccgggaat	GhCPK27
ttctagctctaaaacTGGATCTCTTACAGGTGAAGtgcaccagccgggaat	GhCPK28
ttctagctctaaaacAGGAATAATACAGCCTCCCCtgcaccagccgggaat	GhCPK29
ttctagctctaaaacTTCAGTGTAGTGTCCCCTTGtgcaccagccgggaat	GhCPK30
ttctagctctaaaacCTTGAACCAAAACATCTCCAtgcaccagccgggaat	GhCPK31
ttctagctctaaaacACTCATCTCTTAGTTCCTCAtgcaccagccgggaat	GhCPK32
ttctagctctaaaacGCTCATCAGCCAACTCATCTtgcaccagccgggaat	GhCPK33
ttctagctctaaaacATGTGAAGTCAGGTTCACCAtgcaccagccgggaat	GhCPK34
ttctagctctaaaacATAATCGATGCTCCCACTATtgcaccagccgggaat	GhCPK35
ttctagctctaaaacTGATGCTGGCTTCATCCACCtgcaccagccgggaat	GhCPK36
ttctagctctaaaacGCAATTTATGCTTGCTGCTAtgcaccagccgggaat	GhCPK37
ttctagctctaaaacATGGTTCACGTCATGAATGGtgcaccagccgggaat	GhCPK38
ttctagctctaaaacTGGCCTTTCAGGGCCTCTAAtgcaccagccgggaat	
ttctagctctaaaacTCGGGGAAGAATATCTTCGAtgcaccagccgggaat	
ttctagctctaaaacCGTAGCTTAGTGTTCCACTTtgcaccagccgggaat	
ttctagctctaaaacCTTGTTTGGTCTCTGGTTTAtgcaccagccgggaat	
ttctagctctaaaacGCTAATGCACATACGCACCGtgcaccagccgggaat	
ttctagctctaaaacGTGTCCGATTCTGGTCGGAAtgcaccagccgggaat	
ttctagctctaaaacTCCTGCAGACCTGAAAGATAtgcaccagccgggaat	
ttctagctctaaaacTGGGTTAGAGAAGGAGGAAAtgcaccagccgggaat	
ttctagctctaaaacACGGGTGTGCTGGAGAAGCTtgcaccagccgggaat	
ttctagctctaaaacGACGCTGACAGGATGTTTACtgcaccagccgggaat	
ttctagctctaaaacAGCTTGCGATCCACGAGGTAtgcaccagccgggaat	
ttctagctctaaaacAGGTGAGTTTGGCATTACCTtgcaccagccgggaat	
ttctagctctaaaacCTATTCGTGTTCGGTGACTCtgcaccagccgggaat	
ttctagctctaaaacCTTCTGGTCCGTAATGCTTGtgcaccagccgggaat	
ttctagctctaaaacATAATCGATGCTCCCACTATtgcaccagccgggaat	
ttctagctctaaaacTCCTGTAGGAAGAACACGGTtgcaccagccgggaat	
ttctagctctaaaacTGTTCTTCCTACAGGATGGAtgcaccagccgggaat	
ttctagctctaaaacCATAATCAATTGTGCCGCTAtgcaccagccgggaat	
ttctagctctaaaacAAATCGGAGACTAGGCAAGAtgcaccagccgggaat	

ttctagctctaaaacTTCTGAAGGGTGAATGCACAtgcaccagccgggaat	GhCPK39
ttctagctctaaaacTGGGTAAGAGCTGAAGTCAAtgcaccagccgggaat	
ttctagctctaaaacAATTATGCACCATTTGGCTGtgcaccagccgggaat	GhCPK40
ttctagctctaaaacAGTCGAAATCTCGGTACATAtgcaccagccgggaat	
ttctagctctaaaacAGTCTGCAGGATGTTAACACtgcaccagccgggaat	GhCPK41
ttctagctctaaaacCTGTTGTCAGTATCGATGTTtgcaccagccgggaat	
ttctagctctaaaacCTGTTCATTGCTTGCCCTATtgcaccagccgggaat	GhCPK42
ttctagctctaaaacCCGCACAATTCCATCACCAAtgcaccagccgggaat	
ttctagctctaaaacTTTCGCGATCGACTTGCAAGtgcaccagccgggaat	GhCPK43
ttctagctctaaaacACGGCCTGCATTAGAGCACGtgcaccagccgggaat	
ttctagctctaaaacGAACTCCTCGTAGTTGATCTtgcaccagccgggaat	GhCPK44
ttctagctctaaaacTACCTTCCATCTTAATCAGGtgcaccagccgggaat	
ttctagctctaaaacACAGAGTGACCGATTTTCTGtgcaccagccgggaat	GhCPK45
ttctagctctaaaacACAGAGTGACCGATTTTCTGtgcaccagccgggaat	
ttctagctctaaaacTCCGCTCATTTTACGGTTGAtgcaccagccgggaat	GhCPK46
ttctagctctaaaacTTGGTTGCTGTTGTCGTCGGtgcaccagccgggaat	
ttctagctctaaaacCCTCGGAAAGACTTTCAGCTtgcaccagccgggaat	GhCPK47
ttctagctctaaaacTATTACGTTGCTCCCGAAGTtgcaccagccgggaat	
ttctagctctaaaacTGTCTATTGCAACGTAAGTGtgcaccagccgggaat	GhCPK48
ttctagctctaaaacACGGCTTTTACGAACTGCTAtgcaccagccgggaat	
ttctagctctaaaacTTGCCATGGTGATGTTGACCtgcaccagccgggaat	GhCPK49
ttctagctctaaaacAATGGACTGCATTGTCATCCtgcaccagccgggaat	
ttctagctctaaaacTTTGACTTCACCCCGAACGCtgcaccagccgggaat	GhCPK50
ttctagctctaaaacGAACGCGTCTCCCAAATTGAtgcaccagccgggaat	
ttctagctctaaaacGGAGAAGAGACCCTAAGCCAtgcaccagccgggaat	GhCPK51
ttctagctctaaaacTTGCTTATTGCACAGACAGGtgcaccagccgggaat	
ttctagctctaaaacTTTTGGCATTGTTCCCTTCCtgcaccagccgggaat	GhCPK52
ttctagctctaaaacTAGCAAGTGCAGCTGCCCTTtgcaccagccgggaat	
ttctagctctaaaacCCCAATGACATGGTTTTTGTtgcaccagccgggaat	GhCPK53
ttctagctctaaaacGAGTTTCCTCTTCGAGATCGtgcaccagccgggaat	
ttctagctctaaaacCGTGGAAAGCTCTGTACACAtgcaccagccgggaat	GhCPK54
ttctagctctaaaacACTTCCTCTTCGATATCGACtgcaccagccgggaat	
ttctagctctaaaacCTTGACAACGCAGTTATATGtgcaccagccgggaat	GhCPK55
ttctagctctaaaacAAGAACTAGAGCAAGCTCTGtgcaccagccgggaat	
ttctagctctaaaacGTGTTCAACTTGAAGAAACCtgcaccagccgggaat	GhCPK56
ttctagctctaaaacCGCTTTCATTCAGGTTTATCtgcaccagccgggaat	
ttctagctctaaaacATGCGTTTGGAATGAAAGACtgcaccagccgggaat	GhCPK57
ttctagctctaaaacCCACAATTCAGCAACAACCAtgcaccagccgggaat	
ttctagctctaaaacTCCGGGATTGAGGATTGGGAtgcaccagccgggaat	GhCPK58

