agronomy

MDPI

Article

Trypsin Inhibitor Assessment with Biochemical and Molecular Markers in a Soybean Germplasm Collection and Hybrid Populations for Seed Quality Improvement

Kulpash Bulatova [1,*], Shynar Mazkirat [1], Svetlana Didorenko [1], Dilyara Babissekova [1], Mukhtar Kudaibergenov [1], Perizat Alchinbayeva [1], Sholpan Khalbayeva [1] and Yuri Shavrukov [2]

[1] Kazakh Scientific Research Institute of Agriculture and Plant Growing, Almaty Region, Kazakhstan; shynarbek.mazkirat@gmail.com (S.M.); svetl_did@mail.ru (S.D.); janeka__88@mail.ru (D.B.); muhtar.sarsenbek@mail.ru (M.K.); piko1206@mail.ru (P.A.); sholpan_2706@mail.ru (S.K.)

[2] College of Science and Engineering, Biological Sciences, Flinders University, Bedford Park 5042, Australia; yuri.shavrukov@flinders.edu.au

* Correspondence: bulatova_k@rambler.ru; Tel.: +77273883925

Received: 12 January 2019; Accepted: 9 February 2019; Published: 11 February 2019

Abstract: A soybean germplasm collection was studied for the identification of accessions with low trypsin inhibitor content in seeds. Twenty-nine accessions, parental plants, and two hybrid populations were selected and analyzed using genetic markers for alleles of the *Ti3* locus, encoding Kunitz trypsin inhibitor (KTI). Most of the accessions had high or very high KTI (49.22–73.53 Trypsin units inhibited (TUI/mg seeds), while the two local Kazakh cultivars, Lastochka and Ivushka, were found to have a moderately high content of KTI, at 54.16–54.87 TUI/mg. In contrast, two soybean cultivars from Italy, Hilario and Ascasubi, showed the lowest levels of trypsin units inhibited, at 25.47–27.87 TUI/mg. Electrophoresis of seed proteins in a non-denaturing system showed a simple discrimination pattern and very clear presence/absence of bands corresponding to KTI. The SSR marker *Satt228* was the most effective diagnostic marker among the three examined, and it confirmed the presence of the homozygous null-allele *ti3/ti3* in cultivars Ascasubi and Hilario, which were used for hybridization with the local cv. Lastochka. Heterozygote F_1 hybrid plants and homozygous *ti3/ti3* lines in F_2 segregating populations were successfully identified using *Satt228*. Finally, through marker-assisted selection with *Satt228*, prospective homozygous *ti3/ti3* lines were produced for further application in the breeding program aimed at improving soybean seed quality in Kazakhstan.

Keywords: Kunitz trypsin inhibitor; molecular markers; non-denaturing electrophoresis; seed quality; seed storage proteins; soybean; SSR (Simple sequence repeat) markers; *Ti* gene

1. Introduction

Soybean [*Glycine max* (L.) Merrill] is one of the most important crops for protein and oil production in Kazakhstan, and the processed seed products are mainly used for the purpose of food and fodder. Soybean breeding programs and seed production have been conducted in Kazakhstan for over 40 years, where 16 cultivars have been developed and introduced into local agribusiness. The area under soybean crop is increasing annually and now accounts for more than 120,000 hectares. However, this is still not enough to provide sufficient raw materials for oil processing and livestock and poultry farms in Kazakhstan.

The soybean germplasm collection and breeding lines (750 accessions) have been assessed using the main breeding indicators for seed productivity, quality and length of the growing season. Seeds were evaluated for quality indicators including protein and oil content. However, the content of anti-nutritional compounds in seeds was not previously taken into consideration by domestic breeders. As a result, parents with high levels of anti-nutritional compounds were unwittingly used for hybridization, which has significantly affected the resulting quality of the processed soybean products. The presence of proteinase inhibitors in seeds is known to be one of the main obstacles in the development and expansion of commercial soybean products. Soybean seeds contain two classes of major proteinase inhibitors, Kunitz trypsin inhibitor (KTI) and Bowman-Birk proteinase inhibitor [1,2]. Trypsin is an essential digestive enzyme found in the vertebrate small intestine that catalyzes the degradation of proteins, enabling their absorption into the bloodstream. KTI contributes up to 80% of the total trypsin inhibitor activity in soybean seeds [3].

Despite the moderate content of KTI in total soybean proteins, feeding of such seeds to livestock inhibits their growth and weight gain [4] and causes pancreatic hypertrophy [5]. For soybean seeds with high KTI, a preliminary heat treatment is required to inactivate the trypsin inhibitor enzyme before it can be fed to poultry and livestock. This heat treatment affectively destroys the undesirable anti-nutritional component, but also leads to an increase in the cost of the final product as well as decreased levels of available amino acids [6].

