




Article

Identification of Gene Associated with Sweetness in Corn (*Zea mays* L.) by Genome-Wide Association Study (GWAS) and Development of a Functional SNP Marker for Predicting Sweet Corn

Vinitchan Ruanjaichon ¹, Kanoporn Khammona ¹, Burin Thunnom ¹, Khundej Suriharn ^{2,3}, Chalong Kerdsri ⁴, Wanchana Aesomnuk ¹, Arweewut Yongsuwan ¹, Naraporn Chaomueang ¹, Paradee Thammapichai ⁵, Siwaret Arikrit ^{6,7}, Samart Wanchana ^{1,*} and Theerayut Toojinda ^{1,*}

¹ National Center for Genetic Engineering and Biotechnology (BIOTEC), 113 Thailand Science Park, Pahonyothin Road, Khlong Nueng, Khlong Luang, Pathum Thani 12120, Thailand; vinitchan.rua@biotec.or.th (V.R.); kanoporn.kha@ncr.nstda.or.th (K.K.); burin.thu@biotec.or.th (B.T.); wanchana.a@ku.th (W.A.); golfarweewut@gmail.com (A.Y.); narapornchouw@gmail.com (N.C.)

² Department of Agronomy, Faculty of Agriculture, Khon Kaen University, Khon Kaen 40002, Thailand; sphala@kku.ac.th

³ Plant Breeding Research Center for Sustainable Agriculture, Khon Kaen University, Khon Kaen 40002, Thailand

⁴ Chai Nat Field Crops Research Center, Chai Nat 17000, Thailand; chalong_maize@live.com

⁵ Department of Plant and Soil Sciences, Faculty of Agriculture, Chiang Mai University, Chiang Mai 50200, Thailand; paradee.thammapichai@cmu.ac.th

⁶ Department of Agronomy, Faculty of Agriculture at Kamphaeng Saen, Kasetsart University Kamphaeng Saen Campus, Nakhon Pathom 73140, Thailand; siwaret.a@ku.th

⁷ Rice Science Center, Kasetsart University Kamphaeng Saen Campus, Nakhon Pathom 73140, Thailand

* Correspondence: samart.wan@biotec.or.th (S.W.); theerayut@biotec.or.th (T.T.)



Citation: Ruanjaichon, V.; Khammona, K.; Thunnom, B.; Suriharn, K.; Kerdsri, C.; Aesomnuk, W.; Yongsuwan, A.; Chaomueang, N.; Thammapichai, P.; Arikrit, S.; et al. Identification of Gene Associated with Sweetness in Corn (*Zea mays* L.) by Genome-Wide Association Study (GWAS) and Development of a Functional SNP Marker for Predicting Sweet Corn. *Plants* **2021**, *10*, 1239. <https://doi.org/10.3390/plants10061239>

Academic Editors:
Ioannis Ganopoulos and
Yong-Gu Cho

Received: 6 May 2021
Accepted: 15 June 2021
Published: 18 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Sweetness is an economically important eating quality trait for sweet-corn breeding. To investigate the genetic control of the sweetness trait, we conducted a genome-wide association study (GWAS) in an association panel consisting of 250 sweet corn and waxy corn inbred and recombinant inbred lines (RILs), together with the genotypes obtained from the high-density 600K maize genotyping single-nucleotide polymorphism (SNP) array. GWAS results identified 12 significantly associated SNPs on chromosomes 3, 4, 5, and 7. The most associated SNP, AX_91849634, was found on chromosome 3 with a highly significant p -value of $\leq 1.53 \times 10^{-14}$. The candidate gene identified within the linkage disequilibrium (LD) of this marker was *shrunken2* (Zm00001d044129; *sh2*), which encodes ADP-glucose pyrophosphorylase (AGPase), a 60 kDa subunit enzyme that affects starch metabolism in the maize endosperm. Several SNP markers specific to variants in *sh2* were developed and validated. According to the validation in a set of 81 inbred, RIL, and popular corn varieties, marker Sh2_rs844805326, which was developed on the basis of the SNP at the position 154 of exon 1, was highly efficient in classifying *sh2*-based sweet corn from other types of corn. This functional marker is extremely useful for marker-assisted breeding in *sh2*-sweet corn improvement and marketable seed production.

Keywords: maize; kernel sweetness; *shrunken2*; genome-wide association study (GWAS); single-nucleotide polymorphism (SNP)

1. Introduction

Sweet corn (*Zea mays* L. var. *rugosa* Bonaf.) is globally an economically important crop, and it is widely used for human consumption from both immature and mature kernels as fresh, canned, and frozen due to its high sugar content [1–3]. The demand for sweet corn has been rising on a yearly basis due to enhanced consumption and the increasing

availability of food-processing industries. Corn kernels come in different sizes and shapes depending on types and varieties. Sweet-corn kernels are usually more yellow than orange-tinted field corn kernels and are visibly shriveled when dried. The wrinkled appearance in dried kernels is due to a sugary gene that retards the conversion of sugar into starch during endosperm development and results in kernel sugar accumulation [4,5].

Sweet corn can be classified on the basis of gene composition in the endosperm into five types: normal, sugary enhanced, super sweet, synergistic, and augmented shrunken [6]. Normal sweet corn contains an allelic mutation in the sugar gene, *sugary1* (*su1*), which increases sugar content, but has a high starch conversion rate. Supersweet corn carries a *shrunken2* (*sh2*) mutant, which maintains high sugar level, has a steady rate of starch accumulation, and is suitable for shipping purposes. *Shrunken2* (*Sh2*) encodes large regulatory subunits of the heterotetrameric ADP-glucose pyrophosphorylase (AGPase) present in the maize-kernel endosperm, which is upstream in the starch biosynthetic pathway [7]. It is a rate-limiting enzyme in starch biosynthesis [8,9]. The mutation at the *Sh2* locus causes corn to be supersweet and has thus been widely used for the development of extrasweet corn cultivars [10]. Maize with the *sh2* mutant has three- and eightfold increases in sugar content (29.9% sucrose) compared with maize with *su1* mutant (10.2% sucrose) and normal corn, respectively [6,11–13]. Sugar-enhanced (*se*) sweet corn with an *se* mutant in an *su1* background has a higher sugar level, more tenderness, and an extended shelf life compared with normal sweet corn [14]. Synergistic corn, possessing a gene combination of *su-1*, *sh2*, and *se*, gives a balance quality between high sugar content and tenderness [6,12]. Molecular markers associated with these genes and other sweetness-related genes could help plant breeders to quickly and accurately select the desirable sweetness combination to develop synergistic hybrids. However, functional markers for selecting the sweetness trait in maize are still limited in the public sector and some private companies. Recently, gene-based markers for *shrunken2-Reference* allele have been developed and used in marker-assisted sweet corn (*Zea mays Sachharata*) breeding program [15].