ttctagctctaaaacGCATCAACATTTGAACCTCAtgcaccagccgggaat	GhCPK58
ttctagctctaaaacTGGAAGCAGCTGCTCTTTCAtgcaccagccgggaat	GhCPK59
ttctagctctaaaacATGGAAGCAGCTGCTCTTTCtgcaccagccgggaat	GhCPK60
ttctagctctaaaacTGGGTCGTTTTCGAAACTAAtgcaccagccgggaat	GhCPK61
ttctagctctaaaacCTCATTCGGGTCCATCGTCAtgcaccagccgggaat	GhCPK62
ttctagctctaaaacCGCAATCACCTACACAACACtgcaccagccgggaat	GhCPK63
ttctagctctaaaacCTATTCAGTGGGTTGGTGATtgcaccagccgggaat	GhCPK64
ttctagctctaaaacTTCACCTGTTACAAAGAGCAtgcaccagccgggaat	GhCPK65
ttctagctctaaaacACATAATATGGGTGAAGTGCtgcaccagccgggaat	GhCPK66
ttctagctctaaaacAATGGATCGGGAAGAACATTtgcaccagccgggaat	GhCPK67
ttctagctctaaaacAATTGTCGAACATTACATCAtgcaccagccgggaat	GhCPK68
ttctagctctaaaacTTAGGACTCTTGGCTTCTTCtgcaccagccgggaat	GhCPK69
ttctagctctaaaacAAGAACTAGAGCAAGCACTAtgcaccagccgggaat	GhCPK70
ttctagctctaaaacCTCGTTCATTCTCTTGCATAtgcaccagccgggaat	GhCPK71
ttctagctctaaaacAGCACTCCAGACATCAACCTtgcaccagccgggaat	GhCPK72
ttctagctctaaaacCTGGTTCTGAGTAATGTGACtgcaccagccgggaat	GhCPK73
ttctagctctaaaacGCTCATCAGCCAACTCATCTtgcaccagccgggaat	GhCPK74
ttctagctctaaaacCGTTCTTTATAGCAGTCTGTtgcaccagccgggaat	GhCPK75
ttctagctctaaaacCTCTTACCAACTTCAGGATTtgcaccagccgggaat	GhCPK76
ttctagctctaaaacACGGTTTATGTTTGCTGCTAtgcaccagccgggaat	GhCPK77
ttctagctctaaaacATTCCATACTCCTTCATAGCtgcaccagccgggaat	
ttctagctctaaaacTTTTCATTGCCTATGGCAACtgcaccagccgggaat	
ttctagctctaaaacATGGCAACGTAGGTGTATCCtgcaccagccgggaat	
ttctagctctaaaacCAACGTTAATCTGGGGGAAAtgcaccagccgggaat	
ttctagctctaaaacTTGTTGTGAACTCGTCGTAGtgcaccagccgggaat	
ttctagctctaaaacACTCTTTTCAGTCCGACTTCTtgcaccagccgggaat	
ttctagctctaaaacCTTTGCAATCGATTTGCAAGtgcaccagccgggaat	
ttctagctctaaaacGGCGGTACCCATCATAATGAtgcaccagccgggaat	
ttctagctctaaaacAGGCGGTACCCATCATAATGtgcaccagccgggaat	
ttctagctctaaaacCTGTTTCTGTGGGACTCTTTtgcaccagccgggaat	
ttctagctctaaaacAAATTCACCTCGGCCGAGCTtgcaccagccgggaat	
ttctagctctaaaacTGCGGGAAGCGTACTGATTTtgcaccagccgggaat	
ttctagctctaaaacTGGAAGATGAAGGAACCACGtgcaccagccgggaat	
ttctagctctaaaacGAGAAATTCGATTTCACGTCtgcaccagccgggaat	
ttctagctctaaaacTTGGAACTCCTGCAAGCACGtgcaccagccgggaat	
ttctagctctaaaacCCTCCGGTCCGTAATGCTTGtgcaccagccgggaat	
ttctagctctaaaacCGTAATCGATGCTCCCGCTAtgcaccagccgggaat	
ttctagctctaaaacTGGATCTGATGCGAAATCAAtgcaccagccgggaat	
ttctagctctaaaacTCCTGTAGGAAGAACACTTAtgcaccagccgggaat	

ttctagctctaaaacCATTACGGCTGAAACAGATTtgcaccagccgggaat	GhCPK78
ttctagctctaaaacCATAATCATTGTGCCGCTATtgcaccagccgggaat	
ttctagctctaaaacAAAGATGTTGTGCATTCGCCtgcaccagccgggaat	GhCPK79
ttctagctctaaaacTCCTAGCATCCTCCACATCCtgcaccagccgggaat	
ttctagctctaaaacTGAGCTTTACACTTTAGGCCtgcaccagccgggaat	GhCPK80
ttctagctctaaaacGGTCAAAATCTCGGTACATAtgcaccagccgggaat	
ttctagctctaaaacGCAGACTAGATCCTTGGCACtgcaccagccgggaat	GhCPK81
ttctagctctaaaacGCATGTTTCTGGCCAACAATTtgcaccagccgggaat	
ttctagctctaaaacCGACCTAGTTCTTTTCCGATtgcaccagccgggaat	GhCPK82
ttctagctctaaaacAAATTGACCACGACCTAGTTtgcaccagccgggaat	
ttctagctctaaaacCTGCCTCTGTGAGCTTTGATtgcaccagccgggaat	GhCPK1 and GhCPK44
ttctagctctaaaacTGGCCTTCTATTTCAGAGAGtgcaccagccgggaat	
ttctagctctaaaacGTACCTTTCTCGATGCAAAGtgcaccagccgggaat	GhCPK2 and GhCPK43
ttctagctctaaaacCTTCGTATCTCTGTCGAGCAtgcaccagccgggaat	
ttctagctctaaaacACGAGAATGGATACTTGTCGtgcaccagccgggaat	GhCPK3 and GhCPK45
ttctagctctaaaacATTTTCTGAAGTCCGTCTCTtgcaccagccgggaat	
ttctagctctaaaacCCGTGCTATCACGCTTGAAGtgcaccagccgggaat	GhCPK4 and GhCPK46
ttctagctctaaaacAGGCACAACCTGATGTTCCAtgcaccagccgggaat	
ttctagctctaaaacGTGGACGGAAGACCATTGCAtgcaccagccgggaat	GhCPK5 and Gh_D03G0701
ttctagctctaaaacACACAAGTTAGCTGCTTTCCtgcaccagccgggaat	
ttctagctctaaaacTCCGGTAATCGGAATACTCAtgcaccagccgggaat	GhCPK6 and GhCPK48
ttctagctctaaaacTTCGTAGTTCCTCCGGAGTTtgcaccagccgggaat	
ttctagctctaaaacCCAGTTGGATCATTTAGAACtgcaccagccgggaat	GhCPK7 and GhCPK49
ttctagctctaaaacTCTTCTCATGAGAAGCTGTTtgcaccagccgggaat	
ttctagctctaaaacGGGGCATCTCTCATGCTTATtgcaccagccgggaat	GhCPK8 and GhCPK47
ttctagctctaaaacCGTCGGTAAAAACTTGGCCTtgcaccagccgggaat	
ttctagctctaaaacTTGTTGGGTTCCACTTCTCAtgcaccagccgggaat	GhCPK9 and GhCPK59
ttctagctctaaaacGCGTTAACATCTTTCGGACAtgcaccagccgggaat	
ttctagctctaaaacTAGCAAGCGCAGCTGCCTTTtgcaccagccgggaat	GhCPK10 and GhCPK52
ttctagctctaaaacGTGTCTTGCATCATAGCTACtgcaccagccgggaat	
ttctagctctaaaacTTTGAGAACCTCAGGAGCCAtgcaccagccgggaat	GhCPK11 and GhCPK53
ttctagctctaaaacGAGCTTTGGGTCTGGTTCCAtgcaccagccgggaat	
ttctagctctaaaacAGGGTCTTCTTACCAAGATCtgcaccagccgggaat	GhCPK12 and GhCPK51
ttctagctctaaaacAAGGATGGCCTGAAAGGAACtgcaccagccgggaat	
ttctagctctaaaacTCCTGAAATAAGGCAGAAAGtgcaccagccgggaat	GhCPK13 and GhCPK54
ttctagctctaaaacCGTACGTGCCTTTAATCGTCtgcaccagccgggaat	
ttctagctctaaaacCAGTGCAGGTTCTAAGTAAGtgcaccagccgggaat	GhCPK14 and GhCPK55
ttctagctctaaaacCTTCTTGGTATGTGCTTCAGtgcaccagccgggaat	
ttctagctctaaaacCTCGAATTCAAGGCATCAGCtgcaccagccgggaat	GhCPK15 and GhCPK56

