EcoTILLING Reveals Natural Allelic Variations in Starch Synthesis Key Gene *TaSSIV* and Its Haplotypes Associated with Higher Thousand Grain Weight

Ahsan Irshad †, Huijun Guo †, Shunlin Zhang, Jiayu Gu, Linshu Zhao, Yongdun Xie, Hongchun Xiong, Shirong Zhao, Yuping Ding, Youzhi Ma and Luxiang Liu *

Institute of Crop Sciences, Chinese Academy of Agricultural Sciences/National Engineering Laboratory of Crop Molecular Breeding/National Center of Space Mutagenesis for Crop Improvement, Beijing 100081, China; ahsanirshad@126.com (A.I.); guohuijun@caas.cn (H.G.); zhangshunlin1994@163.com (S.Z.); gujiayu@caas.cn (J.G.); zhaolinshu@caas.cn (L.Z.); xieyongdun@caas.cn (Y.X.); xionghongchun@caas.cn (H.X.); zhaoshirong@caas.cn (S.Z.); dingyuping@caas.cn (Y.D.); mayouzhi@caas.cn (Y.M.)
* Correspondence: liuluxiang@caas.cn
† These authors contributed equally to this work.

Abstract: Wheat is a staple food commodity grown worldwide, and wheat starch is a valuable source of energy and carbon that constitutes 80% of the grain weight. Manipulation of genes involved in starch synthesis significantly affects wheat grain weight and yield. *TaSSIV* plays an important role in starch synthesis and its main function is granule formation. To mine and stack more favorable alleles, single nucleotide polymorphisms (SNPs) of *TaSSIV-A*, *B*, and *D* were investigated across 362 wheat accessions by Ecotype-Targeting Induced Local Lesions IN Genome (EcoTILLING). As a result, a total of 38 SNPs in the amplified regions of three *TaSSIV* genes were identified, of which 10, 15, and 13 were in *TaSSIV-A*, *B*, and *D*, respectively. These 38 SNPs were evaluated by using KASP and six SNPs showed an allele frequency >5% whereas the rest were <5%, i.e., considered to be minor alleles. In the Chinese mini core collection, three haplotypes were detected for *TaSSIV–A* and three for *TaSSIV–B*. The results of an association study in the Chinese mini core collection with thousand grain weight (TGW) and spike length (SPL) showed that *Hap-2-1A* was significantly associated with TGW and *Hap-3-1B* with SPL. Allelic frequency and geographic distribution indicated that the favored haplotype (*Hap-2-1A*) has been positively selected in Chinese wheat breeding. These results suggested that the Kompetitive Allele Specific PCR (KASP) markers can be applied in starch improvement to ultimately improve wheat yield by marker assisted selection in wheat breeding.

Keywords: starch; *TaSSIV*; EcoTILLING; polymorphism; KASP; haplotype

1. Introduction

Breeding for yield potential of cereal crops due to a surging population is an important priority. Wheat is a staple food commodity for the ever-growing worldwide population. Wheat is central to food security and has a large share of required calories globally. Starch is a valuable source of carbon and energy. In leaves, starch is synthesized during daytime from photosynthetically-fixed carbon and is mobilized in the night time. It is also synthesized in other transient organs such as root cap cells and meristems, but major storage organs are fruits, tubers, seeds, and roots [1]. At the physiological level, seed crop yields are mainly determined by source and sink relationships [2]. Strength of source for photoassimilates is dictated by both the photosynthetic rate and the rate of photoassimilate remobilization from source tissues. In wheat, sink capacity is more important than

source accumulation [3], hence, exploitation of enzymes involved in starch synthesis will be more effective in breeding for increased wheat yield [4].

In plants, starch is degraded at night during the process of respiration and helps the formation of sucrose. Starch consists of amylose and amylopectin. Amylopectin is the major part of starch, contributing about 75% of starch granules [5]. Different enzymes are involved in the formation of starch such as ADP-glucose pyrophosphorylases (AGPase) and starch synthases (SS) [6]. SS is controlled by different gene classes such as granule bound starch synthase (GBSS), and starch synthases I, II, III, and IV [5,7]. GBSS completely binds the starch granules and is only responsible for amylose synthesis. SSI, SSII, and SSIII are involved in amylopectin elongation [8], while SSIV controls the number of starch granules in the leaves of *Arabidopsis* [9,10]. The presence of SSIV in thylakoid membranes in *Arabidopsis* suggests that starch granules are initiated at a specific area of the chloroplast. Exon and intron regions of SSIII and SSIV are highly conserved in *Arabidopsis*, rice, and wheat but are different from SSI, SSII, and GBSS [11,12]. The *Arabidopsis* starch mutant plant (*esv1*) cannot synthesize or degrade starch at night and was dwarfed [13]. Overexpression of SSIV increases the level of starch accumulated in the leaves of *Arabidopsis* and potato, which also show higher rates of growth. Overexpression increased starch content in both photosynthetic and sink organs [14,15]. *OsSSIVa* and *OsSSIVb* are mainly expressed in endosperm and leaves, respectively. Starch granule synthesis initiation in rice endosperm does not solely depend on *OsSSIIIa* and *OsSSIVb*, because suppression of these genes did not stop the granule initiation process. These results are different from *Arabidopsis* double mutant in which suppression of these two genes totally stopped granule production [16].

In wheat, *TaSSIV* is closely related to *TaSSIII* according to phylogenetic analysis and these genes share a similar arrangement of exons and introns [7]. The chromosomal location of *TaSSIV* is on the long arm of homoeologous group 1 chromosome and is syntonic with *OsSSIVb* located on chromosome 5 [7]. The presence of *TaSSIII* and *TaSSIV* on the same chromosome but opposite arms and phylogenetic analysis indicates that the evolutionary histories of these genes are related [17]. *TaSSIV-D* mutation of wheat through EMS (ethyl methane sulphonate) showed that the chloroplasts containing 0–2 granules increased significantly while the chloroplasts containing 3–4 granules decreased [18].

EcoTILLING (Ecotype Targeting Induced Local Lesions in Genomes), is an inexpensive, robust, and reliable allele mining approach. This approach is capable of allele mining and identifying genotypes of novel functional and natural variation of known genes [19–25]. In rice, this technique has been used to characterize five genes of salt tolerance [26]. Similarly, eight genes in tomato related to sucrose synthesis were characterized [27]. In wheat, the EcoTILLLING approach has also been deployed to determine allelic diversity of genes such as kernel hardness genes, *Pin a* and *Pin b* [28]. In the future, new cultivars could be explored by EcoTILLING that contain the alleles originating from their wild relatives.

Single nucleotide polymorphisms (SNPs) have been widely used in marker assisted selection (MAS). SNPs and deletions/insertions (InDels) are nucleotide variations in natural populations that can be significantly associated with agronomic parameters [29]. Comprehensive SNP analysis of genes of particular interest are helpful for association analysis between allelic variations with phenotypic differences and the subsequent MAS of the associated traits [30]. To date, DNA polymorphisms of *TaSSIV* have not been systematically surveyed in wheat. Natural polymorphism can be used for further breeding programmes and to understand the function of *TaSSIV*. There is also a need to know the allelic variation of *TaSSIV* in *A*, *B*, and *D* genomes.

