# Identification of a new variant allele of *CaGLK* that is responsible for white color fruit and development of a functional marker based on this variation in pepper

Pingping Fang[1] (iD), Wenqi Chen[2], Ting Zhang[1], Jing Wei[1], Xueyan Wang[1], Huiqin Liu[1], Chaosen Li[1], Xiaomin Xiang[1], Dongfeng Zhao[1], Jia Liu[2,3], Qinwei Guo[1*] and Hongjian Wan[2*]

[1] *Quzhou Key Laboratory for Germplasm Innovation and Utilization of Crop, Institute of Vegetables, Quzhou Academy of Agricultural and Forestry Sciences, Quzhou 324000, China*
[2] *State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-Products, Institute of Vegetables, China-Australia Research Centre for Crop Improvement, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China*
[3] *Wulanchabu Academy of Agricultural and Forestry Sciences, Wulanchabu 012200, China*
* Corresponding authors, E-mail: kongxinzhu0530@163.com; wanhongjian@sina.com

## Abstract

Pepper (*Capsicum annuum* L.) fruits at immature stage (IS) display colors that range from dark green to light green and white. Due to the influence of genetic factors and consumption habits, white-fruited pepper germplasms are rare in nature and depend on artificial creation. However, the unclear complex genetic basis, coupled with the absence of developed functional markers cosegregation with white fruit trait, has impeded the efficiency of breeding white-fruited peppers. Herein, an $F_2$ segregation population consisting of 3,000 individuals is derived from a cross between B1-2 with white fruit and D50 with dark green fruit at IS. Using bulked segregant analysis, a previously reported gene *CaGLK* (*Capana10g000333*) is identified as a candidate that is responsible for white/green fruit color. A single nucleotide polymorphism (SNP) (G/A) at the position of the last base of the second intron of *CaGLK*, resulting in alternative splicing at the junction of the second intron and the third exon, causing the exclusion of 31 bp in cDNA sequence of B1-2, is associated with white fruit trait. A KASP marker is developed and shown to distinguish white-fruited pepper from the $F_2$ population. The distinguishing efficiency of this marker reached up to 78.3% when tested in a natural population. This first report of a new allele *CaGLK^A* that controlled white color provided a new handle to develop pepper cultivars with white color fruit.

## Introduction

Pepper (*Capsicum annuum* L.), a member of the Solanaceae family, is one of the most important crops in the word. Native to South America, specifically southern Brazil and Bolivia, pepper has now spread to many parts of the world, including China[1,2]. The pepper exhibits extreme diversity in fruit traits, including shape, size, color, and phytochemical compounds composition[3]. Among these traits, fruit color is influenced by the relative proportions of various pigments such as chlorophyll, carotenoids, anthocyanins, and flavonoids[3]. These pigments partially contribute to the nutritional, flavor values and health benefits, making fruit color as one of the most important traits. Since the fleshy fruit of peppers are primarily consumed at immature fruit stage, the color of fruits at this stage has become an important reference for consumers to accept the fruits. Typically, peppers at immature stage (IS) exhibit various shades of green, yellow green, white, purple, and nearly black[4]. Among these, white-fruited pepper at IS stage was popular in some regions of China, including Zhejiang, Jiangxi, Fujian, Hainan and Guangdong provinces. Particularly, the 'Qu Jiao' series, created by Quzhou Academy of Agricultural and Forestry Sciences, is renowned for its delightful combination of spiciness and fresh flavor[5,6]. Understanding the genetic basis underlying white fruit color attributes and developing new white-fruited pepper germplasms have become a necessity.

Reportedly, the major locus qualitatively differentiating between white color and dark green color is detected as *pc8.1* (now renamed *pc1*) and *pc10.1*, located on chromosome 8 and 10, respectively[7,8]. However, QTL *pc8.1* was revised to *pc1* because fine mapping placed this QTL on chromosome 1[8]. It is the zinc finger transcription factor *LSD ONE LIKE1* (*CcLOL1*), that serves as the target gene of QTL *pc1*[9]. Analysis of segregation, sequencing and expression conducted by Brand et al. revealed that *CaGLK2*, a *GOLDEN*-like transcription factor that controls chloroplast natural variation in chlorophyll content and pepper color at IS, is co-localized with *pc10*[8]. Null alleles at *CaGLK2* are related to reduced chlorophyll content in light green fruits. In addition, an *ARABIDOPSIS PSEUDO RESPONSE REGULATOR2* homologous gene in pepper (*CaAPRR2*) was closely associated with variations in phenotype of immature fruit color and was in the same position as *pc1* on the genetic map[10]. However, whether there are other QTLs or alleles of the three known genes, *CaGLK2*, *CaAPRR2*, and *CcLOL1*, controlling natural variation in immature fruit color remains unknown.