ttctagctctaaaacCGCGGCTCATGTTTTCCTCTtgcaccagccgggaat	GhCPK15 and GhCPK56
ttctagctctaaaacCAGAGAGTCTCTCTGCTATGtgcaccagccgggaat	GhCPK16 and GhCPK57
ttctagctctaaaacCTGGATTCCGCTGTATTATCtgcaccagccgggaat	GhCPK17 and GhCPK58
ttctagctctaaaacTCTTTAAGCTCCTCCAACGAtgcaccagccgggaat	GhCPK18 and GhCPK50
ttctagctctaaaacCGTATTACCATCCTGACAAAtgcaccagccgggaat	GhCPK20 and GhCPK60
ttctagctctaaaacCGGTGCACTTGAATAAACTGtgcaccagccgggaat	GhCPK21 and GhCPK61
ttctagctctaaaacACCAGATTCTCTTCCCTTTCtgcaccagccgggaat	GhCPK22 and GhCPK62
ttctagctctaaaacCGGAGATTATTTGCTTGATGtgcaccagccgggaat	GhCPK23 and GhCPK63
ttctagctctaaaacAAGAACTTCGGGTGCTATGTtgcaccagccgggaat	GhCPK24 and GhCPK64
ttctagctctaaaacTCTGCTTGCTTTGGCGCATCtgcaccagccgggaat	GhCPK25 and GhCPK65
ttctagctctaaaacTTCTGGGCAGGTTAGTAATCtgcaccagccgggaat	GhCPK26 and GhCPK66
ttctagctctaaaacGAGTTGGGATCAGTACACAAtgcaccagccgggaat	GhCPK27 and GhCPK67
ttctagctctaaaacTCAAGGGGGAAGCTTCCTTCtgcaccagccgggaat	GhCPK28 and GhCPK68
ttctagctctaaaacCTGTTATCGGAATCTATGGTtgcaccagccgggaat	GhCPK29 and GhCPK69
ttctagctctaaaacATAAACATCCTGCAATGACCtgcaccagccgggaat	GhCPK30 and GhCPK70
ttctagctctaaaacCCTCTTTGGGTTTCCTTTTCtgcaccagccgggaat	GhCPK31 and GhCPK71
ttctagctctaaaacCTTTGGGTTTCCTTTTCTCAtgcaccagccgggaat	GhCPK32 and GhCPK72
ttctagctctaaaacTTAGTGCTCTCTTCTTCAGCtgcaccagccgggaat	GhCPK33 and GhCPK74
ttctagctctaaaacGCTACTTCCTCAACTGACAAtgcaccagccgggaat	GhCPK34 and GhCPK75
ttctagctctaaaacCAGCACTCCAAACATCCACTtgcaccagccgggaat	GhCPK35 and GhCPK73
ttctagctctaaaacTCCTGCACGTTCAACAATCAtgcaccagccgggaat	
ttctagctctaaaacCAGTTGGGTTCAGGCTGATGtgcaccagccgggaat	
ttctagctctaaaacCTTGATTCTTTAGTCCTAGGtgcaccagccgggaat	
ttctagctctaaaacCGCTACCTATTGGCTTGTCTtgcaccagccgggaat	
ttctagctctaaaacCGTAGTTAATTCTCCCGTCCtgcaccagccgggaat	
ttctagctctaaaacCGGGCTAGACTGACTGCTGCtgcaccagccgggaat	
ttctagctctaaaacTGGCCAACAATTAGTGACAGtgcaccagccgggaat	
ttctagctctaaaacGGGTCGTCTTGTAGTAGAGCtgcaccagccgggaat	
ttctagctctaaaacTTTGGTACCGGATCTATCATtgcaccagccgggaat	
ttctagctctaaaacCTGCATCCTCATAAGCTCCTtgcaccagccgggaat	
ttctagctctaaaacCCAAAGAATGGCAAGCTTCCtgcaccagccgggaat	
ttctagctctaaaacTGGGATCTTGTCGTAGCATCtgcaccagccgggaat	
ttctagctctaaaacTTCCTTGGGATCTTGTCGTAtgcaccagccgggaat	
ttctagctctaaaacCGTGTCCTTGAGATTCTTCAtgcaccagccgggaat	
ttctagctctaaaacCCGACTTCTACGAACTGCAAtgcaccagccgggaat	
ttctagctctaaaacTTGCAAGAGAAGACGTGAAAtgcaccagccgggaat	
ttctagctctaaaacCGGCAGCTCTTTCAGTGTAAtgcaccagccgggaat	
ttctagctctaaaacACGCGAGGTTGCTATAATGTtgcaccagccgggaat	
ttctagctctaaaacGTTCTAACATCTGCCGAACAtgcaccagccgggaat	

ttctagctctaaaacCGTTTTGGGAAATCCATGGAtgcaccagccgggaat	GhCPK36 amd GhCPK76
ttctagctctaaaacCATGGTTCGTCTTCCGATTCtgcaccagccgggaat	
ttctagctctaaaacCAATACTTCTGGAGCTACATtgcaccagccgggaat	GhCPK37 and GhCPK77
ttctagctctaaaacTCACGATCTGCCGGCACAAAtgcaccagccgggaat	
ttctagctctaaaacCTTCACATGAGGAGGCTTAGtgcaccagccgggaat	GhCPK38 and GhCPK78
ttctagctctaaaacTGCAAGACTTGGTCGAATATtgcaccagccgggaat	
ttctagctctaaaacTTCAGAACGGTTGACAGCTCtgcaccagccgggaat	GhCPK39 and GhCPK79
ttctagctctaaaacTCCCGCTGTTTACAGTTCCTtgcaccagccgggaat	
ttctagctctaaaacCGCGTTCTAGTTTATTGAGGtgcaccagccgggaat	GhCPK40 and GhCPK80
ttctagctctaaaacTAAACATTTCTCGCAAGCCAtgcaccagccgggaat	
ttctagctctaaaacCGGTTTTATTTTGCAGGCTGtgcaccagccgggaat	GhCPK41 and GhCPK81
ttctagctctaaaacCCGGGCTTGTTATCAGATAGtgcaccagccgggaat	
ttctagctctaaaacGGGCAGCTGTTAACCTCCTTtgcaccagccgggaat	GhCPK42 and GhCPK82
ttctagctctaaaacTGCCTGAGTTCTTCAAGTGTtgcaccagccgggaat	

