



Article

# Heteroalleles in Common Wheat: Multiple Differences between Allelic Variants of the *Gli-B1* Locus

Eugene Metakovsky <sup>1,\*</sup>, Laura Pascual <sup>1,\*</sup>, Patrizia Vaccino <sup>2</sup>, Viktor Melnik <sup>3</sup>, Marta Rodriguez-Quijano <sup>1</sup>, Yulia Popovych <sup>4</sup>, Sabina Chebotar <sup>4</sup> and William John Rogers <sup>5</sup>

<sup>1</sup> Department of Biotechnology-Plant Biology, School of Agricultural, Food and Biosystems Engineering, Universidad Politécnica de Madrid, 28040 Madrid, Spain; emetakovsky@gmail.com (E.M.); marta.rurquiaga@upm.es (M.R.-Q.)

<sup>2</sup> Consiglio per la ricerca in agricoltura e l'analisi dell'economia Agraria, Research Centre for Cereal and Industrial Crops, 13100 Vercelli, Italy; patrizia.vaccino@crea.gov.it

<sup>3</sup> Vavilov Institute of General Genetics RAS, 117971 Moscow, Russia; meller@vigg.ru

<sup>4</sup> Department of Genetics and Molecular Biology, National I.I. Mechnikov University, 65058 Odessa, Ukraine; popovych1818@gmail.com (Y.P.); s.v.chebotar@onu.edu.ua (S.C.)

<sup>5</sup> Departamento de Biología Aplicada, CIISAS, CIC-BIOLAB, CONICET-INBIOTEC, CRESCA, Facultad de Agronomía, Universidad Nacional del Centro de la Provincia Buenos Aires, Azul 7300, Provincia de Buenos Aires, Argentina; rogers@faa.unicen.edu.ar

\* Correspondence: laura.pascual@upm.es

**Citation:** Metakovsky, E.; Pascual, L.; Vaccino, P.; Melnik, V.; Rodriguez-Quijano, M.; Popovych, Y.; Chebotar, S.; Rogers, W. J. Heteroalleles in Common Wheat: Multiple Differences between Allelic Variants of the *Gli-B1* Locus. *Int. J. Mol. Sci.* **2021**, *22*, 1832. <https://doi.org/10.3390/ijms22041832>

Academic Editor: Richard R.-C. Wang

Received: 30 November 2020

Accepted: 7 February 2021

Published: 12 February 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 (<http://creativecommons.org/licenses/by/4.0/>).

**Abstract:** The *Gli-B1*-encoded  $\gamma$ -gliadins and non-coding  $\gamma$ -gliadin DNA sequences for 15 different alleles of common wheat have been compared using seven tests: electrophoretic mobility (EM) and molecular weight (MW) of the encoded major  $\gamma$ -gliadin, restriction fragment length polymorphism patterns (RFLPs) (three different markers), *Gli-B1*- $\gamma$ -gliadin-pseudogene known SNP markers (Single nucleotide polymorphisms) and sequencing the pseudogene *GAG56B*. It was discovered that encoded  $\gamma$ -gliadins, with contrasting EM, had similar MWs. However, seven allelic variants (designated from I to VII) differed among them in the other six tests: I (alleles *Gli-B1i*, *k*, *m*, *o*), II (*Gli-B1n*, *q*, *s*), III (*Gli-B1b*), IV (*Gli-B1e*, *f*, *g*), V (*Gli-B1h*), VI (*Gli-B1d*) and VII (*Gli-B1a*). Allele *Gli-B1c* (variant VIII) was identical to the alleles from group IV in four of the tests. Some tests might show a fine difference between alleles belonging to the same variant. Our results attest in favor of the independent origin of at least seven variants at the *Gli-B1* locus that might originate from deeply diverged genotypes of the donor(s) of the B genome in hexaploid wheat and therefore might be called “heteroallelic”. The donor’s particularities at the *Gli-B1* locus might be conserved since that time and decisively contribute to the current high genetic diversity of common wheat.

**Keywords:** *Triticum aestivum*; *Gli-B1*;  $\gamma$ -gliadin polymorphism; APAGE; PCR; DNA sequencing

## 1. Introduction

Wheat is globally one of the most important crops. Currently, with world production at 728.5 million tonnes and a sown area covering 214 million hectares in 2019, wheat includes a quarter of total cereal production [1]. Wheat production includes two different polyploidy species derived from natural hybridization events [2], the tetraploid species (AABB) durum wheat (*Triticum turgidum* L.) and the hexaploid species (AABBDD) bread wheat (*Triticum aestivum* L.). Bread wheat represents roughly 90 to 95% of the total production; hence, any advances related to knowledge or breeding in this species will have a major worldwide impact.

Wheat provides 20% of the daily protein and food calorie intakes worldwide and represents one of the major protein components of the human diet. In recent decades, there has been an unprecedented increase in knowledge of the chemistry, genetics and functionality of the main proteins of the wheat seed, gliadin and glutenin. For example,

there are classifications of different alleles at the seed-storage-protein-encoding loci in their relation to dough quality. This knowledge helps the production of new genotypes with improved characteristics (including field, dough-quality and medical aspects). However, the immense complexity of the bread wheat genome ( $2n = 6x = 42$ , 17Gb) and of the genetic loci encoding the synthesis of the gliadins and glutenins leaves many questions unanswered. Such as the fact the same protein alleles might have different relationships to dough quality. Such findings raise the question of whether these alleles, identified by protein analysis in unrelated cultivars, are indeed identical. Moreover, the degree of difference at the DNA level of allelic variants differing on the protein level at the same seed-storage-protein-encoding locus is unknown.

A common wheat genotype produces more than 50 different polypeptides of gliadin (gliadins) [3,4]. The primary structure of several hundred gliadin-coding DNA sequences from different wheat species has been studied, and three main types of gliadin polypeptides differing between them in primary structure and biochemical properties, the so-called  $\alpha$ -,  $\gamma$ - and  $\omega$ -gliadins, have been described [5]. Moreover, each type, for example, the  $\gamma$ -gliadins [6,7], is composed of a number of families, meaning that members of a given family are more similar than members of different families.

About 20–25 bands differing in electrophoretic mobility (EM) may be observed in a one-dimensional electrophoregram (acid polyacrylamide gel electrophoresis, APAGE) of a single grain, encoded in common wheat by six major (*Gli-A1*, *Gli-B1*, *Gli-D1*, *Gli-A2*, *Gli-B2*, *Gli-D2*) [8] and several minor [9] *Gli* loci. Analysis of segregating progenies from crosses between wheat genotypes showed that any *Gli* allele encodes two or more gliadin polypeptides (a block of electrophoretic bands) inherited together as a Mendelian unit. Alleles at one *Gli* locus differ in the number and/or EM of the encoded gliadin polypeptides of different gliadin types [10]. In total, there are currently about 180 known alleles at the six major loci [11].

A fundamental characteristic of the seed-storage-protein-coding loci (for example, of the tightly linked *Gli-1* and *Glu-3* loci mapped in the distal part of the short arms of the homoeologous group 1 chromosomes [8]), is their great complexity. The size of one storage protein gene, including its regulatory sequences, is known to be about 0.8–1.5 Kb [12–14], but the distance between these genes has been shown to be very large, reaching 55–120 Kb per 1 cM of genetic distance [15,16]. The average distance between active genes was about 81 Kb [14], and the distance of 10 Kb between two gliadin genes, tenfold the length of a gliadin gene itself, may be considered as exceptionally short, besides the fact that the genes can be considered to be clustered [16]. The size of an entire *Gli* locus encoding several gliadin polypeptides may exceed 2000 Kb as suggested for the *Gli-A2* locus [17].

Hence all expressed gliadin genes together occupy only a minor part of the DNA of a *Gli* locus. The largest part of a *Gli* locus was shown to consist of non-coding DNA sequences represented mainly by different types of transposable elements [16,18,19], but also by gliadin pseudogenes and gene fragments [20] for which inter-variety polymorphism was described [21] and even some non-gliadin genes and sequences with open reading frames [16,22,23]. Therefore, allelic variants of a *Gli* locus may differ in the number and types of active genes, in the distance between them, in the quota of pseudogenes among all the gliadin sequences and, probably, in the presence of non-gliadin genes harbored within a given *Gli* locus. The polymorphism of an entire *Gli* locus could be further augmented by the presence of single nucleotide polymorphisms (SNP).

Alleles at a *Gli* locus are routinely identified by the use of APAGE of gliadin proteins, thereby revealing sets of transcribed genes within the locus. However, an allele identified by APAGE may differ, in different genotypes, in the non-coding part of the locus. Hence alleles identified by APAGE may only be considered as markers of corresponding allelic variants of *Gli* loci.

There are few possibilities to compare allelic variants of an entire *Gli* locus, including its non-coding sequences. Restriction fragment length polymorphism (RFLP) analysis

may disclose a difference between the alleles in the distribution of gliadin-coding sequences (including pseudogenes) along the whole locus. Two sets of molecular probes were also made available for the analysis of gliadin-coding DNA sequences. Firstly, the *GAG56B*  $\gamma$ -gliadin pseudogene was detected with the gene-specific PCR primer pair GAG29/30, and one of the variants of the pseudogene sequence (p-aes) was proven to be a characteristic of common wheat *T. aestivum* [24]. Secondly, two allele-specific PCR markers were developed for the *Gli-B1*- $\gamma$ -gliadin sequence (probable pseudogene). Australian common wheat cultivars tested with these markers gave one PCR product with one pair of primers of either 369 or 397 bp [25].

