Genome-wide Association Study Reveals that *PvGUX1_1* is Associated with Pod Stringlessness in Snap Bean (*Phaseolus vulgaris* L.)

Zhiyuan Liu†, Shuo Gao†, Helong Zhang, Zhaosheng Xu, Wei Qian*

Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing, China

† These authors contributed equally to this work.

*** Correspondence:**

Wei Qian, qianwei@caas.cn

Abstract

The suture strings is a particularly important pod trait that determines the quality and texture of snap bean (*Phaseolus vulgaris* L.). The *St* locus on chromosome 2 has been described as a major locus associated with suture strings. However, the gene and genetic basis underlying this locus remain unknown. Here, we investigated the suture strings of 138 snap bean accessions across two years. A total of 3.66 million single-nucleotide polymorphisms (SNPs) were obtained by deep resequencing. Based on these SNPs, we identified a strong association signal on Chr02 and a promising candidate gene, *PvGUX1_1*. Further analysis revealed that the 2-bp deletion in exon of *PvGUX1_1* was significantly associated with stringlessness. Comparative mapping indicated that *PvGUX1_1* was a domesticated locus and diverged from *PvGUX1_2* during an early stage. Our study provides important insights into the genetic mechanism of suture string formation and useful information for snap bean improvement.

Keywords : Snap bean, Genome-wide association study, Candidate gene, Pod stringlessness, Syntenic analysis

Introduction

Snap bean (*Phaseolus vulgaris* L.) is a type of common bean that is harvested before the seeds mature and eaten as a vegetable. Immature snap bean pods are succulent and rich in protein, carbohydrates, vitamin C, vitamin K, and carotenoids (Myers and Kmiecik, 2017). Therefore, the whole pods of snap bean are used for cooking, or preserved for freezing and canning (Hagerty et al.,, 2016). The snap bean is mainly consumed in North America, Europe, the Middle East, Africa, and Asia. In recent year, China has become the first producer of snap beans in the world (Zhang et al., 2008) .

Improving pod quality is a major objective for snap bean breeding. Some pod characteristics, including pod length, pod shape, spur length, and the absence or presence of suture strings, have made the snap bean more palatable than the dry bean (another type of common bean in which the mature seed is consumed). A snap bean with the straight, smooth pod, and lacks suture strings is preferred in the fresh market. The fiber string along the suture is usually discarded before being eaten. Thus, the absence of suture strings is more popular with consumers and facilitates the commercial processing of snap bean.

Reducing suture strings in snap bean is crucial, and the key to this effort lies in understanding the genetic basis of the formation and development of suture string. The inheritance analysis of suture strings revealed that stringlessness was governed by a dominant gene, *St*, in common bean (Prakken, 1934). Quantitative trait locus (QTL)

analysis located the *St* gene on chromosome Pv02 in common bean (Koinange et al., 1996). Working with a recombinant inbred line derived from dry bean and snap bean, a strong QTL, PST2.2, was also found on Pv02, accounting for 32% of total genetic variation in a recombinant inbred line (Hagerty et al., 2016). As the reduction of suture strings and pod wall fibers commonly lead to pod indehiscence in common bean, the indehiscent gene *PvIND* (a homolog of the Arabidopsis INDEHISCENT gene, *IND*) mapped near the St locus was predicted to be the *St* gene. However, there was incomplete co-segregation between *PvIND* and the St locus and a lack of polymorphisms with dehiscent/indehiscent phenotypes, suggesting that *PvIND* was not the gene *St* (Gioia et al., 2012). Recently, a single QTL, qPD5.1-PV, determining pod indehiscence was identified on chromosome Pv05 (Rau et al., 2019). In the attempt to identify the candidate gene underlying the QTL, a BC4/F4 introgression line population was used to narrow down the QTL in a 22.5 kb region and identified *PvMYB26* was the best candidate gene based on mapping and gene expression pattern (Di Vittori et al., 2020). In addition, several genes or QTLs were also discovered to be associated with pod dehiscence, such as *PvPdh1* on chromosome Pv03, QTLs on Pv08, Pv05 and Pv09 (Parker et al., 2020).

The first common bean reference genome was published in 2014 (Schmutz et al., 2014), which made it possible to use different strategies to identify candidate genes and molecular markers for important agronomic traits. Genome-wide association study (GWAS) is an approach based on using the numbers of single-nucleotide polymorphisms (SNPs) to test the association of desired traits. Due to the reduced cost of resequencing, and the repeatability of SNPs in the genome, GWAS has been performed using various landraces and breeding lines in common bean. These studies have focused on grain yield (Kamfwa et al., 2015; Moghaddam et al., 2016; Wu et al., 2020), flowering time (Raggi et al., 2019), resistance to disease (Wu et al., 2017), resistance to pod shattering (Parker et al., 2020), grain mineral content (Delfini et al., 2021), drought resistance (Wu et al., 2021), and abiotic stress (Soltani et al., 2018). However, few studies have focused on specific traits in snap bean (Myers et al., 2019). Pod stringlessness is particularly crucial in snap bean. Therefore, the objective of this study was to identify the candidate gene associated with this trait as a basis for further improving the quality of snap bean.

Materials and Methods

Plant material and resequencing

One hundred and thirty-eight snap bean accessions collected from the Institute of Vegetables and Flowers at the Chinese Academy of Agriculture Sciences (CAAS),

including landraces, elite lines, and breeding lines, were grown between March and June in 2019 and 2020 (Supplementary Table 1). These seeds were grown in mixed nutrient soil at a greenhouse in Beijing(40° N, 116° E). The plants were watered using automatic drip irrigation every 2-3 days throughout entire growth period. The field away from plant was covered with a mulching plastic sheet to reduce the weed.

Young leaves at the unifoliate growth stage from each accession were collected, flash-frozen in liquid nitrogen and stroed in an ultra-low-temperatue freezer. Genomic DNAs were isolated for each genotype using Plant Genomic DNA kit (Tiangen, Beijing) following to the instructions. The integrity of gDNA was determined on 1% agarose gels. The concentration and quality of gDNA were measured using a NanoDrop2000 Spectrophotometer (Thermo Fisher Scientific). The DNA library were constructed accroding to the manufacturer's instructions for the TruSeq nano DNA kit (Illumina). The libraries were genotyped using an Illumina HiSeq 2000 (125PE) sequencer at the facilities of Berry Genomics Co. Ltd, Beijing, China, as described by Wu et al. (2020).

Measurement of pod sutures

At the green mature stage, 10 fresh pods from different plants of each accession were randomly chosen to measure the pod suture strings. The 10 pods from 10 individuals served as technical replications. The strings were evaluated as both a qualitative trait and a quantitative trait. As a qualitative trait, the pod strings were scaled 0–1 (0 = no suture strings, 1 = presence of suture strings). As a quantitative trait, the ratio (string length/pod length) of each pod was measured. The average ratio value of 10 pods and the scale rating of pods were both used for GWAS analysis.

Expression analysis of *PvGUX1_1*

Three stages (T1 for pod elongation, T2 for pod development, and T3 for pod maturity) of R02 (non-suture pod) and R05 (suture pod) were sampled. The total RNA of the three stages of pods in suture pods and non-suture pods was extracted and converted to cDNA using a Reverse Transcription Kit (TransGen Biotech, Beijing, China) according to the manufacturer's instructions. Quantitative real-time PCR was performed with SYBR Green (Vazyme Biotech, Nanjing, China), and the data collection was performed on QuantStudioTM 6 Flex system(ABI, Life, USA) according to the manufacturer's instructions. The primers were synthesized by Sangon Biotech(Shanghai, China). The relative expression levels of *PvGUX1_1* were compared with that of β-actin, and the expression fold changes were calculated using the $2^{-\triangle\triangle Ct}$ method. Each qRT-PCR reaction was performed in triplicate. Sequences of the primers used for qRT-PCR in this study are shown in **Supplementary Table 2**.

Variant calling and annotation

The raw paired-end reads were initially filtered by fastp (v0.20.0) software with the following parameters: -q 30. Next, the clean reads were aligned with the common bean reference genome V2.1 (Schmutz et al.,2014) using MEM algorithm of BWA (v0.7.17-r1188) (Li et al., 2009).

The tools SortSam and MarkDuplicates in PICARD (v1.127) were used to sort mapping results and mark the duplicate reads (https://broadinstitute.github.io/picard/). In addition, realignment around InDels was conducted by RealignerTargetCreator and IndelRealigner in GATK (v3.2) (McKenna et al., 2010). The variant was called by the UnifiedGenotyper module in GATK and SAMTOOLS (v1.6-3-g200708f) (Li et al., 2009). The two variant results were combined and further filtered to obtain a credible variant dataset using the GATK subcomponents SelectVariants and VariantFiltration. The credible variant dataset was employed to recalibrate and realign results using the BaseRecalibrator and PrintReads of GATK. The SNP and InDel were again called against the recalibrated results. Finally, a vcf file including all samples and variants was generated and further filtered using vcftools software (0.1.15) with the following parameters: -max-missing 0.95 -maf 0.05 -min-alleles 2 -max-alleles 2 -recode -recode-INFO-all.

The functional annotation of variants was performed using the software ANNOVAR (Version:2017-07-17) (Wang et al., 2010).

Population genetics analyses

To analyze the population structure, the reduced SNPs were employed based on the value of the correlation coefficient (r^2), where SNPs with strong linkage disequilibrium (LD) ($r^2 > 0.2$) within a 50-kb window were discarded using plink (v1.90b6) software with the following parameters: -indep-pairwise 50 10 0.2. To estimate the most optimal sub-population, a cross-validation procedure was conducted with ADMIXTURE (v1.3.0) runs from K= 2 to 16 (Alexander et al., 2011). A neighbor-joining tree of 138 snap bean accessions was constructed using Phylip 3.68 (Felsenstein et al., 1989) software based on a distance matrix. The bar plots of sub-populations and the phylogenetic tree were plotted using the itol website (https://itol.embl.de/).