Table S4 Validation of vector library construction for pool 1

Targeted gene ID	ırgeted gene nar	sgRNA sequence	Count	Gene covered
Gh_A01G0621	GhCPK1	TCCTCTTCTCCGTTTCTGTG	4	Yes
		TGGTGGAATTGGGGTTGAAC	6	
Gh_A01G1119	GhCPK2	AATGAAGCAGACCATCTAGT	2	Yes
		ACAAGGAAGGGTCCAAGACA	0	
Gh_A02G0144	GhCPK3	GCAAGTTGTCGGCAACCACC	5	Yes
		CCTTCCCTAAGTGTTTGAAT	4	
Gh_A02G0609	GhCPK4	GACAGCAGCAACCAAAGGAT	2	Yes
		CGACAGCAGCAACCAAAGGA	1	
Gh_A02G1029	GhCPK5	CAGTGGAACAATTAGCTTCG	2	Yes
		ATTGCACCACAGATCATCAA	5	
Gh_A02G1635	GhCPK6	TCTTAGGGTAAAAGCACCAA	1	Yes
		AGTGCACCGTGACATGAAAC	4	
Gh_A02G1796	GhCPK7	GCAAGCTTAGGTAGAACTTA	0	Yes
		TCACACCCAAGTTCATATCT	1	
Gh_A03G1505	GhCPK8	GCAAATATGGAGAGATCCGT	1	Yes
		CAGTCCATATTACGTTGCTC	3	
Gh_A04G0148	GhCPK9	GTTGGACACAAGTATGTGAA	3	Yes
		ATTCACAATTGCCCTGCACA	3	
Gh_A04G0467	GhCPK10	GAACAATGCCAATAAGGCTT	3	Yes
		GGTTGGAAGTCCATATTATG	4	
Gh_A04G0780	GhCPK11	GTAGTAATGAACCAACCCGT	1	Yes
		GAGGAAACTCAGAACTGCCG	1	
Gh_A04G1372	GhCPK12	GGTCCTAAACCAGCCGAAGA	2	Yes
		TTGCAGAAAACTGCTTCATG	1	
Gh_A04G1429	GhCPK13	GTCGATATCGAAGACGAAGT	0	Yes
		GCTAAGAAGAAGATGGTTGG	3	
Gh_A05G0617	GhCPK14	CGCTGCTTCTTTGCTTCGCA	2	Yes
		GTCTAATATTGTCGAGCTCA	2	
Gh_A05G1571	GhCPK15	CTAACCACTGACAAGGACGT	3	Yes
		GTTATTTGAATCACGGCCAA	3	
Gh_A05G2355	GhCPK16	TTCAGCATAGCCGATGCCTT	1	Yes
		TGTTGGGTCTTTCATTCCAA	2	
Gh_A05G2859	GhCPK17	TAACATAAGACAACGTACCG	3	Yes
		ATCCAAAGCCAAGCTGCGGA	2	
Gh_A05G3246	GhCPK18	GGGTTCTGAACTTATGGAAT	5	Yes
		GTTCGGAGTGACGGCAAATG	5	
Gh_A06G0013	GhCPK19	GATTTGTTGGACAGGATGGG	2	Yes
		ACAGAAGCAAGCCCGATTGA	5	
Gh_A06G1772	GhCPK20	AGTGAGGAAGATGTTGACAA	0	Yes
		CTTGTTTTGTTGGCTAATGG	3	

Table S5 Primers for sgRNA high throughput detection

Primer name	Primer sequence
GTL-1-F	GACgcctagAGAAGCATCAGATGGGCAAA
GTL-2-F	GACgcgctaAGAAGCATCAGATGGGCAAA
GTL-3-F	GACgcggatAGAAGCATCAGATGGGCAAA
GTL-4-F	GACgcgtcaAGAAGCATCAGATGGGCAAA
GTL-5-F	GACgctacgAGAAGCATCAGATGGGCAAA
GTL-6-F	GACgtacgaAGAAGCATCAGATGGGCAAA
GTL-7-F	GACgtcacgAGAAGCATCAGATGGGCAAA
GTL-8-F	GACgtcgctAGAAGCATCAGATGGGCAAA
GTL-9-F	GACgtcgtaAGAAGCATCAGATGGGCAAA
GTL-10-F	GACgtggacAGAAGCATCAGATGGGCAAA
GTL-11-F	GACgacatgAGAAGCATCAGATGGGCAAA
GTL-12-F	GACgacgtaAGAAGCATCAGATGGGCAAA
GTL-13-F	GACgactgtAGAAGCATCAGATGGGCAAA
GTL-14-F	GACgagtcaAGAAGCATCAGATGGGCAAA
GTL-15-F	GACgatcgaAGAAGCATCAGATGGGCAAA
GTL-16-F	GACgcaccgAGAAGCATCAGATGGGCAAA
GTL-1-R	GACacgtcaGACCCGAATTTGTGGACCTG
GTL-2-R	GACaccatgGACCCGAATTTGTGGACCTG
GTL-3-R	GACacgagcGACCCGAATTTGTGGACCTG
GTL-4-R	GACacagcgGACCCGAATTTGTGGACCTG
GTL-5-R	GACatgatgGACCCGAATTTGTGGACCTG
GTL-6-R	GACatcatcGACCCGAATTTGTGGACCTG
GTL-7-R	GACatggtcGACCCGAATTTGTGGACCTG
GTL-8-R	GACatgctgGACCCGAATTTGTGGACCTG
GTL-9-R	GACagcacgGACCCGAATTTGTGGACCTG
GTL-10-R	GACagctcaGACCCGAATTTGTGGACCTG
GTL-11-R	GACagatgtGACCCGAATTTGTGGACCTG
GTL-12-R	GACagacgaGACCCGAATTTGTGGACCTG

Table S6 Primers for gene editing detection of *GhCPKs* mutant library materials

Primer sequence	sgRNA	Gene ID	Gene name
TCCATTTTGGGCAGGTAA AGGTCCTTGGCACTCTCTG	TCCTCTTCTCCGTTTCTGTG	Gh_A01G0621	GhCPK1
AATGACCCAAAGACAACAATC GTGGTCCCGAACTGTCCT	ACAAGGAAGGGTCCAAGACA	Gh_A01G1119	GhCPK2
TAATCAATCATCATCAATACCGA TCGTAATCTTCACGGCATAA	GACAGCAGCAACCAAAGGAT	Gh_A02G0609	GhCPK4
TAATCAATCATCATCAATACCGA TCGTAATCTTCACGGCATAA	CGACAGCAGCAACCAAAGGA	Gh_A02G0609	GhCPK4
AGAAGAAATTATTGGTTTGAAGG GTAAATTGTTTAAAAGCACAGTGA	CAGTGGAACAATTAGCTTCG	Gh_A02G1029	GhCPK5
AATCTTAGTTTTCCTTCAAGTTTC GAGCACTTTACCTTTAGAGCAA	ATTGCACCACAGATCATCAA	Gh_A02G1029	GhCPK5
AAAGTTGCGGCTGAGTGTC AGTCTGGACGTGTTGACTTAAA	AGTGCACCGTGACATGAAAC	Gh_A02G1635	GhCPK6
GATGGATACTGCCAATAGAGG GTCTGTGCCGCCAAATATC	GCAAGCTTAGGTAGAACTTA	Gh_A02G1796	GhCPK7
GAAAAAAGAAGCAAAATCCATT CAATGTCAACGTCAGTCCTTAG	TCACACCCAAGTTCATATCT	Gh_A02G1796	GhCPK7
ATAATGAAACCGATGTCGTCTG ATACATCTGCTTCTGGTCCATAAT	CAGTCCATATTACGTTGCTC	Gh_A03G1505	GhCPK8
TCAACTCAAGCCCTAGGTACTT CATAAACCGCATAATTTCAATC	GTTGGACACAAGTATGTGAA	Gh_A04G0148	GhCPK9
ATGTGCTGGTGGTGAGTTGT CGAAATCAGTAGCCTTCAAAA	ATTCACAATTGCCCTGCACA	Gh_A04G0148	GhCPK9
TGAGAACACTGCCCCTGTAC CGAATTGTCCTTGACCTAACTT	GAACAATGCCAATAAGGCTT	Gh_A04G0467	GhCPK10
GAGAGTTCGGCGTTACTTATCT AACCGCATTGTCGTCCTC	GAGGAAACTCAGAACTGCCG	Gh_A04G0780	GhCPK11
ATGGGAAATGTTTGTGCTACA GCCTTTGGCTTAGGGTCT	GGTCCTAAACCAGCCGAAGA	Gh_A04G1372	GhCPK12
CCAGCGAACAAGAAAGCA TCCCTCCTAACATCATCCAC	GTCGATATCGAAGACGAAGT	Gh_A04G1429	GhCPK13
AAGCATTCCGTTCATTTGG	CGCTGCTTCTTTGCTTCGCA	Gh_A05G0617	GhCPK14