In this study, EcoTILLING was used to detect allelic variation in the targeted region of *TaSSIV-A*, *B*, and *D* genomes in 362 wheat accessions collected from diverse geographical areas of China and Pakistan to identify nucleotide diversity. The developed Kompetitive allele-specific PCR (KASP) markers were applied to screen 262 Chinese mini core collection (MCC) and 100 Pakistani wheat accessions for particular nucleotide variation. SNPs were identified with distinguishable haplotypes in the germplasm collection. In addition, association analysis between the haplotypes and phenotypic traits of the MCC was also performed to reveal marker trait associations (MTAs).

2. Materials and Methods

2.1. Plant Material and DNA Extraction

A panel of 362 wheat accessions was used for EcoTILLING. This panel mainly consisted of 262 mini core collection (MCC) accessions from China and 100 accessions originating from Pakistan (spring wheat) (Table S1). MCC represents 70% of the total wheat genetic diversity available in China [4]. MCC accessions were planted at CAAS Luoyang Experimental Station in Henan Province in 2002 and 2005. Morphological data was collected at the maturity stage, for thousand grain weight (TGW) and spike length (SPL) [31].

2.2. DNA Isolation and Molecular Analyses

Genomic DNA (gDNA) was isolated from young seedling leaves using the standard CTAB procedure [32]. Extracted gDNA was checked by using 1% agarose gel electrophoresis for quality and quantified using a NanoDrop ND-1000 spectrophotometer.

2.3. Primers Development

The target gene regions with maximum probability for missense variants were predicted using the CODDLE bioinformatics pipeline (http://blocks.fhcrc.org/proweb/). Genomic sequences of the three *TaSSIV* genes were obtained from the GenBank database (http://www.wheatgenome.org/). Multiple alignment of nucleotide sequences was conducted by Bioedit 7.0. Primers were designed using Primer Premier 5.0 (Premier Biosoft International, Palo Alto, CA). Genome-specific primers for *TaSSIV-A* and *TaSSIV-B* were designed on the basis of SNPs and/or InDels; primers were already available for *TaSSIV-D* [18]. Specificity of primers for target genes was tested by PCR amplification in Chinese Spring (CS), Jing411, and nulli-tetrasomic lines of CS. Detailed information for primers is given in Tables S2, S3 and Figure S1.

2.4. EcoTILLING

For EcoTILLING, each accession and Chinese Spring (WT) were pooled in a 1:1 ratio after diluting the samples. PCRs were performed in 10 μL final volumes on 100 ng μL^{-1} DNA using 0.4 U/reaction of Taq polymerase (TaKaRa Ex TaqTM). The mixture contained a 4:1 ratio of IRD800-labeled to unlabeled reverse primers and 3:2 ratio of IRD700-labeled to unlabeled forward primers. The EcoTILLING method was followed as described in [33]. After the denaturation step, each sample was loaded on the 6.5% polyacrylamide gel with a 96 well comb. The gel images were analyzed using Gel buddy software and polymorphisms were identified.

2.5. Sequencing of Variants

Resultant PCR products were directly sequenced in both directions by Sangon Biotech Co., Ltd., Beijing, China, to confirm the sequence variations. Identified sequence variants were analyzed by the PARSESNP (http://blocks.fhcrc.org/proweb/) and SIFT (Sorting Intolerant from Tolerant) (http://sift.jcvi.org/) programs with default parameters [34].

2.6. Development of Functional Markers (FMs) Using KASP

For high-throughput genotyping, KASP primers were developed on the variants of *TaSSIV-A*, *-B*, and *-D* by following standard KASP guidelines (http://www.lgcgenomics.com). The allele-specific primers were developed using the standard FAM (5′ GAAGGTGACCAAGTTCATGCT 3′) and HEX (5′ GAAGGT CGGAGTCAACGGATT 3′) tails with a targeted SNP at the 3′ end. KASP markers were then applied across the entire population. The KASP markers were developed by following standard KASP guidelines [35]. A scatter plot for KASP assays showed clustering of accessions on the *X-(FAM)* and *Y-(HEX)* signals. Detailed information for primers is given in Tables S2, S3 and Figure S1.

2.7. Association between SNPs and Agronomic Traits

Descriptive statistics and estimates of variance were conducted using Microsoft Excel 2013. The effect of haplotypes on agronomic traits were analyzed using Student's *t*-tests at $p < 0.05$ (even 0.01).

3. Results

3.1. Natural Variation of TaSSIV in Accessions

Gene structure analysis showed that *TaSSIV-A* consisted of 13 exons and 12 introns while *TaSSIV-B*, and *TaSSIV-D* each consisted of 16 exons and 15 introns. Each gene was exploited in the wheat population for sequence polymorphism by EcoTILLING. A total of 38 putative natural variation sites were identified across the population (Table 1 and Figure 1). Twelve sequence polymorphic sites were identified in the coding region, while 26 were identified in the non-coding region. Out of 38 DNA polymorphisms, 31 were SNPs while seven were InDels. Among the identified sequence polymorphisms, two missense polymorphisms were identified in *TaSSIV-A* and *TaSSIV-B* each, while *TaSSIV-D* possessed three missense polymorphisms. *TaSSIV-B* and *TaSSIV-D* also possessed one and three silent polymorphisms, respectively.

Table 1. List of nucleotide polymorphism in *TaSSIV* with their effects on codons and variation in amino acid.

No.	Gene	Genotype	Allele	Mutation Type	Variation in Amino Acid
1	*TaSSIV-A*	Laoqiaomai	A1447G		Non-coding
2		Sankecun	T1455G		Non-coding
3		Wenmai-6	C1563T		Non-coding
4		Yangmai-158	A1673T		Non-coding
5		Yangmai-158	C2436T	Missense	R77C
6		Guinong-10	T2073G		Non-coding
6		Zhongzhou-741	A2403C	Missense	I66L
8		Mahuaban	G5081A		Intron
9		Huoliyan	C5952T		Intron
10		Huangshuibai	T6013:		Intron
11		Guinong-11	G2143A		Non-coding
12	*TaSSIV-B*	Dingxingzhai	C1529T	Missense	S149L
13		Pingyuan-50	C1560T	Silent	F159=
14		Dabaimai	A2234G	Missense	K384R
15		Lumai-1	T3314G		Intron
16		Lumai-1	T3325G		Intron
17		Xinmai-18	A3478ATT		Intron
18		Baipixiaomai	C6107T	Silent	I822=
19		Xiaobaimai	G6329A		Intron
20		Shahkar-95	A2950:		Intron
21		Drawar-97	A2991G		Intron
22		Blue silver	C3457T		Intron
23		Lasani	G4115CTT		Intron
24		Kohistan	C4176T		Intron
25		Marvi	T3268G		Intron
26	*TaSSIV-D*	Zijuhong	T5904:		Intron
27		Hongchumai	A6168C		Intron
28		Hongmai	G6485:		Intron
29		Hongmai	T6440C		Intron
30		Kabka-3	A3387:		Intron
31		Chumai	C3593G	Silent	G454=
32		Yangmai	C3899G	Silent	V505=
33		Hongpidongmai	T3981A	Missense	S533T
34		FSD-08	C5849T		Intron
35		BWP-6309	G6164T		Intron
36		Manthar	G6070A	Missense	V800I
37		Kohistan-97	A6540G	Missense	T841A
38		Sehar	T4016C	Silent	Y533=