Functional markers (FMs) are DNA markers that can identify the functionality of alleles within near-isogenic lines, based on the identification of single nucleotide polymorphism loci[11]. Since these markers have a direct correlation with phenotypic traits and do not require the establishment of genetic maps, they have been used as ideal markers for identifying target trait alleles and marker-assisted breeding[12]. However, the absence of FMs associated with white color, together with the complex genetic basis, has impeded the efficiency of breeding white-fruited peppers.

In the present study, BSA-Seq was performed in an $F_2$ segregation population with 3,000 individuals, which was generated from a cross between two inbred lines with white fruit and dark green fruit. A new variant allele of *CaGLK*, responsible for white fruit color, was identified from candidate genes in the candidate genomic region. Additionally, a KASP marker with linked characteristics was developed, and its efficiency was tested in the $F_2$ population and a natural population. These findings will contribute to a better understanding of the molecular mechanisms that underpin white fruit color formation during IS and provide a new approach to enhance the efficiency of white-fruited pepper breeding.

## Materials and methods

### Plant materials and population construction

B1-2 (fruit with white color at the GI stage) and D50 (fruit with dark green color at IS, two pepper inbred lines was provided by the vegetable research group at Quzhou Academy of Agricultural and Forestry Sciences (Quzhou, China). Both of them are landraces from Zhejiang Province, China. $F_1$ population was derived from a cross between female parent B1-2 and male parent D50. Subsequently, the $F_2$ segregation population consisting of 3,000 individuals was obtained through self-crossing the $F_1$ population. Additionally, the natural population, comprising 113 germplasm accessions with white color or green color fruits, primarily collected from different provinces of China (Supplementary Table S1), was applied for further genotypic validation. All plants employed in this study were grown in plastic tunnels under natural sunlight in the experimental field of Quzhou Academy of Agricultural Science (29°10' N, 119°03' E).

### Identification of individuals with fruit phenotypes

The color of pepper fruit at the IS was observed and classified according to the classifications listed in Fig. 1. The classification was based on a color card comparison using the PANTONE color systems. Seedlings classified in Grade I were identified as the extremely white (EW) individuals, while seedlings classified to Grade VI were identified as the extremely green (EG) individuals.
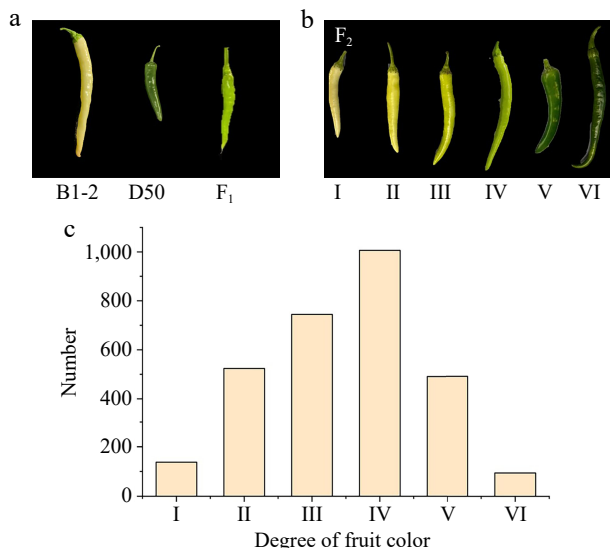


**Fig. 1** Color variation of immature fruits in the $F_2$ population. (a) The fruit performance of two parental lines and $F_1$ population. (b) The classification of fruit color in $F_2$ population. (c) Frequency distribution for fruit color showing phenotypic variation in $F_2$ population.

### DNA extraction and BSA-seq

DNA was extracted from individuals in the two selected EW (30 plants) and EG (30 plants) populations, using the CTAB method[13]. After validating the quality and concentration of the extracted DNA, an equal amount of DNA was taken to build the extreme phenotype mixing pools, i.e., the EW pool and EG pool. Using the CTAB method, DNA from parent materials B1-2 and D50 was also isolated. Four genomic DNA libraries (B1-2, D50, EW-pool, and EG-pool) were prepared for sequencing.

Sequencing was performed by BGI Genomics Co., Ltd. (Shenzhen, China) using the Illumina HiSeq platform, at a sequencing depth of 50×. The raw sequencing reads were filtered using the Soapnuke software[14]. The resulting valid reads were aligned to the reference genome *Capsicum annuum* L. Zunla (https://solgenomics.net/ftp/genomes/Capsicum_annuum/C.annuum_zunla/) using BWA software (Parameters: mem-t 4-M-R)[15]. All high throughput sequencing data have been deposited in the SRA database with accession number SRP (PRJNA1091429).