Quantitative trait locus (QTL) analysis is extensively used to improve the quality of numerous crops such as rice, wheat, tomato, and soybean [16–19]. Most QTL analyses are performed by linkage mapping on the basis of a segregating population from a biparental cross. Alternatively, genome-wide association study (GWAS) is a powerful tool to find associations between natural phenotypic diversity and underlying genetic variants [20]. GWAS is a powerful approach to associate natural phenotypic diversity with underlying genetic variants. It has since gained popularity as a valuable tool in the quest for QTLs or candidate genes from selected genetic resources without having to develop mapping populations, which can be labor- and time-intensive. Since the release of the maize B73 reference genome [21], GWAS is routinely used to detect single-nucleotide polymorphisms (SNPs) associated with various developmental and agronomic traits in maize such as plant height [22], tassel architecture [23], major ear [24], and kernel nutritional traits [25,26].

In this study, we identify the gene associated with the kernel sweetness trait through a GWAS approach using a collection of 250 sweet and waxy corn inbred and recombination inbred lines (RILs). The set of SNPs, obtained from an Axiom® Maize 600k Genotyping Array, were used to evaluate linkage-disequilibrium (LD) decays, population structure, and SNPs associated with kernel sweetness. Additionally, gene-based SNP markers were developed and validated. SNP markers are extremely useful for maize-breeding programs for sweetness traits.

2. Results

2.1. Trait Evaluation, Genotyping, and Population Study in Panel of 250 Maize Lines

A total of 250 maize inbred and recombinant inbred lines (RILs) were used to evaluate the kernel trait (Table S1). The phenotype of these maize lines was determined as sweet or waxy corn on the basis of sucrose content and the structure of dried kernels (Figure 1A; Figure S1; Table S1). Among the 250 lines, 86 lines were classified as sweet corn, and 164 lines as waxy corn (Table S1). The sucrose content in the sweet corn group varied

from 23.33 to 88.71 mg/g fresh weight with an average of 43.61 mg/g, while the content in waxy-corn group varied from 3.46 to 15.59 fresh weight with an average of 7.70 mg/g (Figure S2). We then genotyped these 250 maize lines using the Axiom[®] Maize 600 K Genotyping Array (616,201 variants) to obtain SNP data for the population study and GWAS analyses. Out of 616,201 variants, 426,858 high-quality SNPs were determined in the 250 lines. All of these SNPs had a minor-allele frequency (MAF) greater than 0.05, and maximal missing rate of less than 20%. These numbers of SNPs were further filtered to obtain a manageable SNP dataset. As a result, a subset of 159,201 SNPs were retained after filtering on the basis of the distance of at least 1 kb between each pair of adjacent SNPs (Figure S3); 19,565 SNPs were retained after linkage-disequilibrium (LD) pruning. Determined on the basis of 159,201 SNPs, heterozygosity among maize lines varied from 0.02 to 0.20, with an average of 0.09 (Table S1).

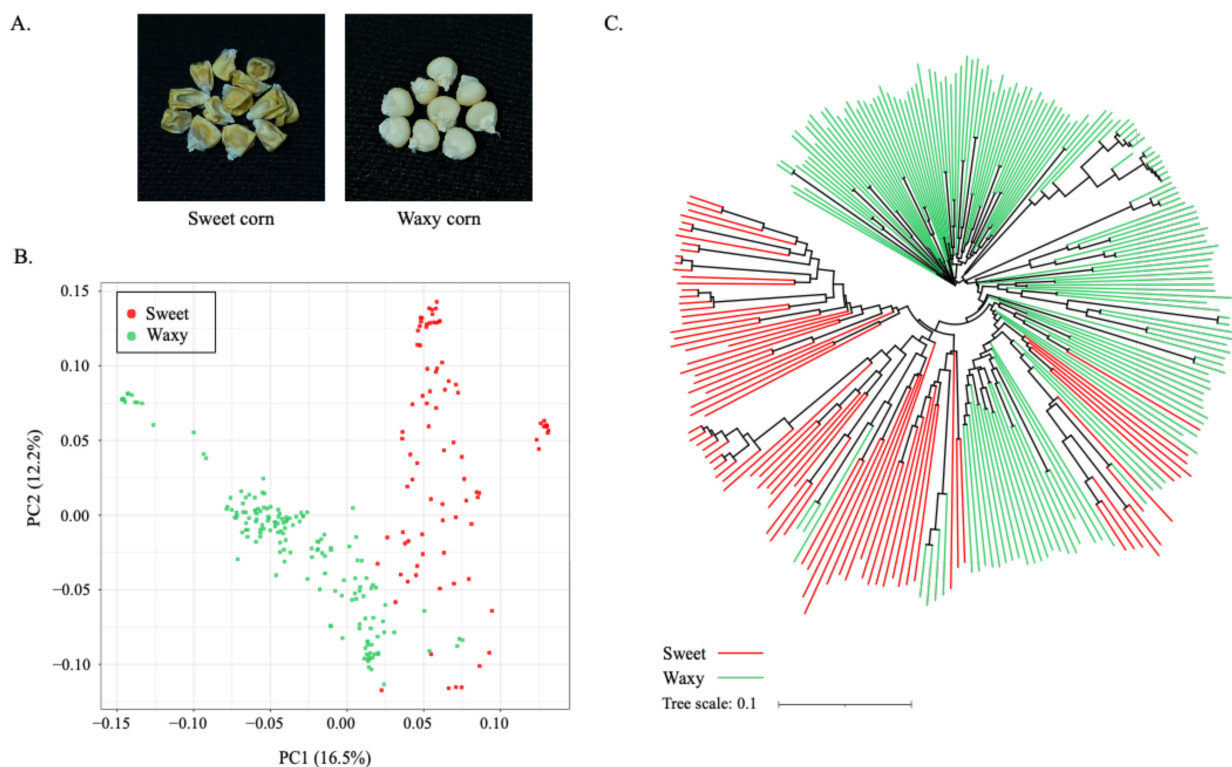


Figure 1. Genetic structure of 250 sweet- and waxy-corn lines. (A) Appearance of kernel structures that was used as a criterion to assign maize lines as sweet corn and waxy corn. (B) Principal-component analysis plots based on 159,201 SNP markers; (C) phylogenetic-tree clustering based on 159,201 SNP markers. The waxy corn and sweet corn accessions in (B,C) are shown in green and red, respectively.

We performed principal-component analysis (PCA) on the basis of the 159,201 SNPs to characterize the population structure, and performed phylogenetic-tree analysis to determine the genetic diversity of the 250 maize lines. PCA revealed clearly separated sweet and waxy corn along the first principal component (PC1; Figure 1B). The phylogenetic tree also identified two major clusters corresponding to waxy and sweet corn, similar to those revealed by PCA (Figure 1C). In addition, on the basis of the 19,565 LD-pruned SNPs, STRUCTURE identified four subpopulations among the 250 maize lines (Figure S4).