Note: only one accession was listed as a representative due to limited space. ":" Deletion; "=" No change in amino acid.
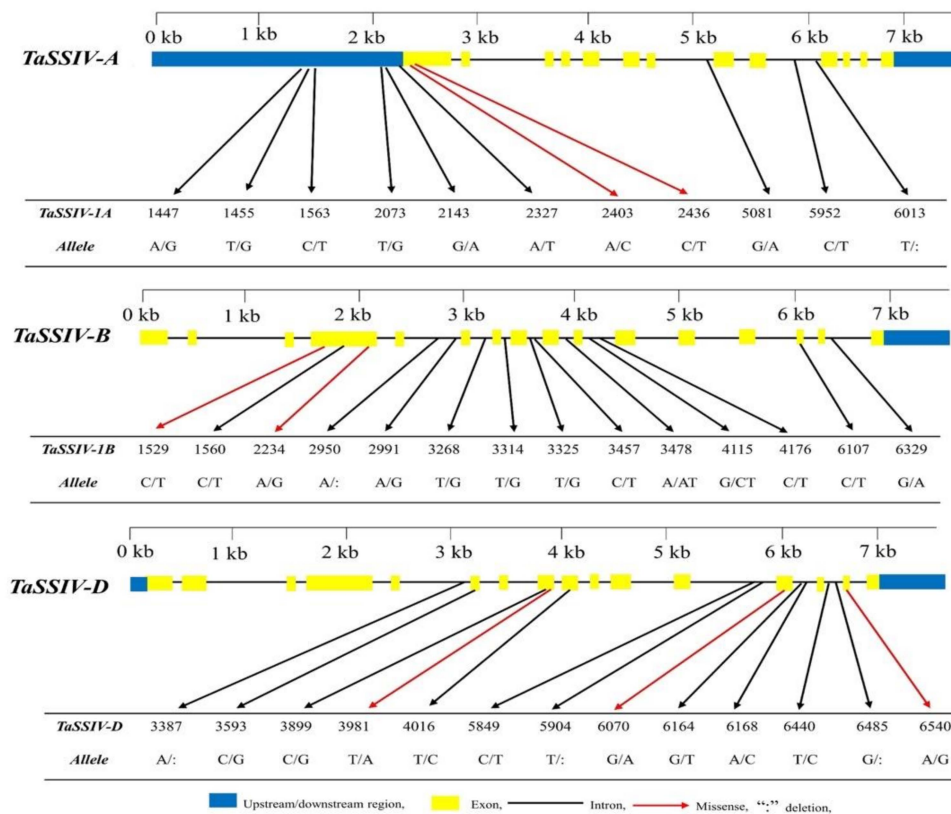
Figure 1. Location of single nucleotide polymorphisms (SNPs) in *TaSSIV-A*, *TaSSIV-B* and *TaSSIV-D*.

The SIFT computer program predicted that only one sequence polymorphism in *TaSSIV-D* led to an amino acid change (V → I at 800 aa), which was located in the GT-5 domain and may affect the protein function. All other missense polymorphisms of *TaSSIV-A*, *B*, and *D* were present in coiled regions of protein structure and not in functional domains.

3.2. KASP Marker Development

KASP markers were developed using identified sequence polymorphisms. Among the KASP markers, four markers (*KASP-A1673T*, *KASP-C5952T*, *KASP-A2403C*, and *KASP-C2436T*) in *TaSSIV-A* and two markers (*KASP-C1560T* and *KASP-C6107T*) in *TaSSIV-B* had allele frequencies over 5%. These results were consistent with EcoTILLING. All SNPs of *TaSSIV-D* had allelic frequencies less than 5% and were therefore excluded from further analysis. For KASP-*A11673T, KASP-C5952T, KASP-A2403C*, and *KASP-C2436T* (Figure 2a–d) accessions in blue circles have *A*, *C*, *A*, and *C* alleles while accessions in red circles have *T*, *T*, *C*, and *T* alleles, respectively, for *TaSSIV-A*. Similarly, for *KASP-C1560T* and *KASP-C6107T* (Figure 2e,f) accessions in blue circles have *C* allele whereas accessions in red circles have *T* allele in *TaSSIV-B* gene (Figure 2).

3.3. Association Between Haplotypes and Yield-Related Traits

There was a total of 14 haplotypes in *TaSSIV-A*. The frequency of each haplotype in MCC was different. Frequencies of the three haplotypes of *TaSSIV-A*, i.e., *Hap-1-1A (AATC)*, *Hap-2-1A (ACTT)*, and *Hap-3-1A* (*ACTC*) were more than 5%. The other eleven haplotypes were less than 5%, marking them as rare haplotypes. Three haplotypes of *TaSSIV-B* i.e., *Hap-1-1B (CC)*, *Hap-2-1B (TC)*, and *Hap-3-1B (CT)* in the wheat population had frequencies more than 5% (Table 2). Association analysis showed that *TaSSIV-A* was associated with TGW; *Hap-2-1A* showed significant difference from the other two haplotypes and had higher TGW. *Hap-3-1B* was the favored haplotype for SPL and was significantly different from *Hap-1-1B* and *Hap-2-1B* (Figure 3). Therefore, these two haplotypes were significantly

associated with yield-related traits, which may have potential as functional markers (FMs) to be used in MAS for starch content breeding.
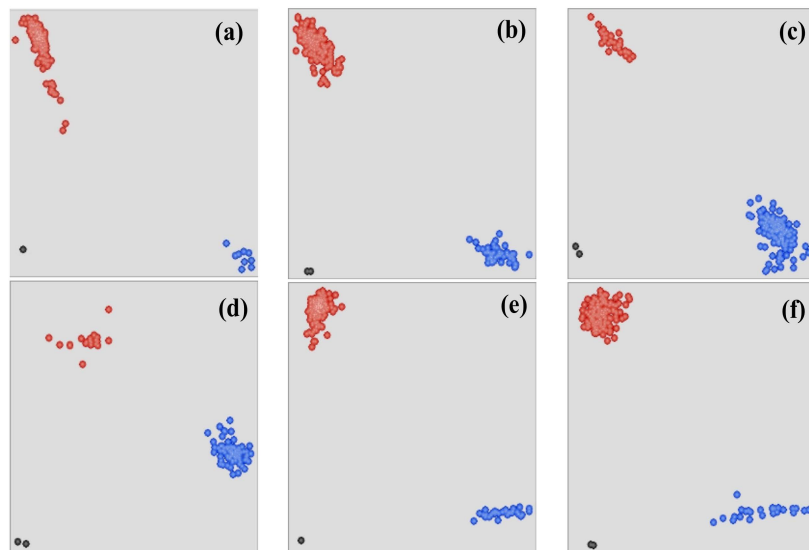

Figure 2. Scatter plot for KASP assays; showing clustering of accessions on the X-(FAM) and Y-(HEX) axes. Accessions colored blue have the FAM- type allele; accessions colored red have the HEX-type allele; black dots represent the NTC (non-template control). (**a**–**d**): *TaSSIV-A*. (**a**) *KASP-A1673T*, (**b**) *KASP-A2403C*, (**c**) *KASP-C2436T*, (**d**) *KASP-C5952T*, (**e**,**f**) *TaSSIV-B*. (**e**) *KASP-C1560T*, (**f**) *KASP-C6107T*.