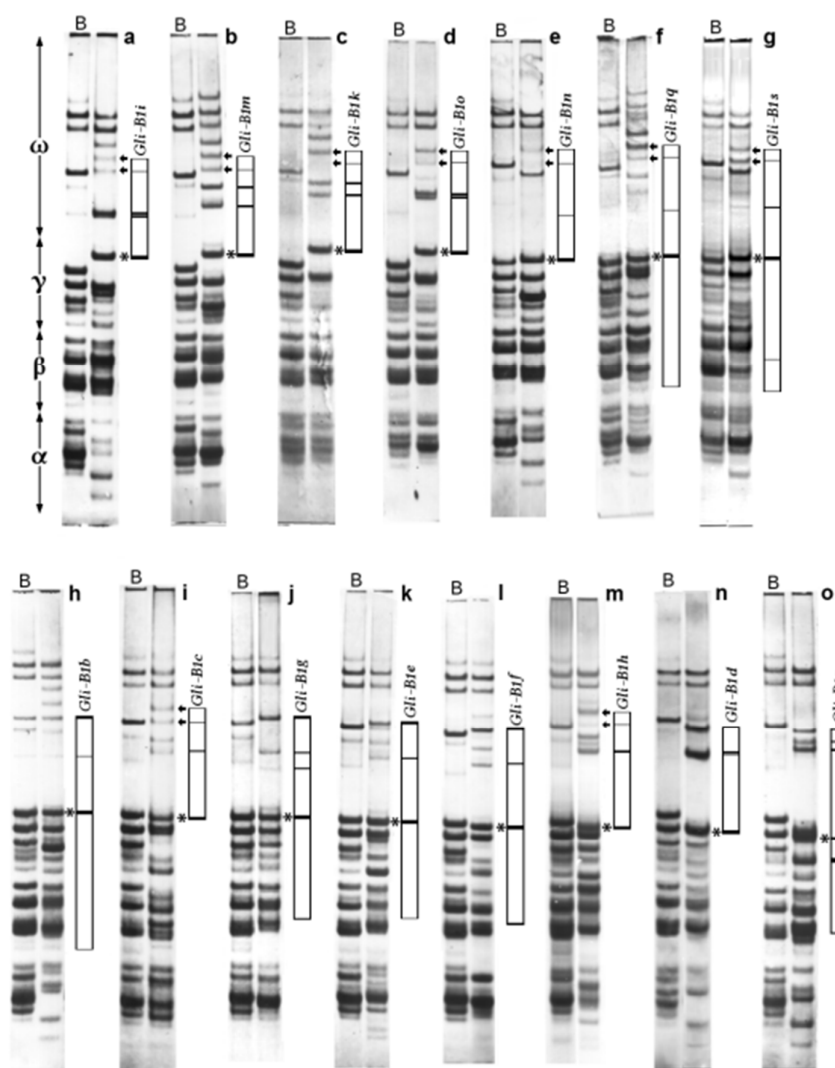
Thus, a *Gli*- locus is a highly complex chromosomal fragment that includes several gliadin-encoding genes, pseudogenes and sequences unrelated to gliadins. Some allelic variants at the *Gli* loci (and at the *Glu-3* loci tightly linked to *Gli-1*) have been related to dough quality, providing useful markers to conduct quality breeding. However, the same allele at the *Gli* locus identified by protein electrophoresis in different cultivars might apparently have different relationships to dough quality, suggesting that this allele might be represented by more than one variant and tampering with the usefulness of gliadins as molecular markers. In order to explore this question, in this work, we compared 15 allelic variants at the *Gli-B1* locus of common wheat (identified as blocks of *Gli-B1*-encoded gliadin polypeptides in APAGE) by means of several approaches: firstly, EM and molecular weight (MW) of encoded  $\gamma$ -gliadins using two-dimensional electrophoresis; secondly, their RFLP patterns (three different restriction enzymes and a  $\gamma$ -gliadin-specific probe); thirdly, the length of the PCR sequence of the  $\gamma$ -gliadin pseudogene; and fourthly, the DNA sequence of another  $\gamma$ -gliadin pseudogene present within the *Gli-B1* locus. It was shown that the alleles studied might be divided into several variants strongly differing between them in their coding, as well as non-coding, sequences. The simultaneous multiple differences (without intermediate forms) among versions of the *Gli-B1* locus may be explained through the suggestion that these variants maintain the particularities of the original genotypes of the corresponding diploid donor(s) in hexaploid wheat.

## 2. Results

### 2.1. Alleles at the *Gli-B1* Locus Identified by APAGE

Twenty-four alleles at the *Gli-B1* locus (including the null-allele and the *Gli-B1l* allele which designates the 1BL.1RS translocation) have been revealed in the wheat germplasm studied [11], and 16 of them, including *Gli-B1l*, were selected for further analysis.

Examples of gliadin electrophoregrams (APAGE) of the common wheat cultivars carrying the alleles studied (Figure 1) and schemes of blocks of jointly inherited gliadin bands (Figure S1) are shown.



**Figure 1.** APAGE electrophoregrams of standard common wheat cultivars carrying different alleles at the *Gli-B1* locus. (a), Insignia (Australia); (b), Silvana (Romania); (c), Pyrotrix-28 (Kazakhstan); (d), Levent (Bulgaria); (e), Intensivnaya (Kirgizstan); (f), Lesostepka-75 (Ukraine); (g), Est-Mottin (Italy); (h), Marquis (Canada); (i), Siete-Cerros-66 (Mexico); (j), Galahad (UK); (k), Solo (Germany); (l), Ducat (France); (m), Barbilla-Leon (Spain); (n), Caprock (USA); (o), Chinese Spring (China). B, universal standard cultivar Bezostaya-1 (*Gli-B1b*) (Russia) [11]. The major  $\gamma$ -gliadin encoded at the *Gli-B1* locus (asterisk) and  $\omega$ -gliadins encoded by the allele *Gli-B5b* (arrows) are indicated.  $\alpha$ ,  $\gamma$  and  $\omega$ , mark the portion of the electrophoregrams where  $\alpha$ -,  $\gamma$ - and  $\omega$ -gliadins migrate respectively. The blocks of jointly inherited gliadin electrophoretic bands are shown schematically.

Alleles producing the slowest-moving  $\gamma$ -gliadin and alleles *Gli-B1n*, *q*, *s* and *h* are commonly accompanied by the allele *Gli-B5b* of the *Gli-B5* locus, tightly linked to *Gli-B1* [26]. The allele *Gli-B5b* encodes a pair of faint  $\omega$ -gliadins, while the allele *Gli-B5a* is null. Allelic variation at *Gli-B5* was accepted as a polymorphism procured by the distal part of chromosome 1B, and  $\omega$ -gliadins encoded by the *Gli-B5b* were considered as a part of the *Gli-B1*-encoded blocks [11].

Each allele at the *Gli-B1* locus of common wheat produces one  $\gamma$ -gliadin (the *Gli-B1a* encodes two  $\gamma$ -gliadins), two or more  $\omega$ -gliadins and some bands in the  $\beta$ -zone of the APAGE electrophoregram (Figure S1). At least seven variants of the EM of the *Gli-B1*-encoded  $\gamma$ -gliadin (therefore, of the locus) were revealed: *Gli-B1i*, *m*, *k*, *o* (variant I, the

slowest EM); *Gli-B1n*, *q*, *s* (II); *Gli-B1b* (III); *Gli-B1e*, *g* (IV); *Gli-B1h* (V); *Gli-B1d* (VI); and *Gli-B1a* (VII, the fastest) (Figure 1). The allele *Gli-B1f* produces a slightly faster  $\gamma$ -gliadin compared to alleles of variant IV (Figure 1l,j,k), but nevertheless may be attributed to this variant due to the similarity of other APAGE bands composing a block. For the same reason (the composition of the encoded block), the allele *Gli-B1b* was assumed to be different from alleles of the group IV.

The allele *Gli-B1c* produces a weaker stained  $\gamma$ -gliadin and is linked to *Gli-B5b* (Figure 1i), thus differing from other alleles of group IV. Hence, the *Gli-B1c* might be considered as some kind of variation of variant IV or even a rather independent variant VIII of the *Gli-B1* locus.