We further analyzed the genome-wide LD decays among the 250 maize lines using the 159,201 SNPs. Results showed that the overall LD decay across the genome of 250 maize lines was 210 kb at a cut-off $r^2 = 0.1$; the mean length of LD decay decreased rapidly to 76 kb at a cut-off of $r^2 = 0.2$ (Figure 2).

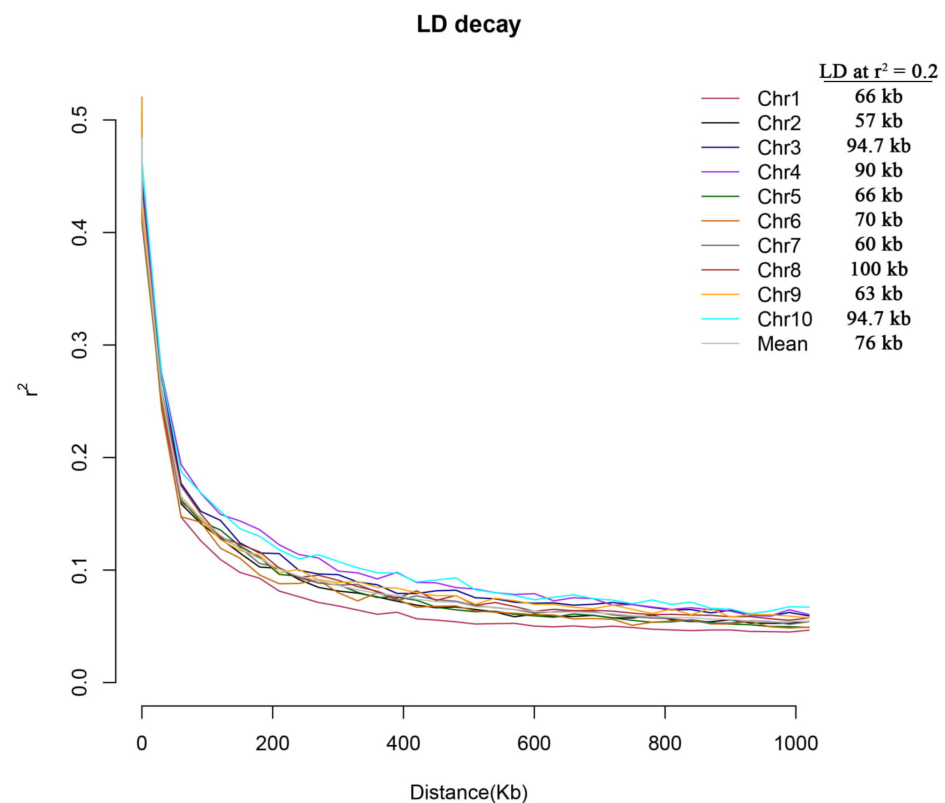


Figure 2. Overall chromosome-wide linkage-disequilibrium (LD) decay estimated from single-nucleotide polymorphism (SNP) genotypes of 250 maize lines. Each line plot represents a smoothed R^2 for all marker pairs on each chromosome depending on the distance between marker pairs.

2.2. Genome-Wide Association Study (GWAS) for Sweetness Trait

We performed a GWAS for the sweetness trait in a panel of 250 inbred and recombinant inbred lines (RILs) on the basis of the 159,201 SNPs using a mixed linear model (MLM) implemented in the Genome Association and Prediction Integrated Tool (GAPIT version 3.0). SNP-based heritability for sweet corn versus waxy corn was 99.97%. The GWAS result identified 12 significant SNPs located on chromosomes 3, 4, 5, and 7 that passed the Bonferroni threshold of $-\log_{10} p\text{-value} = 7.20$ (Figure 3; Table 1). The most associated SNP (AX_91848634) with a $-\log_{10} p\text{-value}$ of 13.81 was located on chromosome 3. Three more significant SNPs located on this chromosome were AX_90850969, AX_91593662, and AX_91593878, which were also detected in the region close to AX_91848634 (Table 1). These four significant SNPs had minor-allele frequency (MAF) varying in the range of 0.31–0.33. The eight other significant SNPs—five of which located on chromosome 4, one of which located on chromosome 5, and two of which located on chromosome 7—mostly had a low MAF value that varied in the range of 0.05–0.09, except for AX_90866628, located on chromosome 4, which had an MAF of 0.48 (Table 1). The defined approximately 100 kb interval of the most associated region on chromosome 3 contained *Shrunken2* (*Sh2*), a gene involved in kernel starch biosynthesis (Table 1). The lead SNP (AX_91848634) was 67.5 kb away from *Sh2*. The gene was located within the LD (210 kb; $r^2 = 0.10$) of the lead SNP. No starch- or sugar-related genes were annotated within the significant regions on the other detected chromosomes (Table S3). Based on these results, *sh2*, which is known to control sweetness in super-sweet corn, was considered as the candidate gene associated with the sweetness trait in this maize panel.

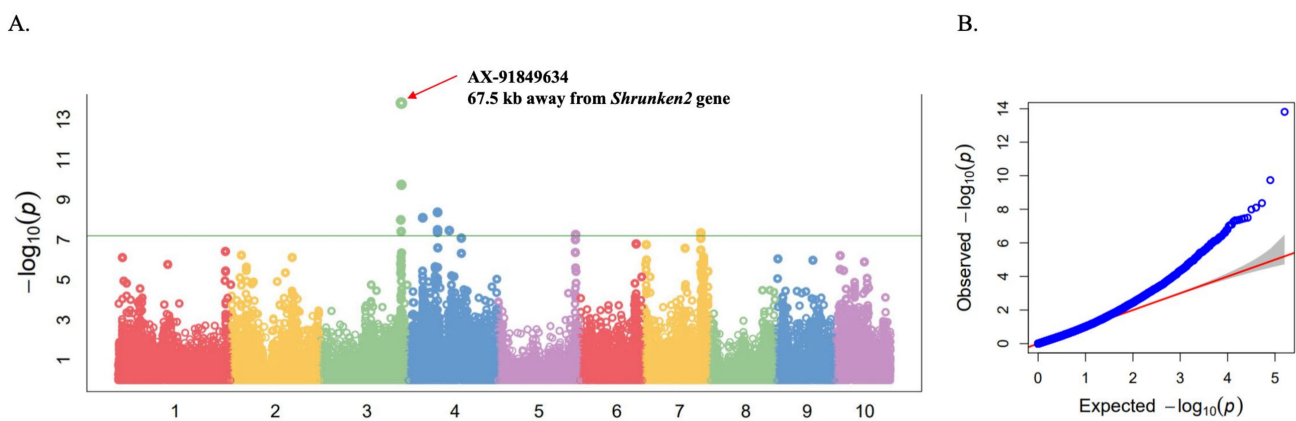


Figure 3. GWAS results for sweetness trait in 250 sweet- and waxy-corn inbred and recombinant inbred lines (RILs), genotyped with 159,201 SNPs, using a mixed linear model implemented in GAPIT. (A) Manhattan plots. Each dot represents a SNP. Bonferroni threshold of $-\log_{10} p$ -value = 13.81 is presented by a green line on Manhattan plots. Most associated SNP AX_91849634, located 67.5 kb on chromosome 3 near the *Shrunken2* gene, is indicated by a red arrow. (B) Quantile–quantile (Q–Q) plots. The plot shows the expected vs. observed $-\log_{10}(p)$ of each marker (blue dots). Red line is a guide for the perfect fit to expected $-\log_{10}(p)$. The gray shaded area shows the 95% confidence interval for the Q–Q plot under the null hypothesis of no association between the SNP and the trait.