Table 2. Nucleotide polymorphisms and functional marker development of *TaSSIV-A* and *TaSSIV-B*.

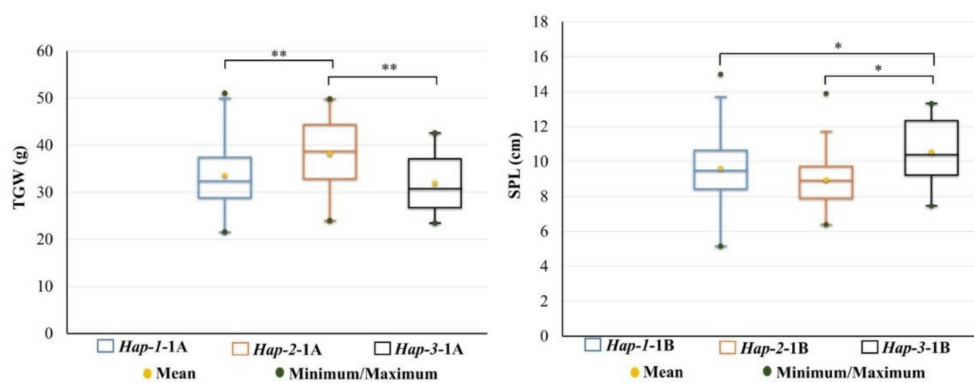| Gene	Haplotype	Position of Nucleotides				Gene	Haplotype	Position of Nucleotides	
TaSSIV-A		1673 nt	2403 nt	2436 nt	5952 nt	*TaSSIV-B*		1560 nt	6107 nt
	Hap-1-1A	A	A	T	C		*Hap-1-1B*	C	C
	Hap-2-1A	A	C	T	T		*Hap-2-1B*	T	C
	Hap-3-1A	A	C	T	C		*Hap-3-1B*	C	T

Figure 3. Favorable haplotypes and their interaction with phenotype. TGW, thousand grain weight; SPL, spike length; * $p < 0.05$, ** $p < 0.01$; Two-year data of 2002 and 2005.

3.4. Geographic Distribution of TaSSIV-A and TaSSIV-B Haplotypes

To determine, whether the favored haplotypes for *TaSSIV-A and TaSSIV-B* were selected in wheat breeding, we investigated the geographic distribution of *TaSSIV-A* and *B* in MCC. On the basis of agro-climatic conditions, China has 10 agro-ecological regions, I to X [36].

Regions I–IV are the major wheat producing regions based on production and cultivation area. In landraces, the frequency of the favored haplotype *Hap-2-1A* for TGW was low and *Hap-1-1A* was dominant in all major areas. In modern cultivars, the favored haplotype *Hap-2-1A* frequency was high (more than 50%) in regions I–IV (Figure 4). The frequency of *Hap-2-1A* was significantly increased from 25% to 50% in region I, 20% to 48% in region II, 11.67% to 33.37% in region III, and 8% to 85.71% in region IV from landraces to modern cultivars, respectively. The results showed that favored haplotypes experienced positive selection in Chinese wheat breeding programs. Similarly, in *TaSSIV-B* haplotypes, the favored haplotype *Hap-3-1B* was significantly different from *Hap-1-1B* and *Hap-2-1B* for SPL but its frequency was balanced or selectively neutral in landraces and modern wheat cultivars. Therefore, the positive alleles of these haplotypes have the potential to be used in wheat breeding to increase in the grain yield.
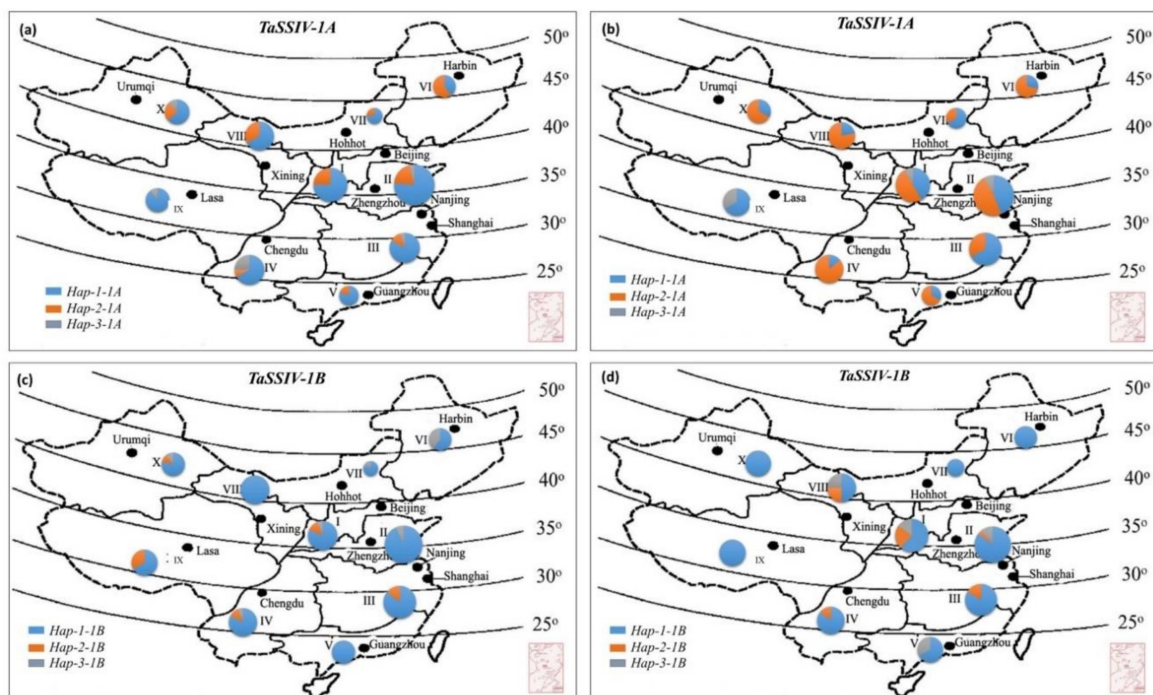

Figure 4. Geographic distribution of *TaSSIV-A and TaSSIV-B* Haplotypes in China. (**a**,**c**) Landraces, (**b**,**d**) modern cultivars. I, Northern winter wheat region; II, Yellow and Huai River valleys winter wheat region; III, Middle and low Yangtze valleys winter wheat region; IV, Southwestern winter wheat region; V, Southern winter wheat region; VI, Northeastern spring wheat region; VII, Northern spring wheat region; VIII, Northwestern spring wheat region; IX, Qinghai-Tibetan spring-winter wheat region; X, Xinjiang winter-spring wheat region. Pie chart size is directly proportional to number of genotypes.