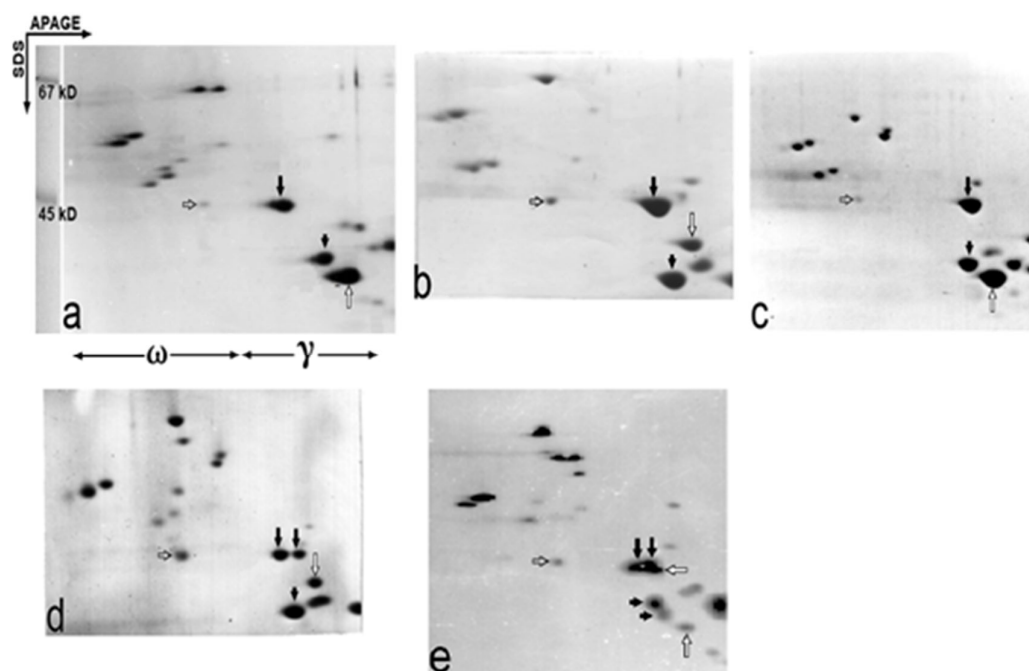
## 2.2. Apparent Molecular Weight of the *Gli-B1*-Encoded $\gamma$ -Gliadins as Studied by Two-Dimensional APAGE $\times$ SDS Electrophoresis

The best way to compare the apparent MWs of the *Gli-B1*-encoded  $\gamma$ -gliadin polypeptides using two-dimensional electrophoresis (TDE) is to analyze a heterozygous seed of a cross between two genotypes carrying the alleles under study. Additionally, there is an “internal marker” present in most of the genotypes, the *Gli-A3*-controlled  $\omega$ -gliadin: allelic variants of this polypeptide show identical apparent MWs of about 41 kD [9]. The major  $\gamma$ -gliadin encoded by the *Gli-D1* locus (apparent MW of about 34 kD, data not shown) may be also used as an internal marker of MW (Figures 2 and S2).

The intriguing result of this approach seems to be a close similarity (identity) of the apparent MWs of  $\gamma$ -gliadins produced by any variant of the *Gli-B1* locus (Figures 2 and S2): *Gli-B1*-encoded  $\gamma$ -gliadins with different (including contrasting) EM had similar MWs of about 41 kD. For example, the MWs of  $\gamma$ -gliadins of alleles of variants I, III, VI (Figure 2a–c) and IV, II, V (Figure S2a–c, respectively) were hardly distinguishable by TDE. The close similarity of the MWs of these  $\gamma$ -gliadins was confirmed by analysis of heterozygous F2 seeds (Figure 2d).

Analysis of TD (APAGE  $\times$  SDS) electrophoregrams published by different authors, as well as the results of our work, showed that alleles *Gli-B1a*, *b*, *c*, *d*, *e*, *f*, *g*, *h*, *k*, *m*, *o*, *q* and *s* each controlled  $\gamma$ -gliadin of similar MWs (Table S1). A new allele, *Gli-B1v* (differing slightly from the *Gli-B1o* in the EM of one  $\omega$ -gliadin), found in Spanish landraces also produced the  $\gamma$ -gliadin with an MW identical to those of variant I (Figure S2d). In durum wheat, known  $\gamma$ -gliadins  $\gamma$ 42 and  $\gamma$ 45 differing in their EM also have similar MWs of about 41 kD (Figure 3 in [27]).

Although SDS-electrophoresis is hardly sensitive enough to document some probable minor differences in the length of compared polypeptides (for example, Figure 2e), it is clear that  $\gamma$ -gliadins controlled by different alleles at the *Gli-B1* locus have a similar (if not identical) MW. Therefore, they did not meet the regularity (higher EM, lower MW) established earlier for the *Gli-A1*- and the *Gli-D1*-encoded  $\gamma$ -gliadins [28].



**Figure 2.** Two-dimensional (APAGE × SDS) separations of gliadin of common wheat cultivars and of F2 seeds of their crosses. (a), Skorospelka-U (*Gli-B1m*); (b), Tarasovskaya-29 (*Gli-B1b*); (c), Dneprovskaya-521 (*Gli-B1d*); (d), F2 seed heterozygous at the *Gli-B1* from the cross between cultivars Rusalka (*Gli-B1d*) and Bezostaya-1 (*Gli-B1b*); (e), F2 seed heterozygous at each of the *Gli-1* loci from the cross between Kharkovskaya-6 (*Gli-A1j*, *Gli-B1e*, *Gli-D1i*) and Skorospelka-U (biotype *Gli-A1m*, *Gli-B1m*, *Gli-D1c*). The long-tailed black arrows indicate the *Gli-B1*-encoded  $\gamma$ -gliadins, and the transparent short-tailed arrow shows the *Gli-A3*-controlled  $\omega$ -gliadin having an MW of 41 kD (“internal marker”). The transparent long-tailed arrows, and the short-tailed black arrows, show the *Gli-A1*-encoded and the *Gli-D1*-encoded  $\gamma$ -gliadins, respectively. Only the  $\gamma$ - and  $\omega$ -zones of the two-dimensional electrophoregram are shown.

A general structure for the  $\gamma$ -gliadin-encoding genes and polypeptides is known [5–7]. The number of repeats and the length of the microsatellite present in the  $\gamma$ -gliadin gene sequence, as well as deletions or insertions, may influence the length (the MW) of the encoded  $\gamma$ -gliadin polypeptide. In contrast, considerable differences in the EM without noticeable alterations in their MW observed in our work, for example, among  $\gamma$ -gliadins of alleles *Gli-B1m*, *b*, *d* (Figure 2a–c, respectively), may depend upon the number of charged amino acids present in these polypeptides of similar length.

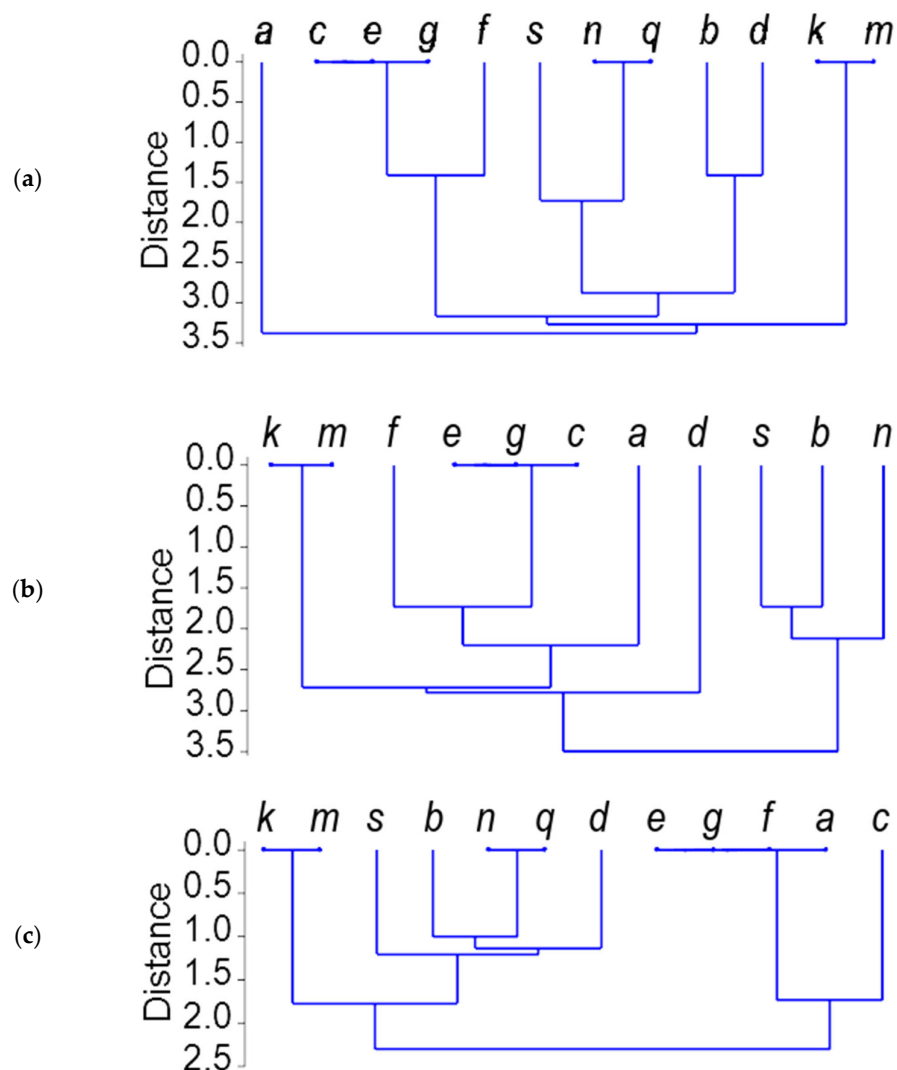
### 2.3. *Gli-B1* Alleles Represented by RFLP Patterns

Allelic variants of the *Gli-B1* locus were studied by the RFLP approach using three different restriction enzymes *TaqI*, *RsaI* or *HaeIII* and the K32 probe, specific for the transcribed  $\gamma$ -gliadin sequences. Previous analysis of aneuploid lines of the cultivar Chinese Spring permitted identification of polymorphic DNA fragments belonging to the allele *Gli-B1a* in the RFLP pattern of this cultivar, and a congruity was established between the presence of definite polymorphic DNA fragments and some *Gli-B1* alleles in wheat genotypes [29].