The alignment results were checked by SAMtools and the duplicate reads were removed. High quality SNPs were detected by filtering SNPs using the UnifiedGenotyper function with the GATK (version 3.4) software[16] according to the following requirements: 1) GQ (Quality Value) ≥ 50, otherwise mark the site as missing; 2) Non-biallelic SNP site; 3) Heterozygous sites in the parent pool; 4) Perform SNP filtering detection for SNP sites missing in the EG-pool and EW-pool. The SNP-index (frequency of SNP) was calculated using sliding window technology, and the SNP-index difference between EW-pool and EG-pool was represented by ΔSNP-index. The Euclidean Distance (ED) was calculated using the following formula:

$$ED = \sqrt{(A_{mut} - A_{wt})^2 + (C_{mut} - C_{wt})^2 + (G_{mut} - G_{wt})^2 + (T_{mut} - T_{wt})^2}$$

where, *mut* represent EW-pool, *wt* represent EG-pool, *A*, *T*, *C*, and *G* represent the proportion of each mutant SNP-index.

### *CaGLK* gene cloning and sequence analysis

Leaf tissues from parent lines B1-2 and D50 were collected for RNA and DNA extraction. The genomic DNA was extracted using the CTAB method, and total RNA was extracted according to the illustratment of the E.Z.N.A®Plant RNA Kit (Omega, New York State, USA). The first strand of cDNA was synthesized using a FastKing cDNA RT Kit (with gDNase) (TIANGEN, Beijing, China). PCR amplification was performed on the 30−50 ng/μL DNA and cDNA samples using following procedure: initial denaturation at 95 °C for 3 min; followed by 35 cycles of amplification (98 °C for 10 s, 55 °C for 10 s, 68 °C for 35 s by using DNA or 10 s by using cDNA), followed a final extension step of 72 °C for 10 min. The PCR products were visualized by 2% agarose gel electrophoresis using Goldview Nucleic Acid Gel Stain (Biohao, Wuhan, China). The specific amplification primers (Supplementary Table S2) were designed based on the gDNA and CDS sequences of the reported *CaGLK* gene (*Capana10g000333*) that were downloaded from the *Capsicum annuum* Zunla-1 cultivar genome, v2.0 https://plantgarden.jp/en/list/t4072/genome/t4072.G002). A p-EASY®-T5 Zero Cloning Kit (Transgen Biotech, Beijing, China) was used for gene cloning and transformation. After successful transformation into *Escherichia coli*., positive clones were selected and subjected to sequencing (Sangon Biotech, Shanghai, China). BioEdit and BioXM software were employed for sequence alignment and protein translation. Additionally, the structure of the *CaGLK* gene was determined using the Gene Structure Display Server (GSDS) 2.0 (http://gsds.gao-lab.org)[17].

### Development of Kompetitive allele-specific PCR (KASP)

Based on the polymorphic site between *CaGLK* alleles, a KASP marker was designed by using LGC proprietary Kraken™. The

allele-specific primers were designed carrying the standard FAM and HEX tails and with the targeted SNP at the 3' end[18]. A common primer was designed so that the total length of PCR products was less than 120 bp. The sequences of allele-specific primers are listed in Supplementary Table S3. KASP assay was carried out in 384-well formats and set up as 1.6 μL reaction mixture. To 0.8 μL 50 ng/μL DNA, 0.8 μL 2x KASP Araay Tape Master mix (LGC Genomics, Hodeston, UK) and 0.022 μL KASP Assay mix that consisted of two tailed primers and one common primer (LGC Genomics, Hodeston, UK), were added. No template control, referred to as NTC, was also established by replacing the DNA sample with ddH$_2$O in the reaction mixture. PCRs and fluorescent readings were performed on IntelliQube automated PCR instrument (LGC Bioserch Technologies, Teddington, UK) following the recommended protocol: hot start at 94 °C for 15 min, followed by 10 touchdown cycles (94 °C for 20 s; touchdown at 60 °C initially and decreasing by −0.6 °C per cycle for 60 s), followed by 26 cycles of amplification (94 °C for 20 s, 55 °C for 1 min). Scatter plots for genotypes was generated through Kluster-Caller software (LGC Genomics, Hoddeston, UK).

## Results

### Color variation of immature fruits in the F$_2$ population

B1-2 is an inbred-line that produces white fruit at the IS. To understand the genetic basis of white color behaviour in B1-2, B1-2 was hybridized with D50, which is a local germplasm resource whose immature fruits are dark green (Fig. 1a). The cross of B1-2 and D50 resulted in plants with green color (intermediate between parent's color) fruits (Fig. 1a). However, segregation of fruit color was observed in the F$_2$ population consisting of 3,000 individuals. The observed fruits color in the F$_2$ population could be classified into six degrees (I–VI) (Fig. 1b). Among these, 3,000 F$_2$ individuals, there were 138, 524, 745, 1,006, 491, and 96 plants belonging to degree I, II, III, IV, V, and VI respectively (Fig. 1c). These reflected that the gene contributing to green fruit color in immature peppers exerts a dominant effect over the gene responsible for white fruit color.