Linkage disequilibrium analysis

The correlation coefficient (r^2) of pairwise SNPs within a 1000-kb window from all chromosomes were used to estimate LD decay, which was calculated and plotted using PopLDdecay software (Zhang et al., 2019). LDBlockShow software was used to

calculate and display LD blocks in candidate regions (https://github.com/BGI-shenzhen/LDBlockShow).

Genome-wide association analysis

The high-quality SNPs were used for GWAS analysis in the R package GAPIT (Tang et al., 2016). To reduce false positives and improve statistical power, the 'Q+K' approach was employed. The kinship matrix (K) was calculated with the default method in GAPIT. The significant threshold ($-\log_{10}P$) was Bonferroni-corrected as $-\log_{10}P$=7.86. The Manhattan plot was run by the CMplot package in R 3.6.1 (https://github.com/YinLiLin/CMplot).

Analyses of collinearity and synteny

The genome sequence information of common bean (*Phaseolus vulgaris* V2.1) and cowpea (*Vigna angularis* V1.2) were downloaded from phytozome 13. The genomes of soybean (*Glycine max* 109) (www.plantgdb.org/XGDB/phplib/download.php?GDB=Gm) and pea (*Pisum sativum*) (https://urgi.versailles.inra.fr/Species/Pisum) were downloaded from public databases. The analysis of collinearity and synteny between the four legumes was implemented with MCScan (Python version) (https://github.com/tanghaibao/jcvi/wiki/MCscan-[Python-version]). The proteins with similarity with over 90% on PvGUX1_1 in common bean, soybean, cowpea, and pea were identified using BLASTP with an e value $<1e^{-5}$. The neighbor-joining tree from the orthologue gene of *PvGUX1_1* was constructed using MEGA X (Kumar et al., 2018) with default parameters.

Results

Pod suture string phenotyping

The pod suture strings of 138 snap bean accessions were investigated based on rating and ratio (Figure 1). For rating, the presence of strings was defined as 1; the absence of strings was defined as 0. The rating were investigated in 2019 (**Figure 1B**) and 2020 (**Figure 1D**). A total of 60 accessions were stringless, whereas 78 accessions had suture strings in 2019 (ST2-2019) (**Figure 1B**). However, five accessions showed different ratings in 2020 (ST2-2020). For ratio, the average ratio values (string length/pod length) of 10 pods in each accession were measured in 2019 (ST1-19) and 2020 (ST1-20) (**Figures 1A,C**). The analysis of correlation for ratio showed that there was a significantly high correlation of 0.93 (P = 0.00015) between 2019 and 2020.

Resequencing of snap bean accessions

The whole-genome resequencing of 138 accessions produced a total of 3.08 billion raw paired-end reads and 0.92 Tb bases, which was approximately 11.4-fold sequence depth, ranging from 10.2- to 13.5-fold. After being filtered, 2.71 billion clean reads were retained. Mapping against the common bean genome V2.1 resulted in 5,130,030 SNPs and 1,524,528 InDels. Further filtering (bi-allelic, missing data < 0.05, minor allele frequency >0.05) identified 3,656,683 high-confidence SNPs and 626,853 InDels. Among these variants, 3,589,978 SNPs and 618,666 InDels were placed on chromosomes; 66,705 SNPs and 8187 InDels were on scaffolds. The distribution of these SNPs across the genome was uneven (**Figure 2**). Most SNPs were located in Chr02 (411,294), and the fewest SNPs were found in Chr06 (238,452). In addition, the frequency of SNPs in Chr02 (8.28 SNPs/kb) was the highest, while the frequency of SNPs in Chr08 (5.97 SNPs/kb) was the lowest **(Supplementary Table 3**).

To investigate distribution regions of these variants across the genome, we carried out variant annotation and found that 146,882, 512,153, 279,102, and 244,805 SNPs and 4180, 53,742, 30,812, and 28,930 InDels were located in exons, introns, upstream, and downstream, respectively. Furthermore, of these SNPs in exons, 65,001 nonsynonymous, 718 stopgain, and 171 stoploss InDels were annotated, which resulted in amino acid changes, premature stop codons, or longer transcripts. Similarly, of these InDels in exons, 697 frameshift insertion, 1091 frameshift deletion, three stoploss, and 49 stopgain InDels were annotated, which also influenced protein sequences.

Population structure and LD analysis

The analysis of population structure allows researchers to understand the genetic relationships and origins of species. After removing the SNPs with strong LD ($r^2 \geq 0.2$), 97,841 SNPs were generated and used to implement population structure analysis with Admixture. The use of K =2 divided the 138 genotypes into two genetic groups, which was in agreement with two domesticated genepools (Andean and Middle American) (**Figure 3**). Among the 138 genotypes, 40 genotypes had predominantly Andean ancestry, and 30 genotypes contained a level of hybridization, suggesting that a high degree of intercrossing between the genepools that has happened within snap beans.

We further analyzed the LD decay across the genome **(Supplementary Figure 1**). The average LD decay of the whole genome was 631.4 kb (r^2 dropped to half of its maximum value), which was faster than that of common bean (107 kb) (Wu et al., 2020), cultivated soybean (150 kb) (Zhou et al., 2015), and cultivated rice (123 kb for indica and 167 kb for japonica) (Huang et al., 2010). In addition, we found that the

rate of LD decay in different chromosomes varied from 184 kb (Chr10) to 976 kb (Chr01) (**Supplementary Table 4**).

Genome-wide association study for pod stringlessness

To find out genetic loci controlling pod stringlessness, we implemented GWAS for four traits (ST1-2019, ST2-2019, ST1-2020, and ST2-2020) using 3,656,683 SNPs (**Figure 4**). The Q2 and relatedness kinship matrix as covariates were taken into account to reduce false positives in GWAS analysis with a compressed mixed linear model. The $-\log10(P)=7.86$ was set as a genome-wide significance threshold based on Bonferroni correction. Strong association signals were used to identify candidate regions and screen candidate genes.

A total of 205 loci were identified with $-\log10(P)>7.86$ for ST1-2019. Of 205 SNPs, 204 were located at Chr02, and one was located at Chr11 (**Supplementary Table 5**). The peak signal was located at Chr02:44026689 ($-\log10(P)=10.08$), accounting for 14.53% of phenotypic variation. The major locus Chr02:44248269 ($-\log10(P)=8.60$) was significantly associated with ST2-2019. Furthermore, strong signals were both found at Chr02:44194640 for ST1-2020 ($-\log10(P)=8.49$) and ST2-2020 ($-\log10(P)=9.62$). Taken together, the peak SNPs for four traits were all located in adjacent physical regions in chromosome 2, which suggested the pod stringlessness was under the control of a major locus.

Identification of candidate genes for pod stringlessness

To identify the candidate regions associated with the significant SNPs, we carried out haplotype analysis in the whole genome. We found that these peak SNPs for the four traits all resided on the same linkage disequilibrium (LD) block (Chr02:43998258–44264446) (**Figure 5**). These genes located in the block were likely responsible for the formation of stringlessness. **Table 1** shows these genes and their homologous genes in Arabidopsis. A total of 23 putative genes were annotated in this block based on the common bean reference genome V2.1. Eighteen out of 23 genes were functionally annotated, and 15 genes had homologous genes in Arabidopsis. Furthermore, 102 SNPs, including 43 nonsynonymous and 59 synonymous SNPs, and 6 InDels, including two frameshift deletions distributed in the coding areas of these genes, were also identified (**Table 2**).

A 2-bp deletion in the exon region was identified in *Phvul.002G270800*, an ortholog to *AtGUX1*, which is responsible for secondary wall deposition in Arabidopsis. The 2-bp deletion introduced a premature stop codon that truncated the protein to 64 amino acids. To verify the deletion, we cloned the gene from two suture and non-suture accessions (**Supplementary Figure 2**). The result was similar to the finding in resequencing. Additionally, the deletion of 2 bp was significantly correlated

with pod stringlessness (2.2×10^{-16}) (**Figure 6**). We identified another a 2-bp deletion in gene *Phvul.002G271600*; however, the function of this gene was unclear.

The most abundant nonsynonymous SNPs were found in *Phvul.002G271700. Phvul.002G271700,* encoding a NAC domain protein, carried eight nonsynonymous SNPs. Among these SNPs, K120I was significantly associated with pod stringlessness (1.39×10^{-8}), whereas other SNPs exhibited weak association.

Three nonsynonymous SNPs, T32C, C604T and C737T, were identified in *PvIND* (*Phvul.002G271000*). The SNPs T32C and C604 showed weak association ($P = 6.42\times10^{-8}$ and 6.79×10^{-6}) with pod stringlessness, while C737T showed no association ($P = 0.24$).

Syntenic analysis of the candidate region between the common bean and other legumes

To identify the function and relation of the candidate gene, we performed syntenic analysis within the candidate region of common bean with three legumes, including soybean (*Glycine max*), cowpea (*Vigna angularis*), and pea (*Pisum sativum*). Common bean, cowpea, and soybean are members of the Phaseoleae tribe, whereas pea belongs to the Fabeae tribe (Dong et al., 2014). Amongst these legumes, the majority of cowpea are stringless, while common bean and pea have stringless and string types. In the Phaseoleae tribe, common bean and cowpea are the two most closely related crop species of the four legumes analyzed. They also exhibited a high degree of synteny (**Figure 7A**), in which 19 of 23 genes were orthologous. Although large-scale synteny with soybean was also observed, the homologous genes in soybean were divided into two regions (Glyma08g15530–Glyma08g15650 and Glyma08g16570–Glyma08g16640) on chromosome 8. In contrast, the pea chromosome exhibited a large rearrangement with common bean.