The RFLP-*TaqI* profiles (Figure S3) for 12 alleles at the *Gli-B1* locus were compared, and 19 polymorphic bands were considered. The tree of relationships among *Gli-B1* alleles drawn from their RFLP patterns (Figure 3) showed that the endonuclease *TaqI* did not distinguish between alleles *Gli-B1m* and *k*, *Gli-B1n* and *q*, and *Gli-B1e* and *g*. Sets of polymorphic bands of alleles *Gli-B1s* and *Gli-B1f* produced by *TaqI* were close to those of variants II and IV, respectively. For example, the difference between the *Gli-B1f* and other alleles of variant IV was in the position of only one RFLP fragment (Figure S3a; lane 1,

band 19, and lane 3, band 16, respectively). The RFLP-*RsaI* profiles produced 25 polymorphic bands (data not shown). The tree of relationships among *Gli-B1* alleles (Figure 3b) was generally similar to that obtained using *TaqI*. However, in contrast to *TaqI*, alleles *Gli-B1b* and *Gli-B1d* had nothing in common when *RsaI* was used (Figure 3a,b).

The RFLP-*HaeIII* profiles produced only nine polymorphic bands, and the relationships among *Gli-B1* alleles were different from those obtained using two other endonucleases. The most important findings were, first, an identity of the allele *Gli-B1a* with allelic variant IV, and, second, a considerable difference of the allele *Gli-B1c* from this variant (Figure 3c).



**Figure 3.** Relationships among alleles at the *Gli-B1* locus based on the restriction fragment length polymorphism (RFLP) data. (a), *TaqI*; (b), *RsaI*; (c), *HaeIII*.

The results obtained with the RFLP approach show considerable differences in DNA sequence (positions of restriction sites) along the *Gli-B1* locus among its alleles. There are six clear-cut allelic variants repeating those revealed by TDE, *Gli-B1k* and *m* (variant I), *Gli-B1n* and *q* (+s) (II), *Gli-B1b* (III), *Gli-B1e* and *g* (+f) (IV), *Gli-B1d* (VI), and *Gli-B1a* (VII). Alleles of different variants do not share or share only a few identical *Gli-B1*-derived polymorphic DNA fragments recognized by the  $\gamma$ -gliadin-specific probe. Furthermore, the

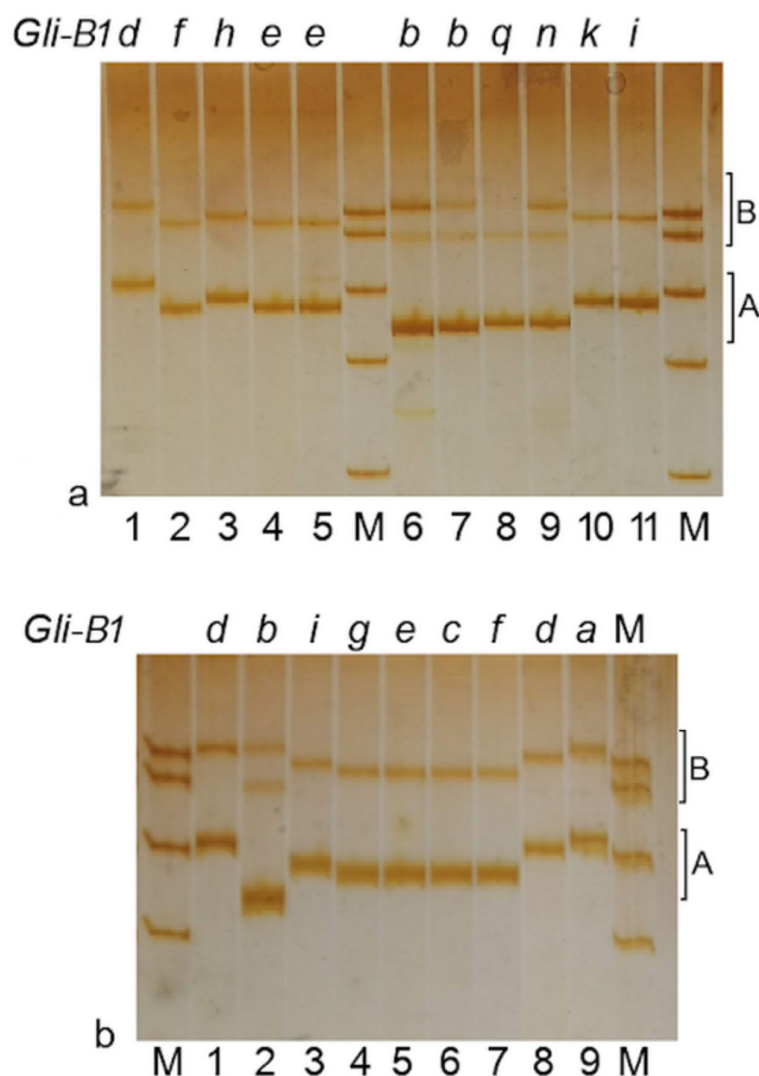
approach has permitted some fine differences to be revealed between alleles belonging to the same variant.

In accordance with the RFLP analysis, the allele *Gli-B1c* may represent one more variant (VIII), related to, but different from, alleles of group IV, as has already been suggested on the basis of APAGE data. The allele *Gli-B1a* seems to be related to variant IV (endonucleases *RsaI* and *HaeIII*).

There was no convincing identity of any RFLP band produced by genotypes carrying the allele *Gli-B1l* (the 1BL.1RS) and any other allele studied.

#### 2.4. Comparison of Alleles at the *Gli-B1* Locus Using $\gamma$ -Gliadin Pseudogene SNP Markers

$\gamma$ -Gliadin SNP markers (PCR primers GliB1.1 and GliB1.2) are specific for a pseudogene located inside the *Gli-B1* locus. Each cultivar tested produced a PCR sequence with either primer GliB1.1 or GliB1.2, whereas cultivars carrying the 1BL.1RS translocation did not produce any amplification [25]. In our work, 16 alleles (including *Gli-B1l*) at the *Gli-B1* locus were analyzed by this test (Figure 4; Table 1), and four main observations were drawn.



**Figure 4.** PCR (primers B1.1 or B1.2) Product Length Polymorphism (PLP) among alleles at the *Gli-B1* locus. A and B represent two differentiated zones in the gel. Alleles at the *Gli-B1* loci for each sample analyzed are indicated in the upper part of the gel. (a) 1, Pavon-F-76; 2, Cappelle-Desprez;



3, Ardec; 4, Escualo; 5, Glenlea; 6, Gabo; 7, Marquis; 8, Goelent; 9, Intensivnaya; 10, Mentana; 11, Insignia; 1–5, primers GliB1.2; 6–11, primers GliB1.1; (b) 1, Laura; 2, Bezostaya-1; 3, Insignia; 4, Galahad; 5, Escualo; 6, Prinqual; 7, Arminda; 8, Suneca; 9, Chinese-Spring; 1, 4–9 primers GliB1.2; 2, 3, primers GliB1.1. Alleles at the *Gli-B1* locus are indicated above the lanes. M, marker pUC19/*MspI* with DNA fragments of length 501, 489, 404, 331 and 242 bp, respectively. The polymorphism of the PCR-derived DNA fragments of higher length (zone B) was not considered in this work.

Firstly, all alleles at the *Gli-B1* locus were divided into two groups, according to [25]. One group gave a PCR product with the GliB1.1 primer and included alleles of variants I (*Gli-B1i*, *k*, *m*, *o*), II (*Gli-B1n*, *q*, *s*) and III (*Gli-B1b*). Another group (PCR product with the GliB1.2) was composed of all other alleles (*Gli-B1a*, *d*, *h*, *c*, *e*, *f*, *g*) (Table 1). As expected, the cultivar Cartaya carrying the 1BL.1RS translocation (allele *Gli-B1l*) did not give a PCR product with any pair of primers.

**Table 1.** Cultivars studied, their alleles at the *Gli-B1* locus, and the approximate length for the product obtained using GliB1.1 or GliB1.2 primers.