Table 1. List of significantly associated SNPs identified by GWAS showing SNP IDs, chromosomes, positions, $-\log_{10}(p)$ values, minor-allele-frequency (MAF) values, marker R^2 values, and candidate gene.

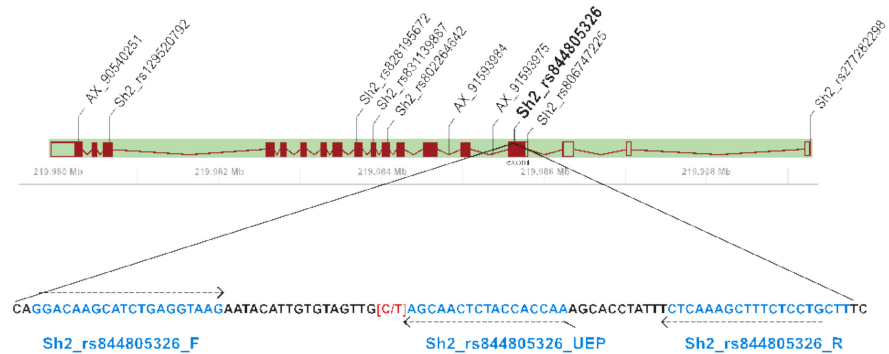
SNP	Chr	Position (V4)	$-\log_{10}(p)$ Value	MAF	R^2	Candidate Gene
AX-91593662	3	218,476,619	7.99	0.33	0.17	-
AX-91593878	3	219,463,930	7.42	0.31	0.16	-
AX-91849634	3	22,005,7126	13.81	0.33	0.18	Zm00001d044129 (<i>Shrunken2</i>)
AX-90850969	3	220,099,694	9.74	0.34	0.17	-
AX-90866628	4	42,950,506	8.1	0.48	0.16	-
AX-90877243	4	83,464,190	7.5	0.05	0.11	-
AX-91612113	4	83,465,341	8.37	0.05	0.16	-
AX-91346854	4	83,501,441	7.37	0.07	0.11	-
AX-91617781	4	115,510,176	7.46	0.05	0.17	-
AX-90975122	5	213,164,020	7.26	0.07	0.10	-
AX-91411958	7	159,043,703	7.34	0.09	0.09	-
AX-91740905	7	159,053,605	7.34	0.09	0.11	-

2.3. Development of SNP Markers Associated with Sweetness Trait Based on *Sh2* Variants and Validation among Different Maize Types

To verify the association between *Sh2* and the sweetness trait in this maize panel, we developed SNP markers on the basis of variants in the *Sh2* gene and genotyped in a subset of the maize lines used in GWAS analysis. Ten of the SNP variants on the gene were selected to develop the SNP markers; seven variants were retrieved from the Gramene (<https://ensembl.gramene.org>, accessed on 1 May 2021) (gene id: Zm00001d044129), and three variants were obtained from the Affymetrix Axiom 600 K genotyping array (Table 2; Figure 4A). Among these, six SNPs had a missense effect, two SNPs were located on the introns, and two SNPs were located upstream the first start codon (Figure 4). These SNPs are different from those used to develop the gene-based markers for *sh2*-Reference allele [15]. The markers were designed on the basis of these 10 selected SNPs and validated in a set of 43 maize lines using a MassARRAY[®] platform (Table S4). According to regression analysis for marker-trait association for these 10 markers, marker Sh2_rs844805326 exhibited the most association with the trait (R^2 value of 1.00, $p < 0.000$) (Table S5). This marker was developed on the basis of the SNP variant (C/T) located at exon 1 at position 154 of the coding sequence (CDS) (Figure 4A). Among the 43 maize lines, 38 lines with the sweet-corn phenotype contained an SNP (C), and five lines with the waxy-corn phenotype contained

an SNP (T). The SNP variant at this position caused an amino acid change from threonine to alanine at position 52 of the amino acid sequence.

A.



B.

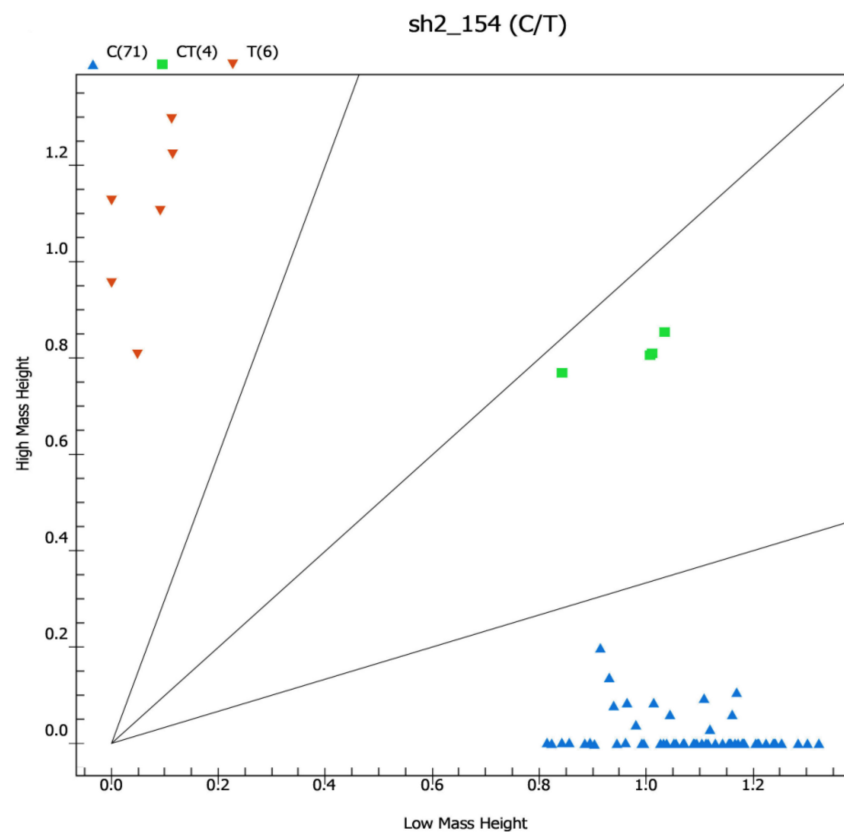


Figure 4. MassArray marker associated with sweetness trait developed based on the SNP in *sh2*. (A) Physical locations of molecular markers developed on the basis of gene *Sh2*. Exons and introns in the gene structure are presented by filled boxes and lines, respectively. Unfilled boxes represent untranslated regions (UTRs). Primers (dashed arrows) are presented along with the target sequence. The letters F and R in the primer's name indicate forward and reverse primers, respectively. UEP indicated an unextended primer used in a single-base primer extension reaction. (B) Allelic discrimination plot of marker Sh2_rs844805326 (SNP 154_CT) validated in panel of 81 lines consisting of different types of corn analyzed by MassARRAY[®] platform. Scatter dots with different colors show clustering of homozygous genotype TT (brown), heterozygous genotype CT (green), and homozygous genotype CC (blue).