Overall, these candidate genes and gene order in common bean were highly conserved and exhibited extensive synteny with cowpea. However, none of orthologs for Phvul.002G270800 in the syntenic block were identified (**Figure 7A**). To identify the orthologous gene of Phvul.002G270800 (PvGUX1_1), the protein of Phvul.002G270800 was used to conduct BLASTP search against cowpea, soybean, pea, and common bean. Specifically, we identified another common bean protein, Phvul.009G148800 (PvGUX1_2), which shared a strong sequence homolog to PvGUX1_1. PvGUX1_2 encoded 636 amino acids, whereas PvGUX1_1 encoded 221 amino acids. The best hit of PvGUX1_1 in cowpea, Vigun03g064600, encoded 629 amino acids, which exhibited large sequence difference with PvGUX1_1. To verify the relationship between PvGUX1_1, PvGUX1_2 and GUX1, we performed a phylogenetic analysis of PvGUX1_1, PvGUX1_2, and other orthologous genes.

PvGUX1_1 and PvGUX1_2 were placed in two different sub-branches (**Figure 7B**). Although the corresponding genes, Glyma08g15640 and Vigun03g064600, in the synteny region were clustered with PvGUX1_1 in one sub-branch, there was large sequence variation between PvGUX1_1 and other orthologs. Collectively, these data demonstrated that PvGUX1_1 and PvGUX1_2 diverged at an early stage in legume evolution, which may have resulted in gene diversification.

Gene expression patterns of PvGUX1_1

The formation of pod sutures is an important agronomic trait. To better reveal the genetic regulation of pod sutures, we conducted qRT-PCR analysis of PvGUX1_1 at the initiation of pod elongation (T1, no suture), pod development (T2, no suture), and pod maturation (T3, sutures were present in sutured pods) for sutured (R05)and non-sutured pods (R02) (**Figure 7D**). The qRT-PCR results indicated that the expression levels of PvGUX1_1 were significantly higher at the T1 and T2 stages in non-sutured pods compared with the sutured pods (**Figure 7E**). Furthermore, the expression level of PvGUX1_1 decreased following the development of pods in non-sutured pods (**Figure 7C**).

Discussion

Understanding the genetic mechanism of suture string development will facilitate the study of domestication and plant breeding in snap bean. Here, we identified a strong signal on Chr02 that determined the formation of pod stringlessness based on large-scale resequencing. Within these putative genes in candidate regions, *PvGUX1_1* was the best candidate gene due to its function and polymorphisms, which was consistent with dominant inheritance.

GWAS analysis for pod stringlessness

As common bean is a selfing species, effective recombination and heterozygosis in common bean are significantly reduced, which results in the generation of large LD and slow LD decay. Generally, LD decay is slower in selfing species than in outcrossing species because of the loss of recombination, which potentially leads to be homozygosity (Morrell et al., 2012). The nature of homozygosity makes common bean access to design GWAS. In particular, once a genotype is sequenced, the phenotype can be investigated in different environments. Moreover, the extensive genetic diversity is advantageous for GWAS analysis in common bean (Blair et al., 2009).

Pod stringlessness was controlled by a major locus

The inheritance of pod stringlessness is complex due to genotype and environmental factors (Ma et al., 2016). Since the stringless trait was observed, various inheritance models for stringless trait in common bean have been proposed. Currence (1930) assumed that two genes (S for dominant, T epistatic to S) regulated the stringless trait. However, more studies revealed that the stringless trait was under the control of a single dominant locus (*St*), which was mapped on chromosome 2 (Koinange et al., 1996; Davis et al., 2006; Gioia et al., 2012). Moreover, there have also been reports that the trait did not fit the ratio of one or more loci, and thus was a quantitative trait (Hagerty et al., 2016). In order to verify the inheritance pattern, qualitative traits and quantitative traits were both used for GWAS analysis. Interestingly, we obtained similar results from the two models. The only strong signal from both models was identified on Chr02, which was in agreement with previous findings, and showed that the formation of suture string was controlled by a major locus.

The formation of suture string is controlled by a single locus, while the level (short versus long) of the string might be more complex. This characteristic was also observed in pod shattering. As suggested by Rau et al., (2019), at least two additional loci were likely relevant to the level and mode of pod shattering. In our study, in addition to Chr02, a SNP located at Chr11 was also associated with stringlessness (**Supplementary Figure 3**). The SNP was about 0.13 Mb from the NAC transcription factor gene *PvCUC2* (*Phvul.011G160400*, Chr11:45614432_45616861). In order to identify more locus, we conducted GWAS only on stringy snap bean for ST1-2019. Strong association signals were identified on Chromosome 7 (Supplementary Figure 4). These locus may be responsible for the level of suture string, along with the St gene. This finding supported the hypothesis by Drijfhout (1978) that a major factor influenced the string formation trait, while other genes led to incomplete strings.

Candidate gene for stringlessness in snap bean

A total of 23 genes within the LD block surrounding the high association signal were identified. Among them, several genes are orthologous genes involved in cell-wall biosynthesis, pod shattering, and fiber formation. *Phvul.002G270800* is the orthologous gene of *AtGUX1* (*AT3G18660*). In Arabidopsis, AttGUX1 belongs to Glycosyltransferase Family 8, which participates in the synthesis of plant cell walls (Yin et al., 2010). *AtGUX1* is responsible for the decoration of xylan, an important component of the secondary cell wall (Lee et al., 2012). Silencing *AtGUX1* led to the decrease of glucuronoxylan content and microsomal xylan in the cell wall (Oikawa et al., 2010). In our study, a 2-bp deletion was found in the exon region of *Phvul.002G270800,* causing a premature stop. The 2-bp deletion was significantly

associated with pod stringlessness. Therefore, we propose *Phvul.002G270800* as the best candidate gene for St locus.

In addition to *Phvul.002G270800,* another gene of interest was *Phvul.002G271000,* the orthologous gene to *AtIND. AtIND,* as a b-HLH transcription factor, plays a crucial role in pod shattering in Arabidopsis (Girin et al., 2010; Kay et al., 2013; Dong and Wang, 2015). However, due to the lack of mutation in *PvIND* (*Phvul.002G270800*), a previous study reported that it was not the *St* gene controlling suture strings (Gioia et al., 2012). Although the present study identified three nonsynonymous SNPs in the exon region of *PvIND*, these SNPs only showed a weak association with the suture strings. Therefore, *PvIND* may not be directly involved in suture development. Other interesting genes included NAC transcription factor *Phvul.002G271700* (*PvVNI1*) and MYB transcription factor *Phvul.002G269900* (*PvMAMYB*). Many studies have suggested that an NAC transcription factor is correlated with pod shattering and secondary cell wall development (Hussey et al., 2011; Yamaguchi et al., 2010; Reusche et al., 2012). In particular, the role of the NAC transcription factor SHA1-5 in regulating pod shattering in soybean has been elucidated in detail (Dong et al., 2014). Likewise, several MYB transcription factors, such as *MYB26* (Wilson et al., 2011), *MYB46* (Kim et al., 2013,2014), and *MYB63* (Zhou et al., 2009), are involved in lignin biosynthesis and secondary cell wall formation in many species (Nakano et al., 2015). Therefore, the functions of *PvVNI1* and *PvMAMYB* need to be further studied in future research to determine whether they are related to suture string development or pod shattering.

Pod stringlessness in other legumes

The loss or presence of suture strings is not an important factor for many legumes in which the dry seeds are consumed. In legumes, reports on the stringless trait are currently only found in common bean and pea. In pea, pod stringlessness arose from spontaneous mutation (Wellensiek, 1971). The recessive gene (*sin-2*) is regarded as the key gene responsible for the stringless trait in pea (McGee and Baggett, 1992; Ma et al., 2016). In contrast, the stringless trait in common bean is governed by the dominant gene *St.* In the synteny block, the orthologs of *GUX1* were not detected in pea, indicating that the genetic mechanism of stringlessness between the two legumes may be different.

Although the same regulation gene may not be shared in common bean and pea, there are many parallels, including seed dormancy, growth habit, and earliness, between common bean and pea that have occurred in the process of crop domestication (Weeden, 2007). The identification of the *St* gene in common bean

would accelerate the mining of *sin-2* and improve the understanding of the genetics of domestication under parallel selection in the future.

Data Availability Statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author Contributions

QW designed the study. LZ and GS conducted the experiments. LZ, ZH, and XZ analyzed the data. LZ wrote the manuscript. QW, XZ, ZH, and GS revised the manuscript.

Funding

This work was performed at the Key Laboratory of Agriculture, Beijing, China, and was supported by the Chinese Academy of Agricultural Sciences Innovation Project (CAAS-ASTIP-IVFCAAS), Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture, Beijing, China.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary Material

The Supplementary Material for this article can be found online.

References

Alexander, D. H., and Lange, K. (2011) Enhancements to the admixture algorithm for individual ancestry estimation. *BMC Bioinformatics.* 12(1): 1-6. doi:10.1186/1471-2105-12-246

Blair, M. W., Astudillo, C., Grusak, M. A., Graham, R., and Beebe, S. E. (2009) Inheritance of seed iron and zinc concentrations in common bean (Phaseolus vulgaris L.). *Molecular Breeding* 23(2):197-207. doi:10.1007/s11032-008-9225-z.