In Pakistani wheat accessions, geographic distribution was investigated in *TaSSIV-A* haplotypes. The frequency of *Hap-2-1A* among five different regions of Pakistan was higher in the Punjab irrigated region (78%) and the Punjab rainfed region (52%). The frequency of this haplotype in Khyber Pakhtonkhwa was 35%. (Figure 5).

3.5. Positive Selection of Hap-2-1A in China's Wheat Breeding Process

On the basis of released time, the wheat population was grouped into six sub-groups, i.e., accessions released before 1951, 1951–1960, 1961–1970, 1971–1980, 1981–1990, and 1991–2001. The frequency of *Hap*-2-1A continuously increased, suggesting progressive selection (Figure 6).
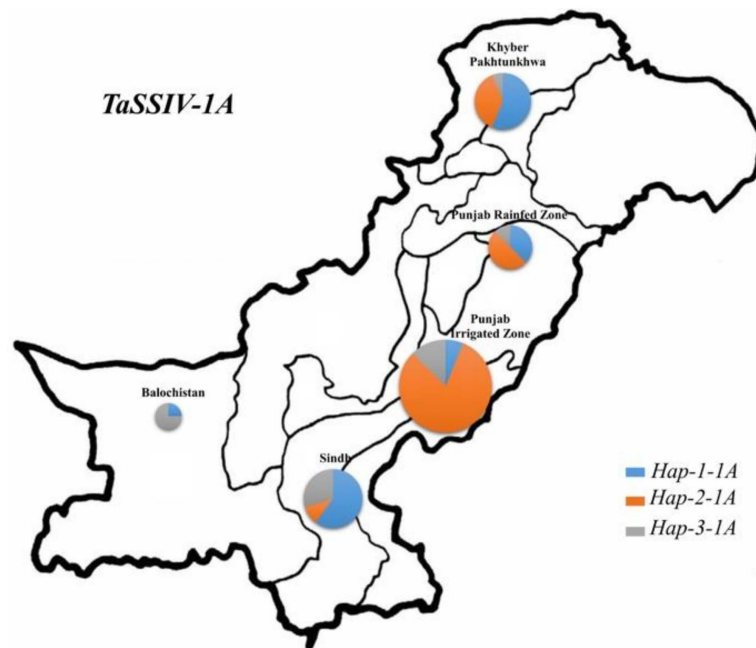
Figure 5. *TaSSIV-A* haplotypes distribution among Pakistani wheat genotypes. I, Punjab rainfed region; II, Punjab irrigated region; III, Sindh region; IV, Balochistan region; V, Khyber Pakhtunkhwa region. Pie chart size is directly proportional to number of genotypes.
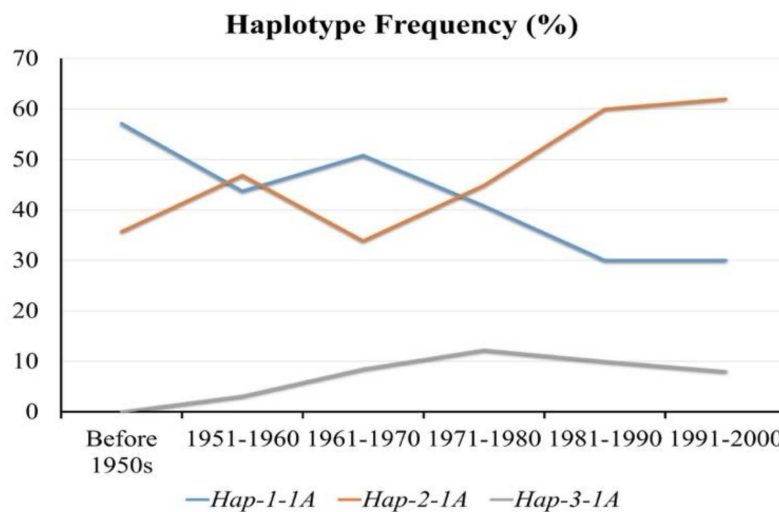

Figure 6. Favored haplotypes of *TaSSIV-A* selected over the history of wheat breeding in China.

4. Discussion

Knowledge-based genetic improvement of crop plants is crucial for global food security in the future [21]. Molecular plant breeding uses phenotypic and genetic variations with contemporary tools to develop new varieties with higher productivity and that are resilient to extreme environments [37].

The EcoTILLING approach has been used to discover natural genetic variation in different crop plants such as *Arabidopsis*, wheat, barley, *Brassica* sp., and rice [20,38,39]. The *SSIV* gene has been characterized in *Arabidopsis*, rice, and wheat; however, its allelic variations were not characterized systematically. Our primary goal was to identify natural genetic variation of *TaSSIV* in wheat and its association with traits contributing to yield. EcoTILLING successfully identified 38 SNPs across three homeologous *TaSSIV* genes in modern cultivars and landraces from China and Pakistan.

Development of high-throughput molecular markers is of utmost importance for use in wheat molecular breeding [40]. Gel-free KASP assays could significantly improve the speed and efficiency of selection in wheat breeding programs. Although EcoTILLING successfully identified all the allelic variation in *TaSSIV* homeologous genes, it is a laborious and costly method for routine genotyping. To overcome this problem, KASP markers were developed for several SNPs for high-throughput and cost-effective genotyping of these variants. The six KASP assays developed in our study can be routinely used for identification of SNPs in *TaSSIV* in other germplasm and breeding populations with extremely high accuracy [41].

The allelic variation and KASP markers developed to distinguish *TaSSIV* haplotypes can be instrumental in MAS, which can be utilized in combination with other FMs for TGW and SPL breeding. Wheat A and B sub-genomes have a relatively broader genetic base compared to the D sub-genome. Additionally, non-coding regions also have relatively higher levels of polymorphism compared to coding regions [42]. The absence of significant polymorphism in *TaSSIV-D* is probably due to the narrow genetic background of the D sub-genome or allele fixation due to an evolution and domestication bottleneck.

A haplotype block combining two or more SNPs in strong linkage disequilibrium are more explanatory than bi-allelic SNPs [43]. Beyond bi-allelic SNP variations, haplotype data can capture associations that evade identification by single SNPs [44]. Haplotype-based analyses are still rare in wheat with few exceptions [45,46]. Identification of haplotypes can also capture epistatic interactions between SNPs. Hence haplotype-based approach could increase prediction accuracies [47].