AGGGGAATCTTCATCCTTATT	CGCTGCTTCTTTGCTTCGCA	Gh_A05G0617	GhCPK14
CAAGGAGGATATTGAGGATGTT	GTCTAATATTGTCGAGCTCA	Gh_A05G0617	GhCPK14
CGTGCGAAGCAAAGAAGC			
ATTGAAAAGGCTACGAAACG	CTAACCACTGACAAGGACGT	Gh_A05G1571	GhCPK15
CATAAGCACCTTTGAACTCCA			
TTGTGGAGTGCCTTCAGC	GTTATTTGAATCACGGCCAA	Gh_A05G1571	GhCPK15
TATCTGGTAAAAGGGAACGAC			
ATTTCTTATTCGGAGGTGTTCT	TGTTGGGTCTTTCATTCCAA	Gh_A05G2355	GhCPK16
TCATTTAATATCCTGTGTTGGG			
GTTTCGACATAAGAACGGGA	ATCCAAAGCCAAGCTGCGGA	Gh_A05G2859	GhCPK17
ATGACCTTTAGCGACGATTC			
AGAAATCGGTGGTCTGAAAG	GGGTTCTGAACTTATGGAAT	Gh_A05G3246	GhCPK18
TAGCAAAGGATCAAAAGTAAGTG			
AGAAATCGGTGGTCTGAAAG	GTTCGGAGTGACGGCAAATG	Gh_A05G3246	GhCPK18
TAGCAAAGGATCAAAAGTAAGTG			
ATGGATCGGGAAGAACATCT	CACTTCGACAAAGACAATAG	Gh_A08G2530	GhCPK23
GCATACCATACTCACGGAGAG			
GCGCATATTATATTGCACCAG	GTATGGCACGTTCATGATGT	Gh_A08G2530	GhCPK23
CCTTTTAGTTTCGCCATGAAT			
TTTTACAAGCCAGGTTGAGAT	CTTCACCTGTAAGAGATCCA	Gh_A09G1033	GhCPK24
TTCAGGTCCATATTTCCTCTTT			
CCATCACGGTGGAACAAG	GGGGAGGCTGTATTATTCCT	Gh_A09G1033	GhCPK24
ATGTTGTAGGTGGCTGTAATGT			
GAAGCATTTGCCCAAGC	CAAGGGGACACTACACTGAA	Gh_A09G1067	GhCPK25
AACCTGAACAACTTCGACAAT			
TCTGCGAGATTCCTTAAACG	TGGAGATGTTTTGGTTCAAG	Gh_A09G1067	GhCPK25
TGCAACAAACTCCTCGTAACT			
TTTACTGCCATCTCCATTCAC	AGATGAGTTGGCTGATGAGC	Gh_A09G1157	GhCPK26
ACCAGTCTAACCTTGTCTGTGTC			
CATTTTGGGGTGGTAAGACT	TGGTGAACCTGACTTCACAT	Gh_A09G1248	GhCPK27
TTGGGGTTTCTAACAAGCAT			
TGCCGCTTGTTTGATGTTA	GGTGGATGAAGCCAGCATCA	Gh_A10G0124	GhCPK28
TGTTCCAAAAGTTTGATTTGC			
ATGGGTTGTTTTAGCAGCA	TAGCAGCAAGCATAAATTGC	Gh_A10G0124	GhCPK28
AAGGCTTCAAAGGGACAGA			

CGGTATAAGCATAGTGGGACTT	TTAGAGGCCCTGAAAGGCCA	Gh_A10G0886	GhCPK29
GAACTGAATGAACAGCCATAAT			
TGTACCGATTCCAAAACATTAT	AAGTGGAACACTAAGCTACG	Gh_A10G1195	GhCPK30
AAGTTCTTCGACTTCGATGTAA			
CCCGAATGCCTGATGACT	TAAACCAGAGACCAAACAAG	Gh_A10G1756	GhCPK31
CTGCACTGGACACCCTTTT			
AAATGCAGAGTGAGGTCAACAT	TATCTTTCAGGTCTGCAGGA	Gh_A11G0213	GhCPK32
CGATATTCATCACCCTTTTCTT			
ACGCTAAGGAAAAGGATGGA	TTTCCTCCTTCTCTAACCCA	Gh_A11G1615	GhCPK33
CGGGAAGACAAACTAGATTACC			
TGTGACGGGATAGACCTTGT	GTAAACATCCTGTCAGCGTC	Gh_A11G2941	GhCPK34
TGGTCTTACTGTGCCTGCTA			
TGGTTGGAAGCCCGTATT	CAAGCATTACGGACCAGAAG	Gh_A12G0109	GhCPK36
ACACATAGAACTGCTTGAAAACTC			
TATTTCATTTGTTCTCAGGCTG	ATAGTGGGAGCATCGATTAT	Gh_A12G0109	GhCPK36
GATGTAGCCACTGTTGTCCTTA			
CATTTCGGATGTGATTCTTTC	ACCGTGTTCTTCCTACAGGA	Gh_A12G2686	GhCPK37
TGCGTCTTCGGTTATTGG			
CATTTCGGATGTGATTCTTTC	TCCATCCTGTAGGAAGAACA	Gh_A12G2686	GhCPK37
TGCGTCTTCGGTTATTGG			
AAGGAATCTTCGCTGCTGT	TGTGCATTCACCCTTCAGAA	Gh_A13G0563	GhCPK39
CATGACCATCAAACAGATAATAAT			
GGACAGGGACAGTTTGGG	CAGCCAAATGGTGCATAATT	Gh_A13G0566	GhCPK40
CTATCAAACAGTTCGCCTCC			
GGACAGGGACAGTTTGGG	TATGTACCGAGATTTCGACT	Gh_A13G0566	GhCPK40
CTATCAAACAGTTCGCCTCC			
AATGCTAATGTGTGACCTGACT	AACATCGATACTGACAACAG	Gh_A13G1164	GhCPK41
TAGATCCCAGTCGAGCTAATC			
GGGACAACCTAATATCGTGG	TTGGTGATGGAATTGTGCGG	Gh_A13G1891	GhCPK42
GCATAATCGTCCGAAGCA			
TGGTGCTAATCTTGCTGAATC	CGTGCTCTAATGCAGGCCGT	Gh_D01G1194	GhCPK43
ATCTGCCTGTGATTCCCAA			
TTGATCTTATCCTTAACCCCTAA	CCTGATTAAGATGGAAGGTA	Gh_D01G2360	GhCPK44
AAACTTCAAAATGCCATAGTCA			
TGATGGACACCGACGAGA	CAGAAAATCGGTCACTCTGT	Gh_D02G0183	GhCPK45

