FUNCTIONAL ANALYSIS OF *PUROINDOLINE* GENES IN WHEAT (*TRITICUM AESTIVUM*)

by

Leila Feiz

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Plant Sciences

MONTANA STATE UNIVERSITY
Bozeman, Montana

November 2008

APPROVAL

of a dissertation submitted by

Leila Feiz

This dissertation has been read by each member of the dissertation committee and has been found to be satisfactory regarding content, English usage, format, citation, bibliographic style, and consistency, and is ready for submission to the Division of Graduate Education.

Dr. Michael J. Giroux

Approved for the Department

Dr. John E. Sherwood

Approved for the Division of Graduate Education

Dr. Carl A. Fox

STATEMENT OF PERMISSION TO USE

In presenting this dissertation in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this dissertation should be referred to ProQuest Information and Learning, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute my dissertation in and from microform along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

Leila Feiz

November 2008

ACKNOWLEDGEMENTS

This is my chance to try to convey the immense gratitude I feel to those people who have supported me and helped make my dreams come true. Their patience is a lesson I would do well to learn. They deserve more thanks than anyone could express and there is absolutely no way I can do them justice, but I will give it a try. Those individuals are:

Dr. Mike Giroux- Advisor (Chair)

Dr. Jack Martin- Advisor (Co-chair)

Dr. Luther Talbert- Committee member

Dr. Andreas Fischer- Committee member

Dr. Michele Hardy- Committee member

Graduate students and friends of Dept. of Plant Science and Plant Pathology especially

Hope Talbert

All my friends and family in Iran and France

Fatimah Erfani- Mother

Navid Movahed- Spouse

&

My father Ali who was my first science teacher

TABLE OF CONTENTS

TABLE OF CONTENTS – CONTINUED

LIST OF TABLES

Table Page

LIST OF FIGURES

ABSTRACT

Grain hardness variation has large effects upon many different end-use properties of wheat (*Triticum aestivum*). The *Hardness (Ha)* locus consisting of the *Puroindoline a* and *b* genes (*Pina* and *Pinb*) controls the majority of grain hardness variation. Starch production is a growing end-use of wheat. The first objective of this study was to estimate the differences in starch yield due to natural and transgenically conditioned grain hardness differences. To accomplish this goal, a small scale wet milling protocol was used to characterize the wet milling properties of two independent groups of isogenic materials varying in grain hardness and in Pin expression level. The results of the first study demonstrate that the *Ha* locus and puroindoline expression are both linked to wet-milling starch yield and that selection for increased *Ha* function increases starch yield via enhanced separation of starch granules and the protein matrix during wet milling. The lack of *Pin* allelic diversity is a major factor limiting *Ha* functional analyses and wheat quality improvement. So the second objective of this study was to create new *Ha* alleles in the soft white spring cultivar Alpowa using ethylmethane sulfonate (EMS) mutagenesis. The M_2 population was screened to identify new alleles of *Pina* and *Pinb*. One hundred and forty eight new *Pin* alleles, including 68 missense alleles, were identified. F_2 populations for 49 of the new *Pin* alleles including 43 unique missense ones were developed after crossing each back to non-mutant Alpowa. Grain hardness was then measured on $F_{2:3}$ seeds and the impact of each allele on grain hardness was quantified. The tested mutations comprised a range of functionality from neutral to function abolishing mutations. Seed weight and vigor of all mutation lines was restored among all of the F_2 populations. The new alleles have the potential to improve end use properties of soft and hard wheats.

CHAPTER 1

THE RELATIONSHIP BETWEEN WHEAT (*TRITICUM AESTIVUM* L.) GRAIN HARDNESS AND WET MILLING QUALITY

Introduction

Maize (*Zea mays* L.), wheat (*Triticum aestivum* L.), rice (*Oryza sativa*) and potato *Solanum tuberosum*) are the four major sources of starch in the world (Sayaslan 2006). Wet milling is an important use of maize grain produced in the USA (Zehr et al. 1996; Parris et al. 2006). The development of high wet milling starch yield maize hybrids is very important for the production of sweeteners and ethanol in USA (Zehr et al. 1996; Parris et al. 2006). However, in Europe, wheat is becoming increasingly important as a starch source material (Bergthaller et al. 2004). From the estimated 593.1 million metric tons (mmt) of annual wheat production in the world (USDA, WASDE-444 2007), 67% is used as food, 20% as feed and 7% as seed (Sayaslan et al. 2006). The remaining 6% of production is used for industrial purposes, which includes wet milling to produce starch and vital gluten (Sayaslan et al. 2006). World wheat starch production was ca. 4.1 million tons in 2000 (LMC International LTD, 2002), originating from ca. 8 million tons of wheat (Van Der Borght et al. 2005).

The two major classes of hexaploid wheat (hard and soft wheat) have different milling and end use product characteristics (reviewed in Morris and Rose; 1996) mainly owing to allelic variation at the *Hardness* (*Ha*) locus located on the short arm of chromosome 5D (Mattern et al. 1973; Law et al. 1978). Mutations in the closely linked

genes *puroindoline a* (*Pina*) or *puroindoline b* (*Pinb*) which comprise the *Ha* locus, have been found in all hard-textured wheat examined (Giroux and Morris 1997; 1998; Lillemo and Morris 2000; Morris et al. 2001b; Cane et al. 2004; Chen et al. 2006). From a biochemical point of view, wheat grain hardness is most likely determined by the degree of adhesion between starch granules and protein matrix. The 15 kDa protein complex friabilin (Greenwell and Schofield, 1986; Morris et al. 1994) consists of the structurally similar proteins puroindoline a and b, (PINA and PINB, respectively) expressed by the *Ha locus* (Giroux and Morris 1997, 1998). Friabilin may functions as a non-stick agent that decreases grain hardness via reduction of the close adherence between starch granules and wheat storage proteins (Anjum and Walker 1991). PINA and PINB are cysteine rich proteins which are unique among plant proteins in having a hydrophobic tryptophan rich domain (Blochet et al. 1993). The tryptophan-rich domain is hypothesized to confer an affinity for lipids (Marion et al. 1994; Kooijman et al. 1997) and may result in PINA and PINB being localized to the glyco- and phospholipid rich surface of the amyloplast membrane (Giroux et al. 2000). These protein-lipid interactions may be disrupted in hard wheats leading to hard endosperm texture. Evidence that PINs affect hardness via binding to polar lipids came first in a report by Morrison et al. (1989) who reported that grain hardness is negatively correlated with free polar lipids content. Similarly, a survey of several soft and hard wheat varieties demonstrated that higher amounts of glyco- and phospholipids were present on soft wheat starch than hard similar to friabilin levels in which soft wheat starch is friabilin rich and hard wheat starch is friabilin deficient (Greenblatt et al. 1995). Defined genetic stocks, that differ in *Ha*

function (*Ha,* soft vs. *ha*, hard) have been valuable not only in studying the biochemical and genetic basis of wheat grain hardness (Giroux and Morris 1997; 1998), but also provided material to study the effect of this gene on flour processing (Martin et al. 2001; Morris and Allan 2001; Morris et al. 2001a; Greffeuille et al. 2006). In addition to natural genetic stocks, transgenic *Pin* addition lines have been successfully used both to prove the function of the *Ha* locus (Beecher et al. 2002; Martin et al. 2006) and to study the impact of this locus on flour processing and different end-use qualities (Hogg et al. 2005; Swan et al. 2006; Martin et al. 2007).

Both small and large scale investigations have suggested that wheat varieties with softer endosperm texture are not only advantageous in dry milling but also in the wet milling process (Bergthaller et al. 2004; Czuchajowska and Pomeranz 1993). Small scale methods to simulate industrial scale wet milling have been used on wheat flour (Czuchajowska and Pomeranz 1993; Sayaslan et al. 2006) or maize intact kernels (Vignaux et al. 2006). In order to test for a linkage between grain hardness and *Pin* expression with starch extractability two independent groups of soft/hard isogenic lines were used in small scale laboratory wet milling tests. These were near isogenic lines (NILs) for the *Ha* locus in two genetic backgrounds, and transgenic isolines in hard red spring wheat 'Hi-Line' over expressing *Pina*, *Pinb* or both *Pina* and *Pinb*. The soft/hard isolines used allow comparison of relatively small (non-transgenic) and large (transgenic) variation in PIN expression level and grain hardness.

Materials and Methods

Hard/Soft NILs

The first group of isolines consisted of two sets of hard/soft NILs chosen from those summarized by Morris and Allan (2001) and Morris et al. (2001a). The first set of NILs were the Australian white spring cultivar Falcon derived NILs that carried either the *Pina-D1a* soft type allele derived from Heron or the Falcon derived *Pina-D1b* hard type *Pina* null allele. Both Falcon and Heron contain the *Pinb-D1a* soft type *Pinb* allele (Giroux and Morris, 1998) (Table 1). Two accessions of hard type Falcon (PI 612556 and PI 612554) and two accessions of soft type Falcon (PI 612555 and PI 612553) formed the first set of NILs. The second set of NILs consisted of sibling accessions of Gamenya, another Australian cultivar classified as hard Gamenya (accessions PI 612548 and PI 612552) carrying the *Pina* null mutation (*Pina-D1b*) and soft Gamenya (accessions PI 612549 and PI 612551) carrying the functional *Pina-D1a* allele (Table 1). All Gamenya lines contain soft type *Pinb* (*Pinb-D1a*). Seeds from this group of genetic material were obtained from single row plots grown in 2006 at the Arthur H. Post Field Research farm near Bozeman, MT under irrigated condition. Each plot was a 3-m row seeded with 4 g with row spacing of 30 cm. Plots received 7.6 cm of water 1 wk before and 1 wk after anthesis. At maturity, plots were cut with a binder (Mitsubishi Agricultural Machinery Co; Ltd, Tokyo, Japan), threshed with a Vogel bundle thresher (Bill's Welding, Pullman, WA), cleaned, and weighed.

Puroindoline Overexpressing Transgenic Isolines

The second group of genetic material included Hi-line (Lanning et al. 1992), a hard red spring wheat cultivar carrying a glycine-to-serine change in the tryptophan-rich domain of PINB (*Pinb*-D1b) (Giroux and Morris 1997) and a soft type *Pina* (*Pina-D1a*), and a selected subset of transgenic lines created in the Hi-line background by adding wild type *Pina, Pinb* or both under the control of glutenin regulatory elements *Dy10* (5') and *Dx*5 (3') (Hogg et al. 2004). In this subset the transgenic line with added *Pina* (HGA3), *Pinb* (HGB12), and both *Pina* and *Pinb* (HGAB18) had intermediate, soft and very soft grain texture, respectively (Table 1). These four genotypes were grown in 2004 in two replications of a randomized block design under both rainfed and irrigated conditions in 12 row plots 25.6 m long for the rainfed trial and 15.2 m long for the irrigated trial, with rows 30 cm apart at the Arthur H. Post Field Research farm near Bozeman, MT (Martin et al. 2007). Each plot was harvested with a plot combine, cleaned, and weighed.

Grain Characterization and Dry Milling

Kernel hardness and seed weight were determined using the Single Kernel Characterization System (SKCS) 4100 (Perten Instruments, Springfield, IL) (Table 1). One hundred-fifty g of Falcon NILs and Gamenya isogenic siblings were milled on a Brabender Quadrumat Jr. flour mill (Brabender GmbH, Duisburg, Germany) as described by Campbell et al. (2007). The seeds were tempered to 14% moisture content for the soft NILs and to 15.5% for the hard NILs, conditioned for 24 hr as per Approved Method 26-50 (AACC, 2003). Flour and bran weights were measured and total flour yield was calculated as (grams of flour)/ (grams of flour and bran).

Dry milling of Hi-Line and the transgenic isolines was done as previously described (Martin et al. 2007) using a Miag Multomat pilot scale mill. The mill produced 10 flour streams and four feed streams from three break and five reduction rolls. Straight grade flour was used in our experiment.

Wet Milling

The wet milling determinations were completed on two independent extractions for Falcon and Gamenya derived NILs and on three independent extractions for Hi-Line and soft transgenic isolines. Wet milling of flour was done by a dough-dispersion and centrifugation method (Sayaslan 2006), an adapted method from Czuchajowska and Pomeranz (1993), with some minor modifications. Flour (50-75 g, 14% mb) was mixed in a ML-33777 N50 Hobart mixer (Troy Ohio) with gradual addition of water (45-50 ml, 25 ºC) until the mixture became a cohesive, stiff dough, and cleaned itself from the mixing bowl (3-4 min). The developed stiff dough was covered with 150 ml water at 25 °C and was rested at room temperature for 30 min. The dough and liquid was transferred to a blender (TSK- 9368AP, China) and dispersed at high speed for 1 min. The slurry was transferred to 250 ml Sorvall centrifuge bottles and centrifuged at 2500 × g at 25 ºC for 15 min using a Sorvall super T21 centrifuge (Kendro Laboratory Products, Newtown, CT) , then the supernatant was weighed and discarded. The top layer which consisted mainly of gluten, insoluble pentosans, damaged starch, and small granular starch, was carefully removed from the bottom layer which consisted of primary prime starch. The primary prime starch was weighed and dried for 2 days at 37 ºC using a forced air incubator. The gluten was hand manipulated in a beaker under three

consecutive 150 ml water washes to obtain the cohesive wet gluten and starch milk. The wet gluten was partly frozen (~30 min at – 20ºC), cut into ~ 2 cm^3 pieces and incubator dried for 3 days at 37 ºC. The starch milk obtained from gluten washing was collected and centrifuged at 2500 × g at 25 ºC for 15 min. The supernatant from the second centrifugation was discarded and the top partly pigmented layer (tailing) was separated from the secondary prime starch. Both fractions were incubator dried for 2 days at 37 ºC. The dried fractions obtained from each independent extraction were coarsely ground using a mortar and pestle and then ground using a Perten Laboratory Mill 3303 (Perten Instruments, Springfield, IL). Starch recovery was determined by dividing the total weight of prime starch including primary prime starch and secondary prime starch over the weight of starch in flour.