The starch synthase family in plants encodes many starch synthesis genes that play essential roles in starch synthesis and ultimately increase yield [48]. There is a strong relationship between starch and grain yield in wheat and maize. Association analysis is a powerful approach to investigate marker trait associations (MTAs) [49]. In the present study, two haplotypes of *TaSSIV* (*Hap-2-1A* in *TaSSIV-A* and *Hap-3-1B* in *TASSIV-B*) were significantly ($p < 0.05$) associated with yield-related traits in wheat.

It has been demonstrated that grain yield in France [50], Italy [51], the UK, and China [52] has largely been attained by improvements in grain number per square meter with significant change in individual grain weight [53,54]. Accessions possessing haplotype *Hap-2-1A* had higher TGW while accessions having *Hap-3-1B* had larger SPL. Yield-related traits of cereal crops are governed by multiple genes and are influenced by environmental factors [55]. Among the reported gene determinants of cereal grain traits, some genes act constitutively under different environments, while others function in specific conditions [56]. In this study, yield-related traits showed significant association with the identified haplotypes, which suggests that use of these haplotypes can be instrumental to improve yield of the wheat crop.

In China, the geographical distribution of two haplotypes is also supportive. Varieties bred in regions I to IV usually have bigger grain size and larger spikes, but fewer tillers to evade canopy structures favoring rust epidemics [57]. Region II covers ~40% of the total national wheat area and accounts for ~45% of total wheat production in China [36]. Average TGW of varieties in this region is 42–43 g. From landraces to modern cultivars, positive selection of haplotype *Hap-2-1A* favored wheat breeding. Regions I, II, and III account for about 64.8% of total national wheat area and greater turnover of wheat cultivars was also reported from these regions [35]. In China, wheat yield increase has largely depended on higher TGW [55]. There was a significant association between haplotype *TaSSIV-1A*-2 and TGW in MCC (Figure 3). In Northern China, 1.3% genetic gain for TGW has been achieved yearly for three decades with concurrent increases in grains per spike in Yellow and Huai River Valleys winter wheat region [52]. Improvement in TGW and grain weight per spike has also been reported in the Southern China winter wheat region since 1949 [52].

Choice of populations plays a critical role to detect MTAs via association analysis. Accessions from Pakistan were also used to evaluate the *TaSSIV-A* haplotypes in five different geographical regions. In these accessions, the favored haplotype (*Hap-2-1A*) was selected in major wheat growing regions of Pakistan especially in Punjab regions (rainfed region and irrigated region) (Figure 5). In both

populations (China and Pakistan) the frequency of the favored haplotype (*Hap-2-1A*) was reasonably high, indicating the positive selection during modern wheat breeding. Although both populations have different population structures, unconscious selection of the favored haplotype is likely due to the high linkage disequilibrium of major yield-related genes selected during selection breeding. However, the frequency analysis in wheat cultivars from other regions like Europe, USA, and Australia would be required to draw the conclusion that selection of favored haplotypes of *TaSSIV* is likely due to the selection of major yield-related genes [35].

In conclusion, we have demonstrated that EcoTILLING is an efficient approach for allele mining of wheat candidate genes and KASP is a robust approach to validate SNPs for further analysis. This study identified 38 unique alleles in *TaSSIV* among 362 wheat accessions collected. Association analysis showed that two haplotypes in *TaSSIV* were significantly associated with TGW and SPL and these haplotypes can be used in future studies to assess their usefulness as selection criteria for improving these yield-related traits.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4425/10/4/307/s1, Table S1: List of modern and land races accessions of China (MCC) and Pakistan, Table S2: List of primers used for EcoTILLING of TaSSIV A, B and D sub genomes, Table S3: KASP markers that are tightly linked to TaSSIV-A and TaSSIV-B and their primer sequences, Figure S1: Specific primer for TaSSIV-A and TaSSIV-B with validation Chinese Spring, nullisomic, tetrasomic lines.

References

1. Martin, C.; Smith, A.M. Starch biosynthesis. *Plant Cell* **1995**, *7*, 971. [CrossRef] [PubMed]

2. Rossi, A.; Kontarakis, Z.; Gerri, C.; Nolte, H.; Hölper, S.; Krüger, M.; Stainier, D.Y. Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature* **2015**, *524*, 230. [CrossRef] [PubMed]

3. Serrago, R.A.; Alzueta, I.; Savin, R.; Slafer, G.A. Understanding grain yield responses to source–sink ratios during grain filling in wheat and barley under contrasting environments. *Field Crops Res.* **2013**, *150*, 42–51. [CrossRef]

4. Hou, J.; Jiang, Q.; Hao, C.; Wang, Y.; Zhang, H.; Zhang, X. Global selection on sucrose synthase haplotypes during a century of wheat breeding. *Plant Physiol.* **2014**, *164*, 1918–1929. [CrossRef] [PubMed]

5. Ball, S.G.; Morell, M.K. From bacterial glycogen to starch: Understanding the biogenesis of the plant starch granule. *Annu. Rev. Plant Biol.* **2003**, *54*, 207–233. [CrossRef]

6. Smith, A.M.; Zeeman, S.C.; Smith, S.M. Starch degradation. *Annu. Rev. Plant Biol.* **2005**, *56*, 73–98. [CrossRef] [PubMed]

7. Leterrier, M.; Holappa, L.D.; Broglie, K.E.; Beckles, D.M. Cloning, characterisation and comparative analysis of a starch synthase IV gene in wheat: Functional and evolutionary implications. *BMC Plant Biol.* **2008**, *8*, 98. [CrossRef] [PubMed]

8. Wang, S.; Wang, J.; Zhang, W.; Li, C.; Yu, J.; Wang, S. Molecular order and functional properties of starches from three waxy wheat varieties grown in China. *Food Chem.* **2015**, *181*, 43–50. [CrossRef]

9. Roldán, I.; Wattebled, F.; Mercedes Lucas, M.; Delvallé, D.; Planchot, V.; Jiménez, S.; Pérez, R.; Ball, S.; D'hulst, C.; Mérida, Á. The phenotype of soluble starch synthase IV defective mutants of *Arabidopsis thaliana* suggests a novel function of elongation enzymes in the control of starch granule formation. *Plant J.* **2007**, *49*, 492–504. [CrossRef]

10. Szydlowski, N.; Ragel, P.; Raynaud, S.; Lucas, M.M.; Roldán, I.; Montero, M.; Muñoz, F.J.; Ovecka, M.; Bahaji, A.; Planchot, V. Starch granule initiation in *Arabidopsis* requires the presence of either class IV or class III starch synthases. *Plant Cell* **2009**, *21*, 2443–2457. [CrossRef]

11. Dian, W.; Jiang, H.; Chen, Q.; Liu, F.; Wu, P. Cloning and characterization of the granule-bound starch synthase II gene in rice: Gene expression is regulated by the nitrogen level, sugar and circadian rhythm. *Planta* **2003**, *218*, 261–268. [CrossRef]

12. Jiang, H.; Dian, W.; Liu, F.; Wu, P. Molecular cloning and expression analysis of three genes encoding starch synthase II in rice. *Planta* **2004**, *218*, 1062–1070. [CrossRef]