Initially, a single gene, *Ti*, encoding a high content of KTI in seeds of soybean, *G. max*, was described [7], where three dominant alleles, Ti^a, Ti^b, and Ti^c, were reported for trypsin inhibitor A_2, and the allele Ti^a was very common and widespread [8]. The multiple alleles were further extended during the study of 1368 germplasm accessions of Korean wild soybean (*G. soja* Sieb. & Zucc.), where the additional two rare dominant alleles, $Tib^{i7\text{-}1}$ and Tib^{i5}, were reported [9]. The *Ti/ti* locus was mapped to linkage group 9, with a genetic distance of 16.2 cM from the acid phosphatase locus, *Ap* [7], and 15.3 cM from the leucine aminopeptidase locus, *Lap1* [10]. Later, a comprehensive molecular analysis revealed 10 genes encoding KTI in soybean, but only four of them, *Ti1, Ti2, Ti3*, and *Ti4*, were transcribed into mRNA [11]. It was confirmed that only one gene, *Ti3* (GenBank ID: S45092.1), represents the major KTI, since its expression level is much more prevalent in soybean seeds compared to the other genes. The *Ti3* gene was reported to be transcribed at the seed maturation stage, producing most of the KTI protein found in soybean seeds [12]. This KTI polypeptide represents a soybean storage protein with a molecular weight of 21–21.5 kDa, with specific activity for trypsin inhibition [13,14]. The sequence of *Ti3* was identical to the common allele Ti^a discovered earlier [11]. Two other genes, *Ti1* and *Ti2*, displaying a lack of KTI activity, are mostly expressed at different stages of leaf, stem, and root development, and contribute a much smaller level of mRNA during embryogenesis and seed development [11]. The remaining *Ti* genes do not encode proteins with KTI activity and are assumed to be 'Pseudogenes'.

A single recessive null-allele, *ti*, resulted in the absence of KTI [8], caused by three nucleotide differences within the *Ti3* coding region [12]. The null-allele was initially originated from soybean germplasm accessions M91-212006, and the segregating analysis of progenies from mapping populations F_2 and F_3 was reported [15]. It was suggested that the absence of a 21.5 kDa protein in the total spectrum of seed storage proteins was related to the null-allele *ti*. Soybean germplasm accessions PI157440 and PI196168 were reported as additional sources of the null-allele *ti*. The breeding lines, based on introgression of the null-allele *ti* into commercial soybean cultivars using a conventional backcrossing program, revealed low KTI and improved seed quality for feeding chickens and young pigs [8]. In a further study, the soybean accession PI542044 with pedigree origin from PI157440 and carrying the null-allele *ti* was backcrossed with recurrent parents. The introgression of the *ti* null-allele at BC_2F_2 and BC_3F_2 generations was controlled by Marker-assisted selection (MAS) with *ti* allele-specific primers and linked SSR markers. Nine and six breeding lines with genetic backgrounds of JS97-52 and DS9712 / DS9814 recurrent parents, respectively, were obtained in Indian breeding

programs [16,17]. The introgressed null-allele *ti* was confirmed in the breeding lines and the KTI content was reduced by 69–84% [16] and an additional seed yield improvement was reported [17].

The aims of this research were: (1) to study a germplasm collection of soybean for the identification of genotypes with low KTI content in seeds; (2) to evaluate diagnostic genetic markers for the null-allele *ti3* in parental forms and produce hybrid populations from crosses between the lowest KTI genotype identified and elite Kazakh soybean cultivars; (3) to employ marker-assisted selection for low KTI genotypes in hybrid populations for seed quality improvement in the Kazakh breeding program.

2. Materials and Methods

The soybean germplasm collection comprising 29 cultivars was selected and received from the Kazakh Scientific Research Institute of Agriculture and Plant Growing, Almaty region, Kazakhstan (Listed in Table 2). Two hybrid combinations, Lastochka × Ascasubi and Lastochka × Hilario, F_1, F_2, and F_3, generations, were produced from manual crosses using a method described recently in the patent [18].

KTI activity was determined according to the method described by Kakade et al. [19], using casein as a substrate and expressed in trypsin units inhibited per milligram of soybean meal.

Isolation of glycinin storage proteins was carried out in phosphate buffer (pH 6.9), with subsequent electrophoretic separation in accordance with the protocol published earlier [20] in the presence of a molecular mass marker set ranging from 10 to 130 kDa (Thermo Fisher Scientific, Vilnius, Lithuania). Fixation and staining of the protein bands was carried out in 12.5% trichloroacetic acid. Quantitative measurement of the spectra components, their molecular mass, and relative mobility was carried out by means of the Quantum ST4 Gel documenting system (Vilber, Collégien, France), and the relative percentage of each band was calculated as a ratio of all components using the 'Quantum Capt' computer software supplemented to the equipment. Electrophoresis in a non-denaturing system was carried out according to the protocol published earlier [21]. Proteins were extracted from 10 mg of flour using 62.5 mM Tris HCl (pH 8.1) buffer for 40 min. The protein probe was prepared by mixing 100 µl of supernatant with 100 µl of 62.5 mM Tris HCl (pH 6.9) buffer mixture containing 20% glycerin and bromophenol blue marker dye. Fixation and staining of the protein bands was performed as mentioned above.

Genomic DNA was extracted using the CTAB method [22] from the first true leaves of individual seedlings grown in greenhouse conditions. Leaf tissue samples, about 200 mg each, were transferred to 2-ml test tubes with 800 µl of CTAB extraction buffer containing 1.0% polyvinylpyrrolidone (PVP40) and 0.2% β-mercaptoethanol, and homogenized using a stainless steel pestle. Extracted and precipitated DNA was re-dissolved in 400 µl of 1 M NaCl solution and treated with 2 µl (10 mg/mL) of RNase A (Thermo Scientific, Waltham, MA, USA) at 37 °C for 30 min. DNA was precipitated with cold 100% ethanol and washed with 70% ethanol. Isolated DNA was then dissolved in 100 µl of sterile water. The concentration and quality of DNA samples was determined at 260 and 280 nm using a spectrophotometer Jenway 6715 (Jenway, Staffordshire, UK). DNA samples were diluted with sterile water to a concentration of 100 ng/µl for use in further experiments.