Moisture, starch and protein contents of the flours and the dried wet milling products were measured after recording dry weights of each fraction. Moisture was determined by AACC Method 44–15A (AACC, 2003). Starch content was measured by a total starch assay (AACC Method 76-13, 2003) using a Megazyme kit (Megazyme International. Bray, Co Wicklow, Ireland), Protein content (N × 5.7) was measured using a Leco FP-2000 (Leco Corp; St. Joseph, MI). Both starch and protein percentages were converted to a dry weight basis after determining as-is moisture content.

Data Analyses

All variables from wet milling were analyzed via analysis of variance using PROC GLM in SAS ((SAS Institute, Inc; Cary, NC). For the Hi-Line and transgenic isolines the model was analogous to a randomized block split plot combined over

environments where genotypes were main plots and independent extractions were subplots. Comparisons between genotypes were made using LSD. The hard/soft near isogenic pairs were analyzed using a model that accounted for independent extractions, genotypes (cultivars x hard vs. soft combination), and sibs within each genotype. Comparisons were made between hard and soft for each cultivar and averaged over cultivars.

Results

Two sets of soft/hard NILs differing in *Pina* function and grain hardness were used for dry and wet milling. The first set consisted of Falcon derived soft/hard NILs with two soft accessions carrying *Pina-D1a* and two hard accessions carrying the *Pina-D1b* allele; the second set consisted of soft/hard Gamenya NILs consisting of two soft *Pina-D1a* and two hard *Pina-D1b* accessions (Morris and Allan 2001; Morris et al. 2001a) (Table 1). All Falcon and Gamenya NILs have the soft-type *Pinb* allele (*Pinb-D1a*). Single kernel characterization system (SKCS) grain hardness was 82 in hard Falcon and Gamenya NILs and 29 and 37.5 for soft Falcon and soft Gamenya NILs, respectively. The Falcon NILs had higher kernel weight than the Gamenya NILs (Table 1). The second group of the genetic material used for starch extractability assay was the transgenic isolines used by Martin et al. (2007) in a study assessing the effect of grain hardness on pilot scale milling quality. This subset included Hi-Line hard red spring wheat, Hi-Line transgenic isolines overexpressing PINA (HGA3), PINB (HGB12) or both PINA and PINB (HGAB18) (Table 1). Grain texture ranged from very soft

(HGAB18) to hard (Hi-Line) (Table 1). Kernel weight was higher in HGB12 and Hi-Line than HGAB18 and HGA3 respectively (Table 1). The flour yield from the hard Falcon and Gamenya NILs were higher than their soft isolines, and that relationship was also seen in the transgenic isolines in which hardness was positively correlated with flour yield (Table 2). Flour protein content did not show any significant relationship with grain hardness in the Gamenya and Falcon NILs. For the Hi-Line isolines, flour protein was highest for the untransformed control variety Hi-Line and lowest in HGB12. Flour starch was not significantly related to hardness for any of the comparison groups (Table 2). All flours were then fractionated via a wet milling dough-dispersion and centrifugation procedure (Czuchajowska and Pomeranz 1993; Sayaslan 2006). The wet milling procedure resulted in the recovery of the four main fractions primary prime starch, secondary prime starch, gluten, and tailings. Primary prime starch resulted from the first centrifugation of the slurry. Separation of the phases after centrifuge gave supernatant, unwashed gluten (gluten with adhering starch, the tailings [mainly cell walls and pentosans]), and primary prime starch fractions. The mean yield (dry weight) of all combined fractions separated from the four NILs is given in Table 3.

Table 1. Kernel characteristics and *Puroindoline* genotypes of soft/hard Falcon and Gamenya near isogenic lines and Hi-Line transgenic isolines.

Genotype	Native *pin*[a]	Added *Pin* coding sequence[b]	SKCS grain hardness[c]	Kernel weight[c] (mg)
Hard Falcon[d]	*Pina-D1b/Pinb-D1a*		82	34.6
Soft Falcon[d]	*Pina-D1a/Pinb-D1a*		29	33.5
Hard Gamenya[d]	*Pina-D1b/Pinb-D1a*		82	29.5
Soft Gamenya[d]	*Pina-D1a/Pinb-D1a*		38	26.6
P value[e]			<.0001	0.078
P value[f]			<.0001	0.007
CV%			4.3	3.8
LSD(0.05)			6.9	3.3
Hi-Line	*Pina-D1a/Pinb-D1b*		73.7	35.0
HGA3	*Pina-D1a/Pinb-D1b*	*Pina-D1a*	42.0	33.6
HGB12	*Pina-D1a/Pinb-D1b*	*Pinb-D1a*	9.7	35.4
HGAB18	*Pina-D1a/Pinb-D1b*	*Pina-D1a/Pinb-D1a*	6.4	34.3
P value[f]			<.0001	0.007
CV%			5.8	1.4
LSD(0.05)			3.3	0.8

[a] Native *pin* refers to the wild-type *Pin* allele residing at the *Ha* locus. *Pina-D1b* and *Pinb-D1b* contain a null and a glycine-serine mutation, respectively, while *Pina-D1a*, and *Pinb-D1a* are the soft type functional alleles.
[b] Coding sequence from alleles used to transform Hi-Line hard red spring wheat (Hogg et al. 2005).
[c] Kernel characteristics for Falcon and Gamenya NILs were determined using the Single Kernel Characterization System. Data for Hi-Line and transgenic isolines were taken from Martin et al. (2007).
[d] The values for these genotypes were averaged over two lab replications and two accessions for each genotype planted in irrigated environments.
[e] *P* value for average of soft vs. hard NILs.
[f] Genotype main effect *P* value.

Table 2. Flour yield and flour protein and starch content for soft/hard Falcon and Gamenya near isogenic lines and Hi-Line transgenic isolines.

Genotype	Flour yield[a]	Flour protein[b, c]	Flour starch[c,d]
	g kg^{-1}	g kg^{-1}	g kg^{-1}
Hard Falcon[e]	741	141	747
Soft Falcon[e]	706	143	755
Hard Gamenya[e]	724	158	762
Soft Gamenya[e]	710	151	798
P value[f]	0.051	0.48	0.233
P value[g]	0.149	0.047	0.258
CV%	1.8	2.7	2.9
LSD(0.05)	37	11	63
Hi-Line[h]	744	155	734
HGA3[h]	732	153	735
HGB12[h]	714	149	738
HGAB18[h]	711	152	768
P value[g]	<0.0001	<0.0001	0.322
CV%	1	0.5	3.7
LSD(0.05)	11	1.3	48

[a] Flour yield for Falcon and Gamenya isogenic lines obtained by a Brabender Quadrumat Jr. flour mill and straight grade flour yield for Hi-Line and transgenic isolines obtained by Martin et al. (2007) using a Miag Multomat pilot scale mill.
[b] Flour protein was obtained using a Leco FP-2000 Nitrogen Determinator.
[c] Means reported on a dry basis.
[d] Flour starch was obtained by Megazyme Amyloglucosidase/α-Amylase method using Megazyme kit.
[e] Values for these genotypes were averaged over two accessions for each genotype planted in irrigated environments.
[f] *P* value for average of soft vs. hard NIL.
[g] Genotype main effect *P* value.
[h] Values averaged over two field replications for rainfed and irrigated environments at Bozeman, MT.

The soft NIL had higher total prime starch yields than the hard isolines when averaged over both genotypes ($P = 0.045$). The difference between hard and soft isolines in total prime starch yield was not significant for Falcon but was significant for Gamenya ($P = 0.05$). Gluten yield was not different between soft and hard NIL in either the Falcon or Gamenya background, although the trend toward soft NIL with increased gluten content was evident. Tailing fraction yield was significantly higher in hard than soft NIL. Total product recovery percentage difference was nonsignificant between soft and hard Falcon NIL but was significant in the Gamenya background in which the soft NIL had higher total product recovery than the hard NIL (77 vs. 72%). Given the trend toward increased starch yield in soft NIL in both the Falcon and Gamenya backgrounds, we explored whether higher transgenically conditioned PIN levels would improve starch extractability relative to a hard control. Wet-milling yield of Hi-Line and three transgenic isolines is given in Table 3. Genotypes differed for yield of primary prime starch, total prime starch, and gluten. Yields of total prime starch were higher in all three transgenic isolines overexpressing one or both PINs relative to Hi-Line ($P < 0.0001$), with the highest yield in the intermediate hardness line HGA3. Starch recovery percentages showed a trend similar to yields of total prime starch. The soft and supersoft textured lines, HGB12 and HGAB18, were not significantly different from each other in starch yield. Interactions between genotype and environment were not detected for total prime starch yield. Although laboratory replicates showed significant differences in total prime starch yields ($P < 0.0001$), genotype interaction with laboratory replicates was not detected ($P = 0.243$). The gluten yield was significantly reduced in HGA3 relative to untransformed Hi-

Line, HGB12, and HGAB18. Environmental effect was highly significant for dry gluten yield ($P = 0.0066$) with flour milled from seed grown under rainfed conditions with more gluten yield than flour milled from seed grown under the irrigated environment (218.7 vs. 194.1 g/kg), but no interaction effect was detected between genotype and environment ($P = 0.824$). Laboratory replicates did not show any significant effect on gluten extractability ($P = 0.149$). Neither genotype (Table 3) nor environment had a significant effect on tailing yield. The weight of this fraction showed differences among laboratory replicates ($P < 0.0001$), but there was no interaction effect between genotype and laboratory replicates for tailing yield ($P = 0.666$). The recovery percentage of total extracted fractions showed significant difference among genotypes; Hi-Line had the lowest recovery percentage. The recovery percentage did not follow the same trend that was observed for total prime starch. The softest genotypes (HGAB18 and HGB12) showed higher recovery compared with HGA3, due to higher gluten yield. Genotype-by environment effect was not detected for total recovery percentage, but rainfed samples had a higher recovery percentage than irrigated samples (0.752 vs. 0.735%, $P = 0.0034$). In fact, the higher total recovery percentage for rainfed versus irrigated environment is a consequence of environmental effect on gluten yield extraction. The higher gluten extractability from rainfed versus irrigated samples is directly related to the environmental effect on the flour protein (145 and 152 g/kg of protein obtained from irrigated and rainfed, respectively; $P < 0.0001$).

Table 3. Yield data (dry basis) for wet milling fractions from flour of soft/hard Falcon and Gamenya near isogenic lines and Hi-Line transgenic isolines.

Genotype	Primary prime starch	Secondary prime starch	Total prime starch	Gluten	Tailing	Total recovery	Starch recovery[a]
	g kg^{-1}	g kg^{-1}	g kg^{-1}	g kg^{-1}	g kg^{-1}	%	%
Hard Falcon[b]	291.5	129	420.5	174.5	109	70	56.3
Soft Falcon[b]	340	103	443	195	82.5	72	58.7
Hard Gamenya[b]	259	118	377	224.5	123.5	72	49.5
Soft Gamenya[b]	315	111	426	238.5	103	77	53.4
P value[c]	0.011	0.086	0.045	0.121	0.002	0.117	0.148
P value[d]	0.036	0.007	0.07	0.009	0.005	0.127	0.05
CV%	10	14	7	9	9	4	7
LSD(0.05)	51.3	27.6	48.9	32.7	17	5	6.38
Hi-Line[e]	304.3	130.2	434.5	203.8	83.5	72	59.3
HGA3[e]	367.1	117.4	484.5	186.3	73.4	74	66
HGB12[e]	365.6	105	471.6	210.5	72.3	75	63.9
HGAB18[e]	336.3	128.3	464.7	218.2	71	75	60.5
P value[d]	0.026	0.407	<.0001	0.025	0.263	0.017	<.0001
CV%	15	33	3.9	11	22	3.4	3.9
LSD(0.05)	41.9	38.3	17	18.6	9.51	1	4

[a] Starch recovery percentage determined by dividing recovered starch (g kg^{-1}) over total starch (g kg^{-1}) in flour.
[b] Values for these genotypes were averaged over two laboratory replicates and two accessions for each genotype planted in irrigated environments.
[c] *P* value for average of soft vs. hard NILs.
[d] Genotype main effect *P* value.
[e] Values for these genotypes were averaged over 3 laboratory replicates and two field replicates for rainfed and irrigated environments at Bozeman, MT.