13. Smith, A.M.; Stitt, M. Coordination of carbon supply and plant growth. *Plant Cell Environ.* **2007**, *30*, 1126–1149. [CrossRef]

14. Gámez-Arjona, F.M.; Li, J.; Raynaud, S.; Baroja-Fernández, E.; Muñoz, F.J.; Ovecka, M.; Ragel, P.; Bahaji, A.; Pozueta-Romero, J.; Mérida, Á. Enhancing the expression of starch synthase class IV results in increased levels of both transitory and long-term storage starch. *Plant Biotechnol. J.* **2011**, *9*, 1049–1060. [CrossRef] [PubMed]

15. Schwarte, S.; Brust, H.; Steup, M.; Tiedemann, R. Intraspecific sequence variation and differential expression in starch synthase genes of *Arabidopsis thaliana*. *BMC Res. Notes* **2013**, *6*, 84. [CrossRef] [PubMed]

16. Toyosawa, Y.; Kawagoe, Y.; Matsushima, R.; Crofts, N.; Ogawa, M.; Fukuda, M.; Kumamaru, T.; Okazaki, Y.; Kusano, M.; Saito, K. Deficiency of starch synthase *IIIa* and *IVb* alters starch granule morphology from polyhedral to spherical in rice endosperm. *Plant Physiol.* **2016**, *170*, 1255–1270. [CrossRef]

17. Li, Z.; Mouille, G.; Kosar-Hashemi, B.; Rahman, S.; Clarke, B.; Gale, K.R.; Appels, R.; Morell, M.K. The structure and expression of the wheat starch synthase III gene. Motifs in the expressed gene define the lineage of the starch synthase III gene family. *Plant Physiol.* **2000**, *123*, 613–624. [CrossRef] [PubMed]

18. Guo, H.; Liu, Y.; Li, X.; Yan, Z.; Xie, Y.; Xiong, H.; Zhao, L.; Gu, J.; Zhao, S.; Liu, L. Novel mutant alleles of the starch synthesis gene *TaSSIVb-D* result in the reduction of starch granule number per chloroplast in wheat. *BMC Genom.* **2017**, *18*, 358. [CrossRef]

19. Mccallum, C.M.; Comai, L.; Greene, E.A.; Henikoff, S. Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. *Plant Physiol.* **2000**, *123*, 439. [CrossRef] [PubMed]

20. Comai, L.; Young, K.; Till, B.J.; Reynolds, S.H.; Greene, E.A.; Codomo, C.A.; Enns, L.C.; Johnson, J.E.; Burtner, C.; Odden, A.R. Efficient discovery of DNA polymorphisms in natural populations by EcoTILLING. *Plant J.* **2004**, *37*, 778–786. [CrossRef] [PubMed]

21. Till, B.J.; Comai, L.; Henikoff, S.; Zerr, T. A protocol for TILLING and EcoTILLING in plants and animals. *Nat. Protoc.* **2006**, *1*, 2465. [CrossRef]

22. Till, B.J.; Cooper, J.; Tai, T.H.; Colowit, P.; Greene, E.A.; Henikoff, S.; Comai, L. Discovery of chemically induced mutations in rice by TILLING. *BMC Plant Biol.* **2007**, *7*, 19. [CrossRef] [PubMed]

23. Raghavan, C.; Naredo, M.E.B.; Wang, H.; Atienza, G.; Liu, B.; Qiu, F.; McNally, K.L.; Leung, H. Rapid method for detecting SNPs on agarose gels and its application in candidate gene mapping. *Mol. Breed.* **2006**, *19*, 87–101. [CrossRef]

24. Wang, T.L.; Uauy, C.; Robson, F.; Till, B. TILLING in extremis. *Plant Biotech. J.* **2012**, *10*, 761–772. [CrossRef] [PubMed]

25. Xia, Y.; Ning, Z.; Bai, G.; Li, R.; Yan, G.; Siddique, K.H.; Baum, M.; Guo, P. Allelic variations of a light harvesting chlorophyll a/b-binding protein gene (*Lhcb1*) associated with agronomic traits in barley. *PLoS ONE* **2012**, *7*, e37573. [CrossRef] [PubMed]

26. Negrao, S.; Almadanim, M.C.; Pires, I.S.; Abreu, I.A.; Maroco, J.; Courtois, B.; Gregorio, G.B.; McNally, K.L.; Oliveira, M.M. New allelic variants found in key rice salt-tolerance genes: An association study. *Plant Biotechnol. J.* **2013**, *11*, 87–100. [CrossRef]

27. Zeng, Y.D.; Sun, J.L.; Bu, S.H.; Deng, K.S.; Tao, T.; Zhang, Y.M.; Zhang, T.Z.; Du, X.M.; Zhou, B.L. EcoTILLING revealed SNPs in *GhSus* genes that are associated with fiber- and seed-related traits in upland cotton. *Sci. Rep.* **2016**, *6*, 29250. [CrossRef] [PubMed]

28. Ma, X.; Sajjad, M.; Wang, J.; Yang, W.; Sun, J.; Li, X.; Zhang, A.; Liu, D. Diversity, distribution of Puroindoline genes and their effect on kernel hardness in a diverse panel of Chinese wheat germplasm. *BMC Plant Biol.* **2017**, *17*, 158. [CrossRef] [PubMed]

29. Russell, J.; Mascher, M.; Dawson, I.K.; Kyriakidis, S.; Calixto, C.; Freund, F.; Bayer, M.; Milne, I.; Marshall-Griffiths, T.; Heinen, S. Exome sequencing of geographically diverse barley landraces and wild relatives gives insights into environmental adaptation. *Nat. Genet.* **2016**, *48*, 1024. [CrossRef] [PubMed]

30. Xu, Y.; Li, P.; Zou, C.; Lu, Y.; Xie, C.; Zhang, X.; Prasanna, B.M.; Olsen, M.S. Enhancing genetic gain in the era of molecular breeding. *J. Exp. Bot.* **2017**, *68*, 2641–2666. [CrossRef]

31. Wang, L.; Ge, H.; Hao, C.; Dong, Y.; Zhang, X. Identifying loci influencing 1,000-kernel weight in wheat by microsatellite screening for evidence of selection during breeding. *PLoS ONE* **2012**, *7*, e29432. [CrossRef]

32. Murray, M.G.; Thompson, W.F. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* **1980**, *8*, 4321–4325. [CrossRef] [PubMed]

33. Wang, J.; Sun, J.; Liu, D.; Yang, W.; Wang, D.; Tong, Y.; Zhang, A. Analysis of *Pina* and *Pinb* alleles in the micro-core collections of Chinese wheat germplasm by EcoTILLING and identification of a novel *Pinb* allele. *J. Cereal Sci.* **2008**, *48*, 836–842. [CrossRef]