### BSA-seq based mapping and identification of SNPs

To identify white color related genes in B1-2, 30 EW, and 30 ED individuals were selected from the F$_2$ population to obtain DNA pools. Four DNA pools consisting of EW-pool, EG-pool, B1-2, and D50 were subsequently subjected to DNA sequencing (Supplementary Table S4). A total of 26.19, 26.21, 26.69, and 21.32 Gb raw data were obtained for the B1-2, D50, EG-pool, and ED-pool libraries. After filtering the raw reads, 861,497,710, 861,251,680, 871,859,126, and 832,814,516 clean reads for B1-2, D50, EG-pool, and ED-pool libraries, respectively, were aligned to the reference genome of 'Zunla-1' with a mean mapping proportion of 98.18%. Among these reads, 766,690,540, 771,839,948, 784,390,422, and 746,836,926 reads were properly paired reads with a mean properly paired ratio of 89.57%, resulting in 48.52 × average coverage depth. The sequencing depth in the extreme bulks was similar to that in the parents, reflecting the accuracy of the BSA-Seq. Moreover, the Q30 value (99.9% accuracy) exceeded 91.46%, and the GC content ranged from 35.20% to 35.42%. These results indicated that the sequencing data were of sufficient quality for further analysis.

SNP detection was subsequently performed in the EW-pool and the EG-pool by using GATK software, and a total of 5,826,852 high-quality SNPs unevenly distributed on 12 chromosomes were identified (Table 1). Among the 12 chromosomes, the highest number of SNPs occurred on chromosome 1 (1,063,601) and the lowest number occurred on chromosome 4 (249,986).

### Identification of white fruit color related QTLs

To identify the white color related QTLs, the correlation algorithms of ΔSNP-index and ED$^4$ were performed between the extreme pools and parents lines using high-quality SNPs. ΔSNP-index plots with 95% and 99% confidence intervals were analyzed using Sliding window (1 Mb interval and 500 kb increment) analysis (Fig. 2a), and ED$^4$ values with a 99% confidence intervals were calculated to indicated the ED association analysis (Fig. 2b). Giving the results of ΔSNP-index with a 99% confidence interval and ED$^4$ values with a 99% confidence interval, a final association interval was found on chromosome 10 with an overall size of 8.94 Mb, containing a total of 83 genes (Supplementary Table S5). Notably, a previously reported candidate gene for *pc10* named *GOLDEN2-LIKE* (*GLK*) that involved in chloroplast development, was one of 83 annotation genes. Therefore, it was speculated that it is *CaGLK* that (*Capana10g000333*) contributes to the white fruit color trait in B1-2.

### Validation of variation in candidate gene *CaGLK* between B1-2 and D50

To validate whether the candidate gene *CaGLK* contributes to the white-fruited phenotype in B1-2, the sequence differences of *CaGLK* genes cloned from B1-2 and D50 were compared. According to the sequencing results, the DNA length of the *CaGLK* was 6,066 bp and 6,058 bp in B1-2 and D 50, respectively. While the coding sequence length was 911 and 942 bp in B1-2 and D50, respectively encoded 173 AAs and 313 AAs. Subsequent multiple sequence alignment revealed that the cDNA sequence cloned from B1-2 had a 31 bp deletion at the third exon compared to the sequence from the reference genome and D50 (Fig. 3a). This resulted in the premature termination in the cDNA of *CaGLK* from B1-2. Then the gDNA sequences of *CaGLK* genes were compared and found that B1-2 carries the 'A' allele (*CaGLK$^A$*) at the last base of the second intron, while D50 carries the 'G' allele (*CaGLK$^G$*) at the same position. This SNP in B1-2 resulted in alternative splicing at the junction of the second intron and the third exon, causing the exclusion of 31 bp in the cDNA sequence (Fig. 3b).

### Development of a KASP marker to rapidly distinguish white-fruited pepper

Based on the single base mutation between alleles *CaGLK$^A$* and CaGLK$^G$, a KASP marker consisted of two competitive allele-specific primers and one common reverse primer was developed (Supplementary Table S3). In addition to the NTC, the expected graphical KASP results should contain another three clusters, corresponding to *CaGLK$^A$* genotype, *CaGLK$^G$* genotype and heterozygous genotype. As shown in Supplementary Fig S1, these three clusters are present. Notably, the cluster represents homozygous *CaGLK$^A$* genotype, corresponding to white color fruit phenotype. These results showed that this developed KASP marker could distinguish white-fruited peppers.