The starch content (dry basis) of fractions separated from the Falcon and Gamenya NIL are given in Table 4. Starch content of primary and secondary prime starch was not

different between hard and soft NIL within and between backgrounds. However, the soft Falcon NIL gluten was higher in starch content than gluten extracted from hard Falcon NIL but the hard and soft Gamenya NIL did not differ for this trait. The starch content in the tailing fraction obtained from the hard Falcon NIL was higher than in the counterparts. The mean value for starch content of primary and secondary prime starch separated from Hi-Line and its transgenic isolines was not different among genotypes, although genotype had a significant effect on starch content of gluten and tailing fractions for this group. Gluten and tailing fraction starch content was lowest in HGA3. The tailing fraction obtained from Hi-Line had a higher amount of starch than all three transgenic isolines. The mean protein content (dry basis) of fractions separated from the Falcon and Gamenya NIL is given in Table 5. Although the differences in protein content of primary and secondary prime starch were not significant, the primary prime starch fraction from the soft NIL had a higher protein content than that from the hard NIL. In addition, the soft version of Falcon had a lower protein content gluten than the hard counterpart, and the Gamenya NIL showed the same trend. The amount of protein present in tailing fractions differed significantly in both Falcon and Gamenya backgrounds in that, in each case, soft NIL had significantly more protein present in the tailing fraction. Protein content of primary prime starch separated from Hi-Line and Hi-Line transgenic isolines was not significantly different among genotypes, even though all transgenics trended higher in protein content. Similarly, protein content of secondary prime starch and tailing fractions from the three transgenics were substantially higher than Hi-Line.

Table 4. Starch content (dry basis) of fractions obtained from wet milling of flours from soft/hard Falcon and Gamenya near isogenic lines and Hi-Line transgenic isolines.

Genotype	Primary prime starch [a]	Secondary prime starch [a]	Gluten [a]	Tailing [a]
		g kg^{-1}		
Hard Falcon[b]	993.2	985.6	383.4	814.2
Soft Falcon[b]	964.5	980.3	476.1	763.2
Hard Gamenya[b]	1000.0	988.2	449.8	865.8
Soft Gamenya[b]	999.6	974.7	470.5	865.4
P value[c]	0.347	0.437	0.019	0.110
P value[d]	0.447	0.845	0.032	0.003
CV%	4	2.3	8.7	3.5
LSD(0.05)	64.4	37.6	62.8	46.6
Hi-Line[e]	957.3	871.3	385.2	774.7
HGA3[e]	993.7	765.1	331.2	708
HGB12[e]	978.0	853.7	408.9	728.7
HGAB18[e]	980	852	429	719.2
P value[d]	0.291	0.208	0.0003	0.008
CV%	4.6	15	11	5.9
LSD(0.05)	89.9	211.5	99	58.7

[a] Determined by Megazyme Amyloglucosidase/α-Amylase method using Megazyme kit.
[b] Values for these genotypes were averaged over two laboratory replicates and two accessions for each genotype planted in irrigated environments.
[c] *P* value for average of soft vs. hard NILs.
[d] Genotype main effect *P* value.
[e] Values for these genotypes were averaged over 3 laboratory replicates and two field replicates for rainfed and irrigated environments at Bozeman, MT.

Table 5. Protein content (dry basis) of fractions obtained from wet milling of flours for soft/hard Falcon and Gamenya near isogenic lines and Hi-Line transgenic isolines.

Genotype	Primary prime starch [a]	Secondary prime starch [a]	Gluten [a]	Tailing [a]
		g kg^{-1}		
Hard Falcon[b]	8.1	5.1	513.8	47.4
Soft Falcon[b]	8.6	7.9	446.5	88.4
Hard Gamenya[b]	5.5	4.6	473	32.2
Soft Gamenya[b]	6.6	6.1	458.9	39
P value[c]	0.230	0.117	0.066	0.008
P value[d]	0.020	0.304	0.150	0.002
CV%	16	41	8.1	26
LSD(0.05)	1.89	4.01	62.4	22.3
Hi-Line[e]	6.7	3.9	535	23.9
HGA3[e]	7.6	8.1	579.9	59.8
HGB12[e]	7	7	502	58.4
HGAB18[e]	8	6.1	489.8	52.3
P value[d]	0.229	0.004	0.0007	0.001
CV%	22	39	8.5	40
LSD(0.05)	2.02	3.91	69.2	23.5

[a] Determined by Leco FP-2000 Nitrogen Determinator.
[b] Values for these genotypes were averaged over two laboratory replicates and two accessions for each genotype planted in irrigated environments.
[c] *P* value for the average of soft vs. hard NILs.
[d] Genotype main effect *P* value.
[e] Values for these genotypes were averaged over 3 laboratory replicates and two field replicates for rainfed and irrigated environments at Bozeman, MT.

Discussion

Czuchajowska and Pomeranz (1993) extracted starch from a set of randomly selected soft and hard wheat varieties. They found the yield of prime starch was higher from the low protein soft wheat flours than from the high-protein hard wheat flours. In order to test whether this difference in starch extractability is linked to the *Ha* locus on the short arm of chromosome 5D and to Pin expression, we used two sets of soft/hard NILs differing in *Pina* function and grain hardness and transgenic *Pin* isolines varying in grain hardness and *Pina* and/or *Pinb* expression level. The soft/hard NILs were created in either the Falcon or Gamenya varieties and in both cases we used two hard type *Pina* null lines (*Pina-D1b*) and two soft type *Pina* functional allele *(Pina-D1a*) lines (Morris and Allan 2001; Morris et al. 2001a) (Table 1). The *Ha* locus did affect total prime starch with the soft having more than the hard counterparts. The *Ha* locus had no effect on starch or protein content of primary or secondary prime starch fractions in either genetic background (Tables 4 and 5 respectively). This fact along with the lack of any significant differences in the starch content of flour between NILs in either genetic background (Table 2) implies that the *Ha* locus directly impacts starch extractability in these NILs since starch recovery is increased while tailings are decreased. Since tailing starch content is not different between hard and soft NILs (Table 4) ($P = 0.11$) the differences in starch extractability is not because of the starch loss via tailing fractions in hard lines. In order to determine the effect of added PIN and grain softness upon starch yield, we selected the transgenic isolines used by Martin et al. (2007) which vary markedly in *Pin* expression level and grain hardness. All three transgenics had increased recovery of

starch with the highest yield seen in the intermediate textured line HGA3 relative to Hi-Line. Significantly, starch content of recovered prime starch fractions was not decreased in the transgenics and gluten starch content was significantly lower in HGA3 relative to Hi-Line. The increased starch extractability seen in HGA3 versus Hi-Line appears to result from greater separation of starch and gluten in that starch content of gluten and tailings are reduced (Table 4) while protein content of both gluten and tailings are increased in HGA3 relative to Hi-Line (Table 5).

Wheat grain hardness is most likely determined by the degree of adhesion between starch granules and protein matrix. Friabilin (PINA and PINB) may decrease grain hardness via reduction of the close adherence between starch granule and gluten (Anjum and Walker 1991). Greater separation of starch and gluten in softer textured wheats is likely achieved via puroindolines coating of starch granules and thus preventing tight adhesion between starch granules and the surrounding protein matrix (Swan et al 2006). Starch loss through supernatant may be an explanation for less starch extractability from Hi-Line than softer genotypes. Higher starch damage in hard wheat than soft wheat due to milling has been reported (Symes 1965; Letang et al. 2001; Van Der Borght et al. 2005). The proportion of damaged starch for Hi-line and the three transgenic isolines used in this study was measured in Martin et al. (2007). The hard textured Hi-Line had the most starch damage followed by the intermediate-textured HGA3, and soft-textured HGB12 and HGAB18. Damaged starch granules have been mentioned as part of squeegee starch that reduce prime starch yield in wet milling (Van Der Borght et al. 2005). Moreover damaged starch fraction increases both water

absorption (up to three times) (Van Der Borght et al. 2005; Tester et al. 2006) and endogenous B-amylase hydrolysis, which generates maltose which is useful in the baking industry (Tester et al. 2006). Unlike native starch granules, which are semicrystalline, insoluble and as a consequence inaccessible to hydrolysis, damaged starch fragments are amorphous, soluble and readily hydrolysed (Tester et al. 2006). The increased water-solubility and susceptibility to hydrolysis due to the level of starch damage may explain the higher rate of starch loss through supernatant in wet milling process of hard wheat versus softer genotypes.

Significant differences among genotypes for tailing yield was not detected (Table 3), but tailing fraction obtained from Hi-Line showed higher starch content and lower protein content than transgenic isolines (Table 4 and 5 respectively). The presence of starch in tailing can be a result of the admixture of secondary prime starch with tailing pentosans during last phase separation. Small starch granules are admixed with pentosans in tailing, whereas large starch granules make the pure starch fractions (Czuchajowska and Pomeranze 1993). The higher rate of starch loss through admixing with tailing fractions in Hi-Line hard wheat may be considered as an explanation for differences in starch extractability between Hi-Line and softer transgenic isolines. There are some evidences that weaken the role of tailing in starch loss. First, secondary prime starch is a small portion of total starch. Second, there were significant differences in primary prime starch yield between Hi-Line and softer textured isolines. And third, the secondary prime starch obtained from genotypes did not show significant differences in starch content, the

expected result if there were significant differences in the admixture between secondary prime starch and tailing due to the phase separation procedure.

Conclusions

In conclusion this study demonstrates that grain hardness is a significant factor in determining the suitability of wheat for wet milling processes. This effect was observed in hard and soft NILs that differed in the state of their *Ha* locus, and also in transgenic isolines having increased dosage of *Pina* and/or *Pinb*. The overexpression of functional *Pins* leads to an increase in starch extractability and increased total recovery with highest yields seen in the transgenic line having intermediate grain texture.

CHAPTER 2

CREATION AND FUNCTIONAL ANALYSIS OF NEW *PUROINDOLINE* ALLELES IN *TRITICUM AESTIVUM*

Introduction

Variation in wheat (*Triticum aestivum* L.) grain hardness is the single most important trait that determines wheat end-use properties (reviewed in Morris and Rose 1996). Grain hardness variation is controlled predominantly by the *Hardness* (*Ha*) locus, located on the extreme distal end of chromosome 5DS (Mattern et al. 1973; Law et al. 1978; Ram et al. 2002; Campbell et al. 1999). The *Ha* locus functionally consists of the *Puroindoline a* and *b* genes (*Pina* and *Pinb* respectively) (Giroux and Morris 1998; Wanjugi et al. 2007a). Soft texture (*Ha*) is the result of both genes being in their wild-type allelic state (*Pina-D1a*, *Pinb-D1a*) while hard texture (*ha*) results from mutations in either *Pina* or *Pinb* (Giroux and Morris 1997, 1998). PINA and PINB are cysteine rich proteins which are unique among plant proteins in having a hydrophobic tryptophan-rich domain (Blochet et al. 1993) and together make up the grain hardness marker protein friabilin. Friabilin was originally described as a 15 kDa protein (Greenwell and Schofield 1986; Morris et al. 1994) present in higher levels on the surface of water washed starch prepared from soft wheats relative to the levels found on hard wheat starch. Similarly, Greenblatt et al. (1995) observed that water washed starch from soft wheats had more glyco- and phospholipids than hard wheat water washed starch. The correlation between friabilin and glyco- and phospholipid levels on the surface of starch granules led to the

hypothesis that PINs are co-localized to the glyco- and phospholipid rich surface of the amyloplast membrane and that this localization may be mediated by their tryptophan-rich domain (Giroux et al. 2000). PINs were conclusively shown to alter grain texture in cereals via transgenic studies. We first demonstrated that PINs can modify grain texture and fungal resistance by transforming rice with both PINs (Krishnamurthy and Giroux 2001, Krishnamurthy et al. 2001). In addition, we demonstrated that *Pin* mutations are causative to hard wheat phenotypes via transgenic complementation studies in which soft phenotypes were achieved when hard wheat *Pin* alleles were complemented with the corresponding wild-type *Pin* allele (Beecher et al. 2002; Martin et al. 2006). Transgenic manipulations have also demonstrated that while both PINA and PINB levels limit grain softness in soft wheats, PINB is a greater limiting factor than PINA (Swan et al. 2006). Further, high levels of either PIN alone can lead to intermediate levels of grain hardness. However, full PIN starch association and a soft grain phenotype is achieved only with both PINs present (Hogg et al. 2004, Wanjugi et al. 2007a). Variation in PIN function affects not only grain hardness but nearly all product quality traits (Hogg et al. 2005, Wanjugi et al. 2007b, Martin et al. 2008) and associated milling effects are also seen in pilot scale milling studies (Martin et al. 2007). While transgenic manipulation could be used to create nearly any desired level of grain hardness, non-transgenic manipulations of the puroindolines in hexaploid wheats are limited due to limited *Pin* allelic variation. The vast majority (>95%) of U.S. hard wheat germplasm contains one of two *Ha* locus mutations (*Pina-D1b* or *Pinb-D1b*) (Morris et al. 2001). Despite this small level of diversity, it is known that hard wheats carrying the *Pinb-D1b* allele wheats are ~7

hardness units softer and superior in milling and bread baking properties to hard wheats carrying the *Pina-D1b* allele (Martin et al. 2001, 2007). Similarly, the addition of an A genome Ha locus to a soft wheat decreased grain hardness and modified milling yield and particle size (Campbell et al. 2007). From an end-product quality point of view the more *Pin* variation found or created the greater chance that some alleles will confer improved end-product quality. The *Pina* and *Pinb* missense alleles discovered in various germplasm surveys are shown in Table 6. While these alleles may vary in function somewhat, they each likely retain very little residual function as all were found in hard wheats. New functional *Pin* alleles could be introduced via direct hybridization between *Ae. tauschii*. and hexaploid wheat (Gill and Raupp 1987; Cox 1998) or the production of synthetic hexaploids (Warburton et al. 2006) followed by crossing of the synthetic hexaploids to *T. aestivum*. While numerous Pin allelic variants have been reported (Gedye et al. 2004, Massa et al. 2004) (Table 6), they are not useful in terms of understanding PINA and PINB function for two reasons. First, a significant amount of the hardness variation is assignable to both *Ae. tauschii* and durum parents and secondly most of the haplotypes contain a large number of amino acid changes relative to *Pina-D1a* and *Pinb-D1a* (Table 6). In addition, all of the *Pin* alleles appear to condition a soft phenotype such that they appear limited as a source of grain hardness variation (Gedye et al. 2004).