GATTATGTCCGCAACCGTA	CAGAAAATCGGTCACTCTGT	Gh_D02G0183	GhCPK45
CTTTACCTGAATACCCCGAC	TCAACCGTAAAATGAGCGGA	Gh_D02G0663	GhCPK46
CCTTCATACAAGCAAGAAAAAT			
TAATCAATCATCATCAATACCGA	CCGACGACAACAGCAACCAA	Gh_D02G0663	GhCPK46
TCGTAATCTTCACGGCATAA			
AGCTGTTTCTCATTAAACATGCT	AGCTGAAAGTCTTTCCGAGG	Gh_D02G1973	GhCPK47
TTCAAGGTAGAGCCGTATCTTC			
TGTGTGGTTTACAATGTTAGGC	ACTTCGGGAGCAACGTAATA	Gh_D02G1973	GhCPK47
ATCAATACCTGCCCAAAATG			
ACCGAAGAAACAAAGTGGG	CACTTACGTTGCAATAGACA	Gh_D03G0087	GhCPK48
ATCCCAGAAGCACAAATCAA			
TTGCCCGTTTTATTAGTAGGA	TAGCAGTTCGTAAAAGCCGT	Gh_D03G0087	GhCPK48
CTACAACTTCTGAGTATTCCGATT			
AGAAAGGAAAAAAGAAGCAGAA	GGTCAACATCACCATGGCAA	Gh_D03G0609	GhCPK49
CAATGTCAACGGCAGTCC			
TGTAATAAACAGGATGGACAAAT	GCGTTCGGGGTGAAGTCAAA	Gh_D04G0366	GhCPK50
TCAAACTGAACTATTCAATTCCTT			
CCGACAACGAAGGTGAGA	TGGCTTAGGGTCTCTTCTCC	Gh_D04G0895	GhCPK51
ATAGGCGTGGACATTGACAT			
GCGTCCGACATGGATATATTA	CCTGTCTGTGCAATAAGCAA	Gh_D04G0895	GhCPK51
GGGCTTCCAACCACATCA			
CTTACGAAGATGCTGTGGCTAT	AAGGGCAGCTGCACTTGCTA	Gh_D04G0900	GhCPK52
ATTTTCGGGCTTGAGGTC			
GAGTTCGGCGTTACTTATCTTT	CGATCTCGAAGAGGAAACTC	Gh_D04G1271	GhCPK53
GAACCGCATTGTCGTCCT			
ATAAGACTCCCAACATTCGTG	TGTGTACAGAGCTTTCCACG	Gh_D04G1486	GhCPK54
ATCTCCCTCCTAACATCATCC			
TGGCTCATTGTTTTGTTTCG	CATATAACTGCGTTGTCAAG	Gh_D05G0748	GhCPK55
CGCAAAGCAACTTTCTTGAA			
AAAGACAACAGCGGGTAATG	CAGAGCTTGCTCTAGTTCTT	Gh_D05G0748	GhCPK55
CAACTTCAGAAACGATTTCCTT			
AAAGCCCTTCAAATCTCCC	GGTTTCTTCAAGTTGAACAC	Gh_D05G1749	GhCPK56
ACTGGACTGCACACGAGCT			
ATTTCTTATTCGGAGGTGTTCT	GTCTTTCATTCCAAACGCAT	Gh_D05G2622	GhCPK57
TCATTTAATATCCTGTGTTGGG			

GGCAGAGGCATTTATGTTGT	TGGTTGTTGCTGAATTGTGG	Gh_D05G2622	GhCPK57
AAGAAGAATGTCCGCTACACA			
CAATTATCGTTGGAGGAGCT	TGAGGTTCAAATGTTGATGC	Gh_D05G3156	GhCPK58
ACGTACACCATTTTCATGTAAAT			
AGGATTATTGCCAAAGGACA	TGAAAGAGCAGCTGCTTCCA	Gh_D05G3567	GhCPK59
CCGATAACCCGAAATCAGTA			
AGGATTATTGCCAAAGGACA	GAAAGAGCAGCTGCTTCCAT	Gh_D05G3567	GhCPK59
CCGATAACCCGAAATCAGTA			
AGATGTTCTGTTCGGATTGTATT	TTAGTTTCGAAAACGACCCA	Gh_D06G2206	GhCPK60
CTCATTCGGGTCCATCGT			
AATCAATAATTTCCTGAAAATTGT	GTGTTGTGTAGGTGATTGCG	Gh_D07G1198	GhCPK61
CGCTGTTATCGGTGTCCA			
TTCTGAAGTTGACATCGATAATG	ATCACCAACCCACTGAATAG	Gh_D07G1198	GhCPK61
TTTCCTCATCATTGCCACA			
CAGATTTGCAGATCCTAGTGG	GCACTTCACCCATATTATGT	Gh_D07G1228	GhCPK62
TGTAAGTGTTCATCATTCGCC			
TAACAAAAGGCGGATGCG	AATGTTCTTCCCGATCCATT	Gh_D08G0142	GhCPK63
ACTATCGTTACCCGCTGTTG			
CATTGAAAACACACCCCAT	CAACAACAAGAACAATTTCA	Gh_D08G0142	GhCPK63
GTTATCGTGCCGCTGTTAT			
TGCCAATGAAAAGGGAGAC	GAAGAAGCCAAGAGTCCTAA	Gh_D09G1054	GhCPK64
CCAATGTTGTAGGTGGCTTTA			
ACAATAGCGGGTAATCATAGTTT	TAGTGCTTGCTCTAGTTCTT	Gh_D09G1054	GhCPK64
TATCAATATCAACTTCGGAAATG			
CGGTTCCGTTGGATTCTT	TATGCAAGAGAATGAACGAG	Gh_D09G1074	GhCPK65
AATTCAGAGTTCCATCACCATC			
TCAATGAGATTGTGGGAAGC	AGGTTGATGTCTGGAGTGCT	Gh_D09G1074	GhCPK65
CTGGATGCACACATTGTCTCT			
GAAGAAGTCATTACTGCCATCA	GTCACATTACTCAGAACCAG	Gh_D09G1163	GhCPK66
TTGCTTCCCATTCTTCTTATT			
TTTACTGCCATCTCCATTCAC	AGATGAGTTGGCTGATGAGC	Gh_D09G1163	GhCPK66
ACCAGTCTAACCTTGTCTGTGTC			
AGGGAACAAGAAGGGAAATC	ACAGACTGCTATAAAGAACG	Gh_D09G1249	GhCPK67
TGCCTCTTAACATTATGGGG			
TGTAAGCCAAACGAGAAGGA	AATCCTGAAGTTGGTAAGAG	Gh_D09G1249	GhCPK67

CAATGCCTCCCTAAATCCA	AATCCTGAAGTTGGTAAGAG	Gh_D09G1249	GhCPK67
CAGTTCAAAGCCAAGCCTA AAAGAATACCTTCATTTTCTCAAT	GTTGCCATAGGCAATGAAAA	Gh_D10G0863	GhCPK69
CAGTTCAAAGCCAAGCCTA AAAGAATACCTTCATTTTCTCAAT	GGATACACCTACGTTGCCAT	Gh_D10G0863	GhCPK69
TTGGGACGACTTTCCTTTG ACAACATGAACTGCCACTGC	CTTGCAAATCGATTGCAAAG	Gh_D10G2029	GhCPK71
ATGGGTAACTGTAACCGTCC CGATGGGTGGAGTGTTGG	TCATTATGATGGGTACCGCC	Gh_D11G0227	GhCPK72
ATGGGTAACTGTAACCGTCC CGATGGGTGGAGTGTTGG	CATTATGATGGGTACCGCCT	Gh_D11G0227	GhCPK72
TGTGTGGGTTTTGGCTTG AAATATCAAGCTGAAATCCATG	AAAGAGTCCCACAGAAACAG	Gh_D11G0314	GhCPK73
TCCGAGTTCTGAAAGATGTGA CTTCCGCTTGGAGATTGAC	AGCTCGGCCGAGGTGAATTT	Gh_D11G0314	GhCPK73
AACCCTCTGCCACCACTG TTCCCAATGGTGTATCGCT	AAATCAGTACGCTTCCCGCA	Gh_D11G1774	GhCPK74
TCACGCCCGTAAAGATTCT GCGGTCGATGCAAAGATA	CGTGCTTGCAGGAGTTCCAA	Gh_D11G3329	GhCPK75
GGTTGGAAGCCCGTATTA AGAACCAAAGAAAAGAAACATC	CAAGCATTACGGACCGGAGG	Gh_D12G0121	GhCPK76
GCTATTTCATTTGTTCTCAGGC GTTATGCTCGGCACAGGC	TAGCGGGAGCATCGATTACG	Gh_D12G0121	GhCPK76
TTAAGCACGCCTATGGAATA ATCTTTGGCACTGTTGGAAA	TTGATTTCGCATCAGATCCA	Gh_D12G2743	GhCPK77
GACACAGACAGGGTATGCTTC CTTGCGTCGTCGGTTATT	TAAGTGTTCTTCCTACAGGA	Gh_D12G2743	GhCPK77
ATGGGTAATACTTGTGTAGGACC TCTTTAGATGCTACTTCATTCCC	AATCTGTTTCAGCCGTAATG	Gh_D13G0033	GhCPK78
CGCTGCTGTGCTAAAAGG ATTGCTTTGTATCAAGGTTCAG	GGCGAATGCACAACATCTTT	Gh_D13G0560	GhCPK79
GACAAGGACAGTTTGGGATTA CGTATAAAGTATCCTCATAAGCG	GGATGTGGAGGATGCTAGGA	Gh_D13G0560	GhCPK79
GAAGAGCCATTGATCACCAA AATCTGAATCTCCCTCCTAACA	GGCCTAAAGTGTAAAGCTCA	Gh_D13G0561	GhCPK80

