Missense alleles of the HMW glutenin subunits Dx5 and Dy10 have small changes in function relative to missense changes in Puroindoline a and b

John M. Martin, Jinrui Zhang, Andrew C. Hogg, Michael J. Giroux

Missense Alleles of the HMW Glutenin Subunits Dx5 and Dy10 Have Small Changes in Function Relative to Missense Changes in Puroindoline a and b

John M. Martin[1,2], Jinrui Zhang[1,3], Andrew C. Hogg[1], & Michael J. Giroux[1]

[1]Department of Plant Sciences and Plant Pathology, 119 Plant Bioscience Building, Montana State University, Bozeman, MT 59717-3150, USA

[2]Corresponding author. Email: jmmartin@montana.edu Phone: (406) 994-5057

[3]Current address: Information, Data and Analytics Capability Center, Shell information technology international BV Postbus 162, 2501 AN Den Haag, the Netherlands.

Funding Information

This project was supported by the Montana Board of Research and Commercialization Technology, the Montana Wheat and Barley Committee, and the Montana Agricultural Experiment Station.

ABSTRACT

Background and Objectives: The high molecular weight glutenin subunit genes (HMW-GS) in wheat (*Triticum aestivum* L.) play a key role in determining dough functionality. More specifically allelic variation for the *Glu-D1* subunit loci is consistently associated with bread making quality. Since much of the hard wheat germplasm is fixed for the more favorable *1Dx5+1Dy10* haplotype, our goal was to identify allelic variation in 1Dx5 and 1Dy10 HMW-GS genes using an ethyl methanesulfonate (EMS) mutagenized population in the soft white spring wheat cultivar 'Alpowa'. The same source population was previously used in screening for create novel alleles in Puroindoline a (*Pina*) and Puroindoline b (*Pinb*) which are responsible for grain hardness variation. Direct sequencing of 384 M_3 families identified 135 mutations equally dispersed across 1Dx5 and 1Dy10. The mutation discovery rate was 1/12.7 kb of DNA, equal to that found in the *Pin* loci. Mutation carrying plants were crossed to the non-mutagenized Alpowa parent to create F_2 populations segregating for the induced EMS HMW Glutenin mutations. Swelling index of gluten (SIG) was measured from seed from field grown plants to assess the impact of each EMS mutation upon Glutenin function.

Findings: Nonsense mutations in both 1Dx5 and 1Dy10 reduced SIG values, indicating both 1Dx5 and 1Dy10 are needed for gluten function. Missense mutations did not alter SIG values compared to their wild type counterparts. In contrast missense mutations within the tryptophan-rich region of both PINS increased grain hardness pointing to the importance of this protein region.. No such region in 1Dx5 or 1Dy10 was found indicating these proteins are structurally stable and tolerant of amino acid substitution.

Conclusions: Results from the *Glu-D1* subunit and *Pin* loci show that useful phenotypic variation can be created using EMS mutagenesis for genes where naturally occurring mutations occur from amino substitutions as is the case for the *Pin* loci.

Significance and Novelty: EMS induced mutations are useful for creating new alleles in storage proteins with much greater impacts in genes that have a defined active site as for the Pin loci than in genes such as the HMW *Glu-1* genes which have a long repetitive region.

KEYWORDS: gluten, puroindoline, semi-dwarf, starch, wheat

1. INTRODUCTION

Milling of wheat (*Triticum aestivum*, L.) grain into flour involves the separation of bran and germ from the endosperm and grinding and sifting endosperm particles into flour. The efficiency of the milling process in terms of proportion of flour recovered and energy required to mill grain has obvious economic implications. Since flour is the basic ingredient for wheat-based products, flour quality dictates end product quality. Wheat breeders have selected cultivars with improved milling and flour quality properties, but breeding programs can only choose from available variation, and for some important traits cultivars that would be useful to improve milling and end product quality are not available. Two genes that have major impacts on the amount and quality of flour are the *Puroindolines*, which influence grain hardness and flour yield, and *Glutenins*, which impact the protein quality in the flour and subsequent dough mixing and handling qualities.

1.1 Grain Hardness and Puroindolines

Grain hardness in wheat is controlled mainly by the *Ha* locus on chromosome 5D, which explains much of the seed texture variation between soft and hard wheats (Campbell et al. 1999). It is comprised of the closely linked genes Puroindoline a (*Pina)* and Puroindoline b (*Pinb)* which code for the proteins PINA and PINB, respectively. Puroindolines are small (~13 Kda), basic, cysteine-rich, endosperm-specific proteins consisting of 148 amino acids (Day et al. 2006; Dubreil et al. 1998) that reside on the surface of starch granules (Feiz et al. 2009c). PINs interact to form friabilin, which is found in greater abundance on the surface of soft than on hard wheat starch (Greenwell and Schofield 1986; Hogg et al. 2004).

Early investigations on the molecular basis for hard grain showed hard grain resulted from a glycine to serine mutation in puroindoline b (allele *PinB-D1b*), or the complete absence

of the puroindoline a protein (allele *Pina-D1b*) (Giroux and Morris, 1997 and 1998). Subsequent studies have shown that all soft hexaploid wheats have functional *Pina* and *Pinb* (*Pina-D1a/Pinb-D1a*), while hard wheat results when there is a sequence alteration in one or both *Pin* genes. An early germplasm survey of North American wheat showed ~75% of hard wheats had the *Pinb-D1b* mutation and ~25% had the *Pina-D1b* mutation (Morris et al. 2001). Numerous germplasm surveys from around the world have identified several mutations in both *Pina* and *Pinb* that confer hard wheat texture. But many of the alleles are rare and are not deployed in widely grown cultivars.

Wheat is classified for marketing purposes as soft or hard. Grain hardness is routinely measured using the Single Kernel Characterization System (SKCS) that provides a unitless hardness index (HI). Soft wheats have an HI below 50, while hard wheats have an HI above 50. Soft wheat flour is used primarily for cakes, cookies, and pastries, while hard wheat flour is used most often for yeast leavened products (reviewed in Pomeranz and Williams, 1990). Even though allelic variation exists within hard wheats, studies using near-isogenic lines for several *Pin* alleles showed the total range in grain hardness within hard wheats is only about 10 HI units with *Pina-D1b* (*Pina* null) ranking as the allele conferring the highest grain hardness (Ma et al. 2009 and Takata et al. 2010). But even small differences in grain hardness can impact milling and end-use quality traits. Martin et al. (2001) showed that seeds containing the *Pina-D1a/Pinb-D1b* haplotype were six HI units softer but showed higher total flour and break flour yield with lower ash and greater loaf volume compared to those containing the harder *Pina-D1b/Pinb-D1a* haplotype. Ma et al. (2009) found similar increased flour yield for these haplotypes using near isogenic lines.