Cultivar	Country	Year	Allele	Primer	Length (bp)
Bezostaya-1	Russia	1959	<i>b</i>	GliB1.1	369
Gabo	Australia	1942	<i>b</i>	GliB1.1	369
Marquis	Canada	1907	<i>b</i>	GliB1.1	369
Insignia	Australia	1946	<i>i</i>	GliB1.1	400
Mentana	Italy	1913	<i>k</i>	GliB1.1	400
Titien	France	1985	<i>m</i>	GliB1.1	400
Aragon-03	Spain	1940	<i>o</i>	GliB1.1	400
Intensivnaya	Kyrgyzstan	1978	<i>n</i>	GliB1.1	372
Goelent	France	1985	<i>q</i>	GliB1.1	372
Salmone	Italy	1980	<i>s</i>	GliB1.1	372
Cartaya	Spain	1984	<i>l</i>	GliB1.1 or GliB1.2	-
Chinese-Spring	China	-	<i>a</i>	GliB1.2	415
Pavon-F-76	Mexico	1971	<i>d</i>	GliB1.2	409
Laura	Canada	1986	<i>d</i>	GliB1.2	409
Rinconada	Spain	1981	<i>d</i>	GliB1.2	409
Suneca	Australia	1981	<i>d</i>	GliB1.2	409
Prinqual	France	1978	<i>c</i>	GliB1.2	397
Escualo	Spain	1981	<i>e</i>	GliB1.2	397
Glenlea	Canada	1972	<i>e</i>	GliB1.2	397
Arminda	Netherlands	1976	<i>f</i>	GliB1.2	397
Cappelle-Desprez	France	1946	<i>f</i>	GliB1.2	397
Galahad	UK	1983	<i>g</i>	GliB1.2	397
Ardec	Belgium	1979	<i>h</i>	GliB1.2	402
Caia	Portugal	-	<i>h</i>	GliB1.2	403

Secondly, there was a polymorphism in the length of the PCR products (Product Length Polymorphism, PLP) (Figure 4, zone A). Three length variants were detected with the GliB1.1 primers (Figure 4a, lines 6–12) and four with GliB1.2 (Figure 4a, lines 1–5; Figure 4b, line 9). A polymorphism observed in zone B of the DNA electrophoregram (Figure 4) that included some PCR-derived DNA fragments of higher length was not considered.

Thirdly, an allele that is present in genotypes of different cultivars produced a PCR sequence of apparent identical length. For example, three cultivars from three countries with the allele *Gli-B1b* and four cultivars from four countries with the allele *Gli-B1d* produced PCR sequences of approximately 369 bp (GliB1.1) and 409 bp (GliB1.2), respectively (Table 1).

Fourthly, alleles encoding slow-moving  $\gamma$ -gliadin (*Gli-B1i*, *k*, *m*, *o*; Figure S1) gave a PCR product of identical length (Figure 4a, lines 10–12; Table 1). In addition, alleles of

variant IV (*Gli-B1e, g, f*) and allele *Gli-B1c* each produced an apparently identical PCR sequence of 397 bp (Figure 4b, lines 4–7).

Therefore, there are seven groups of alleles identified in accordance with the PLP. These groups are identical to those revealed by the APAGE, TDE and RFLP approaches: alleles *Gli-B1i, k, m, o* (variant I); *Gli-B1n, q, s* (II); *Gli-B1b* (III); *Gli-B1e, g, f* (IV), *Gli-B1h* (V); *Gli-B1d* (VI); and *Gli-B1a* (VII). The allele *Gli-B1c* belongs, in this test, to variant IV.

## 2.5. DNA Sequence of the $\gamma$ -Gliadin Pseudogene GAG56B in Allelic Variants of the *Gli-B1* Locus

Seven cultivars studied for polymorphism in their  $\gamma$ -gliadin pseudogene GAG56B of the *Gli-B1* locus were divided into two groups by DNA sequence [24]. In the present work, von Büren's primers GAG29 and GAG30 were used and the resulting products were sequenced. In total, 30 cultivars from 15 countries representing 15 alleles at the *Gli-B1* locus were compared (Table 2).

**Table 2.** Variants of DNA sequences (haplotypes) of the pseudogene GAG56B in cultivars having different alleles at the *Gli-B1* locus.

Cultivar	Country	Year	Allele	Haplotype (Figure 5)	EM of $\gamma$ -gliadin <sup>1</sup>
Insignia	Australia	1946	<i>i</i>	1	I
Pane-247	Spain	1960	<i>k</i>	1	I
Etoile-de-Choisy	France	1950	<i>m</i>	1	I
Aragon-03	Spain	1940	<i>o</i>	1	I
Intensivnaya	Kyrgyzstan	1978	<i>n</i>	2	II
Spada	Italy	1985	<i>n</i>	2	II
Lesostepka-75	Ukraine	1945	<i>q</i>	2	II
Salmone	Italy	1980	<i>s</i>	2*	II
Alcalá	Spain	1984	<i>b</i>	3	III
Anza	Mexico, USA	1971	<i>b</i>	3	III
Glenlea	Canada	1972	<i>e</i>	4	IV
Marius	France	1976	<i>f</i>	4	IV
Pernel	France	1983	<i>f</i>	4	IV
Adalid	Spain	1987	<i>g</i>	4	IV
Calodine	Italy	1991	<i>g</i>	4	IV
Diego	Spain	-	<i>c</i>	4	VIII
Prinqual	France	1978	<i>c</i>	4	VIII
Siete-Cerros-66	Mexico	1966	<i>c</i>	4	VIII
Ardec	Belgium	1979	<i>h</i>	5	V
Caia	Portugal	-	<i>h</i>	5	V
Pepital	Netherlands	1989	<i>h</i>	5	V
Cajeme-71	Mexico	1971	<i>d</i>	6	VI
Chopin	France	1984	<i>d</i>	6	VI
Katepwa	Canada	1981	<i>d</i>	6	VI
Chinese-Spring	China	-	<i>a</i>	7	VII
Pyrotrix-28	Kazakhstan	1973	<i>m</i>	8	I
Titien	France	1985	<i>m</i>	9	I
Astral	France	1972	<i>f</i>	10	IV
Floreal	France	1984	<i>f</i>	10	V
Saratovskaya-39	Russia	1968	<i>e</i>	11	IV

<sup>1</sup> Variant of the electrophoretic mobility (APAGE) of the *Gli-B1*-encoded  $\gamma$ -gliadin (variant I, the slowest; variant VII, the fastest).

We were able to obtain a high-quality sequence between 170 bp from the forward primer and 800bp for 29 cultivars. In Salmone, a high-quality sequence ranged from 200 bp to 800 bp (Figure S4). Aligning against the reference genebank accession M13712, we

identified 14 positions with SNP, one of them (position 259) with three different possible nucleotides. Furthermore, the sequence included a single sequence repeat (SSR) CAA for which 10 possible variants of the length were detected. All these polymorphisms were combined in 11 different haplotypes (Figure 5; Table 2).

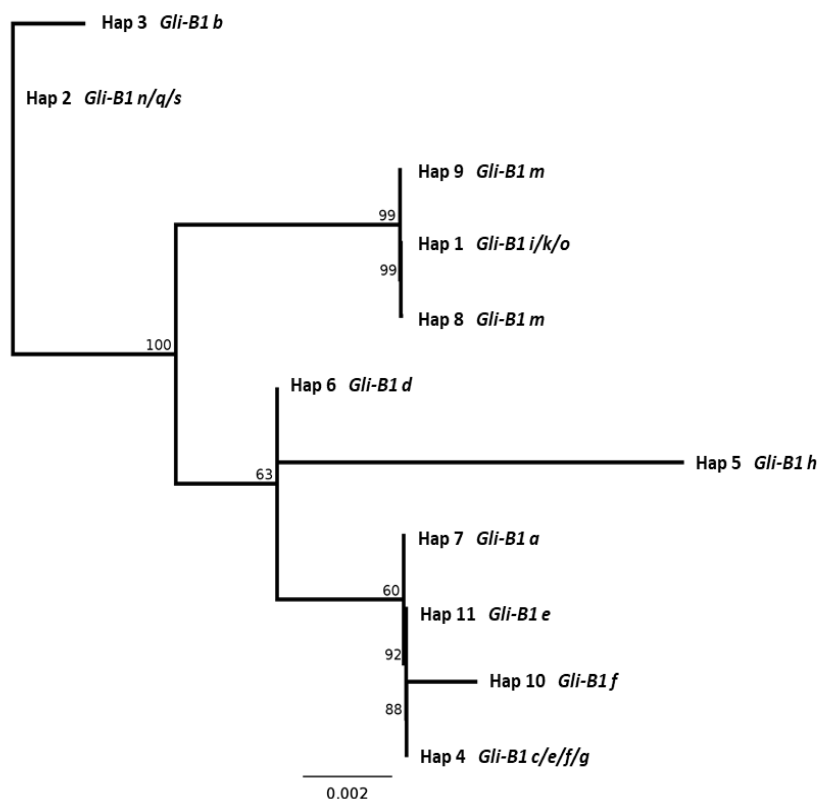
Four main observations have been drawn from these results.

nt position	Hap_1	Hap_2	Hap_3	Hap_4	Hap_5	Hap_6	Hap_7	Hap_8	Hap_9	Hap_10	Hap_11
173	C	C	C	C	T	C	C	C	C	C	C
182	C	C	T	C	C	C	C	C	C	C	C
259	G	T	T	G	A	G	G	G	G	G	G
279	A	C	C	C	C	C	C	A	A	C	C
348	C	C	C	C	C	C	C	C	C	G	C
413	C	C	C	T	C	C	T	C	C	T	T
438	C	C	C	C	T	C	C	C	C	C	C
474	T	C	C	T	T	T	T	T	T	T	T
495	A	A	A	A	T	A	A	A	A	A	A
524	T	C	C	C	C	C	C	T	T	C	C
532	C	C	C	C	T	C	C	C	C	C	C
585	A	A	A	C	C	C	C	A	A	C	C
602	C	C	C	C	A	C	C	C	C	C	C
694	(CAA)18	(CAA)8	(CAA)7	(CAA)15	(CAA)19	(CAA)26	(CAA)31	(CAA)17	(CAA)20	(CAA)15	(CAA)16
755	T	C	C	C	C	C	C	T	T	C	C

**Figure 5.** Differences in nucleotide sequence among 11 haplotypes of the pseudogene *GAG56B*. Differences with respect to haplotype 1 are marked in grey. Positions are counted from the beginning of the primer employed in the DNA amplification.