Currence, T. M. (1930) Inheritance studies in *Phaseolus vulgaris*. *Tech Bull Minn Agric Exp Stn* 68:1-28

Davis, J. W., Kean, D., Yorgey, B., Fourie, D., Miklas, P. N., and Myers, J. R. (2006) A molecular marker linkage map of snap bean (Phaseolus vulgaris). *ANNUAL REPORT-BEAN IMPROVEMENT COOPERATIVE* 49:73

474 Delfini, J., Moda-Cirino, V., dos Santos Neto, J., Zeffa, D. M., Nogueira, A. F., and
475 Ribeiro, L. A. B., et al. (2021) Genome-wide association study for grain mineral
476 content in a Brazilian common bean diversity panel. *Theor Appl Genet*
477 134:2795-2811. doi:10.1007/s00122-021-03859-2
478 Di Vittori, V., Bitocchi, E., Rodriguez, M., Alseekh, S., Bellucci, E., and Nanni, L., et
479 al. (2020) Pod indehiscence in common bean is associated to the fine regulation
480 of PvMYB26 and a non-functional abscission layer. *Journal of experimental*
481 *botany* 72(5):1617-1633. doi:10.1093/jxb/eraa553
482 Dong, Y., and Wang, Y. Z. (2015) Seed shattering: from models to crops. *Frontiers in*
483 *Plant Science* 6: 476. doi: 10.3389/fpls.2015.00476
484 Dong, Y., Yang, X., Liu, J., Wang, B-H., Liu, B-L., and Wang, Y-Z. (2014) Pod
485 shattering resistance associated with domestication is mediated by a NAC gene
486 in soybean. *Nat. Commun* 5:3352. doi:10.1038/ncomms4352
487 Drijfhout, E. (1978) Inheritance of temperature-dependent string formation in
488 common bean (*Phaseolus vulgaris* L.). *Netherlands Journal of Agricultural*
489 *Science* 26(1):99-105. doi:10.18174/njas.v26i1.17114
490 Felsenstein, J. (1989) PHYLIP–Phylogeny Inference Package (Version 3.2).
491 *Cladistics* 5:164–166. https://www.jstor.org/stable/2830216
492 Gioia, T., Logozzo, G., Kami, J., Zeuli, P. S., and Gepts, P. (2012) Identification and
493 characterization of a homologue to the Arabidopsis INDEHISCENT gene in
494 common bean. *Journal of Heredity* 104(2):273-286. doi:10.1093/jhered/ess102
495 Girin, T., Stephenson, P., Goldsack, C. M., Kempin, S. A., Perez, A., and Pires, N., et
496 al. (2010) Brassicaceae INDEHISCENT genes specify valve margin cell fate and
497 repress replum formation. *The Plant Journal* 63(2):329-338.
498 doi:10.1111/j.1365-313X.2010.04244.x
499 Hagerty, C, H., Cuesta-Marcos, A., Cregan, P., Song, Q., McClean, P., and Myers, J. R.
500 (2016) Mapping snap bean pod and color traits, in a dry bean × snap bean
501 recombinant inbred population. *Journal of the American Society for*
502 *Horticultural Science* 141(2):131-138. doi:10.21273/JASHS.141.2.131
503 Huang, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., and Li, C., et al. (2010)
504 Genome-wide association studies of 14 agronomic traits in rice landraces. Nature
505 genetics 42(11):961-967. doi:10.1038/ng.695
506 Hussey, S. G., Mizrachi, E., Spokevicius, A. V., Bossinger, G., Berger, D. K., and
507 Myburg, A. A. (2011) SND2, a NAC transcription factor gene, regulates genes
508 involved in secondary cell wall development in Arabidopsis fibres and increases
509 fibre cell area in Eucalyptus. *BMC plant biology* 11(1): 173.
510 doi:10.1186/1471-2229-11-173
511 Kamfwa, K., Cichy, K. A., and Kelly, J. D. (2015) Genome-wide association study of
512 agronomic traits in common bean. *The Plant Genome* 8(2):1-12.
513 doi:10.3835/plantgenome2014.09.0059
514 Kay, P., Groszmann, M., Ross, J. J., Parish, R. W., and Swain, S. M. (2013)
515 Modifications of a conserved regulatory network involving INDEHISCENT
516 controls multiple aspects of reproductive tissue development in Arabidopsis. *New*
517 *Phytologist* 197(1):73-87. doi:10.1111/j.1469-8137.2012.04373.x

518 Kim, W. C., Kim, J. Y., Ko, J. H., Kang, H., and Han, K. H.(2014) Identification of
519 direct targets of transcription factor MYB46 provides insights into the
520 transcriptional regulation of secondary wall biosynthesis. *Plant molecular*
521 *biology* 85(6):589-599. doi:10.1007/s11103-014-0205-x
522 Kim, W. C., Ko, J. H., Kim, J. Y., Kim, J., Bae, H. J., and Han, K. H. (2013) MYB 46
523 directly regulates the gene expression of secondary wall-associated cellulose
524 synthases in A rabidopsis. *The Plant Journal* 73(1):26-36.
525 doi:10.1111/j.1365-313x.2012.05124.x
526 Koinange, E. M., Singh, S. P., and Gepts, P. (1996) Genetic control of the
527 domestication syndrome in common bean. *Crop Science* 36(4):1037-1045.
528 doi:10.2135/cropsci1996.0011183X003600040037x
529 Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018) MEGA X:
530 molecular evolutionary genetics analysis across computing platforms. *Molecular*
531 *biology and evolution*, 35(6):1547-1549. doi:10.1093/molbev/msy096
532 Lee, C., Teng, Q., Zhong, R., and Ye, Z. H. (2012) Arabidopsis GUX proteins are
533 glucuronyltransferases responsible for the addition of glucuronic acid side chains
534 onto xylan. *Plant and Cell Physiology* 53(7):1204-1216. doi:10.1093/pcp/pcs064
535 Li, H., and Durbin, R. (2009) Fast and accurate short read alignment with
536 Burrows-Wheeler transform. *bioinformatics* 25(14):1754-1760.
537 doi:10.1093/bioinformatics/btp324
538 Ma, Y., Hu, J., Myers, J. R., Mazourek, M., Coyne, C. J., and Main, D., et al. (2016)
539 Development of SCAR markers linked to sin-2, the stringless pod trait in pea
540 (Pisum sativum L.). *Molecular Breeding* 36(7):1-10.
541 doi:10.1007/s11032-016-0525-4
542 McGee, R. J., and Baggett, J. R. (1992) Inheritance of Stringless Pod in Pisum
543 sativum L. *Journal of the American Society for Horticultural*
544 *Science*, 117(4):628-632. doi:10.21273/JASHS.117.4.628
545 McKenna, A., Hann,a M., Banks, E., Sivachenko, A., Cibulskis, K., and Kernytsky, A.,
546 et al. (2010) The Genome Analysis Toolkit: a MapReduce framework for
547 analyzing next-generation DNA sequencing data. *Genome*
548 *research* 20(9):1297-1303. doi:10.1101/gr.107524.110
549 Moghaddam, S. M., Mamidi, S., Osorno, J. M., Lee, R., Brick, M., and Kelly, J., et al.
550 (2016) Genome-wide association study identifies candidate loci underlying
551 agronomic traits in a Middle American diversity panel of common bean. *The*
552 *Plant Genome* 9(3):1-21. doi:10.3835/plantgenome2016.02.0012.
553 Morrell, P. L., Buckler, E. S., and Ross-Ibarra, J. (2012) Crop genomics: advances and
554 applications. *Nature Reviews Genetics* 13(2):85-96. doi:10.1038/nrg3097
555 Myers, J. R., and Kmiecik, K. (2017) Economic importance and relevance to
556 biological science research. In: *The common bean genome*. Springer, Berlin, pp
557 1-20. doi:10.1007/978-3-319-63526-2_1
558 Myers, J. R., Wallace, L. T., Moghaddam, S. M., Kleintop, A. E., Echeverria, D., and
559 Thompson, H. J, et al. (2019) Improving the Health Benefits of Snap Bean:
560 Genome-Wide Association Studies of Total Phenolic
561 Content. *Nutrients* 11(10):2509. doi:10.3390/nu11102509

562 Nakano, Y., Yamaguchi, M., Endo, H., Rejab, N. A., and Ohtani, M. (2015)
563 NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis
564 in land plants. *Frontiers in plant science*, 6:288. doi:10.3389/fpls.2015.00288
565 Oikawa, A., Joshi, H. J., Rennie, E. A., Ebert, B., Manisseri, C., and Heazlewood, J.
566 L., et al. (2010) An integrative approach to the identification of Arabidopsis and
567 rice genes involved in xylan and secondary wall development. *PLoS One*
568 5(11):e15481. doi:10.1371/journal.pone.0015481
569 Parker, T. A., Berny, Mier y Teran, J. C., Palkovic, A., Jernstedt, J., and Gepts, P.
570 (2020) Pod indehiscence is a domestication and aridity resilience trait in
571 common bean. *New Phytol* 225:558–570.doi:10.1111/nph.16164
572 Prakken, R. (1934) Inheritance of colours and pod characters in *Phaseolus vulgaris*
573 L. *Genetica* 16(3-4):177-296.
574 Raggi, L., Caproni, L., Carboni, A., and Negri, V. (2019) Genome-wide association
575 study reveals candidate genes for flowering time variation in common bean
576 (*Phaseolus vulgaris* L.). *Frontiers in plant science* 10:962.
577 doi:10.3389/fpls.2019.00962
578 Rau, D., Murgia, M. L., Rodriguez, M., Bitocchi, E., Bellucci, E., Fois, D., et al.
579 (2019) Genomic dissection of pod shattering in common bean: mutations at
580 non-orthologous loci at the basis of convergent phenotypic evolution under
581 domestication of leguminous species. *The Plant Journal* 97(4):693-714.
582 doi:10.1111/tpj.14155
583 Reusche, M., Thole, K., Janz, D., Truskina, J., Rindfleisch, S., and Drübert, C., et al.
584 (2012) Verticillium infection triggers VASCULAR-RELATED NAC
585 DOMAIN7-dependent de novo xylem formation and enhances drought tolerance
586 in Arabidopsis. *The Plant Cell* 24(9):3823-3837. doi: 10.1105/tpc.112.103374.
587 doi:10.1105/tpc.112.103374
588 Schmutz, J., McClean, P. E., Mamidi, S., Wu, G. A., Cannon, S. B., and Grimwood, J.,
589 et al. (2014) A reference genome for common bean and genome-wide analysis of
590 dual domestications. *Nature genetics* 46(7):707-713. doi:10.1038/ng.3008
591 Soltani, A., MafiMoghaddam, S., Oladzad-Abbasabadi, A., Walter, K., Kearns, P. J.,
592 and Vasquez-Guzman, J., et al. (2018) Genetic analysis of flooding tolerance in
593 an Andean diversity panel of dry bean (Phaseolus vulgaris L.). *Frontiers in plant*
594 *science* 9:767. doi:10.3389/fpls.2018.00767
595 Tang, Y., Liu, X., Wang, J., Li, M., Wang, Q., and Tian, F., et al. (2016) GAPIT
596 version 2: an enhanced integrated tool for genomic association and
597 prediction. *The plant genome* 9(2):1-9. doi:10.3835/plantgenome2015.11.0120
598 Wang, K., Li, M., and Hakonarson, H. (2010) ANNOVAR: functional annotation of
599 genetic variants from high-throughput sequencing data. *Nucleic acids research*
600 38(16):e164-e164. doi:10.1093/nar/gkq603
601 Weeden, N. F. (2007) Genetic changes accompanying the domestication of Pisum
602 sativum: is there a common genetic basis to the 'domestication syndrome' for
603 legumes?. *Annals of botany* 100(5):1017-1025. doi:10.1093/aob/mcm122
604 Wellensiek, S. J. (1971) Lamprecht's gene sin for stringless. *Pisum Newsletter* 3:48.