1.2 Glutenins and Dough Mixing Properties

Gluten has the largest impact on dough and bread making properties in hexaploid wheat. Gluten is comprised of multiple storage proteins: gliadins and high molecular weight (HMW) and low molecular weight (LMW) glutenins. The HMW glutenins have been shown to have a large influence on wheat dough quality (Blechl et al. 2007; Branlard et al. 2001; Branlard and Bellot 1983; and Graybosch et al. 2011). The HMW glutenin subunits (HMW-GS) are encoded by genes residing at loci designated *Glu-1* present on the long arm of the group 1 chromosomes 1A, 1B, and 1D (Payne et al. 1980; 1981; 1984; and 1987). Each HMW glutenin locus is comprised of both a larger *x*-subunit and a smaller *y*-subunit based on their molecular weight (Harberd et al. 1987), where the *x* and *y* subunit genes are inseparably linked. These genes code for large (67-90 Kda), endosperm specific proteins. The primary structure of all HMW-GS proteins is composed of non-repetitive N- and C-terminal domains and a central repetitive domain (Shewry et al. 2003). Three to five conserved cysteine residues are present on the N-terminus and one on the C-terminus. The repetitive domain varies in length and number of repeated fragments with an abundance of proline, glycine, and glutamic acid, and is low in lysine. Dough strength and elasticity are both affected by the size of the repetitive domain, and cysteine residues provide disulfide bonds between HMW and LMW-GS to form protein polymers (Halford et al. 1987; Tatham et al. 1990; Gianibelli et al. 2001).

Naturally occurring allelic variants have been identified for all three *Glu-1* loci. See website (http://wheat.pw.usda.gov/ggpages/wgc/98/Protein3a.htm#Glutenins verified October, 2022). Allelic variation at the *Glu-1* loci results from the number and size of repeated fragments rather from point mutations as for the Puroindolines. Differences in dough handling properties are attributed to the presence of allelic variation at each of the three *Glu-1* loci (Shewry et al.

2003). The presence or absence of subunits and the variation within glutenin subunits have all been investigated to show the strong influence that glutenins have on bread quality (Payne et al. 1979; Shewry et al. 2003; Weegles et al. 1996). The *Glu-D1* locus has the largest effect on bread/dough properties. More specifically, the *1Dx5+1Dy10* haplotype is superior for bread baking traits (Lukow et al. 1989). Yue et al. (2008) successfully silenced *1Dx5* expression via RNAi interference, which resulted in dramatic decreases in dough strength and gluten function. The *1Dx5+1Dy10* pair have been shown to impart longer mix time, greater mixing tolerance, longer dough development time, and larger loaf volume than other common *Glu-D1* haplotypes (Dong et al. 1991; Eagles et al. 2006; Cane et al. 2008; Zheng et al. 2009). In a study of transgenic addition of the *1Dx5+1Dy10* subunits, Blechl et al. (2007) reported that overexpression generally increased dough mixing times and tolerances. Furthermore, Blechl et al. (2007) found that the *1Dx5* subunit has a larger effect on most dough parameters than does *1Dy10*, but that both subunits together have an additive effect on glutenin function. Several tests exist that measure both gluten quality and quantity. The SDS-sedimentation test has been used to screen large numbers of genotypes because it requires small quantities of seed and can give rapid results. SDS-sedimentation measures protein quality by measuring the volume of water/SDS-lactic acid absorbed by whole meal or milled flour (Morris et al 2007). With *1Dx5+1Dy10* providing ideal dough properties, it is no surprise that breeders have selected for the *1Dx5+1Dy10* haplotype. The *1Dx5+1Dy10* haplotype was present in 90% of hard winter wheat varieties released between 1991 and 2005 (Shan et al. 2007). Considering that the *1Dx5+1Dy10* subunits provide the best dough traits, developing new genetic variation in these subunits may provide useful allelic variation.

The *Pin* and *Glu-D1* loci share some similarities. Both are often nearly fixed in regional breeding programs for the most favorable alleles. And the two *Pin* and the two *Glu-D1* subunit genes are closely linked, meaning members of a haplotype occur together. Additional allelic variation in these genes would be useful for expanding the range of existing phenotypic variation and for investigating gene function. An approach to creating new novel alleles is to mutagenize seeds with ethyl methanesulfonate (EMS). It has been used successfully to create point mutations at relatively high frequency in wheat (Slade et al. 2005). The *Pin* genes are ideal candidates for EMS mutagenesis because they are small (447 bp) without introns and occur only on the D genome. Plus, the grain hardness phenotype is easy to measure and is highly heritable. Feiz et al. (2009a; 2009b) described the application of EMS mutagenesis with the soft spring wheat cultivar 'Alpowa' (PI 566596) to develop new alleles in the two puroindoline genes (*Pina/Pinb*). Feiz et al. (2009a; 2009b) demonstrated that PIN EMS missense alleles expanded available hardness variation. In addition, function-abolishing mutations occurred at or near the Tryptophan-rich region of each gene, pointing to this as the function-determining region of the gene. After the development of near-isogenic lines with a subset of the new *Pin* alleles reported in Feiz et al (2009a; 2009b), Kammeraad et al. (2016) and Martin et al. (2019) reported the unique end-product quality traits associated with the new alleles; including unique milling yields and particle size distribution. The benefits of EMS mutagenesis are two-fold; first being the development of novel alleles that are not available naturally, and second is to gain insight into the impact that a single amino acid change has on protein structure/function.

As stated above allelic variation at the *Glu-1* loci is due largely to the size and number of repeated fragments. But Lombardo et al. (2017) generated mutations via EMS in a bread wheat cultivar that were identified based on lack of or changes in mobility of protein bands on sodium

dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). No measure of function was given. The impact of missense mutations dispersed across the *Glu-1 gene* is not known. The goal here was to report the functional impact (as measured by swelling power of gluten) of new EMS induced HMW-GS *1Dx5+1Dy10* mutations from the same Alpowa EMS population as Feiz (2009a and 2009b). The *1Dx5+1Dy10* subunits were the focus because of their high influence on dough quality and their prevalence in hard wheat cultivars used for bread. These results will be compared with *Pina* and *Pinb* mutations from Feiz (2009a and 2009b).

2. Materials and Methods

2.1. Creation of an EMS-Induced Population

The EMS treated seed of the soft white spring wheat variety Alpowa (PI 566596) described in Feiz et al (2009a; 2009b) was used as the source for glutenin mutants. Alpowa contains the *Glu-A1* null, *Glu-Bx7+By9*, *Glu-Dx5+Dy10* alleles (Park et al. 2003). Approximately 10,000 M_0 seeds of Alpowa were treated with EMS and single heads were harvested from ~3,000 fertile M_1 plants. In May 2007 2,000 M_2 head rows were planted in the field and genotyped via Sanger sequencing. Seed was harvested from 1,700 fertile M_2 rows that were screened for the 1Dx5+1Dy10 mutants. Eight M_3 seeds from each $M_{2:3}$ head row were grown in the greenhouse. Mutation carrying plants were used as pollen sources in crossing to non-mutagenized Alpowa. F_1 plants were grown in the greenhouse to produce F_2 plants. Sixty seeds from each of the 47 F_2 populations and the Alpowa parental seed were planted in the greenhouse, and the F_2 plants were genotyped via sanger sequencing. $F_{2:3}$ seed from four F_2 plants homozygous for the new HMW-GS or wild type Alpowa alleles were bulked to have a mutant and wild type pair for each population. F_3 seed from the 47 populations were planted in two replications of a randomized block design in May 2010. Each plot consisted of four rows: two mutant rows and two paired

wild type rows. The experimental plots also included the non-mutagenized Alpowa parent. The seeding rate when planting was 3 g/row and plots were 3 m long with 30 cm between rows, with grain harvested after maturity. This trial was grown in separate, adjacent rain fed and irrigated environments at the Arthur H. Post Field Research Center near Bozeman, MT (latitude 45.67N, longitude 111.15W, elevation 1,455 m, soil type is Amsterdam silt loam). One hundred eighty five kg/ha 46-0-0 fertilizer was added prior to planting. The irrigated trial received 6.25 cm water two weeks prior to anthesis. Additionally, nineteen F_2 populations were planted in the field in spring 2010. Each F_2 population had sixty F_2 plants in three rows with 15 cm between each plant and 30 cm space between rows with the same irrigated conditions as above. . Genotyping for these F_2 plants was performed as described above. Upon genotyping 1Dx5+1Dy10, homozygous mutants and wild types were selected and threshed at maturity.