Firstly, in most cases, any allele at the *Gli-B1* locus (identified through APAGE analysis of the set of gliadin polypeptides) occurring in different cultivars has an identical sequence for the  $\gamma$ -gliadin pseudogene *GAG56B*. For example, haplotype 2 was registered in two unrelated cultivars from two countries with allele *Gli-B1n*, haplotype 4 occurred in three cultivars with *Gli-B1c*, and haplotype 5 in three cultivars from three countries with *Gli-B1h* (Table 2). Therefore, there was a clear congruity of the polymorphisms in the expressed (*Gli-B1*-encoded  $\gamma$ -gliadin) and noncoding ( $\gamma$ -gliadin pseudogene) parts of the *Gli-B1* locus. The allele *Gli-B1c* belongs, in this test, to variant IV.

Secondly, 11 haplotypes of the pseudogene *GAG56B* were divided into seven deeply diverging groups similar to those revealed by other methods in our work: alleles *Gli-B1i*, *k*, *m*, *o* (variant I: haplotype 1); *Gli-B1n*, *q*, *s* (II: 2); *Gli-B1b* (III: 3); *Gli-B1e*, *g*, *f*, *c* (IV: 4); *Gli-B1h* (V: 5); *Gli-B1d* (VI: 6); and *Gli-B1a* (VII: 7) (Figure 5). The smallest differences were observed between alleles of groups II and III (substitution at position 182 and a minimal difference in the length of the SSR sequence) and of the groups IV and VII (a difference in the length of the SSR sequence). More differences were found between variants I, IV + VII and II + III, while the greatest difference was observed between variants I and V (substitution in 10 positions and in the length of the SSR) (Figure 6).



**Figure 6.** Relationships among allelic variants of the *Gli-B1* locus based upon the sequence of the pseudogene *GAG56B* located within the *Gli-B1* locus.

Thirdly, there was one case of a fine difference between alleles belonging to the same group (as has been found using the RFLP approach): the allele *Gli-B1m* differed slightly from other alleles of variant I by the length of its SSR sequence (Figure 5, Table 2).

Fourthly, there were three alleles (identified by APAGE) for which the *GAG56B* sequence differed among wheat genotypes: *Gli-B1m* (haplotypes 1, 8 and 9), *Gli-B1e* (4 and 11) and *Gli-B1f* (4 and 10) (Figure 6, Table 2). The difference might be caused by a mutational change in a given allelic variant of the *Gli-B1* locus causing the appearance of a new haplotype of the pseudogene studied. For example, an expansion of the CAA repeats present inside the *GAG56B* pseudogene converts allelic variant *Gli-B1e* (the cultivar Glenlea, haplotype 4, usual for alleles of variant IV) into the new variant of the same allele (the cultivar Saratovskaya-36, haplotype 11). Three variants of the allele *Gli-B1m* differed in the number of the CAA repeats (18, 17, 20), while the only substitution in position 348 distinguished haplotype 4 from 10 (Figure 5, S4). Any allele at *Gli-B1* (for example, *Gli-B1m*), harboring different haplotypes of the *GAG56B* pseudogene in different cultivars, always produces the same set (a block) of encoded gliadins (data not shown).

Finally, a close similarity was noticed between two different trees of genetic distances, based, firstly, on the RFLP-*RsaI* data (Figure 3b), and, secondly, on the sequencing of the *GAG56B* pseudogene (Figure 6).

### 3. Discussion

Differentiation of homologous chromosomes exists among genotypes of self-pollinated species. In *Triticum aestivum*, differences between homologs in their heterochromatin banding were documented [30]. Intraspecific polymorphism was mostly expressed for chromosomes of the B genome [31]. Due to chromosome-specific variation, the number of chiasmata and pairing of homologous chromosomes in F1 hybrids between cultivars was

reduced compared to that of their parental inbred lines [32]. The observed recombination frequency between genetic markers depends upon the level of differences between the chromosomes of the parental cultivars involved in the cross [32,33].

Homologous chromosomes present in different wheat genotypes were called “heterohomologous” to differentiate them from euhomologous (veritably homologous) chromosomes present in a given genotype. The differences among heterohomologues are distributed along the chromosomes involved and are assumed to be due to differences in the non-coding portion of DNA sequences [34,35]. Deletions of small DNA sequences [36], the presence of different sets of transposable elements and variation in the number of their blocks located between genes [14,16,22,37], as well as SNP [38,39], may contribute to the differences between heterohomologous chromosomes.

The existence of several quite different variants of the DNA sequence at the *Gli-B1* locus of common wheat has been documented in our work. These variants could be designated, analogously to heterohomologous chromosomes, as “heteroallelic”, and alleles belonging to the same variant would be “homoallelic”.

RFLP seems to be especially useful for the comparison of allelic versions of a *Gli* locus, because restriction sites for a given endonuclease may occur at any location along the whole long gliadin locus, including expressed genes and non-coding sequences. Using the RFLP approach permitted several distinct variants of DNA sequences of the *Gli-B1* locus to be documented, strongly differing (without intermediate forms) in their restriction site positions. Moreover, some fine differences between alleles belonging to the same variant were revealed.

The length of the PCR products of the  $\gamma$ -gliadin SNP markers and sequencing the PCR products of the pseudogene *GAG56B* were used for more detailed analysis of two particular regions of the non-coding DNA inside the *Gli-B1* locus. In general, analysis of coding and non-coding parts of 15 allelic variants of the *Gli-B1* locus revealed at least seven “heteroallelic” variants of the *Gli-B1* locus (I–VII) without intermediate forms between them. Two pairs of allelic variants (II and III; IV and VII) were more closely related than the others. Nevertheless, allele *Gli-B1a* (variant VII, which was only documented in the genotype of the local variety Chinese-Spring), for example, differed from alleles of variant IV by the EM (APAGE) of the encoded  $\gamma$ -gliadin (Figure 1, S1), by the RFLP-*TaqI* test (Figure 3a), by the length of the PCR product (Figure 4) and by the sequence of the pseudogene *GAG56B* (Figures 5 and 6).

The allele *Gli-B1c* was identical to other alleles of group IV in the RFLP test (endonucleases *TaqI* and *RsaI*) and in PLP and *GAG56B* tests but differed from them in another RFLP test (*HaeIII*) and in APAGE. Therefore, we assume *Gli-B1c* is one more variant (VIII) related to group IV.

Earlier, six contrasting variants without intermediate forms were described for DNA sequences of the *Glu-B1* locus (B genome) among 58 cultivars of tetraploid and hexaploid wheat. It was suggested that these variants were present in genotypes of the diploid donors of this genome [13].

The pathways of origin of hexaploid wheat through successive crosses between diploid donors of the A and B genomes, and later, between AB and the D genome, are known in detail [40]. Naturally, heteroallelic variants of the *Gli-B1* locus might originate from different genotypes of the donor(s) of the B genome in hexaploid wheat, as suggested by [41], and inherit particularities of the *Gli-B1* locus of these donors. It is suggested that genotypes of the diploid donors of genomes A, B and D were already highly polymorphic before the origin of allopolyploids [42] and that crosses between wild tetraploids and young forms of hexaploid wheat [43] might also have contributed to the currently observed polymorphism of common wheat. A recent study showed great variation in common wheat germplasm (extensive structural rearrangements of DNA sequences and gene content) resulting from the history of wheat breeding aimed at improving different wheat characteristics, such as grain yield and quality, resistance to stresses and diseases, and adaptation to conditions of growth [44].

Analysis of the gliadin genotypes in about 1000 common wheat worldwide cultivars of the 20th century (Table S1 in [11]) showed that five main versions of heteroalleles (I, III, IV+VIII, V, VI) at the *Gli-B1* locus not only dominate in the germplasm studied but are also spatially structured around the world and probably adapted to regional eco-climatic environments [45]. We suggest that genetic diversity at the *Gli-B1* locus of common wheat germplasm studied is based mainly on the differences between variants of genotypes inherited by hexaploid wheat from its diploid donor(s) of the B genome. The donor(s) supplied common wheat with its own genetic variation of the distal part of chromosome 1B, which is now observed as heteroallelic variants at the *Gli-B1* locus. A considerable downturn in the level of genetic diversity of common wheat for this chromosome region would be caused only by the disappearance (impossible) of one of these variants.

If the level of genetic diversity of wheat germplasm of the 20th century calculated for the *Gli-B1* locus depended decisively upon several heteroallelic variants, it would not be as great as that calculated using all (more than 20 [11]) known individual alleles present in registered wheat cultivars. Indeed, several allelic members of one variant might have a mutational origin (for example, alleles of variant I) and not differ among them in any test applied in this work. Although the loss of rare alleles over time may be considered a display of genetic erosion [46], the disappearance of a rare allele (a member of one of the principal variants) would have only a minor influence on the level of genetic diversity.