Table 2. Details of SNP markers for *Sh2* gene used for genotyping by MassARRAY® platform.

Markers	Chr	Position (V.4)	SNPs	Unextended Primer (UEP) Sequence (5'–3') ^a	UEP Mass (Da)	Call 1	Mass 1 (Da)	Call 2	Mass 2 (Da)
AX_90540251	3	219,980,272	G/A	tGTGATCTTGAAGAATGCAA	6180	G	6427.2	A	6507.1
Sh2_rs129520792	3	219,980,689	A/G	ccccgGCTTCTTTCAGTTTCATAG	7839.1	A	8110.3	G	8126.3
Sh2_rs828195672	3	219,983,718	C/G	AATTCCTTGAAAAACCAAAGG	6446.2	G	6693.4	C	6733.5
Sh2_rs831139887	3	219,983,940	C/T	gtatcGTGATATCAGCATCGTCCT	7318.8	C	7565.9	T	7645.9
Sh2_rs802264642	3	219,984,110	C/T	CCTGTATTCTTTAGGATTATTACA	7612	T	7883.2	C	7899.2
AX_91593984	3	219,984,862	T/A	cccgACCACTTCGATTATGG	6052.9	A	6324.2	T	6380
AX_91593975	3	219,985,438	A/T	cccTCAAATAGGGTCATCT	5747.8	T	6019	A	6074.9
Sh2_rs844805326	3	219,985,637	C/T	gTTGGTGGTAGAGTTGCT	5616.6	T	5887.9	C	5903.9
Sh2_rs806747225	3	219,985,795	A/G	AGTGCAAACCTGCATATCT	5482.6	C	5729.8	T	5809.7
Sh2_rs277282298	3	219,989,285	G/A	ACCACTACCACCAACA	4748.1	C	4995.3	T	5075.2

^a The small letters are the bases added to the 5' end to adjust the mass of UEPs.

The SNP marker was then used to genotype a set of 81 diverse maize composing of hybrids, inbred lines, and RILs (Table S2). As expected, this marker clearly separated the *sh2*-based sweet corn from waxy corn, field corn, and the sweet corn that was based on the other genes, i.e., *su1*, *se1*, and *bt2* (Figure 4B). The sweet-corn hybrids (*sh2/se*) exhibited a heterozygous genotype using this marker.

3. Discussion

Sweet corn is a globally popular crop for human food and the seed-production industry. Sweetness in corn kernels is controlled by six recessive mutant genes expressed in the endosperm: *brittle1* (*bt1*), *brittle2* (*bt2*), *shrunkened1* (*sh1*), *shrunkened2* (*sh2*), *sugary1* (*su1*), and *sugary enhancer1* (*se1*) [1–5]. One gene in particular, *shrunkened2* (*sh2*), which causes a significant increase in endosperm sugar level, is extensively used in sweet-corn cultivar development. In this study, we performed GWAS analysis against the sweetness and waxy characteristics in a set of maize inbred and recombinant inbred lines. The used population in this study clearly showed distinct kernel phenotypes as sweet corn with a shriveled kernel, and waxy corn with an opaque kernel. Therefore, the two characteristics were simply determined as a qualitative trait using a binary code of 0 and 1 for sweetness and waxy, respectively, for GWAS analysis. Although GWAS is usually aimed at quantitative traits, GWAS based on qualitative traits were also reported [27]. Individuals in the sweet-corn group could contain a mutant in *sh2*, *se1*, *su1*, and *bt2*, and individuals in the waxy-corn group possibly contained a mutation in the *waxy* gene. However, the exact mutation in these individuals was not clearly characterized.

Due to the strong population structure and a degree of individual relatedness present in our association panel, we performed association tests on the basis of the mixed linear model (MLM) that incorporated the PCA and kinship matrices. The GWAS results accordingly revealed that SNP AX_91849634 was significantly associated with kernel type and was located 67.5 kb away from *Sh2* on chromosome 3. However, no significant SNP was detected for other genes related to the maize endosperm, such as *brittle2* (*bt2*), *sugary1* (*su1*), *sugary enhancer1* (*se1*), and *waxy* (*wx*). The lack of a significant association might be compressed by the amount of sweet-corn type in the genetic resource panel that is commonly controlled by the *sh2* gene. This could be due to the fact that the maize lines were preferentially selected and maintained for sweetness as well as other agronomic traits, as these maize lines are used for genetic mapping and breeding purposes. Therefore, it is possible that the lines with *sh2* background were more desirable and hence preferentially selected.

Shrunkened2 (*sh2*) is extensively used in sweet-corn cultivar development [13]. The *Sh2* gene was previously reported to be involved in starch biosynthesis in the maize endosperm, which encodes the large subunit of the heterotetrameric enzyme, adenosine diphosphate glucose pyrophosphorylases (AGP), and a rate-limiting enzyme [28,29]. A naturally occurring allele, *sh2*, is a loss-of-function mutation that causes a significant reduction in AGP activity [29]. Thus, sucrose could not be converted into ADP glucose, a precursor in starch

biosynthesis pathway, resulting in low levels of starch and endosperm sugar accumulation in sweet corn [30].

Molecular markers are used for the characterization of germplasms, identification and mapping of genes and QTLs for various agriculturally important traits in maize [31–36]. A previous study identified four simple sequence repeat (SSR) markers linked to the *sh2* gene via bulk-segregant analysis (BSA) in three mapping populations [37]. Ideally, markers located on the gene are useful for improving such traits via marker-assisted breeding. However, there is no report of functional markers associated with genes or QTLs underlying kernel sweetness in maize. Chhabra et al. (2020) recently reported gene-based PCR markers that resided in the 5'UTR and intron-12 regions of the *sh2* gene [15]. In this study, we extended the search for gene-based SNPs around our most associated SNP near the *sh2* gene on chromosome 3. One SNP was validated to differentiate sweet corn from other corn types by possessing a *sh2* recessive allele (T). This SNP was located within the coding region in exon 1 at position 154 of the coding sequence, and represents protein sequence conversion from threonine to alanine at position 52 of the amino acid sequence. This might render an enzyme defective due to a change in polarity from a nonpolar, hydrophobic to a polar, uncharged amino acid, thereby affecting starch synthesis in the maize endosperm. The identified SNPs on the *Sh2* gene could be used as selection markers for marker-assisted selection in maize-breeding programs. Moreover, the functional SNP identified on *Sh2* would offer essential information for marker-based breeding in sweet-corn production and seed industry. The gene-based SNP markers identified thus far could effectively save breeders cost and valuable time from lengthy field trials, and aid in accelerating sweet-corn breeding programs.