2.2 *Glu-1Dx5+Dy10* Primer Design/PCR/Sequencing

Sequencing of the *1Dx5+1Dy10* genes was carried out to select homozygous mutant and wild type lines. All primers were designed using Genbank accessions X12928 and X12929 for *1Dx5* and *1Dy10*, respectively. Primer sets were designed to amplify multiple ~700 bp overlapping segments of the *1Dx5* and *1Dy10* genes (three segments in *1Dy10*, and 4 segments in *1Dx5*). Genotype selection began with collecting leaf tissue at the two- to three- leaf stage and the DNA was extracted following Riede and Anderson (1996). PCR reactions consisted of ~100 ng genomic DNA, 15 pmol of primer, 200 uM of dNTP, 1X PCR buffer, 2 mM $MgCl_2$ (Promega, Madison, WI), and 0.5 U *Taq* DNA polymerase (Promega, Madison, WI) per 20 µL reaction volume. Segment amplification was verified via 1% agarose gel electrophoresis. PCR products were submitted to Beckman Coulter Genomics (Danvers, MA) with the forward amplification primers for Sanger sequencing. Sequence analysis was performed using DNAStar-

SeqMan Pro package v12.3.1, 2015 (http://www.dnastar.com verified September 30, 2022). Mutations were identified by comparing anomalous peak heights of trace sequences to the corresponding Genbank reference sequences. Genotyping was dependent on peak height and number; homozygote: single peak replacement, heterozygote: two peaks of half height.

2.3 Gluten Quality Testing of all HMW-GS Alleles

The swelling index of gluten (SIG) test was used as an indicator of gluten strength (Wang and Kovacks, 2002; Uthayakumaran et al. 2007). Initial gluten quality tests were done on the 47 mutant populations grown in the replicated field trials, and the 19 mutation populations grown as spaced plants in the field. The 47 populations from the replicated field trial consisted of F_4 seed from each replication, and the nineteen F_2 populations consisted of F_3 seed from approximately 10 individual mutant and wildtype plants. Five wheat kernels were first ground in a coffee mill for 5 s, then put into a 2 mL microcentrifuge tube with two steel ¼" diameter balls. Ground meal was agitated in the tubes for 20 s in a Mini-BeadBeater-96 (BioSpec, Bartlesville, OK). 40 mg of the agitated ground meal was then transferred into pre-weighed 2 ml tubes. 500 uL of room temperature distilled water was added to the 40 mg ground meal and mixed on an Eppendorf MixMate (Eppendorf, Hauppauge, NY) at 1,400 rpm for 20 min. Then 500 uL of 3% SDS-lactic acid solution was added and mixed at 1,400 rpm for another 20 min. After mixing, samples were immediately centrifuged at 300 g for 5 min and the resulting supernatant was aspirated. Tubes and remaining precipitate were then weighed, and Swelling index of gluten (SIG) values were calculated by dividing the weight of precipitate by the original weight of ground meal (40 mg). Each sample was done twice and averaged for analysis. Kernel weight was measured on the same seed source as the SIG test for the 47 mutation populations grown in replicated trials using the Single Kernel Characterization System (SKCS) (Perten Instruments, Springfield, IL).

2.4 Data Analysis

We used the difference between mutant and wild type as a measure of function for each plot in each segregating F_2 population. The differences values for SIG value and kernel weight for the 47 F_2 populations grown in two environments were analyzed via analysis of variance using PROC MIXED in SAS v9.4 (SAS Institute, 2014). The model included environment, replication within environment, mutation populations, and environment x mutation population interaction. The replications within environments were considered random while all other effects were considered fixed. For the population grown as spaced plants data for SIG values were analyzed using PROC MIXED in SAS including mutation populations and plants within mutation populations in the model.

3. RESULTS

3.1 *Glu-D1x5* and *Glu-D1y10* Mutations

An total of 135 mutations were identified in the 384 M_2 Alpowa families by sequencing individual *Glu-D1x5* and *Glu-D1y10* PCR products. Gene specific primers were designed to amplify individual overlapping segments for each gene (Table 1). There were 80 missense, 22 nonsense, and 33 silent mutations roughly equally distributed in both *1Dx5* and *1Dy10* genes. Each *Glu-D1* gene included at least nine mutations in each of the mutation types (Table 2). Except for one transversion mutation, all others were transition mutations (G to A or C to T). The mutation rate for the combined 1Dx5+1Dy10 mutants was 1/26 Kb. Table 2 shows the heterozygous mutation discovery rate in the segregating population as 1/12.7 kb, which means the true discovery rate is halved at being 1/26 kb. The distribution of mutations was uniform across the terminal and the middle repetitive region of each gene, with the exception of having no mutations in the C-terminus of Dx5 (Figure 1 and 2). We identified 69 and 66 mutations in

1Dx5 and 1Dy10, respectively. Of the 69 mutations in Dx5, 35 were missense, 13 nonsense, and 21 silent. In Dy10, the 66 mutations consisted of 45 missense, 9 nonsense, and 12 silent. Most of the F_2 populations (58 out of 62) segregated in the anticipated 1:2:1 ratio (P>0.05).

3.2 Swelling Index of Gluten

Swelling index of gluten (SIG) was used as a measure of gene function. Figure 2 shows the difference in SIG data from $F_{2:3}$ seed pools from at least three F_2 plants carrying the introduced EMS glutenin allele versus $F_{2:3}$ seeds pools from at least three F_2 plants carrying the wild type *Glu-1* allele. Of the 26 Dx5 missense alleles characterized, only two significantly reduced SIG, while none increased SIG. Missense mutations G355R and G627S, both in *1Dx5* segment 2, showed significant ($p<0.05$) loss of function. However, all 5 of the *1Dx5* stop codon mutations significantly reduced SIG, by an average of 0.91. In *1Dy10* two of the missense alleles (A9T/G432R and G130D) significantly increased function while the remaining 28 missense alleles had no change on function. All four nonsense alleles identified in *1Dy10* significantly reduced SIG by an average of 0.63. The A9T/G432R mutant was derived from an M_2:M_3 family with two separate mutations, both within *1Dy10*. There was no detectable difference in loss of gluten function between the *1Dx5* and *1Dy10* nonsense mutations, however the *1Dx5* nonsense alleles were on average much more severe in terms of reducing SIG values. Measures of individual kernel weight showed the HMW-GS mutant average was not different than the wild type sister line average (data not shown).