Wilson, Z. A., Song, J., Taylor, B., and Yang, C. (2011) The final split: the regulation of anther dehiscence. *Journal of experimental botany* 62(5):1633-1649. doi:10.1093/jxb/err014

Wu, L., Chang, Y., Wang, L., Wang, S., and Wu, J. (2021) The aquaporin gene PvXIP1;2 conferring drought resistance identified by GWAS at seedling stage in common bean. *Theor Appl Genet* 1:16. doi:10.1007/s00122-021-03978-w

Wu, J., Wang, L., Fu, J., Chen, J., Wei, S., and Zhang, S., et al. (2020) Resequencing of 683 common bean genotypes identifies yield component trait associations across a north-south cline. *Nature Genetics* 52(1):118-125. doi:10.1038/s41588-019-0546-0

Wu, J., Zhu, J., Wang, L., and Wang, S. (2017) Genome-wide association study identifies NBS-LRR-encoding genes related with anthracnose and common bacterial blight in the common bean. *Frontiers in plant science* 8:1398. doi:10.3389/fpls.2017.01398

Yamaguchi, M., Goué, N., Igarashi, H., Ohtani, M., Nakano, Y., and Mortimer, J. C., (2010) VASCULAR-RELATED NAC-DOMAIN6 and VASCULAR-RELATED NAC-DOMAIN7 effectively induce transdifferentiation into xylem vessel elements under control of an induction system. *Plant physiology* 153(3):906-914. doi:10.1104/pp.110.154013

Yin, Y., Chen, H., Hahn, M. G., Mohnen, D., and Xu, Y. (2010) Evolution and function of the plant cell wall synthesis-related glycosyltransferase family 8. *Plant Physiology* 153(4):1729-1746. doi:10.1104/pp.110.154229

Zhang, X., Blair, M.W., and Wang, S. (2008) Genetic diversity of Chinese common bean (Phaseolus vulgaris L.) landraces assessed with simple sequence repeat markers. *Theor Appl Genet* 117: 629–640. doi:10.1007/s00122-008-0807-2

Zhang, C., Dong, S., Xu, J., He, W., and Yang, T. (2019) PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics*, 35(10):1786-1788. doi:10.1093/bioinformatics/bty875

Zhou, Z., Jiang, Y., Wang, Z., Gou, Z., Lyu, J., and Li, W. (2015) Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nature biotechnology* 33(4):408-414. doi:10.1038/nbt.3096

Zhou, J., Lee, C., Zhong, R., and Ye, Z-H. (2009) MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in Arabidopsis. *The Plant Cell* 21(1):248-266.doi:10.1105/tpc.108.063321

Figure legends

Figure. 1 Histograms of pod suture strings in 138 snap bean accessions. (A) The ratio (string length/pod length) was measured in 2019. (B) The rating (1 for the presence of a string, and 0 for the absence of a string) was counted in 2019. (C) The ratio was measured in 2020. (D) The rating was counted in 2020.

Figure. 2 The number of SNPs within a 1-Mb window size across common bean chromosomes.

Figure. 3 Neighbor-joining tree and population structure analysis using 97,841 single-nucleotide polymorphisms (SNPs). The genepools are colored with red for Andean and blue for Mesoamerican ancestry.

Figure. 4 Circular Manhattan plots of a genome-wide association study (GWAS) for pod stringlessness. The inner circle to the outer circle represent ST1-2019, ST2-2019, ST1-2020, and ST2-2020, respectively. The red dashed lines of each circle indicate the threshold (7.86). Single-nucleotide polymorphisms (SNPs) over the threshold are highlighted.

Figure. 5 Manhattan plots and linkage disequilibrium (LD) heatmap over 266.29 kb around significant single-nucleotide polymorphisms (SNPs) on chromosome 2. (A) Manhattan plots of ST1-2019. The red dashed line represents the threshold (7.86). SNPs over the threshold are highlighted in red. (B) Annotated genes in the region. These CDS, introns, UTR, and intergenic regions are shown in green, light blue, pink, and orange, respectively. (C) The LD heatmap over the region. Colors are coded according to the r^2 color key.

Figure. 6 The identification of a 2-bp deletion in Phvul.002G270800. (a) Structure of Phvul.002G270800. The red base represents the 2-bp deletion in Phvul.002G270800. (b) Box plot of the ratio of string length/pod length for the 2-bp deletion in Phvul.002G270800.

Figure. 7 The analysis of phylogeny and expression of PvGUX1. A. Syntenic analysis of the candidate region between common bean and other legumes. B. Phylogenetic tree for PvGUX1. Colors located at the right side of each sequence represent their amino-acid composition. C. Gene expression of *PvGUX1_1* at different pod development stages for string and stringless snap bean. D. Morphology of stringless and string pod development stages T1–T3. E. The gene expression of PvGUX1_1 at three pod development stages. *P* values were calculated using student's t test (* $p < 0.05$, ** $p < 0.01$).

687 Table 1. Putative genes in the 266.19 kb of candidate region
688

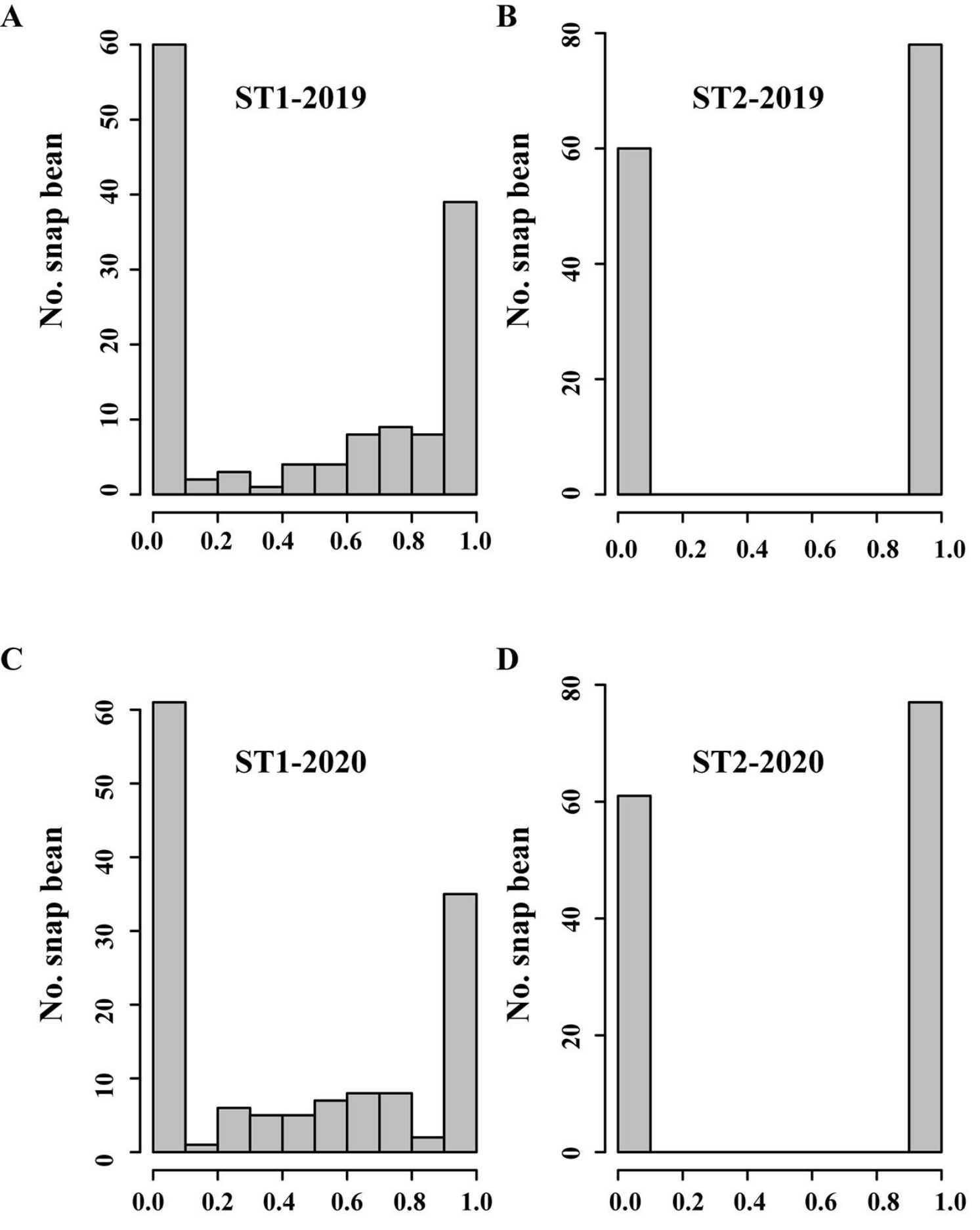
Gene	Positon	Homologs of Arabidopsis	Functional annotation
Phvul.002G269700	Chr02:44022772_44024106		Unknown
Phvul.002G269800	Chr02:44033313_44034850	AT4G08250	GRAS family transcription factor
Phvul.002G269900	Chr02:44036772_44037686	AT5G45420	MYB transcription factor
Phvul.002G270000	Chr02:44037878_44041667	AT3G18750	WNK family of protein kinases
Phvul.002G270100	Chr02:44042454_44059878	AT3G18730	Involved in cell division control and plant morphogenesis
Phvul.002G270200	Chr02:44066265_44067018	AT5G64667	Involved in floral organ abscission
Phvul.002G270300	Chr02:44080456_44082336	AT5G64660	CYS, MET, PRO, and GLY protein
Phvul.002G270400	Chr02:44125511_44131328	AT5G24320	Transducin/WD40 repeat-like superfamily protein
Phvul.002G270500	Chr02:44133980_44137288	AT3G18680	UMP Kinase
Phvul.002G270600	Chr02:44139105_44139713	AT3G18690	Involved in mediating responses to pathogens
Phvul.002G270700	Chr02:44142989_44144630	AT4G14620	Unknown
Phvul.002G270800	Chr02:44150261_44150926	AT3G18660	Encodes a glucuronyltransferase responsible for the addition of GlcA residues onto xylan and for secondary wall deposition.
Phvul.002G270900	Chr02:44155318_44157803	AT5G09760	Plant invertase/pectin methylesterase inhibitor
Phvul.002G271000	Chr02:44186987_44188326	AT4G00120	IND(basic helix-loop-helix (bHLH) DNA-binding superfamily protein)
Phvul.002G271100	Chr02:44199222_44205529	AT1G48880	Encodes a member of the TBL
Phvul.002G271200	Chr02:44205946_44210215	AT5G64630	Involved in organization of the shoot and root apical meristems.
Phvul.002G271300	Chr02:44216969_44222195		Unknown
Phvul.002G271400	Chr02:44224184_44228271	AT1G08490	Chloroplastic NifS-like protein
Phvul.002G271500	Chr02:44229799_44246330	AT5G64070	Encodes a phosphatidylinositol 4-OH kinase
Phvul.002G271600	Chr02:44232557_44233756		Unknown
Phvul.002G271700	Chr02:44247536_44251023	AT5G09330	NAC domain containing protein
Phvul.002G271800	Chr02:44251436_44254178	AT2G13690	PRLI-interacting factor
Phvul.002G271900	Chr02:44257054_44258132		Unknown