To further test whether *CaGLK$^A$* can be used for the identification of white fruit color, the KASP molecular marker was used to identify the genotypes of randomly selected 184 F$_2$ individual plants

**Table 1.** The distribution of high-quality SNPs on 12 chromosomes.

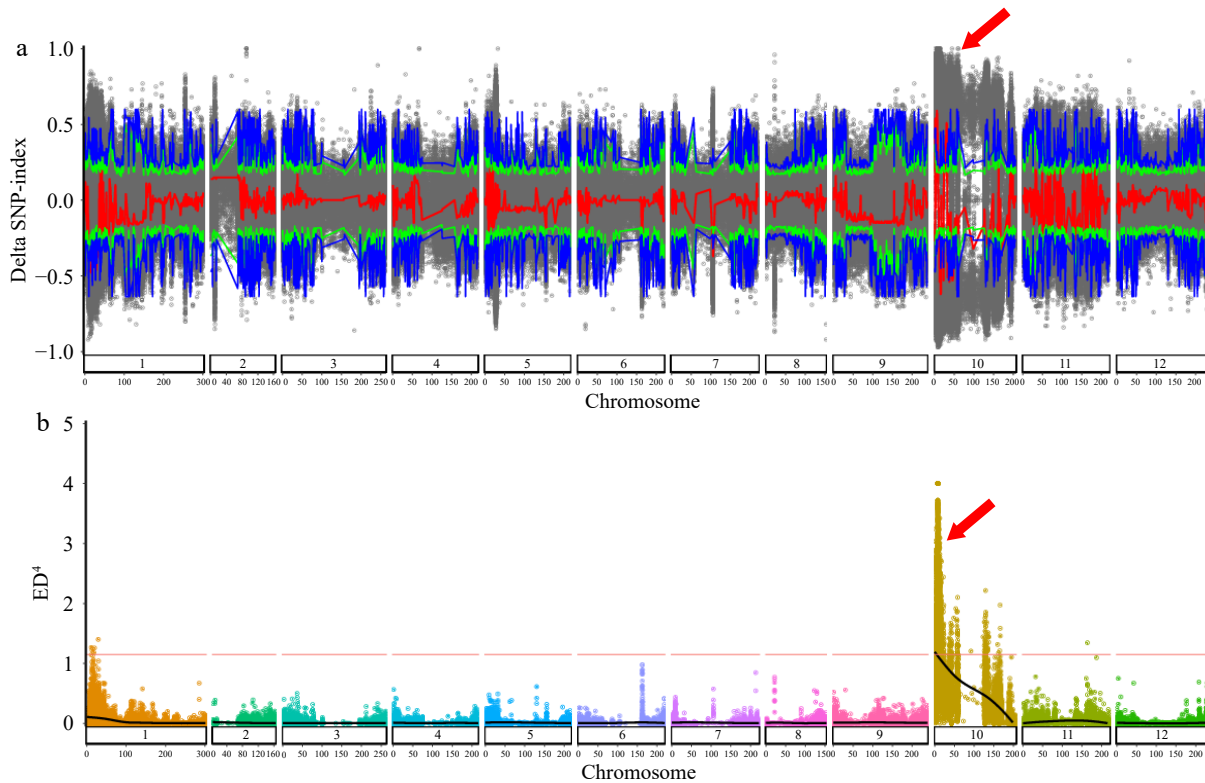| Chromosome name | SNP number | Chromosome name | SNP number |
|---|---|---|---|
| Chr01 | 1,063,601 | Chr07 | 308,508 |
| Chr02 | 308,087 | Chr08 | 270,086 |
| Chr03 | 472,929 | Chr09 | 955,711 |
| Chr04 | 249,986 | Chr10 | 447,555 |
| Chr05 | 279,766 | Chr11 | 631,473 |
| Chr06 | 417,558 | Chr12 | 414,592 |
| Total | | | 5,826,852 |

**Fig. 2** Distribution of ΔSNP-index and ED⁴ values on 12 chromosomes. (a) Distribution of ΔSNP-index on chromosomes of EW-pool and EG-pool. x-axis: the length of chromosomes; y-axis: the value of ΔSNP-index; blue and green lines: 95% and 99% confidence intervals, respectively; gray dot: ΔSNP-index; red: the results of sliding window analysis. (b) Distribution of ED⁴ values on 12 chromosomes. X-axis: the position of the chromosome; Y-axis: the value of ED⁴. Red lines: 99% confidence intervals.
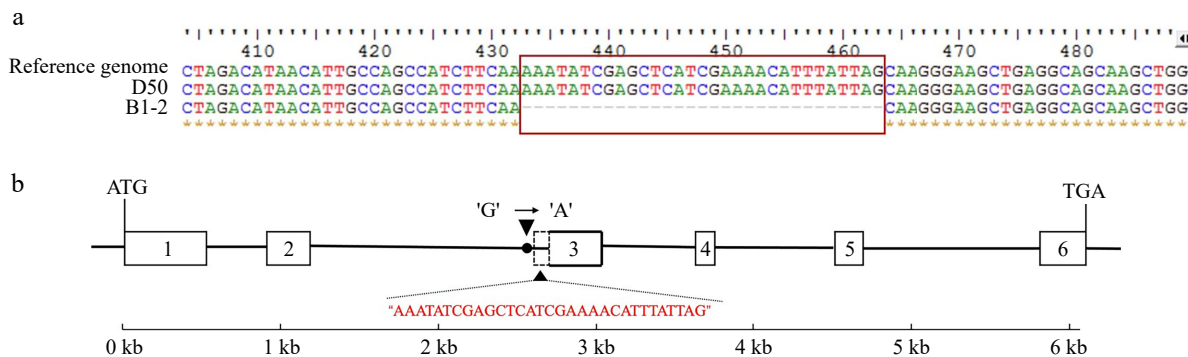