3.3 Summary of Puroindoline mutation screening

We previously reported the creation of novel *Pina* and *Pinb* alleles in the soft white spring wheat Alpowa (Feiz et al. 2009a and 2009b). The results are briefly summarized for comparison to our 1Dx5 and 1Dy10 results. The frequency of mutations across *Pina* and *Pinb*

was 1/11.5 Kb, which was computed from sequencing of both genes. F_2 populations segregating for mutant and wild type were created by crossing 25 *Pina* and 21 *Pinb* missense-carrying M_3 plants back to non-mutagenized Alpowa. The 46 F_2 populations were grown in the field and grain hardness was measured on harvested grain using the single-kernel characterization system (SKCS) (Perten Instruments, Springfield, IL). As with the *Glu-D1* mutations we used the difference between mutant and wild type classes as a measure of function. All stop mutations (W71Stop and Q93Stop for PINA and Q20Stop and W116Stop for PINB) increased grain hardness, but PINA stop mutations increased grain hardness more than the PINB stop mutations (37.8 vs 31.8 hardness units). Nearly all PINA and PINB missense mutations had less impact on grain hardness than the stop mutations (Fig. 4, stops not shown). All severe mutations (>28 hardness unit increase), except one PINA mutation (C132Y), occurred within the tryptophan-rich domain (Fig. 4). One PINB mutation (P41S) that reduced grain hardness and two that had intermediate (G111D, R126K) function were located outside the tryptophan-rich domain. These results point to the tryptophan-rich domain as being the active site for the Pin genes.

DISCUSSION

Here we report the creation of novel alleles in the *1Dx5* and *1Dy10* Glu-D1 subunits in the same EMS mutagenized Alpowa population that was used to create novel *Pin* mutations. These genes were chosen because they have large impacts on wheat end use properties, and most hard wheat varieties have reached near fixation for the most favorable haplotype alleles.

To identify missense mutations, we sequenced the *1Dx5* and *1Dy10* subunit genes from 384 mutagenized families. This identified 135 mutations of which 80 were missense mutations about equally divided between the *1Dx5* and *1Dy10* subunit genes (Table 2). To sequence the complete *Glu-D1* subunit the genes were divided into overlapping segments (four segments for *1Dx5* and

three segments for *1Dy10*) (Table 1). This strategy identified mutations representing all regions of each gene (Figures 1 and 2). The frequency of mutations combined over both subunit genes was 1/12.7 kb.

Results from segregating F_2 populations showed stop mutations in either *1Dx5* or *1Dy10* reduced function as measured by the SIG test, but missense mutations had essentially no impact on function. Previous results have shown that loss of function of *Glu-D1* subunits reduces dough strength. MacRitchie and Lafiandra (2001) showed near isogenic lines that were null for the 1Dx5+1Dy10 subunits had lower mixograph dough development time than wild type 1Dx5+1Dy10 lines. Yue et al (2008) used RNA interference (RNAi) to silence the *1Dx5* gene. Dough strength was reduced in the line showing no expression (null) of *1Dx5*. Neither of these studies provided a comparison of the relative impact of loss of 1Dx5 versus 1Dy10. Our results show that loss of function of either 1Dx5 or 1Dy10 had about equal negative impact on SIG values, an index of dough strength.

Comparisons with Puroindoline Results

Mutations were identified dispersed across both *Pin* genes (Fig. 4 and Feiz et al. 2009b Supplemental Tale 1) and *1Dx5* and 1*Dy10* (Fig. 1 and Fig 2). We found the mutation rate was similar across the *Pina* and *Pinb* alleles and the *1Dx5* and *Dy10* alleles (1/11.5 for *Pins* and 1/12.7 for the *Glu-D1* loci). Similar mutation rates across these loci were expected, since the same mutagenized Alpowa source was used in both cases. The proportion of missense mutations was higher for *Glu-D1* loci (0.59) than for *Pin* loci (0.46), but more similar for nonsense mutations (0.12 for *Pins* and 0.16 for Glu-D1). Mutation rates are difficult to compare across studies because EMS concentrations vary. But the mutation rate reported here is about twice that reported in Slade et al. 2009 for *Wx* loci.

Feiz (2009b) showed missense mutations within or near the tryptophan-rich region in *Pina* and *Pinb* caused most severe loss of function, indicating this region of the gene is most important. It is worth noting that naturally occurring *Pin* mutations showing loss of function also result from single amino acid changes within the tryptophan-rich region. Missense mutations in *Glu-D1* subunit genes did not identify a target region as nearly all had no impact on function as measured by SIG. Several reasons may account for these differences. A phenotype measure was used to measure function in both cases. Grain hardness is easily measured and highly heritable.

It is not surprising that the impact of missense mutations differed between the *Pin* and *Glu-D1* loci. The *Pin* and *Glu-D1* loci differ in size and structure. The *Pin* genes are small, 450 bp genes. By comparison the *Glu-D1* loci are much larger (~ 2100 bp) consisting of a non-repetitive N- and C-terminal domains and a central repetitive domain (Shewry et al. 2003). Grain hardness phenotype is highly heritable and is measured easily and precisely with SKCS technology. Our early phenotypic screen for *1Dx5* and *1Dy10* loci function was the SIG test. Although SIG values were correlated with measures of dough strength in soft wheat (Uthayakumaran et al., 2007), other measures of dough strength may have shown more differentiation among mutant populations. Naurally occurring alleles in the *Pi*n genes result from single amino changes, whereas naturally occurring alleles at *1Dx* or *1Dy* loci differ in number of amino acid residues. Finally, the fact that missense mutations in *Glu-D1* loci showed little impact on function may imply that these proteins are stable and very tolerant of amino acid substitutions as compared to the PIN proteins. The cysteine residues are believed to be crucial for intermolecular bonds between HMW and LMW glutenins (Gianibelli et al., 2001; Li et al., 2021: Shewry, et al., 2003). The cysteine residues occur mainly in the N-terminal domain. The *1Dx5* protein has 3 and *1Dy10* has 5 cysteines in this region (Shewry et al., 2003). The N-terminal

domain comprises 89 and 104 residues for *1Dx5* and *1Dy10*, respectively. Although mutations occurred in the N-terminal region in both subunit genes, none of our mutations targeted a cysteine residue. An amino acid substitution involving a cysteine may have a larger effect on dough strength than those in our study. Future work targeting mutations in one or more of the cysteine residues may be warranted.

CONCLUSIONS

The *Glu-D1* HWM-GS loci have large impacts on end use quality. Selection for end use quality in hard wheats has moved the most favorable *Glu-D1* haplotype, *1Dx5+1Dy10* to near fixation in hard wheat germplasm. New allelic variation is important for continued genetic progress. The missense mutations created in *1Dx5* and *1Dy10* did not produce useful allelic variation, as they did not impact function as measured by SIG value. However, stop mutations in both *1Dx5* and *1Dy10* showed reduced SIG values. These results were in contrast to *Pin* mutations created in the same EMS population where missense mutations identified the Tryptophan-rich region in *Pina* and *Pinb* as being the function-determining region of these genes. EMS generated stop mutations can be a useful tool to determine impacts when gene function is lost. Finally, useful phenotypic variation can be created using EMS mutagenesis for genes where naturally occurring mutations occur from amino substitutions as is the case for the *Pin* loci.

