Waxy genes from spelt wheat: new alleles for modern wheat breeding and new phylogenetic inferences about the origin of this species

Carlos Guzmán*, Leonor Caballero†, Luis M. Martín and Juan B. Alvarez

Departamento de Genética, Escuela Técnica Superior de Ingeniería Agronómica y de Montes, Edificio Gregor Mendel, Campus de Rabanales, Universidad de Córdoba, CeiA3, ES-14071 Córdoba, Spain

* For correspondence. E-mail ge2gugac@uco.es

†Present address: Megaseed S.A., Estrada 624, 5800 Río Cuarto, Provincia de Córdoba, República Argentina.

Background and Aims Waxy proteins are responsible for amylose synthesis in wheat seeds, being encoded by three waxy genes (Wx-A1, Wx-B1 and Wx-D1) in hexaploid wheat. In addition to their role in starch quality, waxy loci have been used to study the phylogeny of wheat. The origin of European spelt (Triticum aestivum ssp. spelta) is not clear. This study compared waxy gene sequences of a Spanish spelt collection with their homologous genes in emmer (T. turgidum ssp. dicoccum), durum (T. turgidum ssp. durum) and common wheat (T. aestivum ssp. aestivum), together with other Asian and European spelt that could be used to determine the origin of European spelt.

Methods waxy genes were amplified and sequenced. Geneious Pro software, DNAsp and MEGA5 were used for sequence, nucleotide diversity and phylogenetic analysis, respectively.

Key Results Three, four and three new alleles were described for the Wx-A1, Wx-B1 and Wx-D1 loci, respectively. Spelt accessions were classified into two groups based on the variation in Wx-B1, which suggests that there were two different origins for the emmer wheat that has been found to be part of the spelt genetic make-up. One of these groups was only detected in Iberian material. No differences were found between the rest of the European spelt and the Asiatic spelt, which suggested that the Iberian material had a different origin from the other spelt sources.

Conclusions The results suggested that the waxy gene variability present in wheat is undervalued. The evaluation of this variability has permitted the detection of ten new waxy alleles that could affect starch quality and thus could be used in modern wheat breeding. In addition, two different classes of Wx-B1 were detected that could be used for evaluating the phylogenetic relationships and the origins of different types of wheat.

Key words: Wheat, Triticum aestivum ssp. spelta, molecular characterization, phylogeny, spelt origin, waxy genes.

INTRODUCTION

The granule-bound starch synthase (GBSSI), or waxy protein, is the enzyme responsible for the synthesis of amylose in wheat grain. As some important technological starch properties, such as gelatinization, pasting and gelation, depend on the amylose:amylopectin ratio (Zeng et al., 1997), waxy protein has been the subject of many studies in recent years. Three waxy proteins are present in hexaploid wheats, which are encoded by three genes: Wx-A1 (located on chromosome 7AS), Wx-B1 (chromosome 4AL translocated from the original 7BS) and Wx-D1 (chromosome 7DS) (Yamamori et al., 1994). Each consists of 11 exons and ten introns, as Murai et al. (1999) showed on sequencing the three genes, coding for three peptides of 604, 605 and 604 amino acid residues, respectively. Given the minor differences noted in their exon sequences, the molecular weights of the three proteins are very similar; this has meant that it has been very difficult to identify allelic variants among them. However, polymorphism studies have been carried out on durum wheat (Yamamori et al., 1995) and on common wheat (Rodríguez-Quijano et al., 1998), permitting the detection of different alleles for these genes, including null ones, which have been used as a basis for breeding programmes focused on the production of amylose-free wheat (Nakamura et al., 1995; Kiribuchi-Otobe et al., 1997).

More recently the search for new alleles has been extended to ancient wheats (Urban et al., 2002; Caballero et al., 2008a; Guzman et al., 2009, 2010, 2011), which have become very important in the search for genes that could be useful in modern wheat breeding programmes since modern agricultural practices have reduced the genetic variability of cultivated wheats. One of these is spelt (Triticum aestivum ssp. spelta; 2n = 6x = 42, AABBDD), currently a minor crop that was widely cultivated in the past (Nesbitt and Samuel, 1996). Recently it has undergone a revival associated with the health food market and low-input agriculture, as this crop can grow without pesticides in marginal areas (Cubadda and Marconi, 2002). Due to the renewed interest, a large collection of Spanish spelt has been analysed for traits related to the use-quality of this species. Caballero et al. (2001, 2004a, b) found considerable variability between seed storage proteins in the Spanish spelt collection. Some of these allelic variants showed an association with high gluten strength (Caballero et al., 2004b).
et al., 2008b). Another important aspect, the starch quality, has also been analysed in this collection through polymorphism assessment of waxy proteins by SDS–PAGE (Guzman et al., 2010). Although some variation was found, most of these accessions (69-52 %) were found to have the Wx-A1a, Wx-B1a and Wx-D1a alleles, which have been catalogued as wild in common and durum wheats. Nevertheless, due to the small differences in size between the waxy proteins, it is possible that the true variability could be higher, mainly through internal differences in the amino acid sequences. This increased variability could generate enzymes with different degrees of functionality. Further evaluation of the accessions that show these wild alleles using molecular characterization could lead to the identification of these mimetic alleles.