#### 4. Materials and Methods

##### 4.1. Plant Materials.

The majority of grain samples of the common wheat (*Triticum aestivum* L.) cultivars used in this work were obtained from genetic and/or breeding laboratories in their countries of origin and studied earlier (Table S1 in [11]). In total, 916 registered cultivars from 18 countries were considered, in which alleles at the *Gli-B1* locus were identified [11].

For this work, some grain samples not studied earlier were taken from local collections in Italy, Spain and Ukraine. Also included in the study were grain samples not studied earlier of several registered cultivars and important genotypes (for example, Chinese Spring), plus several Spanish landraces. The widest possible, first, allelic variation at the *Gli-B1* locus and, second, origin of cultivars were thereby covered.

Names of cultivars, country of origin, year of registration and spring/winter habit were taken from the CIMMYT GRES database [47].

##### 4.2. Molecular Analysis

In the current work, some grain samples from local collections were analyzed with the aim of confirming their gliadin genotypes. To do this, an acid polyacrylamide gel electrophoresis (APAGE) routine procedure [11] was used. Two-dimensional electrophoresis (APAGE  $\times$  SDS-electrophoresis, TDE) was performed as previously described [28].

For RFLP analysis, DNA was extracted from seven-day-old seedlings) using a modified CTAB procedure [48]. DNA digestion, gel electrophoresis, blotting and hybridization were carried out as described by Gebhardt et al. [29,49]. For the identification of the  $\gamma$ -gliadin DNA sequences, the probe K32 (the cDNA clone) specifically recognizing these sequences was applied [50].

DNA was extracted from dry single seeds using the CTAB method [51]. DNA qualities were evaluated on agarose gels, and DNA concentrations were determined spectrophotometrically using an ND-1000 instrument (Thermo Fisher Scientific, Wilmington, North Carolina, USA). All extractions were diluted to 100ng/ $\mu$ l and stored at  $-20^{\circ}\text{C}$  until being used.

PCR amplification was performed in the thermocycler Analytik Jena (Flex Cycler, Jena, Germany) using GliB1.1 and GliB1.2 primers [25]. Products of amplification for each cultivar (three or more single seeds per cultivar were studied) were fractionated in 7% polyacrylamide gels (280 V) for two hours and stained as proposed by the Silver sequence

TM DNA Sequencing System Technical Manual (Promega, Madison, USA). The lengths of the amplification products were compared in neighboring lanes of the same slab gel. The approximate length of the product of amplification for each cultivar was calculated using standard markers with fragments of known length.

$\gamma$ -Gliadin gene *GAG56B* was amplified with GAG29 and GAG30 primers [24]. PCR-amplified fragments were purified with sepharose columns and sequenced by capillary electrophoresis at Macrogen (Macrogen Europe, Amsterdam, The Netherlands). Sequences were analyzed with Sequencher® version 5.0 sequence analysis software (Gene Codes Corporation, Ann Arbor, MI, USA) and aligned with acn M13712 in order to detect polymorphisms. The information obtained was used to reconstruct *GAG56B* for each cultivar and to define haplotypes. DNA sequences were aligned using clustalW [52]. Phylogenetic trees from the nucleotide alignment were calculated employing the Neighbor-joining method, with Jukes-Centre distance and 100 bootstrap.

The trees of the relationship between alleles at the *Gli-B1* locus based on the RFLP data were also designed by the Neighbor-joining method.

**Supplementary Materials:** Supplementary materials can be found at [www.mdpi.com/1422-0067/22/4/1832/s1](http://www.mdpi.com/1422-0067/22/4/1832/s1). Figure S1: Schemes of blocks of jointly inherited gliadin bands (APAGE) controlled by alleles at the *Gli B1* locus studied, Figure S2: Two dimensional (APAGE  $\times$  SDS) gliadin separations of Spanish landraces, Figure S3: The RFLP patterns (restriction Taq I) of some common wheat cultivars, Figure S4: Nucleotide sequences of the  $\gamma$  gliadin pseudogene *GAG56B* in the common wheat cultivars studied, Table S1: Cultivars with different alleles at the *Gli B1* locus studied by two dimensional (APAGE  $\times$  SDS) electrophoresis.

**Author Contributions:** E.M. designed the study; E.M., L.P., P.V., V.M., M.R.Q., Y.P., and S.C. performed experimental work and analyzed the results. E.M. and W.J.R. valorized the results and wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Spanish Ministerio de Ciencia e Innovación, grant number PID2019-109089RB-C32.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

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

## References

1. FAOSTAT (Food and Agriculture Organization of the United Nations). Available online: <http://www.fao.org/fao-stat/es/#data/QC> (accessed on 25 January 2021).
2. Marcussen, T.; Sandve, S.R.; Heier, L.; Spannagl, M.; Pfeifer, M.; Jakobsen, K.S.; Wulff, B.B.; Steuernagel, B.; Mayer, K.F.; Olsen, O.A.; International Wheat Genome Sequencing Consortium. Ancient hybridizations among the ancestral genomes of bread wheat. *Science* **2014**, *345*, 1250092.
3. Redaelli, R.; Metakovsky, E.; Davydov, S.D.; Pogna, N.E. Two-dimensional mapping of gliadins using biotypes and null mutants of common wheat cultivar Saratovskaya 29. *Hereditas* **1994**, *121*, 131–137.
4. Gobaa, S.; Bancel, E.; Branlard, G.; Kleijer, G.; Stamp, P. Proteomic analysis of wheat recombinant inbred lines: Variations in prolamin and dough rheology. *J. Cereal Sci.* **2008**, *47*, 610–619.
5. Barak, S.; Mudgil, D.; Khatkar, B.S. Biochemical and functional properties of wheat gliadins: A review. *Crit. Rev. Food Sci. Nutr.* **2015**, *55*, 357–368.
6. Qi, P.F.; Wei, Y.M.; Ouellet, T.; Chen, Q.; Tan, X.; Zheng, Y.L. The  $\gamma$ -gliadin multigene family in common wheat (*Triticum aestivum*) and its closely related species. *BMC Genom.* **2009**, *10*, 168–181.
7. Anderson, O.D.; Huo, N.; Gu, J.Q. The gene space in wheat: The complete  $\gamma$ -gliadin gene family from the wheat cultivar Chinese Spring. *Funct. Integr. Genom.* **2013**, *13*, 261–273.
8. Payne, P.I.; Holt, L.M.; Lawrence, G.J.; Law, C.N. The genetics of gliadin and glutenin, the major storage proteins in the wheat endosperm. *Qual. Plant. Plant Foods Human Nutr.* **1982**, *31*, 229–241.
9. Metakovsky, E.; Chernakov, V.M.; Upelniek, V.P.; Redaelli, R.; Dardavet, M.; Branlard, G.; Pogna, N.E. Minor  $\omega$ -gliadin-coding loci on chromosome 1A of common wheat: A revision. *J. Genet. Breed.* **1996**, *50*, 277–286.
10. Sozinov, A.A.; Poperelya, F.A. Genetic classification of prolamins and its use for plant breeding. *Ann. Technol. Agric.* **1980**, *29*, 229–245.