ACKNOWLEDGEMENTS

REFERENCES

Blechl, A., Lin, J., Nguyen, S., Chan, R., Anderson, and O.D., Dupont, F.M. (2007). Transgenic wheats with elevated levels of Dx5 and/or Dy10 high-molecular-weight glutenin subunits yield doughs with increased mixing strength and tolerance. *Journal of Cereal Science*, 45,172-183.

Blochet, J.E., Chevalier, C., Forest, E., Pebaypeyroula, E., Gautier, M.F., Joudrier, P., Pezolet, M., Marion, D. (1993). Complete amino-acid-sequence of puroindoline, a new basic and cystine-rich protein with a unique tryptophan-rich domain, isolated from wheat endosperm by triton X-114 phase partitioning. *FEBS Letters*, 329, 336–340.

Branlard, G., & Bellot, P. (1983). Improvement of bread wheat gluten quality. Qual Plant Foods Hum Nutr. 33, 121-126.

Branlard, G., Dardevet, M., Saccomano, R., Lagoutte, F., & Gourdon, J. (2001). Genetic diversity of wheat storage proteins and bread wheat quality. *Euphytica*, 119, 59-67.

Campbell, K. G., C. J. Bergman, D. G. Gaulberto, J. A. Anderson, M. J. Giroux, G. Hareland, R. G. Fulcher, M. E. Sorrells, & P. L. Finney. (1999). Quantitative trait loci associated with kernel traits in a soft x hard wheat cross. *Crop Science*, 39, 1184-1195.

Cane, K., Sharp, P.J., Eagles, H.A., Eastwood, R.F., Hollamby, G.J., Kuchel, H., & Martin, P.J. (2008). The effects on grain quality of a grain serpin protein and the VPM1 segment on southern Australian wheat breeding. *Australian Journal of Agricultural Research*, 59, 883-890.

Day, L., Bhandari, D.G., Greenwell, P., Leonard, S.A., & Schofield, J.D. (2006). Characterization of wheat puroindoline proteins. *FEBS Journal*, 273, 5358–5373.

Dong, H., Cox, T.S., Sears, R.G., & Lookhart, G.L. (1991). High molecular weight glutenin genes: effects on quality in wheat. *Crop Science*, 31, 974-979.

Dubreil, L., Gaborit, T., Bouchet, B., Gallant, D., Broekaert, W., Quillien, L., & Marion, D. (1998). Spatial and temporal distribution of the major isoforms of puroindolines (puroindoline-a and puroindoline-b) and non specific lipid transfer protein (ns-LTPle(1)) of *Triticum aestivum* seeds: relationships with their in vitro antifungal properties. *Plant Science*, 138, 121-135.

Eagles, H.A., Cane, K., Eastwood, R.F., Hollamby, G.J., Kuchel, H., Martin, P.J., & Cornish, G.B. (2006). Contributions of glutenin and puroindoline grain to grain quality traits in southern Australia. *Australian Journal of Agricultural Research*, 57, 179-186.

Feiz, L., Beecher, B., Martin, J., & Giroux, M. (2009a). In planta mutagenesis determines the functional regions of the wheat puroindoline proteins. *Genetics*, 183, 853-860.

Feiz, L., Martin, J., & Giroux, M. (2009b). Creation and functional analysis of new *puroindoline* alleles in *Triticum aestivum*. *Theoretical and Applied Genetics*, 118, 247-257.

Feiz, L., Wanjugi, H., Melnyk, C., Altosaar, I., Martin, J., & Giroux, M. (2009c). Puroindolines co-localize to the starch granule surface and increase seed bound polar lipid content. *Journal of Cereal Science*, 50, 91-98.

Gianibelli, M.C., Larroque, O.R., MacRitchie, F., & Wrigley, C.W. (2001). Biochemical, genetic, and molecular characterization of wheat glutenin and its component subunits. *Cereal Chemistry*, 78, 635-646.

Giroux, M., & Morris, C. 1997. A glycine to serine change in puroindoline b is associated with wheat grain hardness and low levels of starch-surface friabilin. *Theoretical and Applied Genetics*, 95, 857-864.

Giroux, M., & Morris, C. 1998. Wheat grain hardness results from highly conserved mutations in the friabilin components puroindoline a and b. *Proceedings of the National Academy of Sciences, U.S.A.*, 95, 6262-6266.

Greenwell, P., & Schofield, J. D. 1986. A starch granule protein associated with endosperm softness in wheat. *Cereal Chemistry*, 63, 379-380.

Graybosch, R.A., Seabourn, B., Chen, Y.R., & Blechl, A.E. 2011. Quality and agronomic effects of three High-Molecular-Weight glutenin subunit transgenic events in winter wheat. *Cereal Chemistry*, 88, 95-102.

Halford, N.G., Forde, J., Anderson, O.D., Greene, F.C., & Shewry, P.R. 1987. The nucleotide and deduced amino acid sequence of an HMW glutenin subunit gene from chromosome 1B of bread wheat (*Triticum aestivum* L.) and comparison with those of genes from chromosomes 1A and 1D. *Theoretical and Applied Genetics*, 75, 117-126.

Harberd, N.P., Flavell, R.B., & Thompson, R.D. 1987. Identification of a transposon-like insertion in a *Glu-1* allele of wheat. *Molecular and General Genetics*, 209, 326-332.

Hogg, A., Sripo, T., Beecher, B., Martin, J., & Giroux, M. 2004. Wheat puroindolines interact to form friabilin and control wheat grain hardness. *Theoretical and Applied Genetics*, 108, 1089-1097.

Li, Y., Fu, F., Shen, Q., & Yang, D. 2021. High-Molecular-Weight glutenin subunits: Genetics, structures, and relation to end use qualities. International Journal of Molecular Sciences 22, 184. ttps://dx.doi.org/10.3390/ijms22010184

Lombardo, L.A., Nisia M. M., Salinesb, N., Ghionea, C.E., & Helguera, M. 2017. Identification of novel high molecular weight glutenin subunit mutants in bread wheat (Triticum aestivum L.). *Cytology and Genetics*, 51:303-314. DOI: 10.3103/S0095452717040028

Lukow, O.M., Payne, P.I., & Tkachuk, R. 1989. The HMW glutenin subunit composition of Canadian wheat cultivars and their association with bread-making quality. *Journal of the Science of Food and Agriculture*, 46: 451-460.

Ma, D., Zhang, Y., Xia, X., Morris, C., & He, Z. 2009. Milling and Chinese raw white noodle qualities of common wheat near-isogenic lines differing in puroindoline b alleles. *Journal of Cereal Science*, 50:126-130.

Martin, J., Frohberg, R., Morris, C., Talbert, L., & Giroux, M. 2001. Milling and bread baking traits associated with puroindoline sequence type in hard red spring wheat. *Crop Science*, 41:228-234.

Martin, J.M., Hogg, A.C., & Giroux, M.J. 2019. Registration of Wheat Lines Carrying Novel Mutagenesis Derived Novel Puroindoline Alleles. *Journal of Plant Registrations*, 13:455-460. https://doi.org/10.3198/jpr2018.11.0074crg

MacRitchie, F, and Lafiandra, D (2001) Use of near-isogenic wheat lines to determine protein composition-functionality relationships. Cereal Chem. 78:501-506

Morris, C., Lillemo, M., Simeone, M., Giroux, M., Babb, S., & Kidwell, K. 2001. Prevalence of puroindoline grain hardness genotypes among historically significant North American spring and winter wheats. *Crop Science*, 41:218-228.