**Fig. 3** Variation of *CaGLK* gene (cDNA sequence and DNA sequence in B1-2). (a) Multiple sequence alignment analysis of *CaGLK* cDNA sequence among reference genome, D50 and B1-2. (b) Structure and variation sites of *CaGLK* gene in the DNA sequence. The grid represents the exon, the black line represents the intron, and the red base is the deletion part.

(Supplementary Table S6). The genotypes of 42 individual plants with white color fruit were completely consistent with B1-2 (Fig. 4, Supplementary Table S6), and genotypes of 142 individual plants with green color were consistent with D50 or F₁. The 142 individuals included 41 homozygous and 101 heterozygous plants (Fig. 4, Supplementary Table S6). These results strongly implicated that this developed KASP marker could successfully distinguish white-fruited peppers at a molecular level, which can be employed in white fruited pepper breeding programs.

## The applicability of the KASP marker for distinguishing fruit colors in pepper breeding programs

To further determine whether the KASP marker has a broad-spectrum capability of distinguishing fruit with white color in pepper accessions, a natural population, comprising 113 germplasm accessions with different fruit colors, primarily collected from different provinces of China, was randomly selected (Supplementary Table S1). In general, the distinguishing efficiency of *CaGLK*^A/G in pepper accessions with white color fruits was up to 78.3%, indicating that this KASP is a broad-spectrum marker to distinguish the white color fruits in pepper (Fig. 5 & Supplementary Table S1). This finding shows the potential of *CaGLK*^A as a key target in white color breeding programs for pepper.

## Discussion

It is well known that pepper fruit color is of primary significance, as the pigments that confer color are related to nutrition, health, and flavor. Compared to white- and green-colored pepper, those with red, orange, and yellow colors contain higher levels of
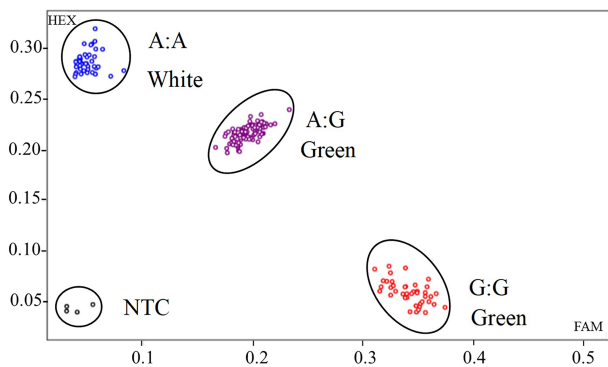
**Fig. 4** KASP genotyping of $GLK^{A/G}$ variant associated with white color fruit for the $F_2$ population. Randomly selected 184 $F_2$ individual plants were classified into three genotypes. Each dot represents a certain genotype. Blue dots represent $CaGLK^A$ genotype, which produce white fruit; red dots represent $CaGLK^G$ genotype, which produce green fruit; and purple dots represent heterozygous genotype. Black dots represent non template control (NTC).
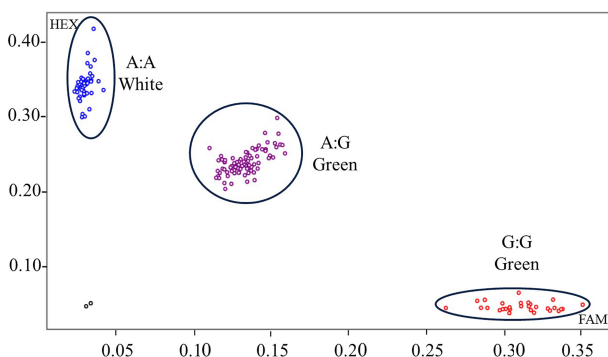


**Fig. 5** KASP genotyping of $GLK^{A/G}$ variant associated with white color fruit for pepper germplasm accessions. Each dot represents a certain genotype. Homozygous genotypes are marked with blue ($CaGLK^A$) and red ($CaGLK^G$) dots, respectively, and heterozygous genotypes are marked with purple dots. Black dots indicates no template control (NTC). A natural population, comprising 113 germplasm accessions were tested.

carotenoids and $\alpha$-tocopheron[19]. However, white pepper contains much more linalool, which was reported to give the fruit a citrus like, fruity, and floral aroma[20,21]. However, pepper germplasms that produce white color fruit at IS are definitely rare in nature. Therefore, development of functional markers that co-segregated with white fruit is crucial in elevating breeding efficiency of white-fruited pepper.