Another interesting question about this crop that has caused a certain amount of controversy is the origin of the spelt grown in Europe. Two main hypotheses have been suggested. The first hypothesis is that European spelt is an ancestor of common wheat that spread from Asia; while the other suggests that Asian spelt is an ancestor of common wheat but European spelt had an independent origin and is derived from a second-ary hybridization between emmer wheat (T. turgidum ssp. dicoccum Schrank) and a cultivated hexaploid wheat, probably T. aestivum L. compactum Host em. Mackey (Liu and Tsuchewaki, 1991; Yan et al., 2003; Blatter et al., 2004; Dedkova et al., 2004). More recently, Dvorak et al. (2012) showed that some forms of the Asiatic spelt could also have their origin in free-threshing wheat. In addition to this, other workers (Dedkova et al., 2004; Elia et al., 2004) have suggested that European spelt could be classified into two geographical groups: the Iberian (pol. ibericum Flsk.b.) and the Bavarian (pol. bavariicum Vav.) geographical groups. These authors suggested that Iberian spelt could have its origin in Asia and be different from the other types of European spelt. Waxy gene sequences have been utilized to study the origin and phylogeny of other Poaceae species, including wheat (Mason-Gamer et al., 1998; Yan et al., 2000, 2003; Mason-Gamer, 2001; Yan and Bhave, 2001; Ingram and Doyle, 2003; Fortune et al., 2007), and, consequently, the characterization of the waxy genes in this Spanish spelt collection could shed light on this question about the origin of Spanish spelt.

The aim of the current study was the molecular characterization of the waxy genes present in this Spanish spelt collection, and to compare them with their homologous genes in emmer, durum and common wheat, together with other spelt examples of Asian and European origin that could be used for phylogenetic studies on the origin of European spelt.

MATERIALS AND METHODS

Plant material

Six Spanish and two Iranian spelt [Triticum aestivum ssp. spelta L. em. Thell.; also considered as Triticum sect. Spelta (Wolff) Dumort.] accessions obtained from three Germplasm Banks were analysed first for their waxy protein composition. The six Spanish accessions (PI348458, PI348471, PI348489, PI348515, PI348595 and PI3487447) were from the National Small Grain Collection (Aberdeen, ID, USA) while the Iranian accessions (CGN8384 and CGN12269) were obtained from the Center for Genetic Resources (The Netherlands). One further spelt accession (CGN11460, Czech Republic origin) and one emmer German accession (CGN16104) were also used in sequencing analysis.

SDS–PAGE analysis

The waxy proteins were extracted from starch granules of mature seeds following the method of Echt and Schwartz (1981) and separated by SDS–PAGE using 12 % polyacrylamide gels with low bisacrylamide concentration (approx. 0.44 %).

DNA extraction and PCR amplification

For DNA extraction, approx. 100 mg of young leaf tissue was excised, immediately frozen in liquid nitrogen and stored at −80 °C. DNA was isolated using the DNAzol® method (Invitrogen, Carlsbad, CA, USA).

The primers designed by Monari et al. (2005) were used to amplify the central region (the region spanning from the middle of exon 3 to exon 6) of the waxy genes: WxBAF (5'-ACTTCCACTGCTACAGCGCGGGGT-3') and WxBAR (5'-GCTGACGTCCATGCGGTGCACT-3'). Each 15 μl reaction included 50 ng of DNA, 1.5 mM MgCl₂, 0.2 μM of each primer, 0.2 mM dNTPs, 1.5 μL of 10× PCR buffer and 0.75 U of DNA polymerase (Promega, Madison, WI, USA). The PCR conditions included an initial denaturation step of 3 min at 94 °C followed by 35 cycles as follows: 45 s at 94 °C, 2 min at 62 °C then 1 min 5 s at 72 °C. After the 35 cycles, a final extension of 5 min at 72 °C was done. Additionally, the PCR products were restricted with the endonuclease BglI (TaKaRa) following the supplier’s instructions. Amplificied and digested products were fractionated in vertical polyacrylamide gels with an 8 % polyacrylamide concentration (w/v, approx. 1.28 %), and the bands were visualized by ethidium bromide staining.