Morris, C.F., Paszczynska, B., Bettge, A.D., & King, G.E. 2007. A critical examination of the sodium dodecyl sulfate (SDS) sedimentation test for wheat meals. *Journal of the Science of Food and Agriculture*, 87: 607-615.

455 Park, C.S., Hong, B.H., & Baik, B.K. 2003. Protein quality of wheat desirable for making fresh
456 white salted noodles and its influences on processing and texture of noodles. *Cereal*
457 *Chemistry*, 80(3): 297-303.
458 Payne, P.I. 1987. Genetics of wheat storage proteins and the effect of allelic variation on bread-
459 making quality. *Annual Review of Plant Physiology*, 38:141-153.
460 Payne, P.I., Corfield, K.G., & Blackman, J.A. 1979. Identification of a high-molecular-weight
461 subunit of glutenin whose presence correlates with bread-making quality in wheats of
462 related pedigree. *Theoretical and Applied Genetics*, 55:153-159.
463 Payne, P.I., Holt, L.M., & Law, C.N. 1981. Structural and genetical studies on the High-
464 Molecular-Weight subunits of wheat glutenin. *Theoretical and Applied Genetics*, 60:
465 229-236.
466 Payne, P.I., Jackson, E.A., Holt, L.M., & Law, C.N. 1984. Genetic linkage between endosperm
467 storage protein genes on each of the short arms of chromosomes 1A and 1B in wheat.
468 *Theoretical and Applied Genetics*, 67: 235-243.
469 Payne, P.I., Law, C.N., & Mudd E.E. 1980. Control by homoeologous group 1 chromosomes of
470 the high-molecular-weight subunits of glutenin, a major protein of wheat endosperm.
471 *Theoretical and Applied Genetics*, 58: 113-120.
472 Payne, P. I., & G. J. Lawrence. 2003. Catalogue of alleles for the complex gene loci Glu-A1,
473 Glu-B1, Glu-D1 which code for high-molecular-weight subunits of glutenin in hexaploid
474 wheat. *Cereal Research Communications*, 11: 29-35.
475 Riede CR, & Anderson JA (1996) Linkage of RFLP markers to an aluminum tolerance gene in
476 wheat. *Crop Science*, 36:905–909.
477 https://doi.org/10.2135/cropsci1996.0011183X0036000400015x

478 Shan, X., Clayshulte, S.R., Haley, S.D., & Byrne, P.F. 2007. Variation for glutenin and waxy
479 alleles in the US hard winter wheat germplasm. *Journal of Cereal Science*, 45: 199-208.
480 Pomeranz, Y., & Williams, P. C. 1990. Wheat hardness: Its genetic, structural, and biochemical
481 background, measurement, and significance. Pages 471-548 in: Advances in Cereal
482 Science and Technology, Vol. 10. Y. Pomeranz, ed. American Association of Cereal
483 Chemists: St. Paul, MN.
484 SAS Institute. 2014. SAS/STAT 9.4 user's guide. SAS Inst., Cary, NC.
485 Slade, A.J., Fuerstenberg, S.I., Loeffler, D., Steine, M.N., & Facciotti, D. 2005. A reverse
486 genetic, nontransgenic approach to wheat crop improvement by tilling, *Nature*
487 *Biotechnology* 23:75–81.
488 Shewry, P.R., Halford, N.G., Tatham, A.S., Popineau, Y., Lafiandra, D., & Belton, P.S. 2003.
489 The high-molecular-weight subunits of wheat glutenin and their role in determining
490 wheat processing properties. *Advances in Food and Nutrition Research*, 45: 220-302.
491 Takata, K., Ikeda, T., Yanaka, M., Matsunaka, H., Seki, M., Ishikawa, N., & Yamauchi, H. 2010.
492 Comparison of five puroindoline alleles on grain hardness and flour properties using near
493 isogenic wheat lines. *Breeding Science*, 60:228-232.
494 Tatham, A.S., Drake, A.F., & Shewry, P.R. 1990. Conformational studies of synthetic peptides
495 corresponding to the repetitive regions of the High-Molecular-Weight (HMW) glutenin
496 subunits of wheat. Journal of Cereal Science, 11: 189-200.
497 Uthayakumaran, S., Barker, N., Batey, I.L., Dines, J., Miskelly, D., & Wrigley, C.W. 2007.
498 Rapid methods to predict soft-wheat dough quality for specific food products. *Cereal*
499 *Chemistry*, 84: 522-526.

Wang, C., & Kovacks, M.I.P. 2002. Swelling index of glutenin test. II. Application in prediction of dough properties and end-use quality. *Cereal Chemistry*, 79: 190-196.

Weegles, P.L., Hamer, R.J., Schofield, J.D. 1996. Critical review: Functional properties of wheat glutenin. *Journal of Cereal Science*, 23:1-18.

Yue, S.J., Li, H., Li, Y.W., Zhu, Y.F., Guo, J.K., Liu, Y.J., Chen, Y., & Jia, X. 2008. Generation of transgenic wheat lines with altered expression levels of 1Dx5 high-molecular weight glutenin subunit by RNA interference. *Journal of Cereal Science*, 47: 153-161.

Zheng S., Byrne, P.F., Bai, G., Shan, Xueyan, S., Reid, S.D., Haley, S.D., & Seabourn, B.W. 2009. Association analysis reveals effects of wheat glutenin alleles and rye translocations on dough-mixing properties. *Journal of Cereal Science*, 50: 283-290.

511 Table 1. Primers used to amplify segments of the 1Dx5 and 1Dy10 HMW-GS subunits.
512

Gene	Segment	Primers	Sequence	Size (bp)	Tm (°C)
1Dx5	Seg 1	Dx55'-(-93)	CCTATAAAAGCCTAGCCAACCTTCAC	912	60.1
		Dx53'-820	CCCTTGTTGTCCTTGTCCG		
	Seg 2	Dx55'-806	AGGTCAGCAGCTCGGACA	426	60.4
		Dx53'-1214	GCTTGTTGCCCTTGTCCTACC		
	Seg 3	Dx55'-1194	GGTAGGACAAGGGCAACAAGC	543	60.3
		Dx53'-1736	ACTTGTGCTGGTTGCTGCC		
	Seg 4	Dx55'-1717	GGCAGCAACCAGCACAAGT	830	63.5
		Dx53'- P16	TCACTGGCTGGCCGACAATGC		
1Dy10	Seg 1	Dy105'-(-93)	CCTATAAAAGCCCAACCAATCTC	1089	59
		Dy103'-996	TGGCTCTTGCTGAGAAGCTG		
	Seg 2	Dy105'-615	CCCAACTTCTCTGCAGCACA	1131	60.8
		Dy103'-4658	TGTCCTGATTGTTGCCCTTGTCCA		
	Seg 3	Dy105'-1494	CTATCCAACTTCTCCACAGCAG	555	60.1
		Dy103'-2027	AACTGTGAACACGCATCACG		

513
514

Table 2. Frequency of EMS-induced mutations in the 1Dx5+1Dy10 HMW-GS subunits. [a]Mutation discovery rate calculated using the frequency of mutations found by sequencing of 384 M_2 Alpowa lines.