4. Materials and Methods

4.1. Plant Materials and Kernel-Trait Evaluation

The association panel consisted of 250 sweet- and waxy-corn inbred and recombinant inbred lines (RILs) originally developed to contain different kernel-trait genes and gene combinations, such as *su*, *se*, *sh2*, *bt*, *wx*, *btwx*, and *btsh2wx*. Out of 250 lines, 60 were developed by Chai Nat Field Crops Research Center (CNFCRC) and 190 were developed by the Plant Breeding Research Center for Sustainable Agriculture, Faculty of Agriculture, Khon Kaen University (KKU; Table S1). The genotyping panel for marker validation consisted of 81 sweet- and waxy-corn inbred and hybrid lines, including field corn (Table S2). Sucrose content was analyzed from immature ears of each line at 24 days after pollination using the method described previously [38]. Based on the sucrose content together with the structure of dried kernels, the kernel phenotype was defined as a simple qualitative trait and the coding numbers 0 as sweet corn (shrunken kernels with high sucrose content) and 1 as waxy corn (non-shrunken and opaque kernels with low sucrose content) were used for phenotyping.

4.2. SNP Array Genotyping and Data Filtering

DNA from the 250 maize lines was extracted from young leaves using DNeasy Plant Mini Kit (QIAGEN) and then genotyped by the Axiom[®] Maize 600k Genotyping Array that contained 616,201 variants (www.affymetrix.com, accessed on 1 May 2021). SNPs with more than 10% missing data and a minor-allele frequency (MAF) less than 0.05 were removed, yielding a total of 426,858 quality SNPs. Resulting SNPs were then further pruned to retain SNPs of at least 1 kb apart between each pair of adjacent SNPs. Lastly, a total of 159,201 remaining SNPs were retained. These SNPs were used for the genome-wide association study (GWAS) for the kernel trait. Another set of pruned SNPs was performed on the basis of linkage disequilibrium (LD) to obtain a subset of unlinked SNPs using a variant pruning tool (-indep-pairwise 50 10 0.1) in PLINK [39]. A total of 19,565 SNPs were retained. This set of SNPs was used for population-structure analysis using STRUCTURE [40]. The genome coordinates of the SNP loci throughout the study were based on B73_RefGen_V4 (maizegdb.org).

4.3. Population-Structure Analysis and Linkage-Disequilibrium (LD) Decay

Principal-component analysis (PCA) was performed to determine the genetic population structure of the samples using TASSEL 5.0 software [41]. The linkage disequilibrium (LD) between SNPs on each chromosome was estimated with r^2 using the same software. Genetic distance was calculated using Nie's standard distance [42], and the phylogenetic tree was created using the CLC Genomics Workbench (QIAGEN, Germany) on the basis of the unweighted pair group with the arithmetic mean (UPGMA) method with 1000 bootstrap replicates. LD-based SNP pruning was also performed to collect representative SNPs that were used for STRUCTURE analysis. STRUCTURE analysis was performed using a Bayesian model-based clustering algorithm implemented in STRUCTURE version 2.3.4 [40], where the admixture model with correlated allele frequencies was used. A total of 10 independent replicates were run for each genetic cluster (K) value ($K = 1-10$), using a burn-in period of 10,000 and a run length of 10,000 iterations. $\text{LnP}(D)$ values were derived for each K and plotted to find the plateau of the ΔK . The final population structure was calculated using the structure harvester [43]. Genome-wide LD decay versus genetic distance was estimated by pairwise analysis of adjacent SNPs within a chromosome using PopLDdecay [44] based on the set of 159,201 SNP markers.

4.4. Genome-Wide Association Study (GWAS)

The GWAS was performed for a simple qualitative trait of kernel structure between sweet versus waxy corn with a mixed linear model (MLM) that incorporated PC and kinship data using R package Genome Association and Prediction Integrated Tool (GAPIT) version 3 [45]. The significant threshold for SNP-trait associations was the Bonferroni correction of $p \leq 0.01/n$, where n represents the number of SNP markers across the entire genome.

4.5. Candidate-Gene Analysis

The genomic regions within the LD block of the significant SNPs (peak SNPs) were selected to identify candidate genes. Gene annotation was based on the B73 reference genome (B73_RefGen_v4). SNP probe sequences of ~150 bp on the Axiom[®] Maize 600 k Genotyping Array were used as queries in a BLAST algorithm-based search against the reference genome sequence in MaizeGDB (<http://www.maizegdb.org/gbrowse>; accessed on 1 May 2021) to identify the positions in the maize genome.

4.6. Gene-Based SNP Marker Development and Validation

Genomic sequences for primer design were obtained from B73 sequence version 4 at Gramene (<https://ensembl.gramene.org>; accessed on 1 May 2021). The primer set for high-throughput MassARRAY was designed and analyzed according to the manufacture of the MassARRAY[®] iPLEX system protocol. The primers for the SNP sites were designed to contain 1 bp mismatch at the second base closest to the 3' end of the forward primer. All markers were validated in the 81 maize accessions, including popular commercial maize varieties (Table S2).

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/plants10061239/s1>, Figure S1: the physical appearance of dried kernels of sweet, waxy and field corns; Figure S2: box plots and statistical test of sucrose content compared between sweet corn and waxy corn groups in the panel of 250 lines; Figure S3: density and distribution of SNPs throughout 12 rice chromosomes; Figure S4: subpopulations of 250 maize lines; Table S1: list of 250 inbred and recombinant inbred lines used for GWAS analysis in this study; Table S2: list of 81 inbred and recombinant inbred lines, and commercial hybrids used for marker validation; Table S3: annotation of genes located within the same LD block as the associated SNPs. Table S4: list of maize samples used to validate MassARRAY markers. Information on types and sources of maize samples, and genotypes of each marker are provided; Table S5: list of MassARRAY markers designed on the basis of 10 variants in Sh2. Marker details and R2 values from single marker analysis validated on 43 maize lines are provided.

Author Contributions: V.R., T.T. and S.W. designed and supervised the research. K.K., K.S., C.K. and V.R. performed the experiments. K.K., A.Y., N.C., B.T., S.A., W.A. and V.R. analyzed the data. S.W., P.T. and V.R. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science and Technology Development Agency (NSTDA), grant number P-17-52167.

Data Availability Statement: Data are contained within the article and the Supplementary Materials.