Cloning of waxy genes and sequencing analysis

To clone almost the entire sequence of the waxy genes, besides primers WxBAF and WxBAR described above, pairs of primers WxF3 (5'-TCTGGTCACGTCCAGTCCGCCAC CT-3')/WxVT1R (5'-ACCCCCCGCTTGGACGCGAAATG-3') (Tm = 64 °C) and WxVT1F (5'-CATCTGCAACGGGATG AGTCTGAC-3')/WxVT1R (5'-CCAGAAGCGACGCTCCCA GTTCTG-3') (Tm = 64 °C) were used to amplify the beginning and the end of the waxy genes, respectively. PCR products were excised from the polyacrylamide gel and cloned into pGEM T-easy vector (Promega, Madison, WI, USA) for sequencing. Inserts were sequenced from at least three different clones using an ABI Prism 310 Genetic Analyzer (Applied Biosystems, Carlsbad, CA, USA).

The sequences reported in the current study were compared with the sequences available in the GenBank database using Geneious Pro ver. 5.0.3 software (Biomatters Ltd).
**Data analysis**

DNA analyses were conducted by DNAsp ver. 5.0 (Librado and Rozas, 2009). Nucleotide diversity was estimated as \( \theta \), the number of segregating (polymorphic) sites (Watterson, 1975), and \( \pi \), the average number of nucleotide differences per site between two sequences (Nei, 1987). Tests of neutrality were performed using Tajima’s \( D \) statistic (1989).

The synonymous (\( K_s \)) and non-synonymous (\( K_a \)) substitution rates and the relationship \( K_a/K_s \) were computed using DNAsp ver. 5.0 (Librado and Rozas, 2009). Divergence times were calculated by the mean divergence times (2.7 million years ago) between the A and D genome estimated by Dvořák and Akhunov (2005).

The phylogenetic tree was constructed with the software MEGA5 (Tamura et al., 2011) using the complete coding regions. A Neighbor–Joining cluster with all sequences analysed was generated using the Maximum Composite Likelihood method (Tamura et al., 2004), and one bootstrap consensus from 1000 replicates was used (Felsenstein, 1985).

**RESULTS**

**Electrophoretic analysis**

As shown by previous results obtained with this collection (Guzmán et al., 2010), no differences could be detected by SDS–PAGE between the different Spanish spelt genotypes and the common wheat (cv. Chinese Spring) used as a standard (Fig. 1). The data obtained by one-dimensional SDS–PAGE separation of the waxy proteins showed that the Spanish spelt (Fig. 1A, lanes 1–2, 5–8) had the same pattern as the common wheat (cv. Chinese Spring), which presented the wild alleles for all three genes (Wx-A1a, Wx-B1a and Wx-D1a). Furthermore, the analysis included two Asiatic spelt of Iranian origin that showed the same pattern of waxy proteins (Fig. 1A, lanes 3–4). This similarity was also observed when the samples were separated by 2-D isoelectric focusing × SDS–PAGE separation (Fig. 1B, C).

However, the comparison of these supposedly wild alleles by PCR amplification of genomic DNA and posterior electrophoretic migration showed two groups within the genotypes analysed. The most important differences between the samples analysed were detected when the central region of the waxy genes was amplified (Fig. 2A). In this figure, the samples used were the same as those shown in Fig. 1, and appear in the same order. Two Spanish and two Iranian spelt samples showed three bands (Fig. 2A, lanes 1–4), each one corresponding to one of the three identified waxy genes, of the expected size of: Wx-D1 (1017 bp), Wx-A1 (958 bp) and Wx-B1 (935 bp), based on the published sequences of Wx-A1a, Wx-B1a and Wx-D1a alleles of cv. Chinese Spring (Murai et al., 1999). This fact meant that they were similar to the common wheat (cv. Chinese Spring; Fig. 2A, lanes 1–3, 4 and 5, respectively). However, the other spelt samples (Fig. 2A, lanes 5–8) had only two bands, as the Wx-A1 and Wx-B1 bands had co-migrated. The first group was called spelt type I and the second spelt type II.

To clarify these band patterns and to confirm the presence of the two spelt groups, the PCR products were digested using the endonuclease, BgII. Based on the published sequences for common wheat (cv. Chinese Spring), only one sequence in the Wx-B1 gene was targeted. Two bands of 897 and 38 bp were expected from the digestion of the Wx-B1 spelt I gene. While the 897 bp band was found in all samples, the 38 bp band was lost due to its small size (Fig. 2B). In the spelt II lines, one additional band was separated from the Wx-A1 band after the digestion (Fig. 2B, lanes 5–8), confirming the presence of a larger Wx-B1 band that had co-migrated with the Wx-A1 band shown in Fig. 2A.

**Sequences of spelt waxy genes and deduced proteins**

In order to characterize the variation found by electrophoretic methods, the three waxy genes from each Spanish spelt type (I and II) and the two Iranian samples were sequenced. Waxy genes from one European spelt accession and one emmer accession were also sequenced. All these sequences were uploaded to the GenBank database (Table 1). Comparisons between these sequences and others from emmer, durum and common wheat previously available in GenBank were undertaken. The criteria used to choose these materials was to compare potential material with the different origins of the spelt (Asian and European) and the putative progenitors that may be involved in the different crossing events that generated spelt (emmer and common wheat).