In the present study, a landrace that produces white-colored fruit was used as the female parent, and was hybridized with a dark-green fruited line D50, to generate large enough $F_2$ populations exhibiting fruit color segregation (Fig. 1). Based on the results of $\Delta$SNP-index and $ED^4$ values with a 99% confidence interval, the candidate genes that control green/white fruit color were mapped to chromosome 10, spanning an overall size of 8.94 Mb and containing a total of 83 genes (Fig. 2, Supplementary Table S5). Among these 83 genes, one gene annotated as GLK (*Capana10g000333*), was previously reported as a transcription factor that belong to GARP super family, modulating natural variation of chlorophyll content and fruit color in pepper fruits[8]. There are several lines of evidence showing that pepper *GLK2* corresponds to reported pc10 QTL that controls chloroplast development and chlorophyll content in pepper[8]. Null alleles at *CaGLK* contributed to light green fruits in PI 225 (6 bp insertion), PI 611(3 bp deletion) and three *C. annuum*

lines (A to G at position 489 of the ORF). The DNA sequence differences of *CaGLK* orthologous genes cloned from B1-2 and D50, together with the capability of the KASP marker developed based on this variation to distinguish white/green color, strongly indicated that *CaGLK* is the candidate gene controlling natural fruit variation in the present study (Figs 3, 4 & 5, Supplementary Tables S3 & S6).

Although no new gene was identified, a new null allele at *CaGLK* associated with reduced green color was found. This null allele was observed in B1-2 due to a single nucleotide change from G to A at the last base of the second intron, resulting in the exclusion of 31 bp in the cDNA sequence and truncated protein of 173 aa (Fig. 3). These results suggest that the newly identified SNP variant is a novel allelic variant of *GLK* transcription factors.

KASP markers were developed based on the competitive allele specific PCR genotyping system, with high efficacy to detect SNP[22–24]. Herein, a KASP marker consisted of two competitive allele-specific primers and one common reverse primer was developed based on the single base mutation between alleles $CaGLK^A$ and $CaGLK^G$. This marker can efficiently detect $CaGLK^A$, $CaGLK^G$, and heterozygous genotype (Supplementary Fig. S1). Additionally, the KASP marker is capable of distinguishing white-fruited pepper from the developed $F_2$ population and the natural population (Figs 4 & 5). The developed functional molecular marker has a great significance for creating new pepper germplasm, and identifying the pepper fruit color variations. It is also a potent tool to expedite the breeding of white fruited peppers.

## Conclusions

This study mapped the QTLs that contribute to natural variation in fruit color at immature stages in pepper, on chromosome 10 with an overall size of 8.94 Mb, contain a total of 83 genes. A previously reported *CaGLK* was suggested as a candidate. A $SNP^{G/A}$ at the position of the last base of the second intron of *CaGLK*, resulting in alternative splicing at the junction of the second intron and the third exon, causing the exclusion of 31 bp in cDNA sequence of B1-2, is associated with the white fruit trait. A KASP marker was developed and shown to distinguish white-fruited pepper contributed by $CaGLK^A$ allele. This work identified a new *CaGLK* allele that controlled white color and provides a new handle to develop pepper cultivars with white color fruit.

## Author contributions

The authors confirm contribution to the paper as follows: study conception and design: Guo Q, Wan H; data collection: Chen W, Zhang T, Wei Jing, Wang X, Zhao D, Xiang X; analysis and interpretation of results: Chen W, Fang P, Li C, Liu H; draft manuscript preparation: Fang P, Liu J; manuscript revision: Fang P, Wan H. All authors reviewed the results and approved the final version of the manuscript.

## Data availability

All high throughput sequencing data have been deposited in the SRA database with accession number SRP(PRJNA1091429). Other datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Acknowledgments

## Conflict of interest

The authors declare that they have no conflict of interest.