PCR was performed in a total volume of 15 µL containing a cocktail with the following final concentrations: 1×PCR buffer, 2.5 mM $MgCl_2$, 200 µM each of dNTPs, 0.5 µM of each forward and reverse primers, BSA (2 mg/mL), and 0.5 U of *Taq* DNA polymerase (GeneLab, Astana, Kazakhstan). The PCR was conducted in a thermocycler (Bio-Rad, iCycler, Portland, ME, USA), where the amplification conditions were as follows: initial denaturation at 94 °C for 5 min, followed by 35 cycles of 94 °C for 20 s, 55 °C for 20 s, 72 °C for 30 s, and with a final extension of 10 min at 72 °C.

The amplification products were separated in polyacrylamide gel (8% acrylamide, 1×TBE buffer), and gels were stained with ethidium bromide for digital imaging by the Quantum ST4 Gel documenting system (Vilber, Collégien, France), as indicated above. The dimensional characteristics of PCR products were determined using the computer software 'Quantum Capt' (Vilber, Collégien, France) to determine the length and intensity of DNA fragments.

Three SSR markers tightly linked to the *Ti3* gene and distinguishing between alleles were selected based on published data [23], and primer sequences are presented in Table 1.

IBM SPSS Statistical software Desktop 25.0.0.0 (IBM, Armonk, NY, USA). was used to calculate and analyze means and standard error. Welch's ANOVA test was applied for comparison of accessions with low and high KTI due to different standard deviations and heteroscedasticity. Observed and expected segregations were analyzed using Chi-square test. One-way ANOVA and post-hoc Tukey–Kramer test with the minimum significant difference were applied for calculation of significant difference among genotypes of parent and hybrids with different KTI.

Table 1. SSR markers and corresponding primer sequences used for the identification of soybean *Ti3/ti3* genotypes.

Markers	Primer Sequences	
Satt228	F	TCATAACGTAAGAGATGGTAAAACT
	R	CATTATAAGAAAACGTGCTAAAGAG
Satt409	F	CCTTAGACCATGAATGTCTCGAAGAA
	R	CTTAAGGACACGTGGAAGATGACTAC
Ti/ti, gene specific marker	F	CTTTTGTGCCTTCACCACCT
	R	GAATTCATCATCAGAAACTCTA

3. Results

Biochemical analyses of soybean seeds in the germplasm collection were carried out for the purpose of characterizing the activity of anti-nutritional components by measuring the inhibition of trypsin by KTI. Our results show that two cultivars originating from Italy—Ascasubi and Hilario—with the lowest trypsin units inhibited, TUI (27.87 and 25.47 units/mg of dry ground seeds, respectively, Table 2), were significantly different to other studied soybean accessions ($p < 0.001$, using Welch's ANOVA test). The best local soybean cultivars from Kazakhstan were Lastochka and Ivushka, with moderately high KTI and showing 54.16 and 54.87 units/mg of TUI, respectively, but these were still significantly higher than Hilario and Ascasubi. (Table 2).

Table 2. Characterization of soybean collection samples according to the activity of the Kunitz trypsin inhibitor. TUI, trypsin units inhibited ± SE (n = 3).

Entry	Sample Name	Country of origin	TUI/mg	Group
1	Hilario	Italy	25.47 ± 2.44	Low
2	Ascasubi	Italy	27.87 ± 0.85	Low
3	Vilana	Russia	49.22 ± 1.56	High
4	Selekta 301	Russia	51.05 ± 0.57	High
5	Blamcos	Italy	51.62 ± 0.57	High
6	Lastochka	Kazakhstan	54.16 ± 0.56	High
7	Ivushka	Kazakhstan	54.87 ± 1.28	High
8	Sava	Serbia	54.87 ± 1.84	High
9	Triumph	Serbia	56.14 ± 0.28	High
10	Galina	Ukraine	56.85 ± 0.15	High
11	Slaviia	Russia	57.41 ± 0.14	High
12	Korsak	Ukraine	57.41 ± 0.42	High
13	Luna	Italy	57.98 ± 0.15	High

Table 2. *Cont.*

Entry	Sample Name	Country of origin	TUI/mg	Group
14	Dekabig	USA	57.98 ± 0.70	High
15	Pamiat IuGK	Kazakhstan	59.24 ± 0.56	High
16	Zhansaya	Kazakhstan	59.96 ± 0.14	High
17	Fora	Russia	61.65 ± 1.72	High
18	Harbin	China	61.93 ± 2.12	High
19	Voevodzhanka	Serbia	62.36 ± 0.85	High
20	Ayzere	Kazakhstan	63.77 ± 0.84	High
21	Perizat	Kazakhstan	64.91 ± 0.56	High
22	Sabira	Kazakhstan	66.18 ± 0.98	High
23	Safrana	France	66.60 ± 2.83	High
24	Delta	Russia	67.45 ± 0.29	High
25	Akky	Kazakhstan	67.59 ± 1.84	High
26	Santana	France	68.30 ± 0.56	High
27	Birlik KV	Kazakhstan	69.85 ± 1.27	High
28	Zen	Switzerland	71.69 ± 1.13	High
29	Atlantic	Italy	73.53 ± 0.71	High

Electrophoresis of seed storage proteins from the studied soybean cultivars revealed the presence of a 21 kDa molecular component, corresponding to the Kunitz trypsin inhibitor, in all analyzed genotypes. Importantly, only two cultivars, Hilario and Ascasubi, with the lowest amounts of KTI, showed the absence of the 21 kDa molecular component. Figure 1a shows the comparison of the KTI spectra in Hilario with three other cultivars as an example of the absence or presence of the KTI band. The relative percentage of KTI component in total seed storage proteins established by densitometry in the cultivar Triumph was only 1.03%, which was clearly absent in Hilario (Figure 1b).