Table 6. Naturally occurring *Puroindoline* missense alleles in *T. aestivum* or *Ae. tauschii*

Allele[a]	Mutation[b]	Reference	Grain Hardness[c]
		T. aestivum	
Pina-D1a	WT	Giroux and Morris (1997)	
Pina-D1f[d]	R86Q	Massa et al. (2004)	
Pina-D1m	P63S	Chen et al. (2006)	
Pinb-D1a	WT	Giroux and Morris (1997)	
Pinb-D1b	G75S	Giroux and Morris (1997)	73±14
Pinb-D1c	L89P	Lillemo and Morris (2000)	80±14
Pinb-D1d	W73R	Lillemo and Morris (2000)	68±14
Pinb-D1l	K45E	Pan et al. (2004)	
Pinb-D1q	W73L	Chen et al. (2005)	
Pinb-D1t	G76R	Chen et al. (2006)	
Pinb-D1v	A8T, L9I	Chang et al. (2006)	
Pinb-D1w	S144I	Chang et al. (2006)	
		Ae. tauschii	
Pina-D1i[e]	R86Q, R49S	Gedye et al. (2004)	
Pina-D1j[e]	R86Q, P136R	Gedye et al. (2004)	
Pinb-D1h[f]	14 missense	Massa et al. (2004)	
Pinb-D1j	9 missense	Massa et al. (2004)	

[a] The updated Allele designations from the Catalogue of Gene Symbols for wheat: McIntosh et al. (2007) were used for this table. Alleles *Pina-D1b*, *Pina-D1k*, *Pina-D1l*, *PinaD1n*, *Pina-D1P*, *Pina-D1q*, *Pinb-D1e*, *Pinb-D1f*, *Pinb-D1g*, *Pinb-D1p*, *Pinb-D1r*, *Pinb-D1s*, *Pinb-D1u*, *Pinb-D1aa*, and *Pinb-D1ab* result from gene deletion, base deletion (frame shift mutation), or nonsense mutations, and so are not included here. Allele *Pina-D1g* carries one silent mutation and encodes an identical protein to *Pina-D1a*.

[b] Alleles are numbered relative to the starting methionine of the proteins.

[c] Single Kernel Characterization System hardness index, average ± standard deviation reported for the near isogenic lines developed from some of the *pinb* natural missense mutations in Alpowa background. The reference Alpowa parent had a SKCS hardness average of 24 +/- 14 (Morris and King 2008).

[d] Alleles *Pina-D1c*, *d*, e, *f*, *h* and *o* encode an identical protein product.

[e] Alleles *Pina-D1i* and *Pina-D1j* were identified in synthetic wheats.

[f] Alleles *Pinb-D1h*, *I*, *k*, *m*, *n* and *o* encode an identical protein product.

Therefore the creation of new *Pin* alleles in one background where each contains a single amino acid change via target-selected mutagenesis is the best approach for creation of new *Puroindoline* alleles in hexaploid wheat. The new alleles would have the potential to improve understanding of PIN function as well as improving wheat end product quality. Direct phenotyping has been used in cases where the target genes confer a detectable phenotype (Zhu et al. 1998; Jander et al. 2003). Our approach described here was to test screening methods useful in identifying new EMS induced *Pin* alleles. We utilized direct grain hardness phenotyping followed by direct sequencing to identify 18 new *Puroindoline* alleles. The results indicate that this approach will successfully create a large resource of new *Puroindoline* alleles to allow us to accomplish two goals. First to reveal the regions of PINA and PINB critical for conferring softness and second to improve wheat end product quality via incorporation of a broader range of PIN alleles into both soft and hard wheat varieties.

Materials and Methods

Creation and Selection of EMS-induced Population

A wheat EMS-induced M_1 population was created using a protocol similar to Slade et al. (2005) with some modifications. Approximately 2000 M_0 wheat seeds of the soft white spring cultivar Alpowa (PI566596) were soaked in 1% EMS for 18 hrs at room temperature followed by decanting of the EMS. Seeds were rinsed under cold running tap water for 5 hrs before planting individually in the greenhouse in 4 x 4 cm cells filled with a peat moss soil mix (Sunshine Mix #1, Sun Gro Horticulture Inc., Bellevue, WA).

Greenhouse conditions consisted of target temperatures of 22°C and 14°C for day and night, respectively, with supplemental lighting providing 400 μE $m^{-2}s^{-1}$ consisting of 1,000-W metal halide lamps on to provide a 16 hr day. Plants were watered as needed with a 100 ppm N-P-K solution (Peters General Purpose Plant Food, The Scotts Company, Marysville, OH). One thousand heads from surviving M_1 plants were planted as head rows in 0.6 m rows spaced 30 cm apart at the Arthur H. Post Field Research farm near Bozeman, MT in April 2006. Single heads from individual plants as well as a bulk harvest of all remaining heads were conducted on the 630 fertile rows by hand sickling followed by threshing using a single plant thresher (Bill's Welding, Pullman, WA). Kernel hardness and seed weight were determined using the Single Kernel Characterization System (SKCS) 4100 (Perten Instruments, Springfield, IL) (Martin et al. 1993) on 50 seeds from each bulk seed sample.

Creation of F_2 Populations

The SKCS data from 630 M_2 families was sorted to give two groups consisting of lines enriched for those containing hard textured seeds. The first group consisted of 121 lines with a grain hardness average above 47 and the second group had grain hardness average less than 47 with a standard deviation above 16. A seed to seed standard deviation above 16 generally indicates a mixture of hardness genotypes (Giroux and Morris 1998). SKCS grain hardness analysis of at least 10 individual seeds from each of 10 individual heads from each of these M_2 families was used to enrich for homozygous *Pin* mutants for each family. A subsample consisting of candidate homozygote mutants from each family was planted in the greenhouse and used for direct sequencing of *Pina*

and *Pinb*. Plants positive for *Pin* mutations were used as pollen source in crosses back to Alpowa non-mutant parent to produce an F_1. One hundred sixty F_1-derived F_2 seeds for each cross along with the twenty M_2:M_4 mutant and Alpowa parental seeds were planted at the Arthur H. Post Field Research farm near Bozeman, MT with within row plant spacing of 15 cm and 30cm between rows in May of 2007. F_2:F_3 seeds were harvested from single F_2 plants and threshed using a single plant thresher (Bill's Welding, Pullman, WA). Twenty F_2 individuals per cross were analyzed by both SKCS and *Pin* genotyping and 1-10 individuals per each M_4 parental line as well as Alpowa non-mutant seeds were analyzed by SKCS. SKCS hardness average and standard deviation was measured on a 50-kernel subsample from each individual F_2 and M_4 parent plant. A 5-10 kernel subsample of the same individuals were bulk planted in the greenhouse for DNA extraction followed by genotyping to identify the genotype of parental line.

DNA Extraction, PCR Amplification, and Sequencing

Leaf tissue was collected at the 2 to 3 leaf stage from individual M_2:M_3 plants and DNA was extracted according to Riede and Anderson (1996). PCR reactions for direct sequencing contained 100 ng of genomic DNA, 15 pmol of each primer, 200 μ*M* of each dNTP, 1X *Taq* DNA polymerase reaction buffer, 0.65 unit of *Taq* DNA polymerase (Promega, Madison, WI), and 2 m*M* of $MgCl_2$ in 20 μl. GenBank accession numbers are X69913 and X69912 respectively, for *Pina* and *Pinb* and are identical to the *Pina-D1a* and *Pinb-D1a* allele sequences present in Alpowa. The PCR primers were the *Pina* and *Pinb* 3' primers of Gautier et al. (1994), which bind to the last 22 and 24 nucleotides of the 3' end of the *Pina-D1* and *Pinb-D1* coding sequence respectively, combined with the

Pina or *Pinb* forward primers designed by Massa et al. (2004) which bind 55 and 61 nucleotides upstream of the *Pina-D1* and *Pinb-D1* start codon's first nucleotide, respectively. The PCR product sizes of *Pina* and *Pinb* genes were 502 and 508 bp respectively and covered the 447 bp uninterrupted reading frame of each gene (Simeone et al. 2006). While, the complete mRNA's are 631 and 598 bp , for *Pina* and *Pinb*, respectively (Gautier et al. 1994) our ~500 bp PCR products allowed us to sequence all but the last 7-8 codons. The sequence for all four primers is as follows:

Pina-D1 Forward: 5'- GGTGTGGCCTCATCTCATCT-3'

PA3: 5'- TCACCAGTAATAGCCAATAGTG-3'

Pinb-D1 Forward: 5'- AATAAAGGGGAGCCTCAACC-3'

PB3: 5'- TCACCAGTAATAGCCACTAGGGAA-3'

Cycling parameters were 94^{o}C for 3 min, followed by forty cycles of 94^{o}C 30 s, 55^{o}C 30 s, 72^{o}C 1 min, followed by a 5-min final extension at 72^{o}C.

PCR products were separated on 1.5% agarose gels followed by gel purification (Qiagene, Valencia, CA) and were then direct sequenced using the forward amplification primer (SeqWright DNA Technology Services, Houston, Texas). The GAP4 and Pregap4 programs from Staden Package v1.6.0, 2004, (http://staden.sourceforge.net/staden_home.html) were used to analyze the sequences. SIFT (Ng and Henikoff 2003) was used to predict the tentative impact of mutations on protein function. The membrane affinity of individual amino acids were as reported by Thorgeirsson et al. (1996). Thorgeirsson et al. (1996) described the transfer free energies of synthetic peptides with different amino acids in the same position (relative to glycine)

from bilayer to water ($\Delta\Delta G_{bilayer}$) and calculated an individual membrane affinity for each amino acid.

Genotyping of F_2 Populations

The same primer pairs and PCR conditions were used to genotype F_2:F_3 lines for four of the crosses between *Puroindoline* mutants and their non-mutagenized parent Alpowa (Table 8). A minimum of four F_3 plants derived from each F_2 line were bulked for DNA extraction followed by genotyping to identify the genotype of each F_2 parent. Polymorphic co-dominant markers were developed based on the differential restriction digestion of *Pinb* mutant alleles *Pinb-D1C146T* and *Pinb-D1C200T* and *Pina* mutant allele *Pina-D1C187T* using restriction enzymes *Hind*III, *Tsp*451 and *Msp*1 respectively. Genotyping of the $F_{2:}$ F_3 lines obtained from the cross between Alpowa and mutant *Pina-D1G212A* was performed using direct sequencing. The association between new alleles and grain hardness was analyzed via analysis of variance using PROC GLM in SAS (SAS Institute, Inc; Cary, NC).

Results

Creation of New *Puroindoline* Alleles

To obtain and identify new *Ha* locus component gene alleles, we developed an EMS-mutagenized wheat population using the soft white spring variety Alpowa . Alpowa carries the functional alleles of *Puroindoline a* and *Puroindoline b* found in all hexaploid soft wheats, *Pina-D1a* and *Pinb-D1a,* respectively. Single Kernel Characterization System (SKCS) analysis of the 630 M_2-derived M_3 bulk seed obtained from an average of 4 M_2 plants per M_2 family showed that the EMS treatment resulted in reduced average seed size (and vigor) and increased grain hardness relative to Alpowa (Figure 1). The SKCS grain hardness mean was 41.1 and 24.4 for 630 M_2-derived M_3 bulk seed and 10 Alpowa plants seed respectively. The individual seed weight averages were 32.8 and 38.4 mg when averaged over the 630 M_2-derived M_3 bulk seed and 10 Alpowa plants, respectively. From the whole population, a 168 member subpopulation was selected based upon grain hardness (121 lines with grain hardness average above 47 and variable standard deviations) and grain hardness standard deviation (47 lines with grain hardness average less than 47 and grain hardness standard deviation greater than 16) for direct sequencing of *Pina* and *Pinb* coding sequences. The grain hardness of 10 individual M_2-derived M_3 single heads from each selected family was measured to identify potential homozygote *Pina* or *Pinb* mutants. Candidate homozygote mutants were selected on the basis of increased grain hardness. *Pina* and *Pinb* were then sequenced from DNA extracted from plants grown from selected heads from each of the

168 lines. In total 18 new *Pin* alleles were identified and 16 of 18 were homozygous for the *Pin* mutation.