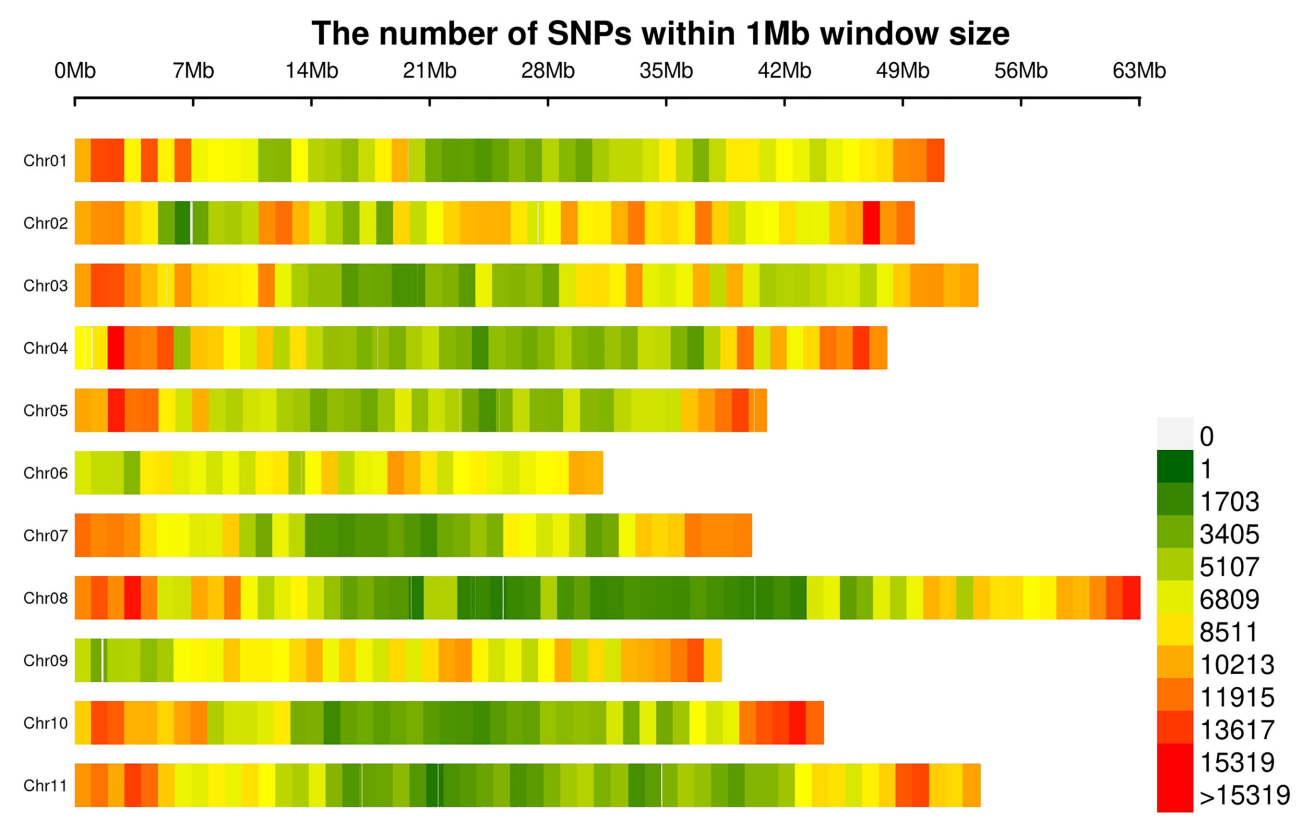
689
690
691
692
693
694
695

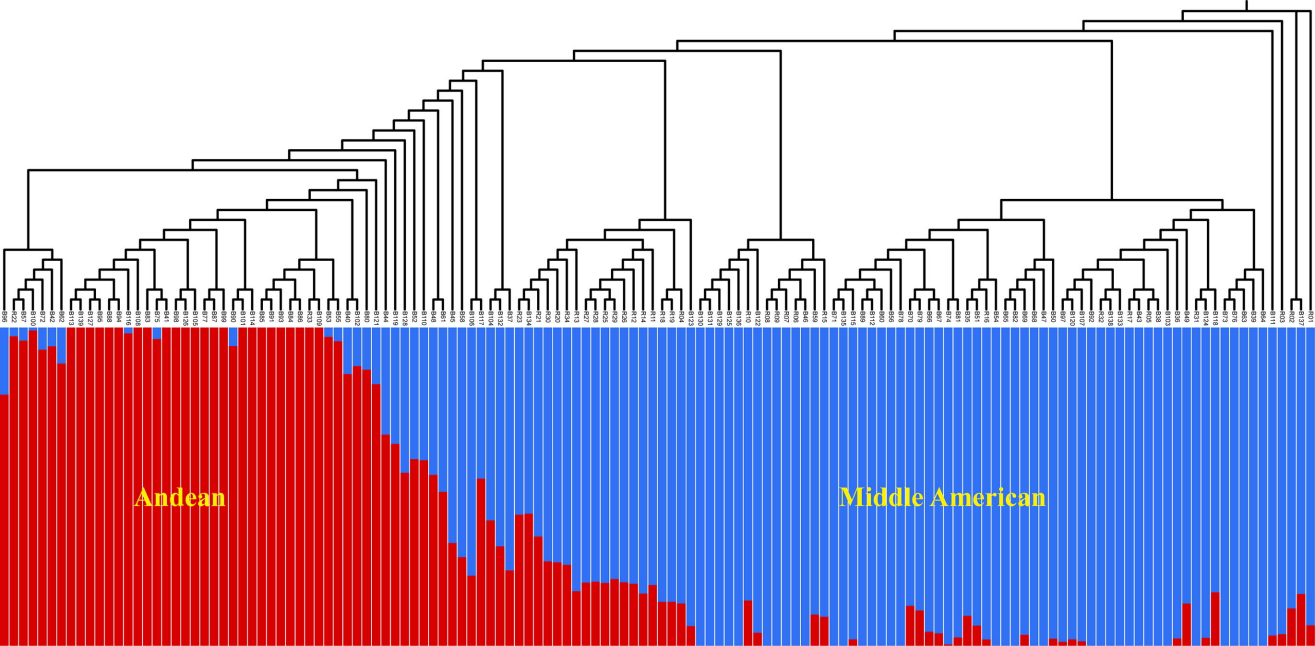
696 Table 2 Functional annotation information of candidate genes