A summary of DNA polymorphism found in the Wx sequences evaluated is shown in Table 2. These data were evaluated using both the complete sequence (exons + introns) and just the coding sequence (exons). In both cases, the Wx-B1 gene displayed the highest level of single nucleotide polymorphisms (SNPs) among the three genes, with 163 polymorphic sites in the complete sequence and 53 in the coding region (Table 2). In the latter, 36 of these nucleotide substitutions were synonymous (silent mutations) while the rest were non-synonymous, which implied changes in the amino acid sequences. The ten alleles identified encoded seven different polypeptides (Table 3). This polymorphism was clear lower down in the coding regions of the other two genes (Tables 2 and 3). The Wx-A1 gene coding region showed 15 nucleotide substitutions (eight synonymous and seven non-synonymous), while the Wx-D1 gene only presented two substitutions that were synonymous. Consequently, only four and one polypeptides were obtained for the seven and two alleles detected in the coding regions of the Wx-A1 and Wx-D1 genes, respectively (Table 3).

Two statistics, \( \pi \) (Nei, 1987) and \( \theta \) (Watterson, 1975), were used to estimate nucleotide diversity (Table 2). Both values were similar in all the cases that were associated with a drift–mutation balance. The values for the Wx-B1 gene were higher than in the other two genes. Tajima’s \( D \) values were not significant for the three genes, which was consistent with a neutral equilibrium (Table 2). The value was negative for the Wx-A1 gene, indicative of an excess of low frequency alleles, while the other two genes showed positive values, which indicated an excess of intermediate frequency alleles.

The data shown in Table 3 suggested that the previous classification of waxy genes in Spanish spelt was incorrect. In all cases, only the class 1 DNA and protein sequences corresponded to the wild allele for each locus. For the Wx-A1 gene, six classes, previously classified as Wx-A1a, could be
reclassified because the nucleotide sequences were different from those in the Wx-A1a allele present in common wheat. However, four of them synthesized a polypeptide with the same amino acid sequence, which included all the Spanish spelt types evaluated (Table 3).

The Wx-D1 gene showed up to four different nucleotide sequences, although only two classes were observed when the coding region was analysed. In all cases, one unique polypeptide was synthesized but the classification of this gene must indicate the presence of the four alleles (Table 3).

The differences in the Wx-B1 gene were larger than in the other two genes. Up to seven nucleotide sequences were detected in the materials previously classified as wild (Wx-B1a). Six were different from the true Wx-B1a allele (class 1) and their translation indicated the presence of four additional polypeptides. Two of these polypeptides were associated with each one of the classes detected in the Spanish spelt (type I and II). It is important to emphasize that although the nucleotide sequences were clearly different, one of these two polypeptides was the same as those detected previously in the BGE-012302 accession of the emmer that had presented the Wx-B1g allele (Guzman et al., 2011).

With respect to the Spanish spelt lines evaluated in the current study, the spelt type I had a homology of 99.5% for...
the Wx-A1 gene compared with the Wx-A1 gene from common wheat. Spelt type II also showed a homology of 99.6% with respect to spelt type I and common wheat, although the SNPs were not the same when compared with spelt type I or with common wheat. Specifically, the deduced Wx-A1 protein from the two spelt types was similar and showed three amino acid differences compared with the Wx-A1 protein from common wheat (Wx-A1a allele): Phe60 → Gly, Asp61 → Asn and Trp454 → Arg. This meant that two alleles were identified within the allele previously catalogued as Wx-A1a. A third polypeptide was detected in the wild emmer that contained two amino acid changes (Glu333 → Arg), while a fourth polypeptide was observed in one German spelt and one Iranian emmer that contained two amino acid changes (Glu333 → Arg, Ile and Gly, and Trp454 → Arg), while a fourth polypeptide was observed in one German spelt and one Iranian emmer that contained two amino acid changes (Glu333 → Arg, Ile and Gly, and Trp454 → Arg). However, for the Wx-D1 gene, both types of spelt showed a similarity of 99.9% to common wheat at the DNA level and no variation was detected in the protein sequences.