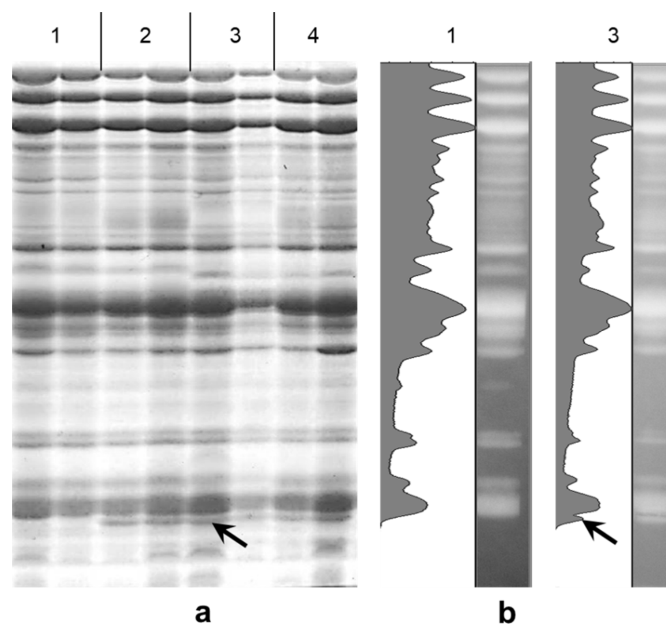

Figure 1. Electrophoregram of polyacrylamide gel (**a**) and densitograms (**b**) of soybean seed storage proteins. Germplasms are labelled above the image as follows: 1, Hilario; 2, Atlantic; 3, Triumph; and 4, Luna. The band for the 21 kDa molecular component KTI is indicated by arrows.

Molecular analysis of *Ti3* locus encoding KTI was carried out with three SSR markers. The markers *Satt228*, *Satt409*, and the gene-specific marker *Ti/ti* are tightly linked to the *Ti3* locus, and they can be perfectly used as diagnostic markers for MAS of genotypes with the null-allele *ti3*. Amplification products of these markers during PCR analysis with DNA from two soybean cultivars with low KTI (Hilario and Ascasubi) and one local cultivar Lastochka with high KTI are shown in Figure 2. The presented data confirmed that both Italian cultivars (Hilario and Ascasubi) have null-allele *ti3* while Kazakh cultivar has dominant allele *Ti3* in the locus.

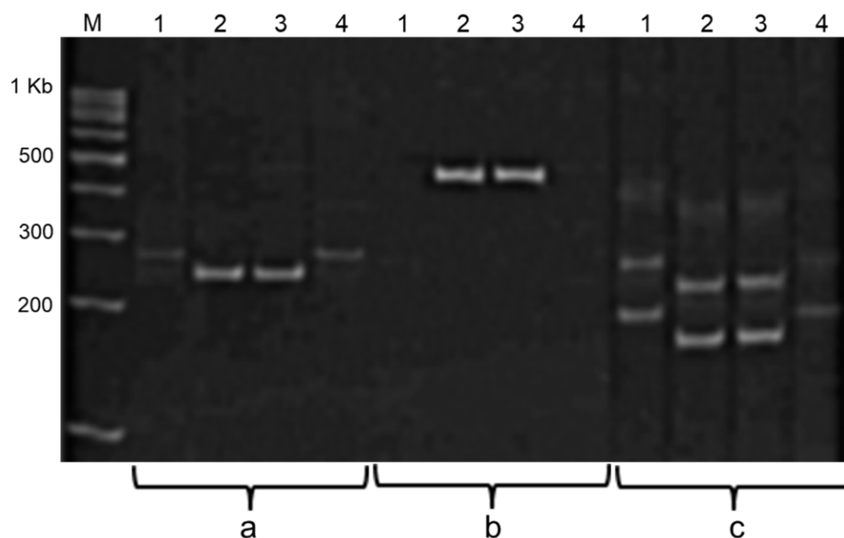

Figure 2. Amplification products of three SSR markers linked to the soybean *Ti3* gene. M, Marker 100 bp; 1 and 4, Lastochka; 2, Hilario; and 3, Ascasubi. The individual marker used is indicated at the bottom as follows: (**a**) *Satt228*; (**b**) *Ti/ti*-gene specific marker; **(c)** *Satt409*.

All markers clearly distinguish between soybean genotypes on the basis of the *Ti3* alleles detected, but the *Ti/ti*-gene-specific marker does not allow for the identification of homo- and heterozygote genotypes with dominant alleles *Ti3*, which makes it difficult to apply this marker in segregating populations. In contrast, the SSR marker *Satt228* was a much more suitable diagnostic marker for the further screening of plants in segregating populations, for the identification of homozygote progenies with null-allele *ti3* and production of the best non-segregating breeding lines with low KTI in seeds.