Phenotypic and Genotypic Characterization of New *Puroindoline* Alleles

We found nine new alleles for both *Pina* and *Pinb*. Of the nine *Pina* alleles, four were missense and five were nonsense mutations. Of nine *Pinb* alleles, four were missense, three were nonsense mutations and two were silent (synonymous) (Table 7 and Figure 2).

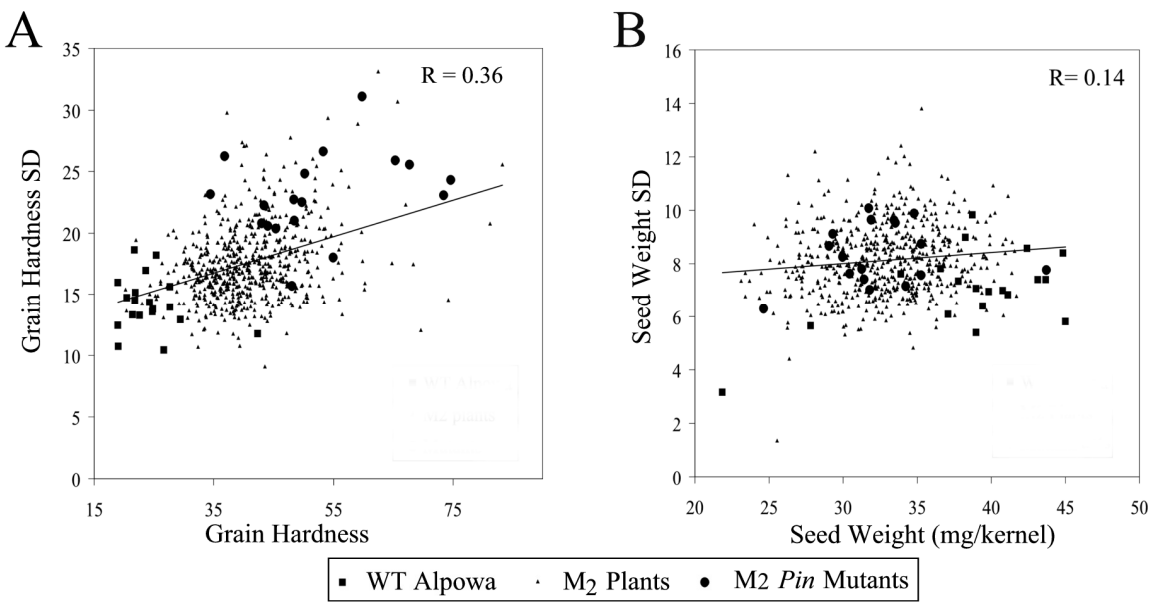

Figure 1. Single Kernel Characterization System (SKCS) hardness (a) and weight (b); SD versus averages measured on 630 M_3 seed bulks obtained from EMS-induced M_2 families.

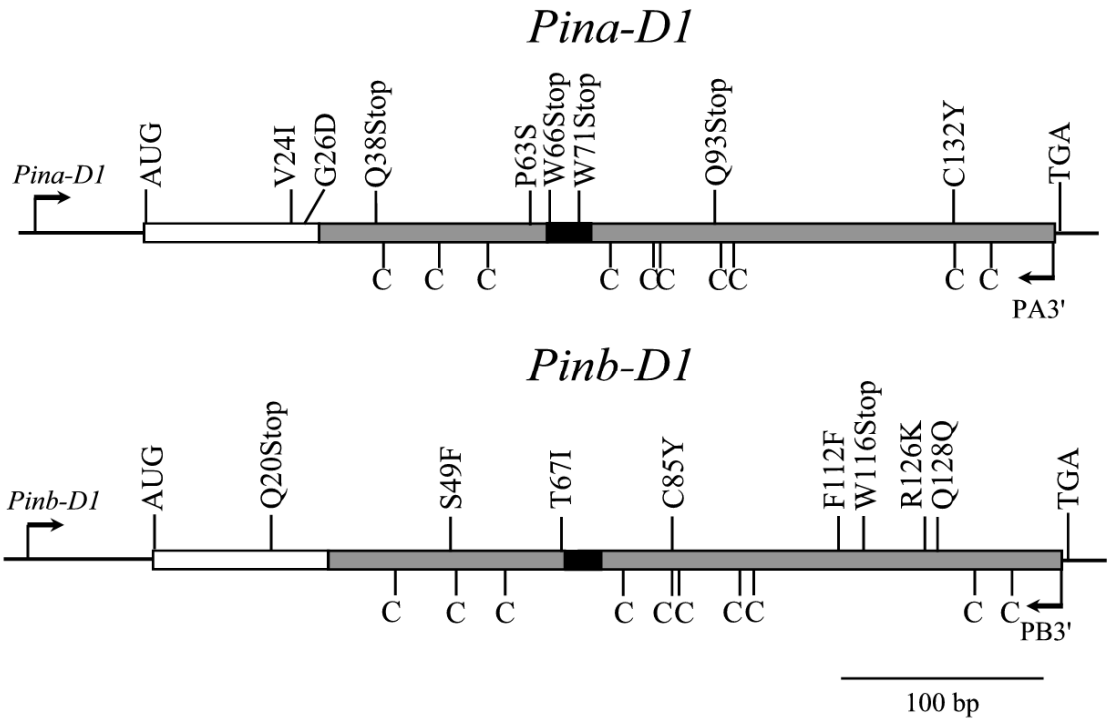

Figure 2. The positions of the induced mutations within the coding regions (boxed region) of *Pina* and *Pinb*. The white boxed area denotes the signal peptide region while the gray/black boxed regions denotes mature peptide sequence with the black boxed region indicating the position of the tryptophan-rich domain. The position of cysteines within PINA and PINB are denoted by 'C's'. The position of the amplification primers is as shown with the forward primers residing within the promoter regions and the reverse primers within the 3' end of the coding sequence.

One of the *Pina* and one of the *Pinb* nonsense mutants were found in duplicate. Also, one of the *Pina* alleles (C187T) had a mutation identical to a naturally occurring *T. aestivum Pina* allele (*Pina-D1m* in Table 6). Grain hardness analysis of the M_2-derived M_3 bulk seeds obtained from these mutants was used as an initial indicator of the mutations impact on the function of these genes (Table 7). The majority of M_2-derived

M_3 bulk seeds obtained from the *Pina* or *Pinb* mutants showed a grain hardness average higher than the grain hardness mean of the whole population (41.2) and a grain hardness standard deviation higher than 20. As expected most of the nonsense mutations led the list of mutants in grain hardness. The fact that two independent nonsense mutants with the same mutations in *Pina* showed a difference of 39 in grain hardness (Table 7, *Pina-D1G198A* found in M_2 lines 376 and 497) is a good indicator that M_2-derived M_3 bulk seeds are not sufficient to test *Pin* gene function or to use for direct phenotypic screening. SIFT (Sorting Intolerant From Tolerant) is a bioinformatics program which uses sequence homology to predict the impact of missense mutations on protein function and potential phenotype (Ng and Henikoff 2003). SIFT scores <0.05 are predicted to be deleterious. Of the eight *Pina* and *Pinb* missense mutations identified, only two were predicted to be tolerated by SIFT (*Pina-D1G70A, Pinb-D1G377A*) (Table 7). Membrane affinities of the mutated and wild type amino acids of PINA and PINB are shown in Table 7 as the individual amino acid affinities reported by Thorgeirsson et al. (1996). The PINA proline to serine change (P63S) in mutant *Pina-D1C187T* PINA resulted in a 0.51 kcal/mol increase in membrane affinity. This value was calculated as 2.15 for the serine to phenylalanine substitution in the *Pinb-C146T* mutant and 2.31 for the threonine 67 to isoleucine substitution in the *PinbD1C200T* mutant. One of the *Pina* mutations and two of the *Pinb* mutations were found in the signal peptides (Figure 2).

Table 7. Allelic series of mutations in *Pina-D1* and *PinbD1* genes found by phenotyping of EMS-mutagenized population.

Mutation[a]	Effect[b]	M_2#	SIFT[c]	Wild type aa polarity / membrane affinity[d]	Mutant aa polarity/membrane affinity[d]	Grain Hardness[e]
***Pina* mutations**						
G70A	V24I	155	0.4	NA[f]	NA[f]	36.8±26
G77A	G26D	402	0.01	NA[f]	NA[f]	45.4±20
C112T	Q38Stop	125				65.3±26
C187T	P63S	363	0.01	Non-polar -0.76	Uncharged polar -0.25	50.1±25
G198A	W66Stop	376				73.4±23
G198A	W66Stop	497				34.4±23
G212A	W71Stop	293				67.0±26
C277T	Q93Stop	271				59.8±31
G395A	C132Y	526	0	Uncharged polar	Uncharged polar 1.11	43.4±22
***Pinb* mutations**						
C58T	Q20Stop	498				43.0±21
C146T	S49F	224	0	Uncharged polar -0.25	Non-polar 2.40	34.0±17
C200T	T67I	421	0	Uncharged polar 0.05	Non-polar 2.36	48.7±23
G254A	C85Y	592	0	Uncharged polar	Uncharged polar 1.11	44.0±20
C336T	F112F (Silent)	599				48.0±16
G348A	W116Stop	451				75.4±24
G348A	W116Stop	529				53.2±27
G377A	R126K	11	0.18	Basic	Basic	55.8±18
G384A	Q128Q (Silent)	303				50.5±22

[a] Nucleotide changes are numbered relative to the starting methionine of each coding sequence. Notation gives original base, position within coding sequence, and altered base. [b] Protein amino acid changes are numbered relative to the starting methionine of the proteins with notation giving original base, position within peptide prior to processing of signal peptide, and altered base. [c] SIFT scores < 0.05 are predicted to be deleterious.[d] Membrane affinity or ΔΔG bilayer to water (kcal/mol) of individuals were derived from Thorgeirsson et al. (1996). The membrane affinity of changed amino acids and cysteine were excluded from the experiment performed by Thorgeirsson et al. (1996). The values were ± 0.15 for all of the amino acids except for Q and N whose values were ± 0.3.[e] The average and standard deviation of kernel hardness was determined using the Single Kernel Characterization System on the bulk of less than 50 M_3 seeds obtained from individual M_2 families in summer 2006.
[f] The values are non-applicable for these mutants due to the occurrence of mutations in signal peptide.

Functional Analysis of New *Puroindoline* Alleles

In order to test the effects of the new *Pin* alleles upon grain hardness as well as to restore plant vigor and seed weight, the M_3 *Pin* mutant plants were crossed back to Alpowa. The *Pin* genotype of each F_1 plant was confirmed via sequencing before planting their F_2 seeds in the field. Twenty F_2 derived F_3 seed pools were evaluated for grain hardness and seed weight parameters for each of the tested *Pin* mutations. Three of the F_2 populations were created using missense alleles (*Pina-D1C187T*, *Pinb-D1C146T* , *Pinb-D1C200T*) along with a single control nonsense allele (*Pina-G212A*). A plot of grain hardness standard deviation against hardness average showed that all new alleles segregated as expected for a single gene with the exception of the *Pinb-D1C146T* allele for which little variation in grain hardness was observed (Figure 3). M_4 mutant and Alpowa parents were used as controls for seed characteristics (Table 8). The F_2 homozygote mutants derived from the *Pina* nonsense mutant *Pina-D1G212A* (a *Pina* nonsense mutant) showed a 39 unit increase in grain hardness vs. their wild type *Pin* allele group (72.1 vs 32.8 hardness units, $P < 0.0001$). Ninety-four percent of the variation in grain hardness in this population was attributable to the *Pina* mutation. Kernel weight of three of the F_2 populations was similar to Alpowa and almost 8 mg more than each parental mutant line. The F_2 *Pina-D1C187T* homozygote mutant group averaged 60 in grain hardness and an increase of 34 in grain hardness over its homozygous wild type control group (60.3 vs. 26.3 hardness units, $P< 0.0001$). Greater than 80% of the grain hardness variation in this population was attributable to the *Pina* nonsense mutation. The *Pinb-D1C200T* homozygous group proved as severe as the *Pina*

nonsense mutation as its F_2 homozygote mutant group averaged 36.5 units harder than its wild type allele groups 35.1 hardness units (71.6 vs. 35.1 hardness units, $P < 0.0001$). The *Pinb-D1C200T* mutation controlled ninety percent of the variation in grain hardness in this population. Although the average kernel weight obtained for the *Pinb-D1C200T* F_2 homozygote mutant group of this cross was comparable to the Alpowa average (~ 1 mg less) the kernel weight average for the other two groups of this population were 4-5 mg less than Alpowa. A 19 mg decrease in the kernel weight of the M_4 parent relative to Alpowa shows that this mutant originally suffered from a severely low kernel weight. The F_2 population derived from *Pinb* missense mutant plant *Pinb-D1C146T* showed a very minor change in grain hardness (5 units) for the homozygote mutant allele group relative to its wild type allele sibling group (34.3 vs wild type 29.3 hardness units, $P < 0.05$) and accounted for only 28% of the grain hardness variation. In summary for all crosses, the average seed weight measured for F_2 plants showed a near complete restoration of seed weight compared to the parental M_4 plants and Alpowa grown under the same field conditions. F_2 populations for the *Pina-D1G212A, Pina-D1C187T,* and the *Pin-D1C146T* mutations showed approximately 1:2:1 segregation ratios (Table 8). The segregation ratio of F_2 population derived from the *Pinb-D1C200T* mutant deviated from expected Mendelian ratios with more homozygote mutant plants than expected.