### Table 1. Lines of wheat used in this study

<table>
<thead>
<tr>
<th>Species</th>
<th>Ploidy</th>
<th>Genome</th>
<th>Cultivar/ accession*</th>
<th>Origin</th>
<th>Gene and GenBank no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Triticum monococcum ssp. monoccum</td>
<td>2x</td>
<td>AABB</td>
<td>AUS22986</td>
<td>Australia</td>
<td>Wx-A1*: AF110373</td>
</tr>
<tr>
<td>Triticum urartu</td>
<td>2x</td>
<td>ABB</td>
<td>MG26992</td>
<td>Iraq</td>
<td>Wx-A1*: JN857937</td>
</tr>
<tr>
<td>Aegilops speltoides</td>
<td>2x</td>
<td>SSS</td>
<td>AUS21638</td>
<td>Iraq</td>
<td>Wx-S1*: AF110374</td>
</tr>
<tr>
<td>Aegilops tauschii</td>
<td>2x</td>
<td>DD</td>
<td>CP1110799</td>
<td>–</td>
<td>Wx-D1*: AF110375</td>
</tr>
<tr>
<td>Triticum turgidum ssp. dicoccoides</td>
<td>4x</td>
<td>AABB</td>
<td>–</td>
<td>–</td>
<td>Wx-A1*: AB029061</td>
</tr>
<tr>
<td>Triticum turgidum ssp. dicoccum</td>
<td>4x</td>
<td>AABB</td>
<td>CGN16104</td>
<td>Germany</td>
<td>Wx-A1*: JN935600</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PI275996</td>
<td>Spain</td>
<td>Wx-A1*: HM751941</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BGE012302</td>
<td>Spain</td>
<td>Wx-A1*: GQ205417</td>
</tr>
<tr>
<td>Triticum turgidum ssp. durum</td>
<td>4x</td>
<td>AABB</td>
<td>–</td>
<td>–</td>
<td>Wx-A1*: AB029063</td>
</tr>
<tr>
<td>Triticum aestivum ssp. aestivum</td>
<td>6x</td>
<td>AABDD</td>
<td>CGN8384</td>
<td>China</td>
<td>Wx-A1*: AB019626</td>
</tr>
<tr>
<td>Triticum aestivum ssp. spelta</td>
<td>6x</td>
<td>AABDD</td>
<td>CGN11460</td>
<td>Czech Rep.</td>
<td>Wx-A1*: JN935596</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CGN11461</td>
<td>Germany</td>
<td>Wx-A1*: JN935598</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CGN112269</td>
<td>Iran</td>
<td>Wx-A1*: JN935591</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PI348458</td>
<td>Spain</td>
<td>Wx-A1*: HQ338723</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PI348471</td>
<td>Spain</td>
<td>Wx-A1*: HQ338714</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PI348489</td>
<td>Spain</td>
<td>Wx-A1*: HQ338717</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PI348515</td>
<td>Spain</td>
<td>Wx-A1*: HQ338726</td>
</tr>
</tbody>
</table>

* BGE, Centro de Recursos Fitogenéticos-INIA (Alcalá de Henares, Spain); CGN, Center for Genetic Resources (Wageningen, The Netherlands); PI, National Small Grain Collection (Aberdeen, ID, USA).

Bold text, sequenced in this study; italic text, sequenced in our laboratory in previous works. The remaining accession numbers were obtained from GenBank.

### Table 2. Summary of DNA polymorphism and test statistics for selection of the 38 sequences from polyploid wheat evaluated

<table>
<thead>
<tr>
<th>Gene</th>
<th>n</th>
<th>k</th>
<th>s</th>
<th>h</th>
<th>θ × 10⁻³</th>
<th>π × 10⁻³</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wx-AI</td>
<td>14</td>
<td>5-67</td>
<td>22</td>
<td>7</td>
<td>2.6</td>
<td>2.1</td>
<td>−0.779 ns</td>
</tr>
<tr>
<td>Wx-BI</td>
<td>15</td>
<td>54-40</td>
<td>163</td>
<td>10</td>
<td>18.9</td>
<td>20.5</td>
<td>0.376 ns</td>
</tr>
<tr>
<td>Wx-BI (type I)</td>
<td>12</td>
<td>5-6</td>
<td>22</td>
<td>8</td>
<td>2.7</td>
<td>2</td>
<td>−1.019 ns</td>
</tr>
<tr>
<td>Wx-BI (type II)</td>
<td>3</td>
<td>6-66</td>
<td>10</td>
<td>2</td>
<td>2.4</td>
<td>2.4</td>
<td>−0.232 ns</td>
</tr>
<tr>
<td>Mean</td>
<td>7.3</td>
<td>7.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>n</th>
<th>k</th>
<th>s</th>
<th>h</th>
<th>θ × 10⁻³</th>
<th>π × 10⁻³</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coding region</td>
<td>14</td>
<td>5-67</td>
<td>22</td>
<td>7</td>
<td>2.6</td>
<td>2.1</td>
<td>−0.779 ns</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>54-40</td>
<td>163</td>
<td>10</td>
<td>18.9</td>
<td>20.5</td>
<td>0.376 ns</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>5-6</td>
<td>22</td>
<td>8</td>
<td>2.7</td>
<td>2</td>
<td>−1.019 ns</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>6-66</td>
<td>10</td>
<td>2</td>
<td>2.4</td>
<td>2.4</td>
<td>−0.232 ns</td>
</tr>
<tr>
<td>Mean</td>
<td>7.3</td>
<td>7.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

n, number of sequences; k, average number of nucleotide differences; s, number of polymorphic sites; h, number of haplotypes; θ, Watterson’s estimate; π, nucleotide diversity; D, Tajima’s estimate D-test; ns, not significant.