GCAGAAACTGAGAAGGGAATAT	GTGCCAAGGATCTAGTCTGC	Gh_D13G1455	GhCPK81
AAATAAAGGCTGCTGTAAAGAA			
TCCCACCGTTACTTCACAT	ATCGGAAAAGAACTAGGTCG	Gh_D13G2278	GhCPK82
CCGTTTCGCTATCGTCTTA			
TCCCACCGTTACTTCACAT	AACTAGGTCGTGGTCAATTT	Gh_D13G2278	GhCPK82
CCGTTTCGCTATCGTCTTA			
ATGTGCAGGGTTGCAGGT	CTTTGCATCGAGAAAGGTAC	Gh_A01G1119 and Gh_D01G1194	GhCPK2 and GhCPK43
TCAGTCGTCGTGAGCTTCC			
TGACTTCTCATCTGATCCCTG	TGCTCGACAGAGATACGAAG	Gh_A01G1119 and Gh_D01G1194	GhCPK2 and GhCPK43
ATTCGGATGGTTTTGATTAGT			
ATTTGATTATGGAACTGTGTGC	GGAAAGCAGCTAACTTGTGT	Gh_A02G1029 and Gh_D03G0701	GhCPK5 and Gh_D03G0701
GAGCCCGAAATCAGTAGGTT			
GACCCTTTACTTGAGGAGGC	TGAGTATTCCGATTACCGGA	Gh_A02G1635 and Gh_D03G0087	GhCPK6 and GhCPK48
GATACCAGGCAAGCACAAA			
TGTTGATTTCACCGAGTTTGT	AACTCCGGAGGAACTACGAA	Gh_A02G1635 and Gh_D03G0087	GhCPK6 and GhCPK48
ATTGTCACAAGAGTGCCAAAT			
AGAAAGGAAAAAAGAAGCAGAA	GTTCTAAATGATCCAACTGG	Gh_A02G1796 and Gh_D03G0609	GhCPK7 and GhCPK49
CAATGTCAACGGCAGTCC			
GCATTTGTTTTTACAGGATGGT	TGAGAAGTGGAACCCAACAA	Gh_A04G0148 and Gh_D05G3567	GhCPK9 and GhCPK59
AGAAAACCTGGAACATCTCTGT			
CTTACGAAGATGCTGTGGCTAT	AAAGGCAGCTGCGCTTGCTA	Gh_A04G0467 and Gh_D04G0900	GhCPK10 and GhCPK52
GGCTTGAGGTCACGATGC			
TCTTGCTGGTTGTAGATTTTATTC	TGGCTCCTGAGGTTCTCAAA	Gh_A04G0780 and Gh_D04G1271	GhCPK11 and GhCPK53
AAAGACACTTACCAGCCCAA			
TCGTGGGTTAATAGATTTCAAA	TGGAACCAGACCCAAAGCTC	Gh_A04G0780 and Gh_D04G1271	GhCPK11 and GhCPK53
ATCCACCTCCCTTCCCTG			
GATTCTTTGGTGTTAGACTATGGA	GTTCCTTTCAGGCCATCCTT	Gh_A04G1372 and Gh_D04G0895	GhCPK12 and GhCPK51
GTCTCAGCGGCAGAAACTTA			
CAAAGAGGATGTGGATGATGTT	GACGATTAAAGGCACGTACG	Gh_A04G1429 and Gh_D04G1486	GhCPK13 and GhCPK54
GTAATGCCCCCTCTGGATAA			
GGGTCACTCTGGCAAATACTA	CATAGCAGAGAGACTCTCTG	Gh_A05G2355 and Gh_D05G2622	GhCPK16 and GhCPK57
ATCGTCCCACTGTTGTCTGT			
ATACAAAGCGAAAGGGAAGG	TTTGTCAGGATGGTAATACG	Gh_A05G2859 and Gh_D05G3156	GhCPK17 and GhCPK58
GCCTCTGCCCAACTCTTTT			
CATTGACAACAACGGGACA	GAAAGGGAAGAGAATCTGGT	Gh_A05G3246 and Gh_D04G0366	GhCPK18 and GhCPK50

ACGTCGCTTAATCCAAACTC	GAAAGGGAAGAGAATCTGGT	Gh_A05G3240 and Gh_D04G0500	GhCPK18 and GhCPK50
AGTGGGTGAGTTCCGAAATA	CATCAAGCAAATAATCTCCG	Gh_A06G1772 and Gh_D06G2206	GhCPK20 and GhCPK60
ATAGAAGAAGTTTCCGAGTTTG	ACATAGCACCCGAAGTTCTT	Gh_A06G1772 and Gh_D06G2206	GhCPK20 and GhCPK60
TTTCGGTTTGTCTGTCTTTATC	GATTACTAACCTGCCCAGAA	Gh_A07G1099 and Gh_D07G1198	GhCPK21 and GhCPK61
TACAAGATAACACCAGCACTCC	GGTCATTGCAGGATGTTTAT	Gh_A08G2530 and Gh_D08G0142	GhCPK23 and GhCPK63
CCAACAGGTGAAGAATTCAAAG	TGAGAAAAGGAAACCCAAAG	Gh_A09G1033 and Gh_D09G1054	GhCPK24 and GhCPK64
AAAAGGAAATGAGACTGGCAA	GCTGAAGAAGAGAGCACTAA	Gh_A09G1067 and Gh_D09G1074	GhCPK25 and GhCPK65
CATTGAAAACACACCCCAT	AGTGGATGTTTGGAGTGCTG	Gh_A09G1157 and Gh_D09G1163	GhCPK26 and GhCPK66
GTTATCGTGCCGCTGTTAT	CCTAGGACTAAAGAATCAAG	Gh_A09G1248 and Gh_D09G1249	GhCPK27 and GhCPK67
TATATCACTCACGGTATATCAAAAA	GGACGGGAGAATTAACTACG	Gh_A10G0124 and Gh_D10G0130	GhCPK28 and GhCPK68
TCATAAAATATAATCAACATCACGC	CTGTCACTAATTGTTGGCCA	Gh_A10G0886 and Gh_D10G0863	GhCPK29 and GhCPK69
GCTACAAAATGCCAAGAAAG	GCTCTACTACAAGACGACCC	Gh_A10G1195 and Gh_D10G1303	GhCPK30 and GhCPK70
CTCAGCAATCACCTGGAAAT	ATGATAGATCCGGTACCAAA	Gh_A10G1195 and Gh_D10G1303	GhCPK30 and GhCPK70
AGGTGAAAGGTTTACGGAGAT	GGAAGCTTGCCATTCTTTGG	Gh_A10G1756 and Gh_D10G2029	GhCPK31 and GhCPK71
GCCCTTGTGCGGTAGAAT	TACGACAAGATCCCAAGGAA	Gh_A11G0213 and Gh_D11G0227	GhCPK32 and GhCPK72
CTGATTATTAGGGTTAGGTTGC	TGAAGAATCTCAAGGACACG	Gh_A11G1615 and Gh_D11G1774	GhCPK33 and GhCPK74
TCACGGGTAGAAACTTACTCTTA	TTGCAGTTCGTAGAAGTCGG	Gh_A11G1615 and Gh_D11G1774	GhCPK33 and GhCPK74
GGCTGTAATCTCTGAAGTGGA	TTTCACGTCTTCTCTTGCAA	Gh_A11G2941 and Gh_D11G3329	GhCPK34 and GhCPK75
CTTTTTCAGTTTGTTGTGTTCC			
ATTTTGTTCACCTTCCAAGTTAC			
GAGCAGCAGTCAGTCTAGCC			
CAGCGACGACCAACTCTCT			
CATGAAACATAAGATCAAAAACCT			
CACAAGCGAAACTAATAGATTCA			
CGCCTATCCCCAACTTCT			
TGTGCTGGTGGAGAACTTTT			
TCCTCCTGCTGATTCACAAA			
TTCCGACAGTGCCAAAGAT			
CAGATGATCTGCAGTTTATTTCTT			
GCTTCTTTGTATTCGTCCTGAT			
CTCCACAACAATCAAAACCAC			
AGGTTCCATTGACCCTTTG			
ATCAGATGAACAAGGTACATTCC			
ATGGGGAATTGTTGTAGATCTC			
TAAGTAACTCCGAACTCACCTC			