The cultivars Hilario and Ascasubi were identified as the genotypes with the lowest amounts of KTI, and they were selected in the results of the biochemical and PCR analyses with the confirmed null-allele *ti3*. Hybrid populations were produced with the null-allele *ti3* from Hilario and Ascasubi introgressed into the genetic background of local Kazakh cultivars.

In the hybrid combinations, Lastochka × Ascasubi and Lastochka × Hilario, PCR analysis of DNA from the F_1 plants and SSR marker *Satt228* confirmed the presence of the null-allele *ti3* in Ascasubi and Hilario (Data not shown). All F_1 plants were heterozygous for the *Ti3* locus with high KTI and were used for further production of the F_2 populations. Biochemical analysis for the spectrum of seed storage proteins using non-denaturing electrophoresis was applied for analyses of Kazakh varieties and F_2 segregating populations originating from the crosses, Lastochka × Ascasubi and Lastochka x Hilario. All local varieties and the parental plants from cultivars Ascasubi and Hilario, as well as segregants with null-alleles, *ti3/ti3*, showed very clear presence/absence of bands in a simple discrimination pattern. This method can be used for the selection of genotypes with the null allele of the KTI locus in seeds. However, homo- and heterozygote genotypes with dominant alleles, *Ti3/Ti3* and *Ti3/ti3*, could not be separated using polyacrylamide gel electrophoresis (Figure 3).

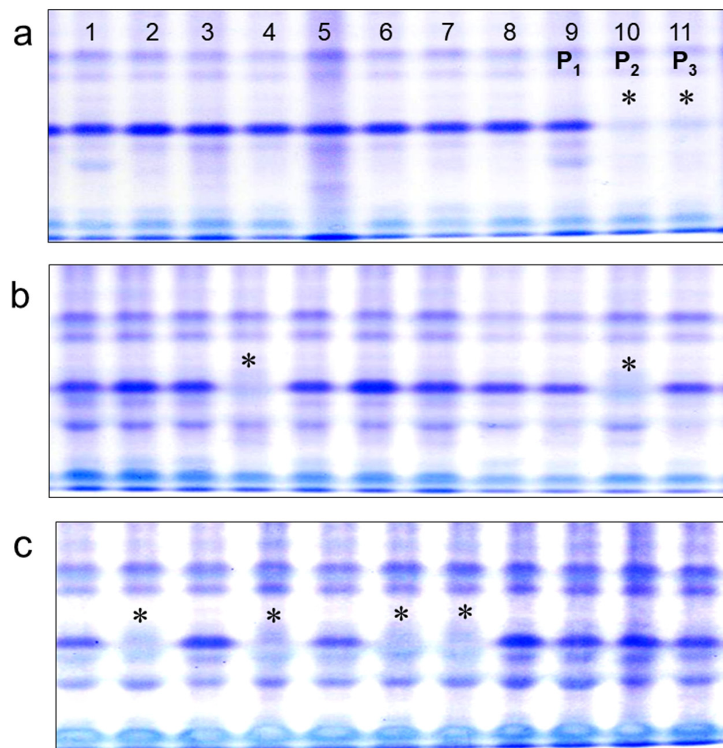

Figure 3. Polyacrylamide gel patterns of soybean seed storage proteins: (**a**) local varieties (1–8) and (9) parental cultivar Lastochka (P_1) with genotypes *Ti3/Ti3*; two other parental forms (10 and 11), Ascasubi (P_2) and Hilario (P_3), respectively, with genotype *ti3/ti3*. Examples of the F_2 segregating populations: (**b**) Lastochka × Ascasubi; and (**c**) Lastochka × Hilario. Homozygote genotypes, *ti3/ti3*, are indicated by asterisks (*).

In contrast, application of the SSR marker *Satt228* could identify all three types of genotypes at the *Ti3* locus. An example of genotyping of soybean plants using this SSR marker is presented in the second cross, Lastochka × Hilario. PCR analysis of the F_2 of hybrid plants in the presence of both parents confirmed that plants of cv. Lastochka have genotypes with the dominant alleles *Ti3/Ti3*, while plants of cv. Hilario are homozygotes with the recessive null-allele *ti3/ti3*. Among progenies of the F_2 hybrid population, all genotypes with the dominant and recessive alleles of the *Ti3* locus were identified (Figure 4).

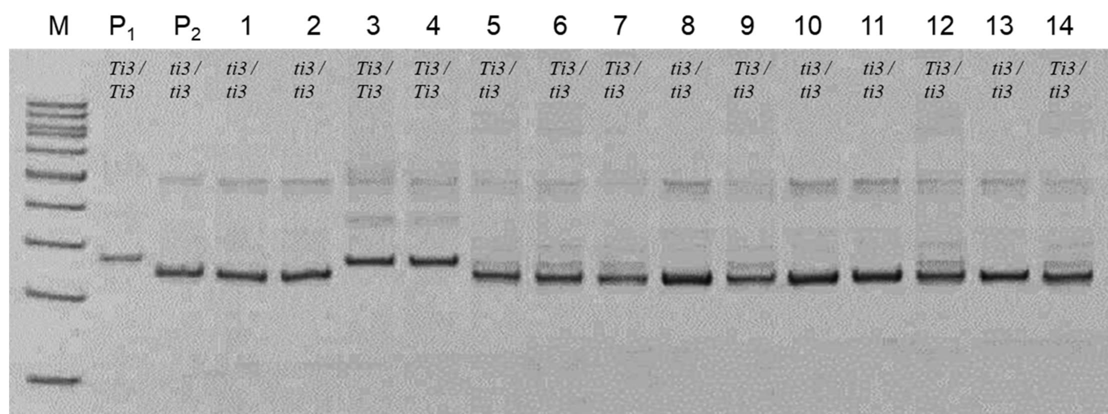

Figure 4. PCR products of parental forms and the F_2 hybrid population of Lastochka × Hilario amplified with *Satt228*. M, Marker 100 bp; P_1, Lastochka; P_2, Hilario; 1–14, Hybrid F_2 plants. Genotypes for the *Ti3* locus are shown for each plant.