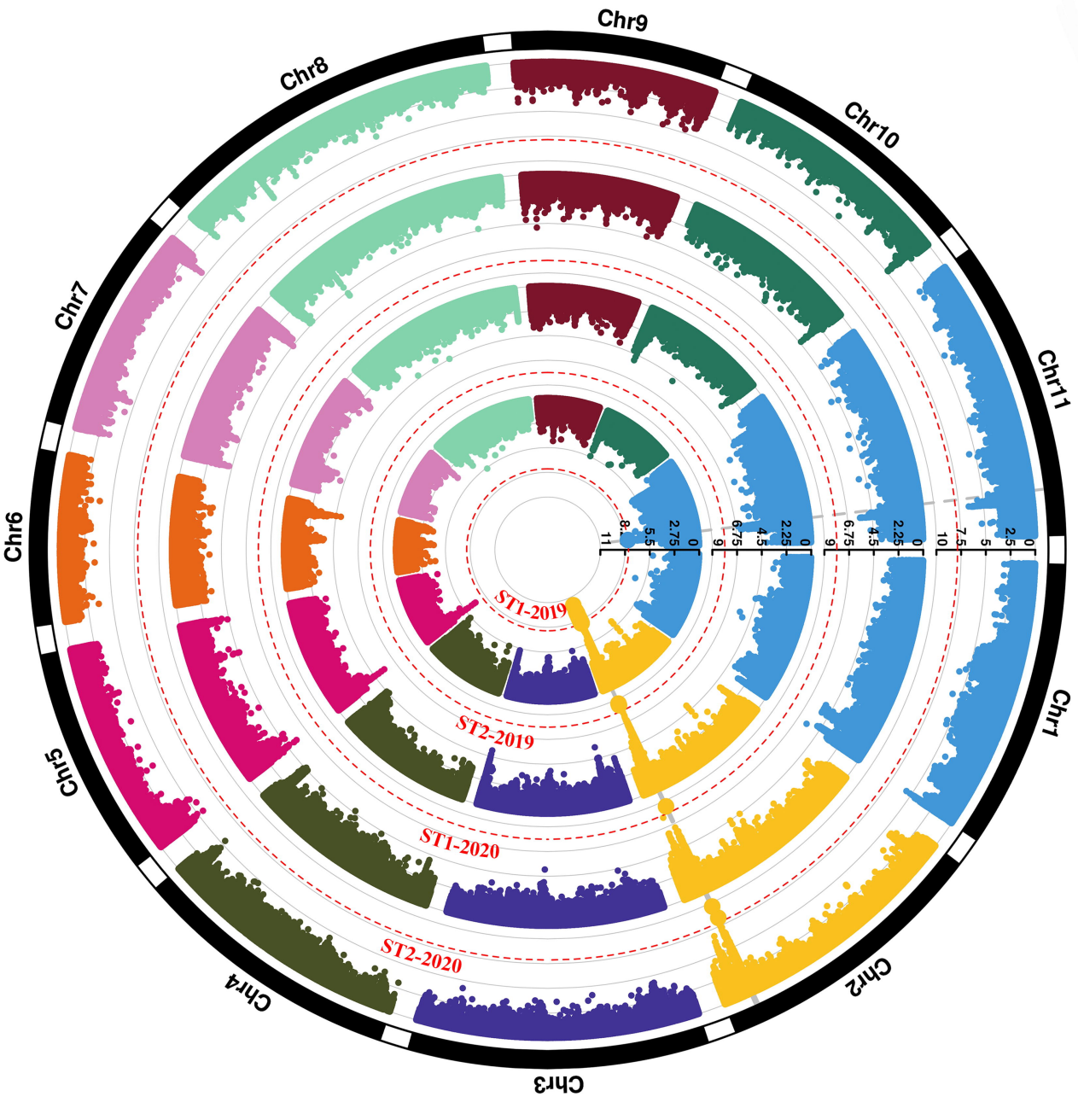
Gene	Variant Type	Base Change	Amino Change
Phvul.002G269700	Nonsynonymous	A491G	E164G
	Nonsynonymous	T1617A	D539E
Phvul.002G270100	Nonsynonymous	A1899T	R633S
	Nonsynonymous	G2092A	D698N
	Nonsynonymous	C3050T	S1017L
Phvul.002G270300	Nonsynonymous	C257T	S86L
	Nonsynonymous	T382C	F128L
Phvul.002G270400	Nonsynonymous	G563A	R188K
	Nonsynonymous	C833A	T278N
Phvul.002G270800	Nonsynonymous	A92G	D31G
	Frameshift deletion	AT163_	
Phvul.002G270900	Nonsynonymous	G1374C	E458D
	Nonsynonymous	T32C	V11A
Phvul.002G271000	Nonsynonymous	C604T	P202S
	Nonsynonymous	C737T	T246M
Phvul.002G271100	Nonsynonymous	A257G	N86S
Phvul.002G271200	Nonsynonymous	G870A	M290I
	Nonsynonymous	G241T	G81C
	Nonsynonymous	A638C	E213A
Phvul.002G271300	Nonsynonymous	A781G	T261A
	Nonsynonymous	T1945G	S649A
	Nonsynonymous	A2008G	N670D
	Nonsynonymous	G234A	M78I
Phvul.002G271400	Nonsynonymous	A335C	K112T
	Nonsynonymous	C902T	A301V
	Nonsynonymous	A1099G	T367A
	Nonsynonymous	T17C	L6S
Phvul.002G271600	Nonsynonymous	A24C	L8F
	Frameshift deletion	GT57_	
	Nonsynonymous	A77G	N26S
	Nonsynonymous	C78G	N26K
	Nonsynonymous	T100G	F34V
Phvul.002G271700	Nonsynonymous	T154C	S52P
	Nonsynonymous	G162C	K54N
	Nonsynonymous	A359T	K120I
	Nonsynonymous	T476C	V159A
	Nonsynonymous	A785C	D262A
	Nonsynonymous	A1346C	Y449S
	Nonsynonymous	A1138C	K380Q
Phvul.002G271800	Nonsynonymous	A772C	N258H
	Nonsynonymous	T563A	F188Y
	Nonsynonymous	A385G	N129D
	Nonsynonymous	C328A	L110M
Phvul.002G271900	Nonsynonymous	T5G	F2C
	Nonsynonymous	T377C	V126A

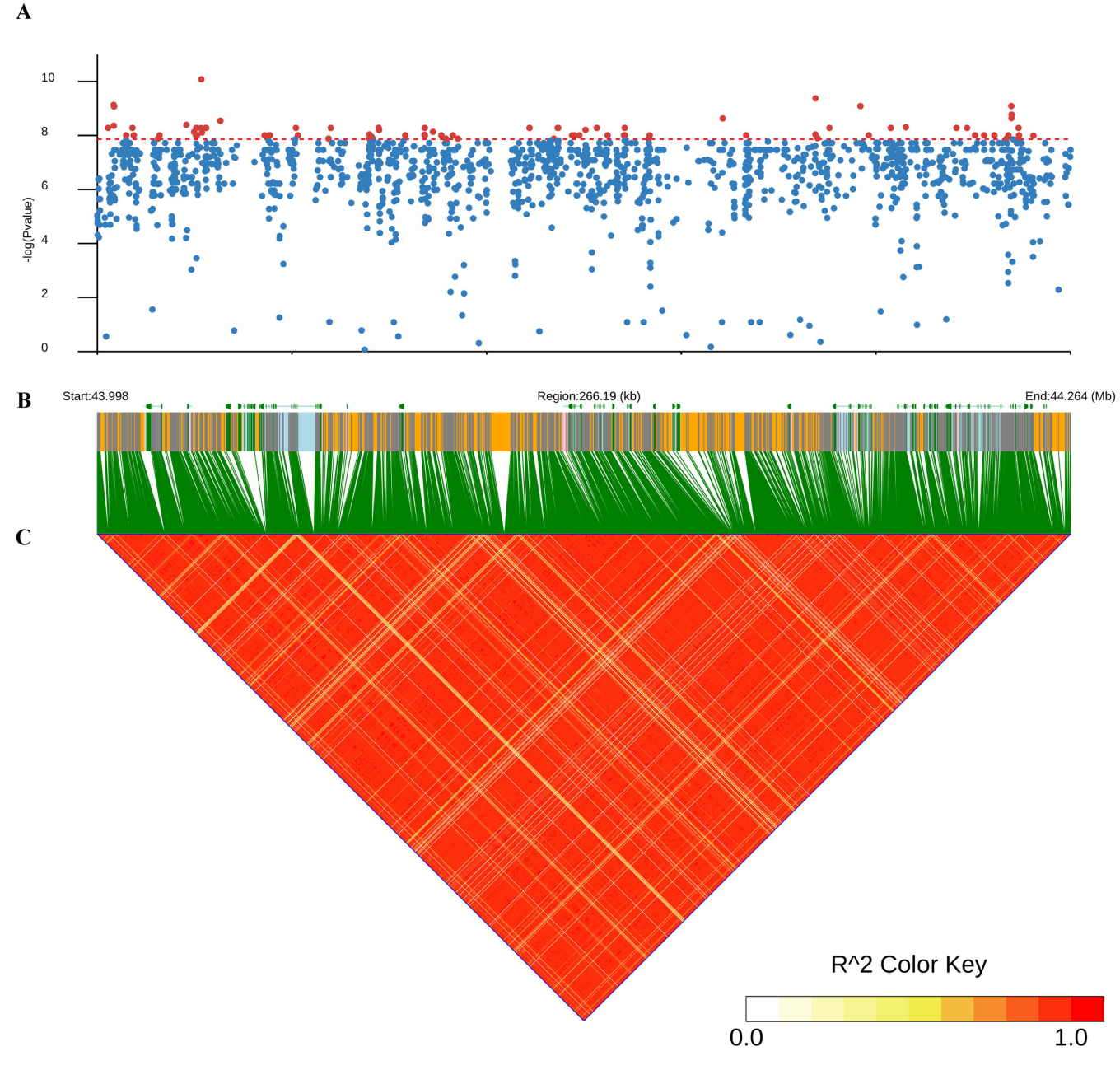
697


A
ST1-2019
No. snap bean
0
10
20
30
40
50
60
0.0
0.2
0.4
0.6
0.8
1.0
B
ST2-2019
No. snap bean
0
20
40
60
80
0.0
0.2
0.4
0.6
0.8
1.0
C
ST1-2020
No. snap bean
0
10
20
30
40
50
60
0.0
0.2
0.4
0.6
0.8
1.0
D
ST2-2020
No. snap bean
0
20
40
60
80
0.0
0.2
0.4
0.6
0.8
1.0