Acknowledgments: The authors thank the Chai Nat Field Crops Research Center and Plant Breeding Research Center for Sustainable Agriculture, Khon Kaen University for providing the plant materials used in the study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tracy, W.F.; Shuler, S.L.; Dodson-Swenson, H. The Use of Endosperm Genes for Sweet Corn Improvement: A Review of Developments in Endosperm Genes in Sweet Corn since the Seminal Publication in *Plant Breeding Reviews*, Volume 1, by Charles Boyer and Jack Shannon (1984). *Plant Breed. Rev.* **2019**, *43*, 215–241.
2. Boyer, C.D.; Shannon, J.C. The use of endosperm genes for sweet corn improvement. In *Plant Breeding Reviews*; Janick, J., Ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1983; pp. 139–161. ISBN 9780870553974.
3. Tracy, W.F. History, genetics, and breeding of supersweet (*shrunk2*) sweet corn. In *Plant Breeding Reviews*; Janick, J., Ed.; John Wiley & Sons, Inc.: Oxford, UK, 1996; pp. 189–236. ISBN 9780470650073.
4. James, M.G.; Robertson, D.S.; Myers, A.M. Characterization of the Maize Gene Sugary1, a Determinant of Starch Composition in Kernels. *Plant Cell* **1995**, *7*, 417. [[CrossRef](#)]
5. Laughnan, J.R. The Effect of the Sh(2) Factor on Carbohydrate Reserves in the Mature Endosperm of Maize. *Genetics* **1953**, *38*, 485–499. [[CrossRef](#)]
6. Singh, I.; Langyan, S.; Yadava, P. Sweet Corn and Corn-Based Sweeteners. *Sugar Tech* **2014**, *16*, 144–149. [[CrossRef](#)]
7. Thévenot, C.; Simond-Côte, E.; Reyss, A.; Manicacci, D.; Trouverie, J.; Le Guilloux, M.; Ginhoux, V.; Sidicina, F.; Prioul, J.-L. QTLs for Enzyme Activities and Soluble Carbohydrates Involved in Starch Accumulation during Grain Filling in Maize. *J. Exp. Bot.* **2005**, *56*, 945–958. [[CrossRef](#)]
8. Espada, J. ADP-glucose pyrophosphorylase from corn grain. In *Complex Carbohydrates; Methods in Enzymology*; Elsevier: Amsterdam, The Netherlands, 1966; Volume 8, pp. 259–262. ISBN 9780121818081.
9. Smidansky, E.D.; Clancy, M.; Meyer, F.D.; Lanning, S.P.; Blake, N.K.; Talbert, L.E.; Giroux, M.J. Enhanced ADP-Glucose Pyrophosphorylase Activity in Wheat Endosperm Increases Seed Yield. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1724–1729. [[CrossRef](#)]
10. Feng, Z.L.; Liu, J.; Fu, F.L.; Li, W.C. Molecular Mechanism of Sweet and Waxy in Maize. *Int. J. Plant Breed. Genet.* **2008**, *2*, 93–100. [[CrossRef](#)]
11. Creech, R.G. Genetic Control of Carbohydrate Synthesis in Maize Endosperm. *Genetics* **1965**, *52*, 1175–1186. [[CrossRef](#)]
12. Szymanek, M.; Tanaś, W.; Kassar, F.H. Kernel Carbohydrates Concentration in Sugary-1, Sugary Enhanced and Shrunken Sweet Corn Kernels. *Agric. Agric. Sci. Procedia* **2015**, *7*, 260–264. [[CrossRef](#)]
13. Lertrat, K.; Pulam, T. Breeding for Increased Sweetness in Sweet Corn. *Int. J. Plant Breed.* **2007**, *1*, 27–30.
14. Ferguson, J.E.; Dickinson, D.B.; Rhodes, A.M. Analysis of Endosperm Sugars in a Sweet Corn Inbred (Illinois 677a) Which Contains the Sugary Enhancer (Se) Gene and Comparison of Se with Other Corn Genotypes. *Plant Physiol.* **1979**, *63*, 416–420. [[CrossRef](#)] [[PubMed](#)]
15. Chhabra, R.; Hossain, F.; Muthusamy, V.; Baveja, A.; Mehta, B.K.; Uttamrao Zunjare, R. Development and Validation of Gene-based Markers for *Shrunken2-Reference* Allele and Their Utilization in Marker-assisted Sweet Corn (*Zea Mays Sachharata*) Breeding Programme. *Plant Breed.* **2020**, *139*, 1135–1144. [[CrossRef](#)]
16. Wang, J.; Wan, X.; Li, H.; Pfeiffer, W.H.; Crouch, J.; Wan, J. Application of Identified QTL-Marker Associations in Rice Quality Improvement through a Design-Breeding Approach. *Appl. Genet.* **2007**, *115*, 87–100. [[CrossRef](#)]
17. Kumar, N.; Kulwal, P.L.; Gaur, A.; Tyagi, A.K.; Khurana, J.P.; Khurana, P.; Balyan, H.S.; Gupta, P.K. QTL Analysis for Grain Weight in Common Wheat. *Euphytica* **2006**, *151*, 135–144. [[CrossRef](#)]
18. Bernacchi, D.; Beck-Bunn, T.; Eshed, Y.; Lopez, J.; Petiard, V.; Uhlig, J.; Zamir, D.; Tanksley, S. Advanced Backcross QTL Analysis in Tomato. I. Identification of QTLs for Traits of Agronomic Importance from *Lycopersicon Hirsutum*. *Appl. Genet.* **1998**, *97*, 381–397. [[CrossRef](#)]
19. Chung, J.; Babka, H.L.; Graef, G.L.; Staswick, P.E.; Lee, D.J.; Cregan, P.B.; Shoemaker, R.C.; Specht, J.E. The Seed Protein, Oil, and Yield QTL on Soybean Linkage Group I. *Crop Sci.* **2003**, *43*, 1053–1067. [[CrossRef](#)]
20. Brachi, B.; Morris, G.P.; Borevitz, J.O. Genome-Wide Association Studies in Plants: The Missing Heritability Is in the Field. *Genome Biol.* **2011**, *12*, 232. [[CrossRef](#)]
21. Schnable, P.S.; Ware, D.; Fulton, R.S.; Stein, J.C.; Wei, F.; Pasternak, S.; Liang, C.; Zhang, J.; Fulton, L.; Graves, T.A.; et al. The B73 Maize Genome: Complexity, Diversity, and Dynamics. *Science* **2009**, *326*, 1112–1115. [[CrossRef](#)] [[PubMed](#)]