The comparison of segregation analyses using seed storage proteins and the molecular SSR marker *Satt228* revealed full consensus and confirmed the Mendelian monogenic-type inheritance in both studied F_2 populations of Lastochka × Ascasubi and Lastochka × Hilario (Table 3).

Table 3. Segregation analyses of genotypes of the *Ti3* locus using seed storage proteins and the molecular SSR marker *Satt228* in F_2 populations of Lastochka × Ascasubi and Lastochka × Hilario. The asterisks (* and **) indicate no significant differences ($p < 0.05$ and $p < 0.01$, respectively) between observed and expected segregations for simple Mendelian ratio (3:1 for seed storage proteins and 1:2:1 for SSR marker *Satt228*), compared to the Chi-square distribution table.

	Genotypes			Total Number of Plants	$\chi^2 < \chi^2_{(Tablel)}$
	Ti3/Ti3	*Ti3/ti3*	*ti3/ti3*		
Lastochka × Ascasubi					
Seed storage proteins, observed segregation	46		24	70	2.78 * < $3.84_{(df = 1)}$
Expected segregation	52.5		17.5	70	
SSR marker *Satt228*, observed segregation	15	31	24	70	3.23 ** < $4.61_{(df = 2)}$
Expected segregation	17.5	35	17.5	70	
Lastochka × Hilario					
Seed storage proteins, observed segregation	38		15	53	0.31 ** < $2.71_{(df = 1)}$
Expected segregation	39.75		13.25	53	
SSR marker *Satt228*, observed segregation	12	26	15	53	0.36 ** < $4.61_{(df = 2)}$
Expected segregation	13.25	26.5	13.25	53	

MAS was applied for segregants with homozygote null-allele genotypes, *ti3/ti3*, based on the results of screening using the SSR marker *Satt228* on F_2 hybrid progenies of Lastochka × Ascasubi and Lastochka × Hilario. The selected homozygote F_2 plants, both with recessive and dominant alleles at the *Ti3* locus, were grown and seeds from F_3 families were finally analyzed for KTI content based on TUI results (presented in Table 4).

Table 4. Kunitz trypsin inhibitor analysis in homozygote F_3 families originating from two hybrid populations after MAS with SSR marker *Satt228* for trypsin units inhibited (TUI) ± SE ($n = 3$). Minimum significant differences ($p > 0.01$) based on one-way ANOVA with post-hoc Tukey–Kramer test are shown by different letters. No statistical differences are shown by identical letters.

Parent/Progeny	Name	Genotype *Ti3/ti3*	TUI/mg of Dry Seeds	Statistical Differences
Lastochka × Ascasubi				
♀P_1	Lastochka	*Ti3/Ti3*	53.2 ± 0.2	a
♂P_2	Ascasubi	*ti3/ti3*	25.2 ± 0.1	c
F_3 family 1	(Lastochka × Ascasubi) − 1	*Ti3/Ti3*	43.0 ± 0.6	b
F_3 family 2	(Lastochka × Ascasubi) − 2	*ti3/ti3*	23.7 ± 1.3	c

Table 4. *Cont.*

Parent/Progeny	Name	Genotype *Ti3/ti3*	TUI/mg of Dry Seeds	Statistical Differences
		Lastochka × Hilario		
♀P_1	Lastochka	*Ti3/Ti3*	53.2 ± 0.2	a
♂P_3	Hilario	*ti3/ti3*	23.2 ± 0.1	c
F_3 family 3	(Lastochka × Hilario) − 3	*Ti3/Ti3*	45.5 ± 0.6	b
F_3 family 4	(Lastochka × Hilario) − 4	*ti3/ti3*	17.4 ± 1.3	d

The presented results show consistent inheritance and very significant differences in KTI in homozygote F_3 families originating from recombinants in both hybrid combinations. In the first hybrid (Lastochka × Ascasubi), the F_3 family showed KTI content statistically similar to the parental form Ascasubi, while in the second hybrid combination (Lastochka × Hilario), statistically reduced KTI was found in the F_3 family (Table 4).

4. Discussion

Soybean is a very widely used legume crop, but improvement of seed quality for lower KTI content is a very important and challenging task representing a critical step for soybean breeding. Therefore, the identification of soybean germplasm resources with low KTI and their introgression in hybridization programs promotes the development of promising soybean breeding lines with improved protein composition.

Among studied local Kazakh soybean cultivars, only two, Ivushka and Lastochka, were found to have more moderate levels of KTI. However, two Italian cultivars, Hilario and Ascasubi, were identified as having the lowest KTI, with significantly less than other studied soybean germplasms at 25.47 and 27.87 units of trypsin inhibited, respectively. Nevertheless, these cultivars with the lowest content of anti-nutritional factors are relatively old, having been developed in Italy in the early 1990s. Therefore, they are most valuable as donors of genetic resources to the modern soybean breeding program.