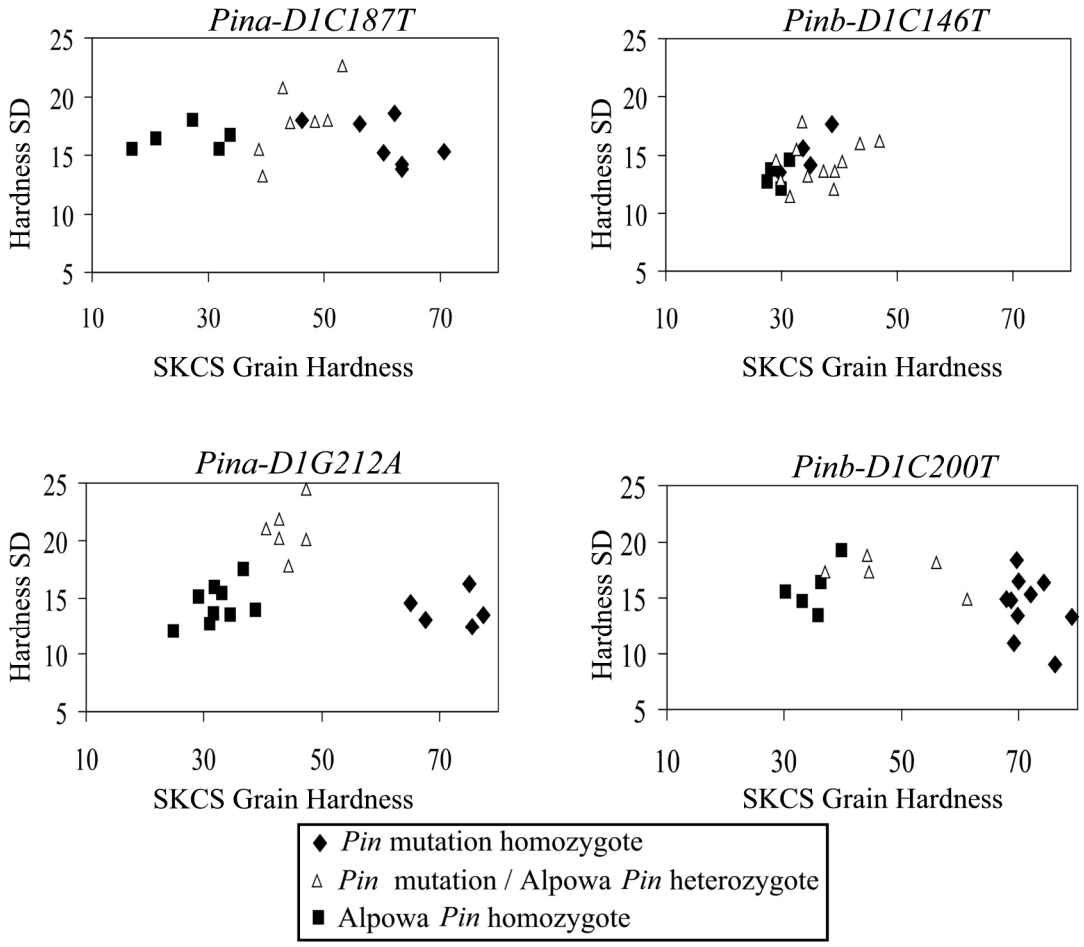

Figure 3. Single Kernel Characterization System (SKCS) standard deviations versus averages measured on 50-kernel sub-samples from 20 individual F_2 plants from crosses of four select *Pin* mutants with WT Alpowa. *Pin* mutant parents are as noted above each graph with each F_2 line classified according to *Pin* allele genotype type as indicated.

Table 8. Kernel characteristics and *Puroindoline* allelic state of parental lines and F_2 families derived from four crosses between Alpowa non - mutant × new *Puroindoline* mutants.

Pin allele group	*Pina-D1* allele	*Pinb-D1* alleles	Grain Hardness[b]	Kernel Wt[b] mg	n[c]	X^2(*P value*)[d]	R^{2e}
F_2 Progeny lines[a]							
Pina-D1G212A	*a/a*	*a/a*	32.8± 4.2	38.3±4.2	9	4.8(0.09)	0.9
	a / G212A	*a/a*	44.2 ±2.7	39.8±3.0	6		
	G212A/G212A	*a/a*	72.1±5.4**	38.6±1.6	5		
Pina-D1C187T	*a/a*	*a/a*	26.3 ±7.1	39.6±5.1	5	1.2(0.54)	0.8
	a / C187T	*a/a*	44.7± 5.3	40.7±2.4	8		
	C187T/C187T	*a/a*	60.3± 7.6**	42.6±3.3	7		
Pinb-D1C200T	*a/a*	*a/a*	35.0 ±3.6	34.8±3.4	5	7.5(0.02)	0.9
	a/a	*a/C200T*	48.5± 9.8	34.1±5.6	5		
	a/a	*C200T/C200T*	71.6 ±3.7**	37.7±2.1	10		
Pinb-D1C146T	*a/a*	*a/a*	29.3 ±1.7	43.2±3.4	4	0.8 (0.67)	0.3
	a/a	*a/C146T*	36.5± 5.6	44.5±2.3	12		
	a/a	*C146T/C146T*	34.3 ±3.7*	40.9±1.9	4		
F_2 Population Parents							
Alpowa	*a/a*	*a/a*	29.6 ±2.4	38.9±1.3	10	-	-
Pina-D1G212A[f]	*G212A/G212A*	*a/a*	81.9 ±6.4	30.6±6.4	10	-	-
Pina-D1C187T[f]	*C187T/C187T*	*a/a*	58.1 ±7.1	35.7±4.4	9	-	-
Pinb-D1C200T[f]	*a/a*	*C200T/C200T*	76.23	20.1	1	-	-
Pinb-D1C146T[f]	*a/a*	*C146T/C146T*	46.7 ±7.1	35.1±4.8	9	-	-

*, ** denote significance at $P \leq 0.05$ and $P < 0.0001$ respectively, in comparisons of the homozygote wild type versus homozygote mutant within a cross.

[a] Three F_2 groups, one parental group from each mutant family and one group of Alpowa non-mutant parent were used to determine kernel characteristics. *Pin* mutation alleles are named for the nucleotide change as listed in Table 7.

[b] Mean of kernel characteristics averaged over number of plants per each genotype planted. Kernel characteristics were determined using the Single Kernel Characterization System.

[c] n is the number of plants analyzed per parental lines and the number of genotypes obtained and analyzed per each F_2 family comprised of 20 individuals.

[d] X^2 was obtained from the segregation pattern of F_2 progenies of the crosses between each mutant and Alpowa non-mutant parent.

[e] R^2 value shows the proportion of the impact of new *Pin* alleles on grain hardness in comparison of homozygote wild type, homozygote mutant and heterozygote groups in each F_2 population.

[f] M_4 parental individuals obtained from M_3 single parent used for sequencing.

Discussion

Induced mutagenesis in adapted elite lines followed by marker-assisted backcrossing is an efficient approach to obtain favorable allelic variants for wheat, the most important staple crop world-wide (Bagge et al. 2007). *Puroindoline* allelic variation provides the genetic basis for most of the phenotypic variation in wheat grain hardness. In order to overcome the lack of *Pin* diversity (Table 6), our goal in this study was to create more allelic diversity for these proteins in hexaploid wheat and then to test the segregation pattern and function of these new alleles via segregating F_2 populations. These new alleles may be useful in improving wheat end use quality when they are incorporated into adapted genotypes. Specifically, hard wheats with reduced grain hardness may be improved in milling yield as suggested by the results of Martin et al. (2001, 2007). Soft wheats with reduced grain hardness may be improved in break flour yield (Hogg et al. 2005, Wanjugi et al. 2007b) or in starch extractability (Feiz et al. 2008). Soft wheats with increased grain hardness would likely have increased particle size but be more readily siftable (Campbell et al. 2007). Our first approach here was to advance an M_1 single head derived M_2 population and measure grain hardness of both bulk seeds and single plants from each M_2 family to identify new *Puroindoline* alleles (Figure 1). Since we were time constrained we then measured grain hardness on M_3 seeds from M_2 plants, however the chimeric nature of M_1 plants makes them less than desirable for direct mutation screening. It would be more efficient to conduct direct screening after advancing all lines to homozygousity via single seed descent (SSD). Population sizes used for screening for new mutations via direct screening or TILLING (Targeting

Induced Local Lesions in Genomes) vary dramatically. The size of an original M_1- single seed derived M_2 library was reported to be 20000 for barley (Caldwell et al. 2004) and 10000 for hexaploid wheat (Slade et al. 2005).

Perhaps the best approach to obtain a high mutation frequency is to use the maximum amount of EMS that still permits plant germination. We started with an M_1 population greater than two times the expected size of the M_2 library and then used the M_1 single head-derived M_2 families (1000) as individual members of the population planted as field head rows. The relatively low survivability we experienced indicates that there is a high risk of losing individuals in the early generations of selfing due to the segregation of deleterious genes (Colbert et al. 2001). Since our M_1:M_2 families were segregating 1:2:1 for a new mutation, a sample over a large M_2 family should yield seeds that are equal parts wild type and mutant allele assuming similar viability and fertility between wild type and mutant alleles.

The increased grain hardness mean of 630 M_2-derived M_3 seeds over Alpowa (Figure 1) was likely the result of reduced seed size and increased protein content. As expected, direct grain hardness phenotyping of the M_2-derived M_3 seeds increased *Pin* mutation frequency among selected lines. However, direct selection for *Pin* mutation containing M_2 lines likely did not detect most silent and missense mutations. EMS-induced mutagenesis is a random event; hence any deviations from randomness are due to bias from selection or mutation (Greene et al. 2003). While the overall predicted frequency of stop codon mutations is 4.7%, identification of 45% stop codon mutations from our pre-screened Alpowa-mutagenized population is a good indicator of the biases

induced by selecting for hard phenotypes (Table 7). Another reason for an increase in stop codon mutations in our population could be the higher tryptophan content of both PINA and PINB relative to other proteins. Tryptophan and methionine are the only amino acids each coded by a single codon and therefore are not as tolerant as other amino acids to single nucleotide changes. A transition mutation in any of the two bases of the tryptophan codon results in a stop codon. If the frequency of nonsense mutations was used to predict the total mutation rate then the total mutation rate is at least ~1 mutation per 4 kb. However, the screening of a larger EMS- mutagenized population of Alpowa accomplished by direct sequencing over 1500 kb of *Pina* and *Pinb* together resulted in a total mutation rate of 1 per 12 kb and a 11.5% rate of nonsense mutations (unpublished data). Our total mutation rate obtained by direct genotyping (unpublished data) is about 13 times higher than that reported for *A. thaliana* (Greene et al. 2003), 22 times higher than rice (Till et al. 2007), 77 times higher than barley (Caldwell et al. 2004) and ~ 2 times higher than previously observed in hexaploid wheat (Slade et al. 2005). The redundant nature of the hexaploid wheat genome is the most likely explanation for the higher mutation frequency observed in wheat relative to diploid plant species (Slade et al. 2005). All of the mutations were transitions (G to A or C to T) (Table 7), in accordance with other EMS-mutagenesis studies, since EMS induces G-residue alkylation (Anderson 1995). Phenotyping of M_2-derived M_3 bulk seeds proved non-conclusive (Table 7) in terms of the functional analysis of *Pin* genes for three reasons. First, because M_2-derived M_3 bulk seeds are composed of homozygote mutant and wild type as well as heterozygote seeds. Second, because there is a chance of inequality between the mutant and wild type

alleles in the M_2-derived M_3 bulk seeds which may result in a grain hardness bias toward wild type or mutant allele. Finally mutation in other genes may affect grain texture independent of the *Puroindoline* locus.