The total length of the sequences was 2695–2690 bp for Wx-A1, 2702 for Wx-BI type I, 2717 for Wx-BI type II and 2771 for Wx-D1; the length of the coding regions analysed was 1724 bp, except for Wx-BI type I that was 1727 bp.
In the case of the Wx-B1 gene, the differences detected were remarkable. Most of the differences observed were found in introns, but 53 of the 163 SNPs were also present in the coding region, which generated up to ten Wx-B1 alleles. The Spanish spelt lines divided into two different groups. Spelt type I presented one polypeptide corresponding to the Spanish spelt lines divided into two different groups. Spelt type II presented one polypeptide corresponding to the Spanish emmer, although the nucleotide sequences showed some differences between them. One similar case was observed with spelt type II and cv. Langdon. The 14 amino acid differences between both types of polypeptides were remarkable (Fig. 3): Ser34 → Thr, Ala76 → Val, Thr45 → Ile, Glu54 → Asp, Thr62 → Ser, Ala76 → Gly, Ser246 → Asn, Arg250 → Met, Ala358 → Thr, Ala365 → Val, Ser451 → Asn, Met510 → Val and Glu554 → Gly. In a previous study based on protein analysis (Yamamori et al., 1995), the Wx allele present in cv. Langdon was classified as Wx-B1a, probably due to a similar problem to the one that arose with the Spanish spelt in the current study. However, the current analysis showed that both alleles were very different. Spelt type I lines showed only one difference with respect to common wheat: Arg520 → His.

In summary, analyses of the Spanish spelt lines that were previously classified as wild genotypes (Wx-A1a, Wx-B1a and Wx-D1a) have shown that, although they presented one unique allele for the Wx-A1 gene, this was not the Wx-A1a allele. Furthermore, the Spanish spelt lines showed two different alleles for the Wx-B1 gene, neither of which was the Wx-B1a allele.

### Phylogenetic analysis

A phenogram based on the Maximum Composite Likelihood method was constructed using all the waxy sequences evaluated in this study, together with the putative donors from the genomes present in tetra- and hexaploid wheats (Fig. 4). Triticum urartu Thum. ex Gandil (A"A", Wx-A¹1: JN857937) and einkorn (T. monococcum L. ssp. monococcum; A"A"M", Wx-A¹m1: AF110373) were included from the A genome; Aegilops speltoides Tausch. (S"S", Wx-S¹1: AF110374) was included from the B genome and A. tauschii Coss. (DD, Wx-D1; AF110375) was included from the D genome. The total nucleotide sequence (exons + introns) was used in all cases.

Seven groups were observed in the dendrogram (Fig. 4). Two of them corresponded to the Wx-A1 and Wx-D1 genes, respectively. With respect to the putative donors from the wheat genomes, only the Wx-D1 gene from A. tauschii showed a narrow relationship with the D genome. The other three species were clearly separated from the genomes with which they had been associated. Both species with the A genome (T. urartu and einkorn) were clearly separated from the Wx-A1 gene present in the tetra- and hexaploid wheats evaluated in the current study.

Another important result was that the Wx-B1 gene showed two separated groups. Both groups have been associated with different types of the Wx-B1 gene detected in the spelt groups (type I and II). Inside each group, the differences between nucleotide sequences were small compared with the differences detected between groups.

The Ks and Ka substitution rates among Wx genes both for homeologous (Wx-A1, Wx-B1 and Wx-D1) as for orthologous (Wx-B1 type I and Wx-B1 type II) genes were calculated by using the coding sequence of the complete gene. The comparison values of the homeologous genes were approx. 3-fold higher than the value obtained for the two types of Wx-B1 genes (Table 4). Consequently, the divergence time between both Wx-B1 gene types was estimated as approx. 0.6 million
years ago (Table 4), based on the mean divergence time for the separation between the A and D genome (2.7 million years ago) estimated by Dvorak and Akhunov (2005).

**DISCUSSION**

The major nutritional component in wheat grains is starch, which is formed by two glucose polymers: amylose and amyllopectin, whose synthesis involves up to five starch synthases (Baldwin, 2001). The variation in the ratio between both polymers, together with their chemical properties, is important for defining the end-use of a specific wheat flour type. Starch with high amylose content could be used to create healthier foods because the amylose is digested more slowly in the small intestine, providing beneficial effects for human health (Topping and Clifton, 2001). However, wheat containing amylose-free starch has been reported to improve noodle quality (Oda et al., 1980) and be more efficient than standard wheat if the grain is used as a substrate for bioethanol production (Wu et al., 2006). Consequently, the search for different forms of starch synthases has increased so that new resources can be made available for breeding programmes focused on starch properties.