Biochemical screening of the germplasm collection for KTI and electrophoretic analysis of the seed protein composition allow for the evaluation of a range of soybean cultivars produced both locally and overseas, thus making it possible to select the most suitable parental forms for crossings and hybrid production. Homozygote progenies with the null-allele *ti3/ti3* can be successfully identified and novel breeding lines can be produced from hybrid populations. However, this process must be improved and significantly sped up via the use of MAS in the initial steps of selections, using a suitable diagnostic marker strongly associated with the null-allele *ti3* to produce new soybean cultivars with improved seed quality.

The presented results show the effective application of both biochemical and molecular markers for the null-allele *ti3* in the studied soybean germplasm collection, and in two segregating populations. Biochemical analysis of seed storage proteins using polyacrylamide gel electrophoresis is very accurate and can be used at the seed stage whilst also saving the viable part of the *ti3/ti3* line for multiplication. Therefore, it is important to enrich the initial pool of recombinant lines with the null allele of the *Ti3* locus because a large part of the lines can be removed at the early stage of the breeding process. The use of the diagnostic SSR marker *Satt228* is based on regular PCR and is very simple and quick. SSR markers are well known as being polymorphic, with codominant inheritance, and therefore can be widely used for genotype identification and selection of desired traits [23]. However, the distance between the SSR marker and the gene of interest can vary depending on the type of population [6,24]. For example, the genetic distance between *Satt228* and *Ti3* varied between 0 and 3.7 cM in two different populations [23]. Therefore, the number of recombinant genotypes in F_2 hybrids of Lastochka

× Ascasubi and Lastochka × Hilario (Table 3) is very small and can be estimated as 0–2.6 and 0–2.0 plants, respectively. Another codominant SSR marker *Satt409* could also be successfully used instead (Figure 2), but it was mapped to a region more genetically distant from the *Ti3* locus at 4.5–21.9 cM [23], meaning that there would be a greater chance of unwanted recombinants. Therefore, either native electrophoresis of storage proteins or SSR markers can support the reliable selection of new promising breeding lines, and the choice will depend on the cost or convenience as preferred by researchers. The applied MAS with *Satt228* was very effective in our experiments with both segregating populations, where homozygote genotypes *ti3/ti3* were identified, isolated, and propagated. The simple Mendelian-type inheritance of the *Ti3* locus has been confirmed [12,15] and helps to estimate the ratio for homozygotes with the null-allele *ti3*. Finally, segregants with *ti3/ti3* genotypes were successfully verified for low KTI content and propagated for further yield analysis and development of prospective breeding lines. Our results were similar to those published earlier on the introgression of the null-allele *ti3* and MAS to select recombinant genotypes with low KTI content in an Indian soybean breeding program [16,17]. Hybridization and transference of the beneficial *ti3* null-alleles determining low KTI in seeds using MAS is very important to enhance the market value and overall soybean seed quality in Kazakhstan.

5. Conclusions

Biochemical analysis of seeds of 29 varieties from soybean germplasm collection revealed only two cultivars, Hilario and Ascasubi, with the lowest activity of the Kunitz trypsin inhibitor (KTI). In contrast, all soybean cultivars currently grown in Kazakhstan have a KTI enzyme activity ranging from 54.16 to 69.85 of trypsin units inhibited per mg of seeds, for Lastochka and Birlik KV, respectively. Using a method employing protein analysis and molecular markers, the null-allele *ti3* was confirmed in the cultivars Hilario and Ascasubi. These genotypes were then used in crosses with domestic soybean cultivars. The SSR marker *Satt228* was identified as the most effective diagnostic marker, confirming the heterozygosity of the F_1 generation and helping to select homozygous lines with the null allele *ti3* in F_2 and in F_3 segregated populations to enable the reduction of KTI and the improvement of seed quality.

Author Contributions: Conceptualization and supervision: K.B.; data collecting and analysis: S.M.; soybean resources and hybrid populations design: S.D.; molecular genotyping: D.B.; agronomic performance: M.K.; protein electrophoresis: P.A.; trypsin inhibitor activity assay: S.K.; review and editing: Y.S.

Funding: This research was funded by the Ministry of Agriculture (Kazakhstan), Scientific Technical Program, BR06249214 (M.K.).

Acknowledgments: We thank Carly Schramm for her critical comments and English editing.

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

References

1. Laskowski, M., Jr.; Kato, I. Protein inhibitors of proteinases. *Annu. Rev. Biochem.* **1980**, *49*, 593–629. [CrossRef] [PubMed]
2. Tan-Wilson, A.L.; Chen, J.C.; Duggan, M.C.; Chapman, C.; Obach, R.S.; Wilson, K.A. Soybean Bowman-Birk trypsin isoinhibitors: Classification and report of a glycine-rich trypsin inhibitor class. *J. Agric. Food Chem.* **1987**, *35*, 974–981. [CrossRef]
3. Kumar, V.; Rani, A.; Rawal, R. Deployment of gene specific marker in development of Kunitz trypsin inhibitor free soybean genotypes. *Indian J. Exp. Biol.* **2013**, *51*, 1125–1129. [PubMed]
4. Palacios, M.; Easter, R.A.; Soltwedel, K.T.; Parsons, C.M.; Douglas, M.W.; Hymowitz, T. Effect of soybean variety and processing on growth performance of young chicks and pigs. *J. Anim. Sci.* **2004**, *82*, 1108–1114. [CrossRef] [PubMed]
5. Liencer, I.E. Possible adverse effects of soybean anticarcinogens. *J. Nutr.* **1995**, *125*, 744–750.