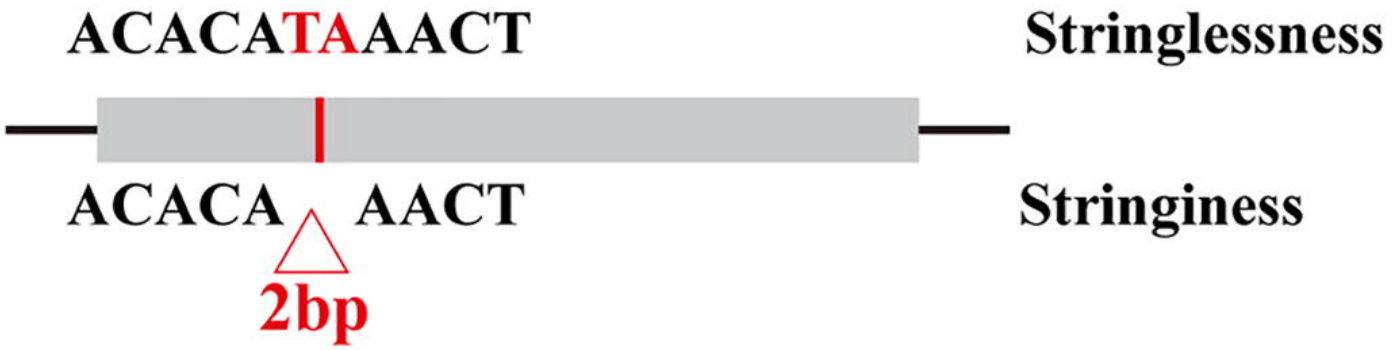

The number of SNPs within 1Mb window size
0Mb
7Mb
14Mb
21Mb
28Mb
35Mb
42Mb
49Mb
56Mb
63Mb
Chr01
Chr02
Chr03
Chr04
Chr05
Chr06
Chr07
Chr08
Chr09
Chr10
Chr11
0
1
1703
3405
5107
6809
8511
10213
11915
13617
15319
>15319


Andean
Middle American

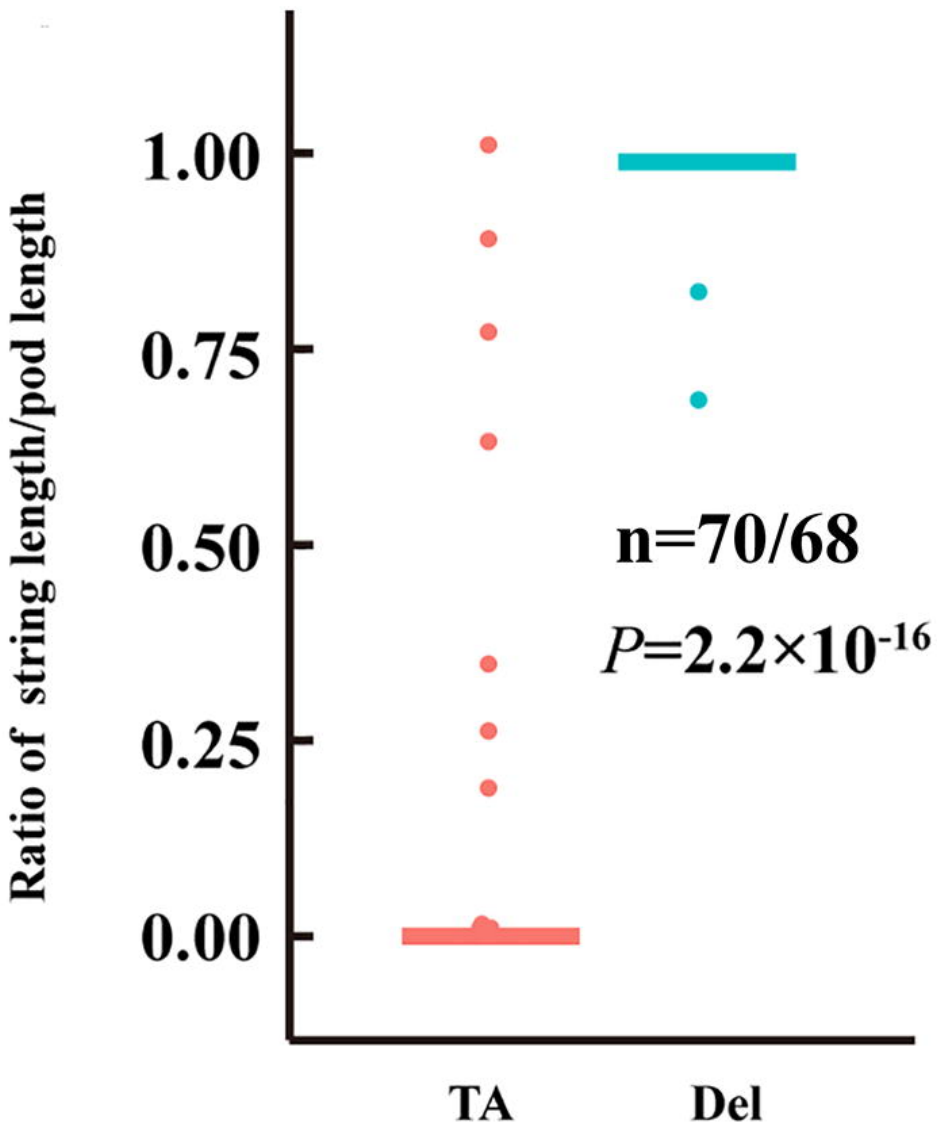

Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chr10
Chr11
ST1-2019
ST2-2019
ST1-2020
ST2-2020

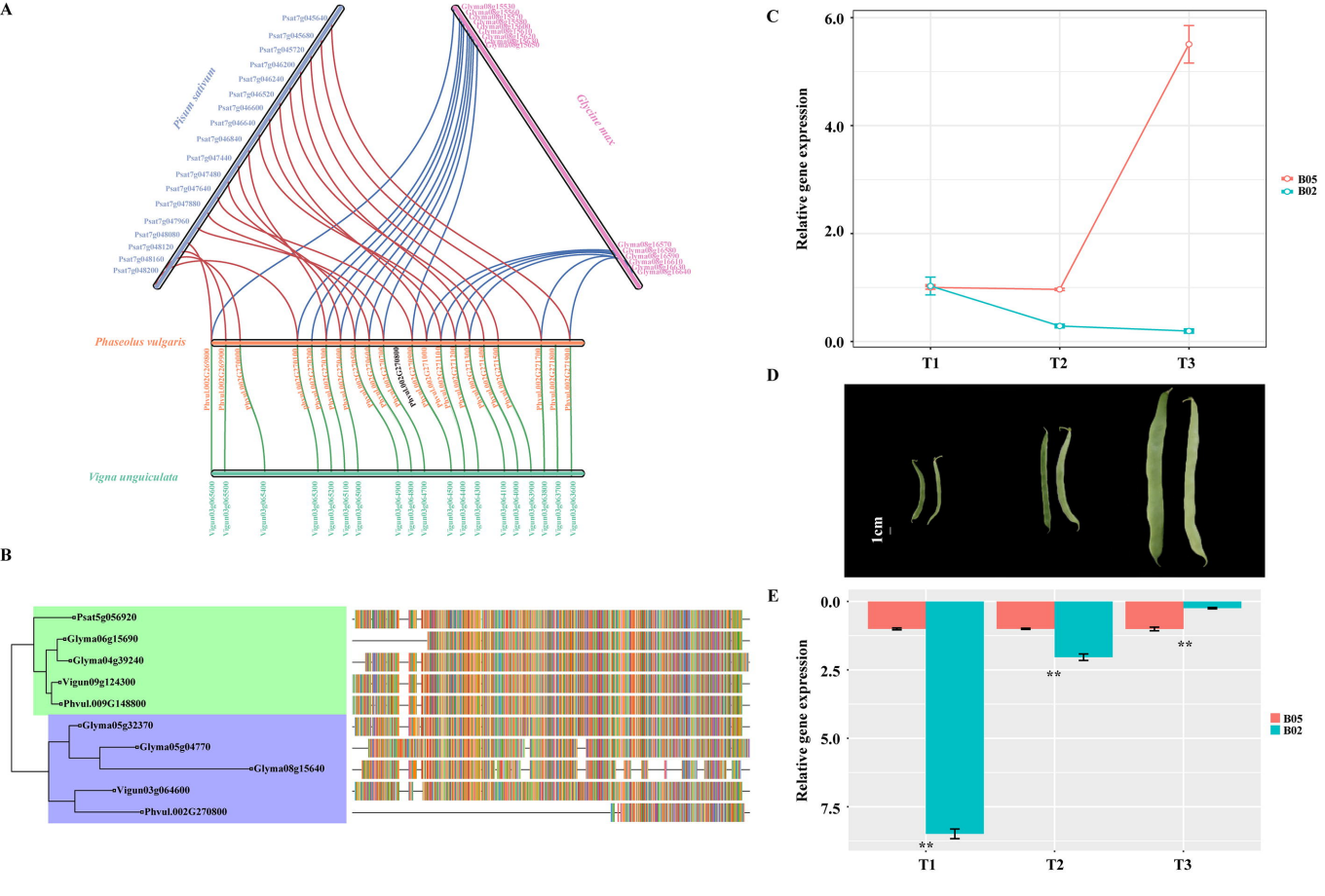

A
10
8
6
4
2
0
-log(Pvalue)
B
Start:43.998
Region:266.19 (kb)
End:44.264 (Mb)
C
R^2 Color Key
0.0
1.0

A


ACACATAAACT
Stringlessness
ACACA
AACT
Stringiness
2bp

B


Ratio of string length/pod length
1.00
0.75
0.50
0.25
0.00
n=70/68
P=2.2×10-16
TA
Del


A
Pisum sativum
Glycine max
Phaseolus vulgaris
Vigna unguiculata
B
Psat5g056920
Glyma06g15690
Glyma04g39240
Vigun09g124300
Phvul.009G148800
Glyma05g32370
Glyma05g04770
Glyma08g15640
Vigun03g064600
Phvul.002G270800
C
Relative gene expression
6.0
4.0
2.0
0.0
T1
T2
T3
B05
B02
D
1cm
E
Relative gene expression
0.0
2.5
5.0
7.5
**
T1
T2
T3
B05
B02