In the last 20 years, the waxy proteins have been the main subject of many studies, mainly focused on the search for null forms of these proteins or mutants showing less activity (Yamamori et al., 1994; Rodríguez-Quijano et al., 1998; Guzmán et al. — *Spelt waxy genes: new alleles and phylogenetic inferences*. 1167

| Common wheat  | 3 LVTSQ/LATSGTVL/GITDFRRAFGQVPRSPADAPL/GMETG/AAAPKQQSRKARH/GTRCLSHVVRATG/NSNLF/VFGAEMAPW/SKTG/LLGDLV |
| Spanish spelt type I | 3 ............................................................. .A.V.I. ........................................... .S ........................................... .G. |
| Spanish spelt type II | 3 ............................................................. .N .....A.V.I. ........................................... .S ........................................... .G. |
| Common wheat  | 100 GGLPAMAANGHRV/NISPRYQYKDAMDTSVSVSE/IAVDEYERVYFHCYRGVDVFVDPHCFL/ERGKTKEIKY/GPDAGTDYEDNQLRSILC |
| Spanish spelt type I | 100 ............................................................. ........................................... ........................................... |
| Spanish spelt type II | 99 ............................................................. ........................................... ........................................... |
| Common wheat  | 197 QAALAP/RILD/LNSNNP/YSFGPGEDVFV/CMDWSTGULLACYLS/NQSSG/IIYRTAKVAP/CINISYQGRFSFDDPAQLNL/PDRFSSDFID/GYDFKP |
| Spanish spelt type I | 197 ............................................................. ........................................... ........................................... |
| Spanish spelt type II | 196 ............................................................. ........................................... ........................................... |
| Common wheat  | 294 VEGSKIN/SKAGILQADFVLTV/SPPYEELIS/GEARGC/ELNIMTLGTGTGIVNGHDE/SWDAKREDKFLAAMYDVTPIAE/GRSKEALQAEVGLPV |
| Spanish spelt type I | 294 ............................................................. ........................................... ........................................... |
| Spanish spelt type II | 293 ............................................................. ............ .T...V. ........................................... |
| Common wheat  | 391 DRKVPLVAFGRIE/EEQGPDVMIAIPEILKEEDQV/ULGTO/GKKEF/ELLKSVE/KEFKPSKV/RAVRPHAPL/AQMAGADVLAVT/SPPFCGL/IQ |
| Spanish spelt type I | 391 ............................................................. ........................................... ........................................... |
| Spanish spelt type II | 390 ............................................................. ........................................... ........................................... |
| Common wheat  | 488 QHRYGT/PCACASTGQLVDITIM/EKTGHMRLSV/DCIVPEADVKKVVT/TLKRA/VXGTPAYHENVKN/CQDLSWKGFAENNEDVL/LELGV |
| Spanish spelt type I | 488 ............................................................. ........................................... ........................................... |
| Spanish spelt type II | 487 ............................................................. ........................................... ........................................... |

**FIG. 3.** Diagrammatic representation of the Wx-B1 protein sequences from the Spanish spelt types together with common wheat.
Yanagisawa et al., 2001). However, the characterization of the apparently functional enzymes has been less well studied because the dosage effect of these proteins makes it difficult to evaluate the specific effect of each variant on the amylose content. In this context, the current study characterized, at a molecular level, the waxy genes from different spelt accessions.
that had not shown any differences when they were previously evaluated for waxy protein polymorphism by SDS–PAGE (Guzman et al., 2010).

Although other studies have found high levels of nucleotide diversity for these genes in common wheat (Huang and Bülé-Babel, 2012), it is important to emphasize that the allelic variation found in the current study was detected in materials that were previously classified as similar when they were analysed by SDS–PAGE. This suggested that a greater part of the Wx gene variation could have emerged due to the similarity between the different synthesized proteins. However, the variation found in the nucleotide sequences was considerably higher. Seven, ten and four different alleles for the Wx-A1, Wx-B1 and Wx-D1 genes, respectively, were detected where only one polypeptide was detected by SDS–PAGE for each gene. This variation at the DNA level was finally confirmed by the existence of three and six novel waxy alleles for the Wx-A1 and Wx-B1 proteins, respectively. The new alleles detected could have a different functionality from that of the wild-type alleles and thus affect the composition (amylose/amyllopectin) and functionality of starch. This fact should be tested in further experiments transferring these alleles to the same genetic background and evaluating their use for modern wheat breeding.

In addition to the potential improvements that these mutations could provide in starch quality, variation in the nucleotide sequence of the Wx genes has been identified as a useful tool in phylogenetic analysis (Mason-Gamer et al., 1998, Yan et al., 2000; Mason-Gamer, 2001; Ingram and Doyle, 2003; Fortune et al., 2007). The Wx-A1, Wx-B1 and Wx-D1 spelt sequences were compared with the homologous genes in the putative donors of A, B and D genomes in polyploid wheats to shed more light on the origin of each of the genomes. The current data only supported the hypothesis that suggested that A. tauschii is the D genome donor (Dvorak et al., 1998), whereas there were no data to support the second hypothesis about the origin of the A and B genomes. The current theory on the origin of the A genome in hexaploid wheat suggests that the wild wheat species (T. urartu) could be the A genome donor (Dvorak et al., 1993). However, the Wx-A*1 sequence obtained in a previous study (Guzman and Alvarez, 2012) was clearly different from the rest of the Wx-A1 sequences (Fig. 4), with the exception of the Wx-A***1 gene sequenced in einkorn by Yan et al. (2000). Similar results were found when the homology between the Wx-S1 gene from A. speltoides, the main candidate for the B genome (Petersen et al., 2006), and the Wx-B1 gene were evaluated. Although these results should be treated with caution, due to the fact that only one sequence of A. speltoides was used, these data suggested that the homology among these sequences was low and this proposed origin was not clear based on the information derived from waxy genes.