6. Chung, J. Identification and confirmation of SSR marker tightly linked to the *Ti* locus in soybean [*Glycine max* (L.) Merr.]. In *Soybean—Genetics and Novel Techniques for Yield Enhancement*; Krezhova, D., Ed.; InTech: Rijeka, Croatia, 2011; pp. 201–212.
7. Hildebrand, D.F.; Orf, J.H.; Hymowitz, T. Inheritance of an acid phosphatase and its linkage with the Kunitz trypsin inhibitor in seed protein of soybeans. *Crop Sci.* **1980**, *20*, 83–85. [CrossRef]
8. Hymowitz, T. Genetics and breeding of soybeans lacking the Kunitz trypsin inhibitor. In *Nutritional and Toxicological Significance of Enzyme Inhibitors in Foods. Advances in Experimental Medicine and Biology*; Friedman, M., Ed.; Springer: Boston, MA, USA, 1986; Volume 199, pp. 291–298.
9. Krishnamurthy, P.; Singh, R.J.; Tsukamoto, C.; Park, J.H.; Lee, J.D.; Chung, G. Kunitz trypsin inhibitor polymorphism in the Korean wild soybean (*Glycine soja* Sieb. & Zucc.). *Plant Breed.* **2013**, *132*, 311–316.
10. Klang, Y.T.; Chiang, Y.C. Genetic linkage of a leucine aminopeptidase locus with the Kunitz trypsin inhibitor locus in soybeans. *J. Hered.* **1986**, *77*, 128–129. [CrossRef]
11. Jofuku, K.D.; Goldberg, R.B. Kunitz trypsin inhibitor genes are differentially expressed during the soybean life cycle and in transformed tobacco plants. *Plant Cell* **1989**, *1*, 1079–1093. [CrossRef] [PubMed]
12. Jofuku, K.D.; Schipper, R.D.; Goldberg, R.B. A frameshift mutation prevents Kunitz trypsin inhibitor mRNA accumulation in soybean embryos. *Plant Cell* **1989**, *1*, 427–435. [CrossRef] [PubMed]
13. Ryan, C.A. Proteinase inhibitors. In *Biochemistry of Plants*; Marcus, A., Ed.; Academic Press: San Diego, CA, USA, 1981; pp. 351–370.
14. Kim, S.H.; Hara, S.; Hase, S.; Ikenaka, T.; Toda, H.; Kitamura, K.; Kaizuma, N. Comparative study on amino acid sequences of Kunitz-type soybean trypsin inhibitors, *Ti, Tib,* and *Tic*. *J. Biochem.* **1985**, *98*, 435–448.
15. Vollmann, J.; Schausberger, H.; Bistrich, H.; Lelley, T. The presence or absence of the soybean Kunitz trypsin inhibitor as a quantitative trait locus for seed protein content. *Plant Breed.* **2002**, *121*, 272–274. [CrossRef]
16. Kumar, V.; Rani, A.; Rawal, R.; Mourya, V. Marker assisted accelerated introgression of null allele of Kunitz trypsin inhibitor in soybean. *Breed. Sci.* **2015**, *65*, 447–452. [CrossRef] [PubMed]
17. Maranna, S.; Verma, K.; Talukdar, A.; Lal, S.K.; Kumar, A.; Mukherjee, K. Introgression of null allele of Kunitz trypsin inhibitor through marker-assisted backcross breeding in soybean (*Glycine max* L. Merr.). *BMC Genetics* **2016**, *17*, 106. [CrossRef] [PubMed]
18. Didorenko, S.V.; Karyagin, Y.G.; Bulatova, K.M. Method of soybean hybridization. Submitted by Kazakh Scientific Research Institute of Agriculture and Plant Growing, Almaty region, Kazakhstan. Republic of Kazakhstan Patent No. 31427, 21 July 2016.
19. Kakade, M.L.; Simons, N.; Liener, I.E. An evaluation of natural vs synthetic substances for measuring the antitryptic activity of soybean samples. *Cereal Chem.* **1974**, *46*, 518–526.
20. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227*, 680–685. [CrossRef] [PubMed]
21. Davis, B.J. Description of discontinuous buffer system for non-denaturing gels in electrophoresis. *Ann. New York Acad. Sci.* **1964**, *209*, 373–381.
22. Murray, M.G.; Thompson, W.F. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* **1980**, *8*, 4321–4325. [CrossRef] [PubMed]
23. Kim, M.S.; Park, M.J.; Jeong, W.H.; Nam, K.C.; Chung, J.I. SSR marker tightly linked to the *Ti* locus in soybean [*Glycine max* (L.) Merr.]. *Euphytica* **2006**, *152*, 361–366. [CrossRef]
24. Cregan, P.B.; Jarvik, T.; Bush, A.L.; Shoemaker, R.C.; Lark, K.G.; Kahler, A.L.; Kaya, N.; VanToai, T.T.; Lohnes, D.G.; Chung, J.; et al. An integrated genetic linkage map of the soybean genome. *Crop Sci.* **1999**, *39*, 1464–1490. [CrossRef]

© 2019 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 (http://creativecommons.org/licenses/by/4.0/).