The bias caused by selection of harder phenotypes resulted in the identification of 16 homozygous Pin mutants among 18 total mutations which made it easier to cross the mutants back to non-mutant Alpowa. The backcross largely restored and seed size as well as allowed us to measure the impact of each mutation on wheat grain texture (Table 8). Grain hardness of seed from individual F_2 plants divided each F_2 population into three distinct classes. The impact of the segregating mutant alleles on grain hardness indicated the efficiency of this approach in creating a broad range of grain hardness phenotypes. The comparison of grain hardness of the F_2 segregating populations with the predicted effects of mutations via SIFT (Table 7) indicates that SIFT was not capable of distinguishing severe mutants from very mild mutations. SIFT predictions of deleterious mutations relies solely on the sequence and the conservation pattern of amino acids (Ng and Henikoff 2003).

The tryptophan-rich domain of puroindolines is hypothesized to confer an affinity for lipids (Marion et al. 1994; Kooijman et al. 1997) and may result in PINA and PINB being localized to the glyco- and phospholipid rich surface of the amyloplast membrane (Giroux et al. 2000). A prediction of secondary structure and measurement of amino acids membrane affinity was used by Giroux and Morris (1997) to hypothesize that the *Pinb-D1b* allele had reduced function due to its glycine 75 to serine mutation. A close examination of our new induced alleles shows that a mutation in threonine 67 to

isoleucine immediately adjacent to the tryptophan rich domain of PINB had a severe effect on grain hardness similar to a nonsense mutation (Tables 7 and 8). All three missense mutations, P63S in PINA and T67I and S49F in PINB resulted in an increase in membrane affinity. However, the mutation in serine 49 to phenylalanine (*Pinb-D1C146T*) resulted in a significant change in membrane affinity comparable to the other PINB missense mutation (Table 7) but resulted in a very small grain hardness change (Table 8).

In conclusion, similar to SIFT the change in membrane affinity of substituted amino acids in *Puroindoline* mutants was not predictive in terms of the functional analysis (Table 7). Phenotype analysis of F_2 segregating groups after crossing the mutant back to the non-mutant parent is the most efficient way to understand the impact of mutation on the gene function. It decreases the effect of other mutations as well as restores the seed weight and vigor. More importantly after seed increase, this F_2 population will be used to analyze the effect of each *Pin* mutation upon a broader range of functions including milling, baking and other wheat end qualities. Further functional analysis of F_2 families carrying a wider range of missense mutations dispersed throughout PINA and PINB will help us better understand the importance of different amino acids and regions of the puroindoline proteins. The selected alleles with desirable end quality features could then be inserted into other elite germplasm adapted to different agricultural regions.

CHAPTER 3

IN PLANTA MUTAGENESIS DETERMINES THE FUNCTIONAL REGIONS OF PUROINDOLINE PROTEINS

Introduction

Variation in wheat (*Triticum aestivum* L.) endosperm texture is a primary determinant of wheat end-product quality (reviewed in Morris and Rose 1996). Endosperm texture variation is controlled predominantly by the *Hardness* (*Ha*) locus, located on the extreme distal end of chromosome 5DS (Mattern et al. 1973; Law et al. 1978; Ram et al. 2002; Campbell et al. 1999). The *Ha* locus functionally consists of the *Puroindoline a* and *b* genes (*Pina* and *Pinb*, respectively) (Giroux and Morris 1998; Wanjugi et al. 2007a). Soft texture (*Ha*) is the result of both genes being in their wild-type allelic state (*Pina-D1a*, *Pinb-D1a*) while hard texture (*ha*) results from mutations in either *Pina* or *Pinb* (Giroux and Morris 1997, 1998). PINA and PINB are cysteine rich proteins which are unique among plant proteins in having a hydrophobic tryptophan-rich domain (Blochet et al. 1993) and together make up the grain hardness marker protein friabilin. Friabilin was originally described as a 15 kDa protein (Greenwell and Schofield 1986; Morris et al. 1994) present in higher levels on the surface of water washed starch prepared from soft wheats relative to the levels found on hard wheat starch. Similarly, Greenblatt et al. (1995) observed that water washed starch from soft wheats had more glyco and phospholipids than hard wheat water washed starch.

The correlation between friabilin and glyco- and phospholipid levels on the surface of starch granules led to the hypothesis that PINs are co-localized to the glyco- and phospholipid rich surface of the amyloplast membrane and that this localization may be mediated by their tryptophan-rich domain (Giroux et al. 2000). Although PINA and PINB show considerable amino acid identity (55%) and similarity (80%) relative to each other, they show no more than 25% identity to any of the members of their known protein family. According to a protein family database, http://pfam.janelia.org (Finn et al. 2007) puroindolines belong to the protease inhibitor/ seed storage/ lipid transfer protein family (PF00234). The shared characteristic of this family is a domain of four helical bundles in a right-handed superhelix with a folded leaf topology, which is stabilized by disulfide bonds. Neither one of the members of this protein family nor any other known plant proteins have a Trp rich region. However the Trp rich region is a characteristic of bacterial anti-microbial peptides (Chan et al. 2006) which interact with microbial membranes. The Trp rich domain of PINs has been predicted to be in a beta sheet conformation and joins the first two α –helices of four PIN helices (Bihan et al. 1996; Giroux and Morris 1997).

PINs were conclusively shown to alter grain texture in cereals via transgenic studies. By transforming rice with both PINs, Krishnamurthy and Giroux (2001) and Krishnamurthy et al. (2001) demonstrated that PINs can modify grain texture and fungal resistance, respectively. In addition, we demonstrated that *Pin* mutations are causative to hard wheat phenotypes via transgenic complementation studies in which soft phenotypes were achieved when hard wheat *Pin* alleles were complemented with the corresponding

soft-type *Pin* allele (Beecher et al. 2002; Martin et al. 2006). Transgenic manipulations have also demonstrated that while both PINA and PINB levels limit grain softness in soft wheats, PINB is a greater limiting factor than PINA (Swan et al. 2006). Further PINA and PINB are not functionally interchangeable, as shown by transgenic analysis (Hogg et al. 2004; Wanjugi et al. 2007a). This shows the specificity in roles conferred by each puroindoline. High levels of either PIN alone can lead to intermediate levels of grain hardness. However, full PIN starch association and a soft grain phenotype is achieved only with both PINs present.

Variation in PIN function affects not only grain hardness but nearly all product quality traits (Hogg et al. 2005; Wanjugi et al. 2007b; Martin et al. 2008; Feiz et al 2008) and associated milling effects are also seen in pilot scale milling studies (Martin et al. 2007). The vast majority (>95%) of US hard wheat germplasm contains one of two *Ha* locus mutations (*Pina-D1b* or *Pinb-D1b*) (Morris et al. 2001). Despite this small level of diversity, it is known that hard wheats carrying the *Pinb-D1b* allele wheats are >7 hardness units softer and superior in milling and bread baking properties relative to hard wheats carrying the *Pina-D1b* allele (Martin et al. 2001, 2007). Similarly, the addition of an A genome *Ha* locus to a soft wheat decreased grain hardness and modified milling yield and particle size (Campbell et al. 2007). From an end-product quality point of view the more *Pin* variation found or created the greater chance that some alleles will confer improved end-product quality. The *Pina* and *Pinb* missense alleles discovered in various hexaploid wheat germplasm surveys may vary in function somewhat, but they each likely retain very little residual function as all were found in hard wheats. While several *Pin*

allelic variants have been reported among *Ae. tauschii* (the diploid donor of D genome to hexaploid wheat) and synthetic wheat germplasm (Gedye et al. 2004; Massa et al. 2004), they are not useful in terms of understanding PINA and PINB function for two reasons. First, a significant amount of the hardness variation is assignable to both the D genome donor (*Ae. tauschii*) and durum parents and secondly most of the haplotypes contain a large number of amino acid changes relative to *Pina-D1a* and *Pinb-D1a*. In addition, all of the alleles introduced from *Ae. tauschii* appear to condition a soft phenotype such that they appear limited as a source of grain hardness variation (Gedye et al. 2004). Therefore, the creation of new *Pin* alleles in one background where each contains a single amino acid change via target-selected mutagenesis is the best approach for creation of new *Puroindoline* alleles in hexaploid wheat. The new alleles would have the potential to improve understanding of PIN function as well as improving wheat end product quality.

Materials and Methods

Creation and Selection of EMS-induced Population

A wheat EMS-induced M_1 population was created using a protocol similar to Slade et al. (2005) with some modifications (Feiz et al. 2009). Approximately, 10,000 M_0 wheat seeds of the soft white spring cultivar Alpowa (PI566596) were EMS-mutagenized to obtain a total of ~ 3,000 heads from surviving M_1 plants. This population was divided into two parts. The first 1,000 M_1: M_2 heads were planted in May 2006. Six hundred thirty surviving M_2 rows were harvested and processed to test the screening methods and

applicability of the EMS-induced mutagenesis in finding new *Pin* alleles (Feiz et al. 2009).

In May 2007, an additional 2000 M_1: M_2 head rows were planted under the same conditions as described for the first population (Feiz et al. 2009). Leaf tissue was collected and pooled at the 2–3 leaf stage from at least four plants per row of the 1,700 surviving M_2 rows. Bulk harvest was conducted on 1,700 rows as explained for the first population (Feiz et al. 2009).

DNA was extracted according to Riede and Anderson (1996). The same PCR conditions and primers pairs used for direct sequencing by Feiz et al. (2009) were used in this study. PCR products were separated on 1.5% agarose gels followed by gel purification (Qiagene, Valencia, CA) and were then direct sequenced using the forward amplification primer (SeqWright DNA Technology Services, Houston, TX). The GAP4 and Pregap4 programs from Staden Package v1.6.0, 2004, (http://staden.sourceforge.net/staden_ home.html) were used to analyze the sequences. Sorting intolerant from tolerant (SIFT) (Ng and Henikoff 2003) was used to predict the tentative impact of mutations on protein function.

Creation of F_2 Populations

Eight plants from each line carrying a *Pin* mutation were planted in the greenhouse and used for direct sequencing of *Pina* and *Pinb*. Plants positive for *Pin* mutations were used as pollen source in crosses back to Alpowa non-mutant parent to produce F_1 seeds. One hundred F_1-derived F_2 seeds from each cross along with Alpowa parental seeds were planted at the Arthur H. Post Field Research farm near Bozeman, MT

with within row plant spacing of 15 and 30 cm between rows in May 2008. Leaves were collected from 48 single F_2 plants from each cross at the 2–3 leaf stage to be used for genotyping.

Genotyping and Phenotyping of F_2 Populations

The same primer pairs and PCR conditions as Feiz et al. (2009) were used to genotype F_2 plants growing in the field. Polymorphic co-dominant markers were developed based on the differential restriction digestion of some of the *Pina* and *Pinb* mutant alleles (Table 9). F_2:F_3 seeds were harvested from homozygote single F_2 plants and threshed using a single plant thresher (Bill's Welding, Pullman, WA). One hundred and fifty F_2: F_3 seeds were bulked from each one of the two groups of homozygote mutant plants and homozygote wild-type plants with each group composed of 5 plants. The bulked seeds for the homozygote mutant F_2:F_3 groups as well as Alpowa non-mutant seeds were analyzed by SKCS. Grain hardness and kernel weight were measured with SKCS using 50 kernels. Each grain sample was measured twice. The same planting, seed bulking and SKCS analysis process was performed on F_3: F_4 seeds derived from the four F_2:F_3 mutant populations analyzed by Feiz et al. (2009).

Analysis of variance was computed by including genotype class by population combinations in the model using PORC GLM in SAS (SAS Institute, 2004). The error represented pooled within population variance. The impact of new alleles on grain characteristics was assessed by comparing difference between mutant and wild type class means for each population.

Results

Creation of Novel *Puroindoline* Alleles

To obtain and identify new alleles for grain hardness gene, we developed an EMS-mutagenized wheat population from the soft white spring wheat variety Alpowa. This population was divided into two groups comprised of 630 and 1700 M2 families, respectively. Direct phenotyping of the first group resulted in the identification of 18 novel *Pin* alleles consisting of 8 missense and 8 nonsense mutations in one of the *Pin*s (Feiz et al. 2009). Direct sequencing of the rest of the 1700 M_2-derived M_3 bulk leaves obtained from an average of four M_2 plants per M_2 family resulted in the identification of 130 new *Pina* and *Pinb* alleles of which 60 were missense, 15 nonsense and 55 silent mutations. All the *Pin* alleles obtained from phenotyping and/or direct sequencing are listed in Table 9. Some of the mutations were identified more than once. A total of 71 alleles of *Pina* were identified using direct phenotyping or sequencing, of which 37 were missense and 11 were nonsense mutations (Table 10). Of the 77 alleles of *Pinb*, we identified 31 missense and 12 nonsense mutations. After exclusion of thirty base pairs from each end of the *Pin* sequences, the mutation density was calculated from the frequency of mutations identified via direct sequencing of 1700 EMS-induced M_2 lines over 1,496 kb of *Pina* and *Pinb* (Table 10). We arrived at a mutation density of 1/11.5 kb which is the highest rate ever reported for an EMS-induced mutagenesis study. The higher percentage of nonsense mutation identified via direct sequencing (11.5% or 15/130) than what is expected from a random mutagenesis study (4.7%) is due to the nature of tryptophan. Tryptophan and methionine are the only amino acids encoded by

single codons and thus are not as tolerance as other amino acids to single nucleotide changes. A transition mutation in any of the two variable bases of the tryptophan codon results in a stop codon. Almost 54.5% (6/11) and 91.7% (11/12) of the total nonsense induced mutations happened in tryptophan amino acids of *Pina* and *Pinb*, respectively (Table 10).