	1Dx5	1Dy10	Both
Missense	35	45	80
Nonsense	13	9	22
Silent	21	12	33
Total	69	66	135
Mutation Discovery Rate[a]	1/14.0 kb	1/11.3 kb	1/12.7 kb

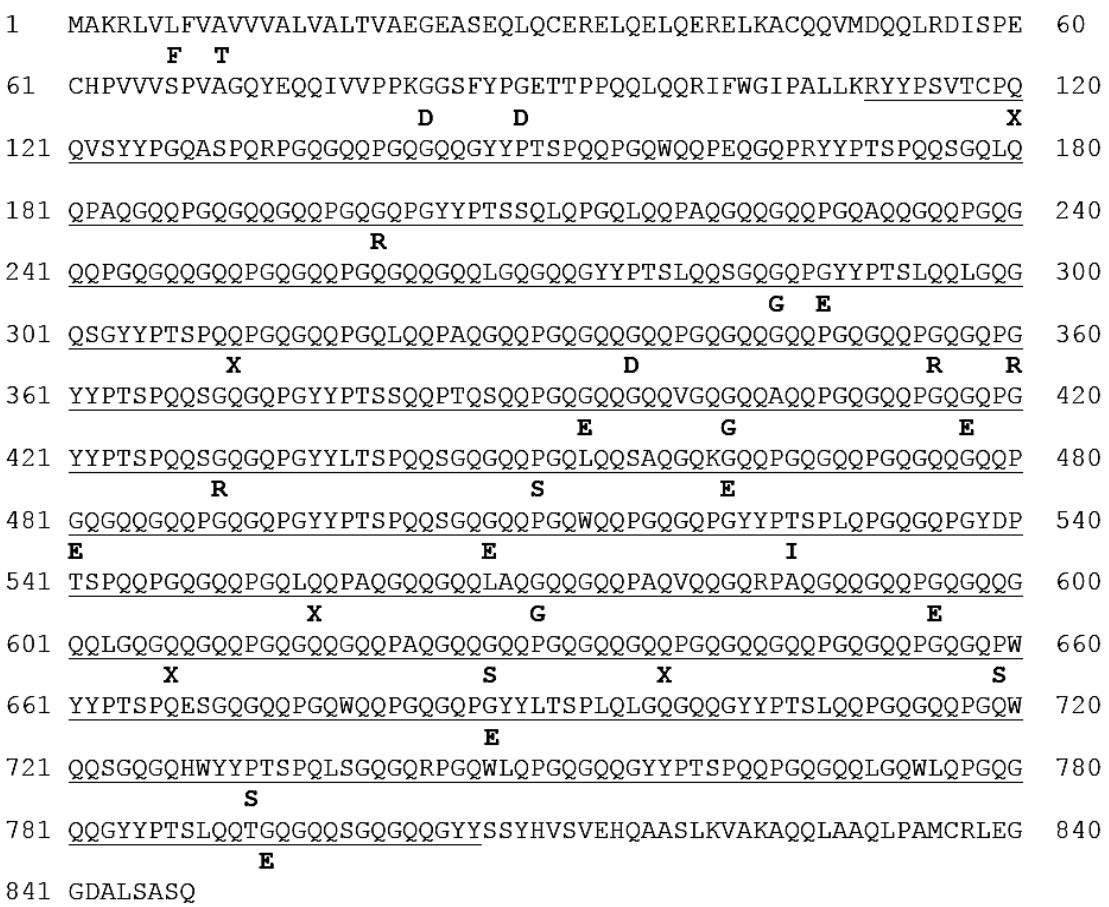

Figure 1. Amino acid sequence of *Glu-D1x5* subunit. Locations of each identified allele are as noted underneath the *Glu-D1x5* coding sequence (Genbank accession X12928). X denotes a stop codon and the underlined region denotes the repetitive domain.

```
1    MAKRLVLFAAVVIALVALTTAEGEASRQLQCERELQESSLEACRQVVDQQLAGRLPWSTG 60
      T      T                                                D      E
61   LQMRCCQQLRDVSAKCRSVAVSQVARQYEQTVVPPKGGSFYPGETTPLQQLQQGIFWGTS 120
                     H                   L
121  SQTVQGYYPGVTSPRQGSYYPGQASPQQPGQGQQPGKWQEPGQGQQWYYPTSLQQPGQGQ 180
               D     XE     X                                  S
181  QIGKGQQGYYPTSLQQPGQGQQGYYPTSLQHTGQRQQPVQGQQPEQGQQPGQWQQGYYPT 240
       R G  R                        X
241  SPQQLGQGQQPRQWQQSGQGQQGHYPTSLQQPGQGQQGHYLASQQQPGQGQQGHYPASQQ 300
                        R
301  QPGQGQQGHYPASQQQPGQGQQGHYPASQQEPGQGQQGQIPASQQQPGQGQQGHYPASLQ 360
                     LR                    S
361  QPGQGQQGHYPTSLQQLGQGQQTGQPGQKQQPGQGQQTGQGQQPEQEQQPGQGQQGYYPT 420
                                           E                R
421  SLQQPGQGQQQGQGQQGYYPTSLQQPGQGQQGHYPASLQQPGQGQPGQRQQPGQGQHPEQ 480
            E   R                                R
481  GKQPGQGQQGYYPTSPQQPGQGQQLGQGQQGYYPTSPQQPGQGQQPGQGQQGHCPTSPQQ 540
                                  E                                X
541  SGQAQQPGQGQQIGQVQQPGQGQQGYYPTSVQQPGQGQQSGQGQQSGQGHQPGQGQQSGQ 600
                                           E          L
601  EQQGYDSPYHVSAEQQAASPMVAKAQQPATQLPTVCRMEGGDALSASQ              648
                     LT N          V           D
528
```

529
530 Figure 2. Amino acid sequence of *Glu-D1y10* subunit. Locations of each identified allele are as
531 noted underneath the *Glu-D1y10* coding sequence (Genbank accession X12929). X denotes a
532 stop codon and the underlined region denotes the repetitive domain.
533
534