34. Ng, P.C.; Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **2003**, *31*, 3812–3814. [CrossRef]

35. Rehman, S.U.; Wang, J.; Chang, X.; Zhang, X.; Mao, X.; Jing, R. A wheat protein kinase gene *TaSnRK2. 9-5A* associated with yield contributing traits. *Theor. Appl. Genet.* **2019**, *132*, 907–919. [CrossRef]

36. He, Z. *A History of Wheat Breeding in China*; Cimmyt: Texcoco, Edo Mex, Mexico, 2001.

37. Moose, S.P.; Mumm, R.H. Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiol.* **2008**, *147*, 969–977. [CrossRef] [PubMed]

38. Wang, N.; Shi, L.; Tian, F.; Ning, H.; Wu, X.; Long, Y.; Meng, J. Assessment of *FAE1* polymorphisms in three Brassica species using EcoTILLING and their association with differences in seed erucic acid contents. *BMC Plant Biol.* **2010**, *10*, 137. [CrossRef] [PubMed]

39. Mejlhede, N.; Kyjovska, Z.; Backes, G.; Burhenne, K.; Rasmussen, S.K.; Jahoor, A. EcoTILLING for the identification of allelic variation in the powdery mildew resistance genes *mlo* and *Mla* of barley. *Plant Breed.* **2006**, *125*, 461–467. [CrossRef]

40. Semagn, K.; Babu, R.; Hearne, S.; Olsen, M. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): Overview of the technology and its application in crop improvement. *Mol. Breed.* **2014**, *33*, 1–14. [CrossRef]

41. Rasheed, A.; Wen, W.; Gao, F.; Zhai, S.; Jin, H.; Liu, J.; Guo, Q.; Zhang, Y.; Dreisigacker, S.; Xia, X. Development and validation of KASP assays for genes underpinning key economic traits in bread wheat. *Theor. Appl. Genet.* **2016**, *129*, 1843–1860. [CrossRef]

42. Nasu, S.; Suzuki, J.; Ohta, R.; Hasegawa, K.; Yui, R.; Kitazawa, N.; Monna, L.; Minobe, Y. Search for and analysis of single nucleotide polymorphisms (SNPs) in rice (*Oryza sativa*, *Oryza rufipogon*) and establishment of SNP markers. *DNA Res.* **2002**, *9*, 163–171. [CrossRef] [PubMed]

43. Stephens, J.C.; Schneider, J.A.; Tanguay, D.A.; Choi, J.; Acharya, T.; Stanley, S.E.; Jiang, R.; Messer, C.J.; Chew, A.; Han, J.H. Haplotype variation and linkage disequilibrium in 313 human genes. *Science* **2001**, *293*, 489–493. [CrossRef] [PubMed]

44. Lorenz, A.J.; Hamblin, M.T.; Jean-Luc, J. Performance of single nucleotide polymorphisms versus haplotypes for genome-wide association analysis in barley. *PLoS ONE* **2010**, *5*, e14079. [CrossRef]

45. Li, B.; Li, Q.; Mao, X.; Li, A.; Wang, J.; Chang, X.; Hao, C.; Zhang, X.; Jing, R. Two Novel *AP2/EREBP* Transcription Factor Genes*TaPARG*Have Pleiotropic Functions on Plant Architecture and Yield-Related Traits in Common Wheat. *Front. Plant Sci.* **2016**, *7*, 1191.

46. Miao, L.; Mao, X.; Wang, J.; Liu, Z.; Zhang, B.; Li, W.; Chang, X.; Reynolds, M.; Wang, Z.; Jing, R. Elite Haplotypes of a Protein Kinase *GeneTaSnRK2.3* Associated with Important Agronomic Traits in Common Wheat. *Front. Plant Sci.* **2017**, *8*, 341. [CrossRef] [PubMed]

47. Bevan, M.W.; Uauy, C.; Wulff, B.B.; Zhou, J.; Krasileva, K.; Clark, M.D. Genomic innovation for crop improvement. *Nature* **2017**, *543*, 346. [CrossRef] [PubMed]

48. Parry, M.A.; Reynolds, M.; Salvucci, M.E.; Raines, C.; Andralojc, P.J.; Zhu, X.-G.; Price, G.D.; Condon, A.G.; Furbank, R.T. Raising yield potential of wheat. II. Increasing photosynthetic capacity and efficiency. *J. Exp. Bot.* **2010**, *62*, 453–467. [CrossRef] [PubMed]

49. Zhu, C.; Gore, M.; Buckler, E.S.; Yu, J. Status and prospects of association mapping in plants. *Plant Genome* **2008**, *1*, 5–20. [CrossRef]

50. Brancourt-Hulmel, M.; Doussinault, G.; Lecomte, C.; Berard, P.; Le Buanec, B.; Trottet, M. Genetic improvement of agronomic traits of winter wheat cultivars released in France from 1946 to 1992. *Crop Sci.* **2003**, *43*, 37–45. [CrossRef]

51. Canevara, M.; Romani, M.; Corbellini, M.; Perenzin, M.; Borghi, B. Evolutionary trends in morphological, physiological, agronomical and qualitative traits of *Triticum aestivum* L. cultivars bred in Italy since 1900. *Euro. J. Agron.* **1994**, *3*, 175–185. [CrossRef]

52. Zhou, Y.; Zhu, H.; Cai, S.; He, Z.; Zhang, X.; Xia, X.; Zhang, G. Genetic improvement of grain yield and associated traits in the southern China winter wheat region: 1949 to 2000. *Euphytica* **2007**, *157*, 465–473. [CrossRef]

53. Gaju, O.; Reynolds, M.; Sparkes, D.; Foulkes, M. Relationships between large-spike phenotype, grain number, and yield potential in spring wheat. *Crop Sci.* **2009**, *49*, 961–973. [CrossRef]

54. Hanif, M.; Gao, F.; Liu, J.; Wen, W.; Zhang, Y.; Rasheed, A.; Cao, S. TaTGW6-A1 an ortholog of rice TGW6 is associated with the grain weight and yield in bread wheat. *Mol. Breed.* **2016**, *36*, 1–8. [CrossRef]

55. Zhang, K.; Wang, J.; Zhang, L.; Rong, C.; Zhao, F.; Peng, T.; Li, H.; Cheng, D.; Liu, X.; Qin, H. Association analysis of genomic loci important for grain weight control in elite common wheat varieties cultivated with variable water and fertiliser supply. *PLoS ONE* **2013**, *8*, e57853. [CrossRef] [PubMed]

56. Dong, L.; Wang, F.; Liu, T.; Dong, Z.; Li, A.; Jing, R.; Mao, L.; Li, Y.; Liu, X.; Zhang, K. Natural variation of *TaGASR7-A1* affects grain length in common wheat under multiple cultivation conditions. *Mol. Breed.* **2014**, *34*, 937–947. [CrossRef]

57. Jiang, Q.; Hou, J.; Hao, C.; Wang, L.; Ge, H.; Dong, Y.; Zhang, X. The wheat (*T. aestivum*) sucrose synthase 2 gene (*TaSus2*) active in endosperm development is associated with yield traits. *Funct. Integr. Genom.* **2011**, *11*, 49–61. [CrossRef]