**Supplementary information** accompanies this paper at (https://www.maxapress.com/article/doi/10.48130/vegres-0024-0040)

## Dates

Received 3 April 2024; Revised 13 September 2024; Accepted 23 September 2024; Published online 20 January 2025

## References

1. Perry L, Dickau R, Zarrillo S, Holst I, Pearsall DM, et al. 2007. Starch fossils and the domestication and dispersal of chili peppers (*Capsicum* spp. L.) in the Americas. *Science* 315:986−88
2. Zou Z, Zou X. 2021. Geographical and ecological differences in pepper cultivation and consumption in China. *Frontiers in Nutrition* 8:718517
3. Wang L, Zhong Y, Liu J, Ma R, Mao Y et al. 2023. Pigment biosynthesis and molecular genetics of fruit color in pepper. *Plants* 12:2156
4. Lightbourn GJ, Griesbach RJ, Novotny JA, Clevidence BA, Rao DD, et al. 2008. Effects of anthocyanin and carotenoid combinations on foliage and immature fruit color of *Capsicum annuum* L. *The Journal of Heredity* 99:105−11
5. Guo Q, Zhang T, Li C, Xiang X, Wang X, et al. 2023. A new pepper F$_1$ hybrid — 'Qujiao No. 5'. *China Vegetables* 2023:106−08
6. Lui H, Zhang T, Li C, Xiang X, Xu J. 2012. A new pepper F$_1$ hybrid — 'Qujiao No.1'. *China Vegetables* 2012:102−03
7. Brand A, Borovsky Y, Meir S, Rogachev I, Aharoni A, et al. 2011. *pc8.1*, a major QTL for pigment content in pepper fruit, is associated with variation in plastid compartment size. *Planta* 235:579−88
8. Brand A, Borovsky Y, Hill TA, Rahman KA, Bellalou A, et al. 2014. *CaGLK2* regulates natural variation of chlorophyll content and fruit color in pepper fruit. *Theoretical and Applied Genetics* 127:2139−48
9. Borovsky Y, Monsonego N, Mohan V, Shabtai S, Kamara I, et al. 2019. The zinc-finger transcription factor *CcLOL1* controls chloroplast development and immature pepper fruit color in *Capsicum chinense* and its function is conserved in tomato. *The Plant Journal* 99:41−55
10. Pan Y, Bradley G, Pyke K, Ball G, Lu C, et al. 2013. Network inference analysis identifies an *APRR2-like* gene linked to pigment accumulation in tomato and pepper fruits. *Plant Physiology* 161:1476−85
11. Andersen JR, Lübberstedt T. 2003. Functional markers in plants. *Trends in Plant Science* 8:554−60
12. Varshney RK, Graner A, Sorrells ME. 2005. Genomics-assisted breeding for crop improvement. *Trends in Plant Science* 10:621−30
13. Porebski S, Bailey LG, Baum BR. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Molecular Biology Reporter* 15:8−15
14. Chen Y, Chen Y, Shi C, Huang Z, Li S, et al. 2018. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *GigaScience* 7:gix120
15. Qin C, Yu C, Shen Y, Fang X, Chen L, et al. 2014. Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. *Proceedings of the National Academy of Sciences of the United States of America* 111:5135−40
16. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* 20:1297−303
17. Hu B, Jin J, Guo A, Zhang H, Luo J, et al. 2015. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31:1296−97
18. He C, Holme J, Anthony J. 2014. SNP genotyping: the KASP assay. In *Crop Breeding*, eds Fleury D, Whitford R. New York, NY: Humana Press. Volume 1145. pp. 75−86. doi: 10.1007/978-1-4939-0446-4_7
19. Matsufuji H, Ishikawa K, Nunomura O, Chino M, Takeda M. 2007. Antioxidant content of different coloured sweet peppers, white, green, yellow, orange and red (*Capsicum annuum* L.). *International Journal of Food Science & Technology* 42:1482−88
20. Xiao ZB, Zhu JC, Feng T, Tian HX, Yu HY, et al. 2010. Comparison of volatile components in Chinese traditional pickled peppers using HS−SPME−GC−MS, GC−O and multivariate analysis. *Natural Product Research* 24:1939−53
21. Mavi K, Hacbekir H, Uzunoğlu F, Türkmen M. 2021. The use of volatile compounds as an alternative method in pepper breeding (*Capsicum baccatum* var. *pendulum*). *Ciência Rural* 51:e0201066
22. Ayalew H, Tsang PW, Chu C, Wang J, Liu S, et al. 2019. Comparison of TaqMan, KASP and rhAmp SNP genotyping platforms in hexaploid wheat. *PLoS One* 14:e0217222
23. Rahman MZ, Hasan MT, Rahman J. 2023. Kompetitive Allele-Specific PCR (KASP): an efficient high-throughput genotyping platform and its applications in crop variety development. In *Molecular Marker Techniques*, ed. Kumar N. Singapore: Springer. pp. 25–54. doi: 10.1007/978-981-99-1612-2_2
24. Wilkes JE, Fallen B, Saski C, Agudelo P. 2023. Development of SNP molecular markers associated with resistance to reniform nematode in soybean using KASP genotyping. *Euphytica* 219:27