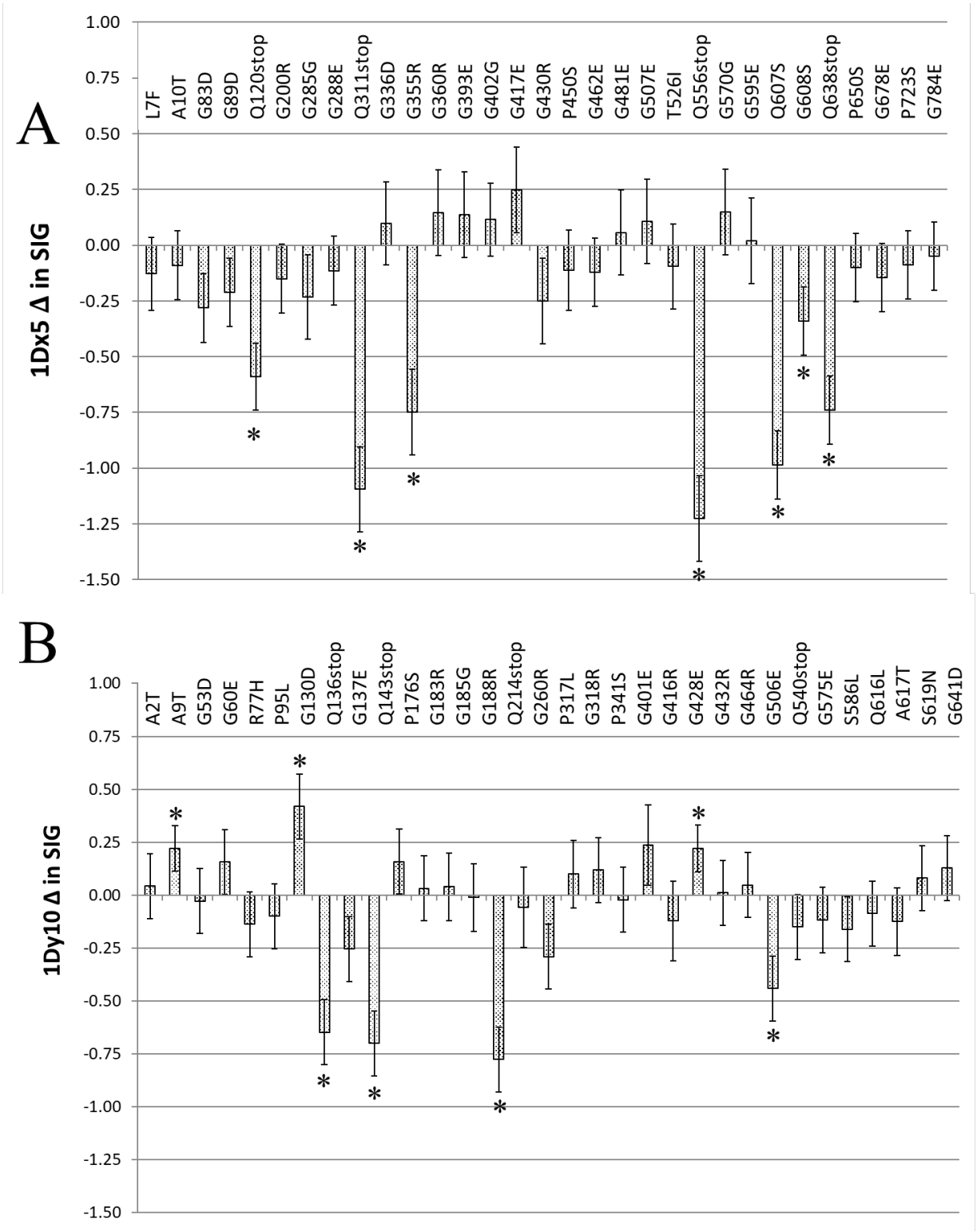

Figure 3. Change in Swelling Index of Gluten for all EMS created alleles in *Glu-D1x5* (A) *and* *Glu-D1y10* (B). Significance ($p<0.05$) is denoted with *.

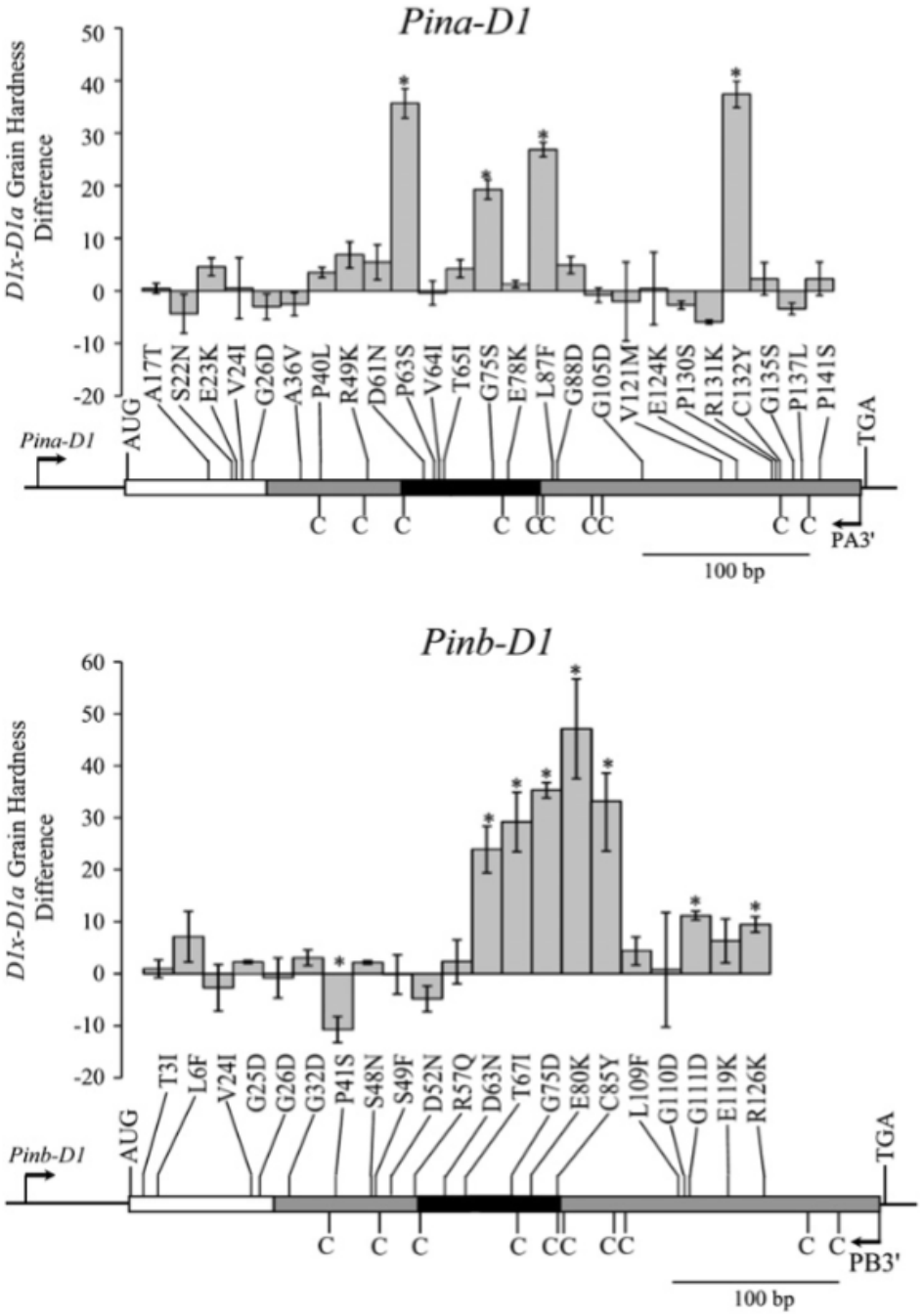

Figure 4. The mean grain hardness difference between lines homozygous positive or negative for individual PIN mutations relative to their location. Asterisks indicate where the hardness difference between an individual Pin mutant and its wild-type group was significant. The open boxed area denotes the signal peptide region while the shaded/solid boxed region denotes mature peptide sequence with the solid boxed region indicating the tryptophan-rich loop region. The positions of cysteines are denoted by C's. The position of the amplification primers is as shown. (From Feiz et al. 2009b)