CGTGCAAGTCAATCTCCAA TTCACAAAGCTCCATAACCA	ACATTATAGCAACCTCGCGT	Gh_A11G3011 and Gh_D11G0314	GhCPK35 and GhCPK73
TGTTGTTTGAAAATGTATCCCA TAGTTTGTTCATTGCCGAGA	TCCATGGATTTCCCAAAACG	Gh_A12G0109 amd Gh_D12G0121	GhCPK36 amd GhCPK76
AGGTAAAGACAGATACTTTTCGC CAAATATCAGCTTCATGTCCAT	ATGTAGCTCCAGAAGTATTG	Gh_A12G2686 and Gh_D12G2743	GhCPK37 and GhCPK77
TAGGGAAACTGGGAATGAAG AAAGGAAAGTTGTCCCGAAT	CTAAGCCTCCTCATGTGAAG	Gh_A13G0017 and Gh_D13G0033	GhCPK38 and GhCPK78
TTTGATCTTCACATTTTTTTCTC ACCGTTTCTTGGGGTCTC	ATATTCGACCAAGTCTTGCA	Gh_A13G0017 and Gh_D13G0033	GhCPK38 and GhCPK78
GCCCCTAATATCCGATAGC ATGATTCAAAAAGAAATGGAGA	GAGCTGTCAACCGTTCTGAA	Gh_A13G0563 and Gh_D13G0560	GhCPK39 and GhCPK79
GCCTGATCCTTTTGTGTTTTC TCCCGTATCTCTGTATCCTTCA	TGGCTTGCGAGAAATGTTTA	Gh_A13G0566 and Gh_D13G0561	GhCPK40 and GhCPK80
AAGAAATAAGGGGATTGAAAGA	ACACTTGAAGAACTCAGGCA	Gh_A13G1891 and Gh_D13G2278	GhCPK42 and GhCPK82

Table S7 The other primers used in this study

Primer	Sequence	Note
GhCPK33-OE-S	ATGGGAGCCTGTCTCTCC	Overexpression
GhCPK33-OE-A	CTATAACTTCCGAGAATTCCG	Overexpression
GhCPK74-OE-S	ATGGGAGCCTGTCTCTCC	Overexpression
GhCPK74-OE-A	CTATAACTTCCGAGAATTCCG	Overexpression
GhCPK33-QRT-S	CACCGCAAACATCAACCCT	qRT-PCR
GhCPK33-QRT-A	GTAGAACTTGGCTTCCCTTTGA	qRT-PCR
GhCPK74-QRT-S	CACCGCAAACATCAACCCT	qRT-PCR
GhCPK74-QRT-A	CGAACTTGGCTTCCCTTTGA	qRT-PCR
GhCPK33-104-S	CAGGTACCCGGGGATCC ATGGGAGCCTGTCTCTCC	BiFC
GhCPK33-104-A	CTGCCACCGCCGTCGAC CTATAACTTCCGAGAATTCCG	BiFC
GhCPK74-104-S	CAGGTACCCGGGGATCC ATGGGAGCCTGTCTCTCC	BiFC
GhCPK74-104-A	CTGCCACCGCCGTCGAC CTATAACTTCCGAGAATTCCG	BiFC
GhCPK33-BD-S	ATGGAGGCCGAATTC ATGGGAGCCTGTCTCTCC	Y2H
GhCPK33-BD-A	GCAGGTCGACGGATCC CTATAACTTCCGAGAATTCCG	Y2H
GhCPK74-BD-S	ATGGAGGCCGAATTC ATGGGAGCCTGTCTCTCC	Y2H
GhCPK74-BD-A	GCAGGTCGACGGATCC CTATAACTTCCGAGAATTCCG	Y2H
GhCPK33-771-S	CTCGGTACCCGGGATCC ATGGGAGCCTGTCTCTCC	LCI
GhCPK33-771-A	TACGAGATCTGGTCGAC CTATAACTTCCGAGAATTCCG	LCI
GhCPK74-771-S	CTCGGTACCCGGGATCC ATGGGAGCCTGTCTCTCC	LCI
GhCPK74-771-A	TACGAGATCTGGTCGAC CTATAACTTCCGAGAATTCCG	LCI
GhCPK33-BP-S	GGGGACAAGTTTGTACAAAAAAGCAGGCTCA ATGGGAGCCTGTCTCTCC	Subcellular Localization
GhCPK33-BP-A	GGGGACCACTTTGTACAAGAAAGCTGGGTA CTATAACTTCCGAGAATTCCC	Subcellular Localization
GhCPK74-BP-S	GGGGACAAGTTTGTACAAAAAAGCAGGCTCA ATGGGAGCCTGTCTCTCC	Subcellular Localization
GhCPK74-BP-A	GGGGACCACTTTGTACAAGAAAGCTGGGTA CTATAACTTCCGAGAATTCCC	Subcellular Localization
GhSAMS1-AD-S	CAGATTACGCTCATATG ATGGAGACCTTTCTATTCACATC	Y2H
GhSAMS1-AD-A	TGCTTGGGTGGAATTC TTAAGATTGGGGCTTGTCC	Y2H
GhSAMS2-AD-S	CAGATTACGCTCATATG ATGGAGACCTTTCTATTCACATCT	Y2H

GhSAMS2-AD-A	TGCTTGGGTGGAATTC TTAAGACTGAGGCTTCTCCCA	Y2H
GhSAMS1-106-S	AGGACGCCGGCGGATCC ATGGAGACCTTTCTATTCACATC	BiFC
GhSAMS1-106-A	AAGCTCTGCAGGTCGAC TTAAGATTGGGGCTTGTCC	BiFC
GhSAMS2-106-S	AGGACGCCGGCGGATCC ATGGAGACCTTTCTATTCACATCT	BiFC
GhSAMS2-106-A	AAGCTCTGCAGGTCGAC TTAAGACTGAGGCTTCTCCCA	BiFC
GhSAMS1-772-S	GGCGGTACCCGGGATCCA ATGGAGACCTTTCTATTCACATC	LCI
GhSAMS1-772-A	AAGCTCTGCAGGTCGAC TTAAGATTGGGGCTTGTCC	LCI
GhSAMS2-772-S	GGCGGTACCCGGGATCCA ATGGAGACCTTTCTATTCACATCT	LCI
GhSAMS2-772-A	AAGCTCTGCAGGTCGAC TTAAGACTGAGGCTTCTCCCA	LCI
GhSAMS1-QRT-S	CATCAAGCCTGTCATCCCT	qRT-PCR
GhSAMS1-QRT-A	CCAGCCACCATAAGTGTCAA	qRT-PCR
GhSAMS2-QRT-S	AAACATGCACCAAGACCAAC	qRT-PCR
GhSAMS2-QRT-A	TCAAGACCCACATCATCAGAA	qRT-PCR
GhSAMS1-VIGS-S	CAAAATGGCATGCCTGCAGACTAGT TATGCCACTGATGAAACCCC	VIGS
GhSAMS1-VIGS-A	GAATTCACTAGACCTAGGGGCGCGCC TTGATGACATGCTCCTTGAGG	VIGS
GhSAMS2-VIGS-S	CAAAATGGCATGCCTGCAGACTAGT GCCGCTCTATCGGATTTGT	VIGS
GhSAMS2-VIGS-A	GAATTCACTAGACCTAGGGGCGCGCC GTGCCATTCTTCCTAACCTCA	VIGS