11. Metakovsky, E.; Melnik, V.A.; Rodriguez-Quijano, M.; Upelniek, V.P.; Carrillo, J.M. A catalog of gliadin alleles: Polymorphism of 20th-century common wheat germplasm. *Crop J.* **2018**, *6*, 629–641.
12. Shewry, P.R.; Halford, N.G. Cereal seed storage proteins: Structures, properties and role in grain utilization. *J. Exp. Bot.* **2000**, *53*, 947–958.
13. Blatter, R.H.E.; Jacomet, S.; Schlumberg, A. About the origin of European spelt (*Triticum spelta* L.): Allelic differentiation of the HMW Glutenin B1-1 and A1-2 subunit genes. *Theor. Appl. Genet.* **2004**, *108*, 360–367.
14. Huang, X.Q.; Cloutier, S. Molecular characterization and genomic organization of low molecular weight glutenin subunit genes at the *Glu-3* loci in hexaploid wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* **2008**, *116*, 953–966.
15. Gill, K.S.; Gill, B.S.; Endo, T.R.; Taylor, T. Identification and high-density mapping of gene-rich regions in chromosome group 1 of wheat. *Genetics* **1996**, *144*, 1883–1891.
16. Gao, S.; Gu, Y.Q.; Wu, J.; Coleman-Derr, D.; Huo, N.; Crossman, C.; Jia, J.; Zuo, Q.; Ren, Z.; Anderson, O.D.; et al. Rapid evolution and complex structural organization in genomic regions harboring multiple prolamin genes in the polyploid wheat genome. *Plant. Molec. Biol.* **2007**, *65*, 189–203.
17. Anderson, O.D.; Litts, J.C.; Greene, F.C. The  $\alpha$ -gliadin gene family. 1. Characterization of ten new wheat  $\alpha$ -gliadin genomic clones, evidence for limited sequence conservation of flanking DNA, and Southern analysis of the gene family. *Theor. Appl. Genet.* **1997**, *95*, 50–58.
18. Li, W.; Zhang, P.; Fellers, J.P.; Friebe, B.; Gill, B.S. Sequence composition, organization, and evolution of the core *Triticeae* genome. *Plant J.* **2004**, *40*, 500–511.
19. Charles, M.; Belcram, H.; Just, J.; Huneau, C.; Violett, A.; Couloux, A.; Segurens, B.; Carter, M.; Huteau, V.; Coriton, O.; et al. Dynamics and differential proliferation of transposable elements during the evolution of the B and A genomes of wheat. *Genetics* **2006**, *180*, 1071–1086.
20. van Herpen, T.W.J.M.; Goryunova, S.V.; van der Schoot, J.; Mitreva, M.; Salentijn, E.; Vorst, O.; Schenk, M.F.; van Veelen, P.A.; Koning, F.; van Soest, L.J.M.; et al. Alpha-gliadin genes from the A, B, and D genomes of wheat contain different sets of celiac disease epitopes. *BMC Genom.* **2006**, *7*, 1–12.
21. Harberd, N.P.; Bartles, D.; Thompson, R.D. Analysis of the gliadin multigene loci in bread wheat using nullisomic-tetrasomic lines. *Molec. Gen. Genet.* **1985**, *198*, 234–242.
22. Gu, Y.Q.; Crossman, C.; Kong, X.; Luo, M.; You, F.M.; Coleman-Derr, D.; Dubcovsky, J.; Anderson, O.D. Genomic organization of the complex  $\alpha$ -gliadin gene loci in wheat. *Theor. Appl. Genet.* **2004**, *109*, 648–657.
23. Anderson, O.D.; Dong, L.; Huo, N.; Gu, Y.Q. A new class of wheat gliadin genes and proteins. *PLoS ONE* **2012**, *7*, e52139.
24. Von Büren, M.; Lüthy, J.; Hübner, P. A spelt-specific  $\gamma$ -gliadin gene: Discovery and detection. *Theor. Appl. Genet.* **2000**, *100*, 271–279.
25. Zhang, W.; Gianibelli, M.C.; Ma, W.; Rampling, L.R.; Gale, K.R. Identification of SNPs and development of allele-specific PCR markers for gamma-gliadin alleles in *Triticum aestivum*. *Theor. Appl. Genet.* **2003**, *107*, 130–138.
26. Pogna, N.E.; Metakovsky, E.; Redaelli, R.; Raineri, F.; Dachkevitch, T. Recombination mapping of *Gli-5*, a new gliadin-coding locus on chromosome 1A and 1B in common wheat. *Theor. Appl. Genet.* **1993**, *87*, 113–121.
27. Pogna, N.E.; Lafiandra, D.; Peillet, P.; Autran, J.C. Evidence for a direct causal effect of low molecular weight subunits of glutenins on gluten viscoelasticity in durum wheats. *J. Cereal. Sci.* **1988**, *7*, 211–214.
28. Metakovsky, E.; Melnik, V.A.; Vaccino, P.; Rodriguez-Quijano, M. Comparison of alleles at the *Gli-1* loci of common wheat by means of two-dimensional electrophoresis of gliadin and RFLP analysis. *Cytol. Genet. Kiev* **2018**, *52*, 11–20.
29. Vaccino, P.; Metakovsky, E. Gliadin alleles in DNA RFLP patterns of common wheat: Implication for analysis of organization and evolution of complex loci. *Theor. Appl. Genet.* **1995**, *90*, 173–181.
30. Endo, T.R.; Gill, B.S. Somatic karyotype, heterochromatin distribution, and nature of chromosome differentiation in common wheat, *Triticum aestivum* L. em Thell. *Chromosome* **1984**, *89*, 361–369.
31. Friebe, B.; Gill, B.S. C-band polymorphism and structural rearrangements detected in common wheat (*Triticum aestivum*). *Euphytica* **1994**, *78*, 1–5.
32. Dvorak, J.; Chen, K.C. Distribution of nonstructural variation between wheat cultivars along chromosome arm 6Bp: Evidence from the linkage map and physical map of the arm. *Genetics* **1984**, *106*, 325–333.
33. Dvorak, J.; Appels, R. Investigation of homologous crossing over and sister chromatid exchange in the wheat *Nor-B2* locus coding for rRNA and *Gli-B2* locus coding for gliadins. *Genetics* **1986**, *113*, 1037–1056.
34. Dvorak, J.; McGuire, P.E. Nonstructural chromosome differentiation among wheat cultivars, with special reference to differentiation of chromosomes in related species. *Genetics* **1981**, *97*, 391–414.
35. Crossway, A.; Dvorak, J. Distribution of nonstructural variation along three chromosome arms between wheat cultivars Chinese Spring and Cheyenne. *Genetics* **1984**, *106*, 309–324.
36. Dvorak, J.; Yang, Z.L.; You, F.M.; Luo, M.C. Deletion polymorphism in wheat chromosome regions with contrasting recombination rates. *Genetics* **2004**, *168*, 1665–1675.
37. Wicker, T.; Yahiaoui, N.; Guyot, R.; Schlagenhauf, E.; Liu, Z.D.; Dubcovsky, J.; Keller, B. Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and A(m) genomes of wheat. *Plant Cell* **2003**, *15*, 1186–1197.
38. Ravel, C.; Praud, S.; Murigneux, A.; Canaguier, A.; Sapet, F.; Samson, D.; Balfourier, F.; Dufour, P.; Chalhoub, B.; Brunel, D.; et al. Single-nucleotide polymorphism frequency in a set of selected lines of bread wheat (*Triticum aestivum* L.). *Genome* **2006**, *49*, 1131–1139.



39. Ram, S.; Sharma, S.; Verma, A.; Tyagi, B.S.; Peña, R.J. Comparative analyses of LMW glutenin alleles in bread wheat using allele-specific PCR and SDS-PAGE. *J. Cereal Sci.* **2011**, *54*, 488–493.
40. Matsuoka, Y. Evolution of polyploid Triticum wheats under cultivation: The role of domestication, natural hybridization and allopolyploid speciation in their diversification. *Plant Cell Physiol.* **2011**, *52*, 750–764.
41. Metakovsky, E.; Kudryavtsev, A.M.; Iakobashvili, Z.A.; Novoselskaya, A.Y. Analysis of phylogenetic relations of durum, cartholicum and common wheats by means of comparison of alleles of gliadin loci. *Theor. Appl. Genet.* **1989**, *77*, 881–887.
42. Dvorak, J.; Luo, M.C.; Akhunov, E.D. N.I. Vavilov theory of centres of diversity in the light of current understanding of wheat diversity, domestication and evolution. *Czech J. Genet. Plant. Breed.* **2011**, *47*, S20–S27.
43. Dvorak, J.; Akhunov, E.D.; Akhunova, A.R.; Deal, K.; Luo, N.C. Molecular characterization of a diagnostic DNA marker for domesticated tetraploid wheat provides evidence for gene flow from wild tetraploid wheat to hexaploid wheat. *Molec. Biol. Evol.* **2006**, *23*, 1386–1396.
44. Walkowiak, S.; Gao, L.; Monat, C.; Haberer, G.; Kassa, M.T.; Brinton, J.; Ramirez-Gonzalez, R.H.; Kolodziej, M.C.; Delorean, E.; Thambugala, D.; et al. Multiple wheat genomes reveal global variation in modern breeding. *Nature* **2020**, doi:10.1038/s41586-020-2961-x.
45. Metakovsky, E.; Melnik, V.A.; Pascual, L.; Wrigley, C.W. Gliadin genotypes worldwide for spring wheats (*Triticum aestivum* L.). 2. Strong differentiation of polymorphism between countries and regions of origin. *J. Cereal. Sci.* **2019**, *87*, 311–317.
46. Fu, Y.B.; Peterson, G.W.; Yu, J.K.; Gao, L.; Jia, J.; Richards, K.W. Impact of plant breeding on genetic diversity of the Canadian hard red spring wheat germplasm as revealed by EST-derived SSR markers. *Theor. Appl. Genet.* **2006**, *112*, 1239–1247.
47. GRIS database (CIMMYT International Maize and Wheat Improvement Center). Available online: <http://wheatpedigree.net/> (accessed on 11 February 2021).
48. Murray, M.; Thompson, W.F. Rapid isolation of high-molecular-weight plant DNA. *Nucleic Acid Res.* **1980**, *8*, 4321–4325.
49. Gebhardt, C.; Ritter, E.; Debener, T.; Schachtschabel, U.; Walkemeier, B.; Uhrig, H.; Salamini, F. RFLP analysis and linkage mapping in *Solanum tuberosum*. *Theor. Appl. Genet.* **1989**, *78*, 65–75.
50. Bartels, D.; Altosaar, I.; Harberd, N.P.; Barker, R.F.; Thompson, R.D. Molecular analysis of gamma-gliadin gene families at the complex *Gli-1* locus of bread wheat (*T.aestivum* L.). *Theor. Appl. Genet.* **1986**, *72*, 845–853.
51. Doyle, J.J.; Doyle, J.L. Isolation of plant DNA from fresh tissue. *Focus* **1990**, *12*, 13–15.
52. Thompson, J.D.; Higgins, J.G.; Gibson, T.J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acid Res.* **1994**, *22*, 4673–4680.