However, it is remarkable that the Wx-B1 genes from the polyploid wheats evaluated in this study clearly separated into two different groups. The Wx-B1 type I group had a greater similarity to the Wx-S1 gene than the type II group. Although the type II group had been associated with one durum wheat cultivar (cv. Langdon), the genealogy of this cultivar [Langdon = Carleton/Mindum/khapli/3/Carleton/Mindum/Stewart/4/Stewart] by Zeven and Zeven-Hissink (1976) indicated that it had been generated using several lines of emmer wheat (Carleton = vernal emmer/Mindum, Khapli = emmer and Stewart = Mindum/2* vernal emmer). This suggested that there were two different origins for the emmer wheat that has been found to be part of the spelt genetic make-up, although, to date, the second origin has only been found in the Spanish spelt. These results confirm the unique nature of the Iberian spelt gene pool indicated by other studies (Dedkova et al., 2004; Elia et al., 2004), who suggested that Iberian spelt (pol. ibericum Flaskh.) could have its origin in Asia and be different from the rest of the European spelts, which are said to derive from a secondary hybridization between emmer and a cultivated hexaploid wheat, probably club wheat, as several workers have proposed (Liu and Tsunewaki, 1991; Yan et al., 2003; Blatter et al., 2004; Dedkova et al., 2004). This last hypothesis has not been totally verified (for a review, see Salamini et al., 2002) as the current data have not revealed major differences between the Wx sequences for the three genomes of European and Asian spelt evaluated. In the same way, a recent hypothesis exposed by Dvorak et al. (2012) that suggested that all spelt (except the Iranian one) was certainly derived from a common hexaploid ancestor could not be proved with our data as the Iranian accessions analysed in this study were very similar to the rest of the accessions. In addition to this, we also found two different groups based on B genome differences, a fact that could disagree with their hypothesis. The same authors also suggested that the hexaploid ancestor could have its origin in the cross of A. tauschii with a free-threshing tetraploid, and not emmer wheat. This fact could also be revised in further surveys with waxy data from more tetraploid species, such as durum wheat. On the other hand, other studies have estimated that the time of the origin of tetraploid wheat could be between 0.36 and 0.50 million years ago (Dvorak and Akhunov, 2005; Huang et al., 2002, respectively). Data obtained in the current study showed that the separation between both types of Wx-B1 gene could have emerged during the synthesis of the tetraploid wheat. Although this hypothesis would be confirmed with further studies, this opens up the possibility that the origin of these wheats could be a consequence of different events that would have effects on the origin of the spelt.

Further studies should be carried out in the future with these genes and others, with emmer, durum, spelt and common wheat accessions from as many locations as possible as well as the supposed ancestral genomes that may have contributed to tetraploid and hexaploid wheat (T. urartu, einkorn and
Aegilops sp.) in order to clarify further the origin of spelt as well as the other cultivated wheats.

Conclusions

The first notable outcome from this study concerning the origin of Iberian spelt is that both spelt I and spelt II showed the same Wx-A1 protein sequence as emmer and durum wheat and varied only slightly with respect to common wheat. The results of this study suggested a single origin for spelt and that common wheat developed subsequently. Nevertheless, the findings in relation to the Wx-B1 genes supported the idea that, at least in the Iberian Peninsula, spelt could have a double phylogenetic origin.

The results from the current survey also suggested that the Wx gene variability seen in wheat could be underestimated because some of these variants were not detected by the traditional classification based on SDS–PAGE. The evaluation of this variability by the current study detected two different classes of Wx-B1 gene, which could be used for evaluating the phylogenetic origins of and relationships between different wheat species.

ACKNOWLEDGEMENTS

This research was supported by grant AGL2010-19643-C02-01 from the Spanish Ministry of Economy and Competitiveness, co-financed with the European Regional Development Fund (FEDER) from the European Union.

C.G. expresses his appreciation to Alessandra di Francesco, Andrea Aglieco, Tülin Sarigul and Ana Moral for their technical assistance.

LITERATURE CITED


Zeven AC, Zeven-Hissink NCh. 1976. Genealogies of 14,000 wheat varieties. Netherlands Cereals Centre (Wageningen)—International Maize and Wheat Improvement Center (Mexico).