22. Peiffer, J.A.; Romay, M.C.; Gore, M.A.; Flint-Garcia, S.A.; Zhang, Z.; Millard, M.J.; Gardner, C.A.C.; McMullen, M.D.; Holland, J.B.; Bradbury, P.J.; et al. The Genetic Architecture of Maize Height. *Genetics* **2014**, *196*, 1337–1356. [[CrossRef](#)] [[PubMed](#)]
23. Wu, X.; Li, Y.; Shi, Y.; Song, Y.; Zhang, D.; Li, C.; Buckler, E.S.; Li, Y.; Zhang, Z.; Wang, T. Joint-Linkage Mapping and GWAS Reveal Extensive Genetic Loci That Regulate Male Inflorescence Size in Maize. *Plant Biotechnol. J.* **2016**, *14*, 1551–1562. [[CrossRef](#)]
24. Zhu, X.-M.; Shao, X.-Y.; Pei, Y.-H.; Guo, X.-M.; Li, J.; Song, X.-Y.; Zhao, M.-A. Genetic Diversity and Genome-Wide Association Study of Major Ear Quantitative Traits Using High-Density SNPs in Maize. *Front. Plant Sci.* **2018**, *9*, 966. [[CrossRef](#)]
25. Baseggio, M.; Murray, M.; Magallanes-Lundback, M.; Kaczmar, N.; Chamness, J.; Buckler, E.S.; Smith, M.E.; DellaPenna, D.; Tracy, W.F.; Gore, M.A. Genome-Wide Association and Genomic Prediction Models of Tocochromanols in Fresh Sweet Corn Kernels. *Plant Genome* **2019**, *12*. [[CrossRef](#)] [[PubMed](#)]
26. Xiao, Y.; Yu, Y.; Li, G.; Xie, L.; Guo, X.; Li, J.; Li, Y.; Hu, J. Genome-Wide Association Study of Vitamin E in Sweet Corn Kernels. *Crop J.* **2020**, *8*, 341–350. [[CrossRef](#)]
27. Bandillo, N.B.; Lorenz, A.J.; Graef, G.L.; Jarquin, D.; Hyten, D.L.; Nelson, R.L.; Specht, J.E. Genome-Wide Association Mapping of Qualitatively Inherited Traits in a Germplasm Collection. *Plant Genome* **2017**, *10*. [[CrossRef](#)]
28. Bhave, M.R.; Lawrence, S.; Barton, C.; Hannah, L.C. Identification and Molecular Characterization of Shrunken-2 CDNA Clones of Maize. *Plant Cell* **1990**, *2*, 581–588. [[CrossRef](#)]
29. Hannah, L.C.; Nelson, O.E. Characterization of ADP-Glucose Pyrophosphorylase from Shrunken-2 and Brittle-2 Mutants of Maize. *Biochem. Genet.* **1976**, *14*, 547–560. [[CrossRef](#)]
30. Boyer, C.D.; Hannah, L.C. Kernel mutants of corn. In *Specialty Corns*; CRC Press: Boca Raton, FL, USA, 2000; pp. 13–44. ISBN 0429127456.
31. Ferreira, F.; Scapim, C.A.; Maldonado, C.; Mora, F. SSR-Based Genetic Analysis of Sweet Corn Inbred Lines Using Artificial Neural Networks. *Crop Breed. Appl. Biotechnol.* **2018**, *18*, 309–313. [[CrossRef](#)]
32. Wang, G.; Zhao, Y.; Mao, W.; Ma, X.; Su, C. QTL Analysis and Fine Mapping of a Major QTL Conferring Kernel Size in Maize (*Zea mays*). *Front. Genet.* **2020**, *11*, 603920. [[CrossRef](#)]
33. Zhao, Y.; Su, C. Mapping Quantitative Trait Loci for Yield-Related Traits and Predicting Candidate Genes for Grain Weight in Maize. *Sci. Rep.* **2019**, *9*, 16112. [[CrossRef](#)]
34. Mahato, A.; Shahi, J.P.; Singh, P.K.; Kumar, M. Genetic Diversity of Sweet Corn Inbreds Using Agro-Morphological Traits and Microsatellite Markers. *3 Biotech* **2018**, *8*, 332. [[CrossRef](#)]
35. Mehta, B.; Hossain, F.; Muthusamy, V.; Baveja, A.; Zunjare, R.; Jha, S.K.; Gupta, H.S. Microsatellite-Based Genetic Diversity Analyses of Sugary1-, Shrunken2- and Double Mutant- Sweet Corn Inbreds for Their Utilization in Breeding Programme. *Physiol. Mol. Biol. Plants* **2017**, *23*, 411–420. [[CrossRef](#)]
36. Kashiani, P.; Saleh, G.; Panandam, J.M.; Abdullah, N.A.P.; Selamat, A. Molecular Characterization of Tropical Sweet Corn Inbred Lines Using Microsatellite Markers. *Maydica* **2012**, *57*, 154–163.
37. Hossain, F.; Nepolean, T.; Vishwakarma, A.K.; Pandey, N.; Prasanna, B.M.; Gupta, H.S. Mapping and Validation of Microsatellite Markers Linked to Sugary1 and Shrunken2 Genes in Maize (*Zea mays* L.). *J. Plant Biochem. Biotechnol.* **2015**, *24*, 135–142. [[CrossRef](#)]
38. Simla, S.; Lertrat, K.; Suriharn, B. Carbohydrate Characters of Six Vegetable Waxy Corn Varieties as Affected by Harvest Time and Storage Duration. *Asian J. Plant Sci.* **2010**, *9*, 463. [[CrossRef](#)]
39. 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. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [[CrossRef](#)]
40. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the Number of Clusters of Individuals Using the Software STRUCTURE: A Simulation Study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [[CrossRef](#)] [[PubMed](#)]
41. Bradbury, P.J.; Zhang, Z.; Kroon, D.E.; Casstevens, T.M.; Ramdoss, Y.; Buckler, E.S. TASSEL: Software for Association Mapping of Complex Traits in Diverse Samples. *Bioinformatics* **2007**, *23*, 2633–2635. [[CrossRef](#)]
42. Nei, M. Genetic Distance between Populations. *Am. Nat.* **1972**, *106*, 283–292. [[CrossRef](#)]
43. Earl, D.A.; vonHoldt, B.M. STRUCTURE HARVESTER: A Website and Program for Visualizing STRUCTURE Output and Implementing the Evanno Method. *Conserv. Genet. Resour.* **2012**, *4*, 359–361. [[CrossRef](#)]
44. Zhang, C.; Dong, S.-S.; Xu, J.-Y.; He, W.-M.; Yang, T.-L. PopLDdecay: A Fast and Effective Tool for Linkage Disequilibrium Decay Analysis Based on Variant Call Format Files. *Bioinformatics* **2019**, *35*, 1786–1788. [[CrossRef](#)]
45. Lipka, A.E.; Tian, F.; Wang, Q.; Peiffer, J.; Li, M.; Bradbury, P.J.; Gore, M.A.; Buckler, E.S.; Zhang, Z. GAPIT: Genome Association and Prediction Integrated Tool. *Bioinformatics* **2012**, *28*, 2397–2399. [[CrossRef](#)]