The bioinformatic program SIFT (Sorting Intolerant From Tolerant) uses sequence homology in a protein family to predict the impact of the missense mutations on protein function and potential phenotype. SIFT scores < 0.05 are predicted to be deleterious (Ng and Henikoff 2003). Via SIFT 47% (14/30) and 54% (13/24) of the total missense mutations were predicted to have severe impact on PINA and PINB function, respectively (Table 9).

Table 9: Allelic series of mutations in *Pina-D1* and *Pinb-D1* genes found by phenotyping and direct sequencing or only direct sequencing of EMS-mutagenized population.

Mutation[a]	Effect[b]	SIFT[d]	Polymorphic marker(s)[e]	Sizes of fragments (bp)		Number of F_2 progeny lines[f]				*P Value* (X^2)[g]
Pina-D1				*D1a*[h]	*D1x*[h]	*D1a*[h]	*D1a*/*D1x*[h]	*D1x*[h]	n[i]	
C15T	F5F									
G23A	G8E	0								
G36A	L12L									
G49A	A17T	0.67	sequenced			7	23	14	2	0.31(2.3)
C58T	Q20stop									
C58T	Q20stop									
G65A	S22N	0	sequenced			10	25	12	1	0.83(0.37)
G67A	E23K	0.01	sequenced			13	21	13		0.67(0.8)
G70A	V24I	0.4								
G70A	V24I									
G70A	V24I		*Tsp*509I	190,12,300	123,67,12,300	18	23	4	1	0.01(9.2)
G77A	G26D	0.01	*Hinc*II,	502	132,370	12	22	10	2	0.91(0.19)
G95A	G32D									
G95A	G32D									
G104A	G35D									
C107T	A36V	0	*Bsi*HKAI	101,62,128,211	101,190,211	11	19	13	3	0.68(0.77)
C109T	Q37stop									
C112T	Q38stop									
C119T	P40L	0.11	*Sfc*I	174,328	502	10	25	9	2	0.65(0.86)
G126A	E42E									
G146A	R49K	1	sequenced			15	23	7	1	0.24(2.8)
G157A	L53I	0.16								

Table 9-continued

G180A	K60K		*Hpy*AV	52,174,39,237	52,213,237	11	232	10	2	0.93(0.14)
G181A	D61N	0								
G181A	D61N		*Apo*I	190,312	190,45,267	13	19	12	2	0.65(0.86)
C187T[j]	P63S	0.01	*Msp*I	242, 235, 25	477, 25	5	8	7		0.54(1.2)
G190A	V64I	0.13	*Msp*I	242, 235, 25	477, 25	16	22	7		0.16(3.66)
C194T	T65I	0								
G198A	W66stop									
G198A	W66stop									
G207A	W68stop									
G207A	W69stop									
G212A[j]	W71stop		sequenced			9	6	5		0.09(4.8)
G212A	W71stop									
G220A	G74R	0.53								
G223A	G75S	0.81	*Mnl*I	18,56, 6,189,202,31	18,56,6,391,31	11	26	8	1	0.48(1.46)
G234A	E78E									
G232A	E78K	0								
G258A	R86R									
G258A	R86R									
C259T	L87F	0	sequenced			10	23	10	5	0.9(0.21)
G263A	G88D	0.35	*Taq*I	217,285	217,98,187	17	22	7		0.12(4.2)
C265T	Q89L	0.09								
C277T	Q93stop		*Hpy*CH4III	502	330,172					
C294T	I98I									
C297T	I99I									
C297T	I99I									
G303A	G101G									
G314A	G105D	0.05	*Sty*I	292,72,138	292,210	7	25	11	3	0.39(1.88)
G314A	G105D									
G315T	G105G									

Table 9-continued

G342A	Q114Q									
G351A	R117R									
G351A	R117R									
G352A	A118T	0.34								
G361A	V121M	0.13	sequenced			15	26	7		0.22(3)
G370A	E124K	0.16	sequenced			12	22	14		0.78(0.49)
G378A	K126K									
C382T	L128L									
C382T	L128L									
C388T	P130S	1	*Bce*AI	502	426,76	8	27	13		0.4(1.8)
G393A	R131R									
G392A	R131K	1	sequenced			13	29	6		0.13(4)
G395A	C132Y	0								
C399T	N133N									
C399T	N133N									
G402A	Q134Q									
C405T	G135G									
G403A	G135S	0.19								
C410T	P137L	0.05	sequenced			13	22	13		0.85(0.32)
C421T	P141S	0	*Bts*CI	348,123,31	348,154	18	17	21		0.07(5.3)
Pinb-D1										
C8T	T3I	0.14	*Bg*1II	508	66,442	13	19	11	2	0.68(0.77)
C9T	T3T									
C16T	L6F	0.09	sequenced			16	26	5	1	0.06(5.6)
C18T	L6L									
C27T	L9L									
G37A	V13I	0								
C58T	Q20stop		*Hin*P1I	116,355,37	471,37					
G70A	V24I	0.03	*Tsp*509I	508	129,379	9	26	9	2	0.48(1.5)

Table 9-continued

G74A	G25D	0.13								
C78T	G26G									
G77A	G26D	0.15	*Bsr*I	254,252,2	143,111,252,2	15	20	12	1	0.49(1.42)
G81A	W27stop									
G95A	G32D	0	*Hin*cII	508	154,354	12	20	12	2	0.83(0.37)
G101A	G34E	0.1								
C121T	P41S	0	sequenced			16	28	4		0.025(7.3)
G129A	E43E									
G138A	K46K									
G138A	K46K									
C139T	L47L									
G143A	S48N	1	*Bl*pI	200,126,182	316,182	11	23	12		0.98(0.04)
G143A	S48N		*Bl*pI	200,126,182	316,182	6	15	12	12	0.29(2.48)
C146T[j]	S49F	0	*Hin*dIII	508	202,306	4	12	4		0.67(0.8)
G153A	K51K									
G154A[j]	D52N	0.91	*Tsp*509I	508	214,294	4	3	1		0.25(2.8)
G168A	E56E									
G168A	E56E									
G170A	R57Q	0.2	*Bsr*	508	236, 272	10	24	9	3	0.71(0.68)
G187A	D63N	0.01	*Tsp*509I	508	247,261	12	20	11	3	0.88(0.26)
C200T[j]	T67I	0	*Bst*N1	258,250	508	5	5	10		0.02(7.5)
C200T	T67I									
G203A	W68stop									
G204A	W68stop									
G204A	W68stop									
G216A	W72stop									
G223A	G75S									
G224A	G75D	0.13	*Aci*I	135,21,26,8,95,223	135,21,26,8,318	18	18	9	1	0.07(5.3)
G224A	G75D		*Aci*I	135,21,26,8,95,223	135,21,26,8,318	9	21	7	8	0.64(0.89)

Table 9-continued

G224A	G75D									
G234A	E78E									
G240A	E80E									
G238A	E80K	0.04	*Nla*III	64,235,170,39	64,405,39	13	23	10		0.82(0.4)
G241A	V81I	0.02								
G246A	R82R									
G252A	K84K									
C255T	C85C									
G254A	C85Y	0	*Sca*I	508	314,194	14	15	17		0.051(5.95)
G254A	C85Y									
G269A	S90N	0.02								
C303T	I101I									
G306A	R102R									
G306A	R102R									
G306A	R102R									
G306A	R102R									
G306A	R102R									
C315T	I105I									
G324A	R108R									
G324A	R108R									
C325T	L109F	0.01	*Bfu*AI	508	372,136	11	26	8	3	0.47(1.5)
G329A	G110D	0	*Taq*I	508	121,387	11	23	11	1	0.99(0.02)
G332A	G111D	0	*Tsp*451	256,252	256,119,133	10	24	14		0.72(0.66)
C336T	F112F									
C336T	F112F									
C342T	G114G									
G348A	W116stop		sequenced							
G348A	W116stop									
G348A	W116stop									

Table 9-continued

G348A	W116stop									
G348A	W116stop									
G348A	W116stop									
G355A	E119K	0.25	*Hph*I	250,175,83	250,258	13	28	6	1	0.15(3.8)
G378A	R126R									
G377A	R126K	0.18	*Apa*I	438,70	508	8	25	11	2	0.54(1.23)
C381T	A127A									
G384A	Q128Q									
G384A	Q128Q									
C402T	C134C									
C417T	D139D									

[a] Nucleotide changes are numbered relative to the starting methionine of each coding sequence. Notation gives original base, position within coding sequence, and altered base.
[b] Protein amino acid changes are numbered relative to the starting methionine of the proteins with notation giving original base, position within peptide prior to processing of signal peptide, and altered base.
[c] The mutations which have the M_2 number from 1 to 630 were identified via phenotyping (Feiz et al. 2009).
[d] SIFT scores <0.05 are predicted to be deleterious.
[e] Polymorphic marker indicates the restriction enzymes which are able to differentially cut *Pin* PCR products between Alpowa wild type allele and mutant allele. These markers were used for genotyping of the F_2 progeny lines derived from crosses between each mutant and Alpowa non-mutant parent. Whenever a mutation site was not located in the restriction site of a restriction enzyme, direct sequencing was used to genotype the F_2 progeny lines
[f] The number of genotypes obtained and analyzed per each F_2 family comprised of 46-48 individuals
[g] X^2 was obtained from the segregation pattern of F_2 progenies of the crosses between each mutant and Alpowa non-mutant parent
[h] D1a, D1a/D1x and D1x indicate the wild type, heterozygote and mutant alleles, respectively.
[i] n indicates the progenies which were failed in genotyping via markers or sequencing.
[j] F_2 Segregation data from crosses of *Pina-D1C187T*, *Pina-D1G212A*, *Pinb-D1C146T* and *Pinb-D1C200T* into Alpowa non-mutant parent are obtained from 20 individuals harvested in summer 2007 and published by Feiz et al. (2009). The number of progenies in F_2 line *Pinb-D1G154A* was less than 10.

Table 10: *Pin* mutation frequency and density in Alpowa EMS-induced population.

	Missense[a]	Nonsense[a]	Silent[a]	Total[a]	Mutation density [b]	% of nonsense mutations[b]	% of nonsense mutations in Trp[c]
Pina	37(33)	11(6)	23	71(62)	1/12 kb	9.6	54.5
Pinb	31(27)	12(9)	34(32)	77(68)	1/11 kb	13.2	91.7
Both *Pin*s	68(60)	23(15)	57(55)	148(130)	1/11.5 kb	11.5	73.9

[a] Total mutation frequency found via phenotyping of 630 or direct sequencing of 1700 M_2 Alpowa lines. The numbers in () are the frequency of mutations found by direct sequencing of 1700 M_2 Alpowa lines.
[b] The mutation density and the nonsense mutation percentage were calculated using the frequency of mutations found via direct sequencing of 1700 M_2 Alpowa lines which is presented in ().
[c] The percentage of nonsense mutations which occurred in Trp amino acids of PINs was calculated using the frequency of total nonsense mutations which obtained either via phenotyping of 630 or direct sequencing of 1700 M_2 Alpowa lines .

Creation of Segregating F_2 Populations for the New *Puroindoline* Alleles

In order to test the effects of the new *Pin* alleles upon grain hardness as well as to restore plant vigor and seed weight, segregating F_2 populations were developed for each unique *Pin* allele. To achieve this, 23 and 20 M_3 mutant plant carrying unique missense mutations in *Pina* and *Pinb*, respectively, were crossed back to Alpowa. Two of the *Pinb* duplicated mutations (identified more than once in the population) and one *Pina* silent mutation was crossed back to Alpowa to be used as controls. The *Pin* genotype of each F_1 plant was confirmed via sequencing before planting their F_2 seeds in the field. Forty six and 48 F_2 plants per population were genotyped via polymorphic markers or sequencing respectively. Ninety-four percent of the F_2 populations showed approximately 1:2:1 segregation ratios (Table 9). The segregation ratio of F_2 population derived from three of the mutants deviated from expected Mendelian ratios. PINAV24I and PINBP41S F_2 populations showed more homozygote wild *Pin* allele plants and PINBT67I showed more mutant plants than expected (Table 9).

Functional Analysis of New *Puroindoline* Alleles

F_2 derived F_3 seeds from each F_2 population were pooled to obtain wild type and mutant genotype classes for each population (Table 11). Alpowa parent, F_2 family derived from PINAK60K (a silent mutation) and F_2 families obtained from the replicates of two *Pinb* mutations (PINBS48N) and PINBG75D were used as controls for seed characteristics. The grain hardness of Alpowa parent was not significantly different from 74% of the F_2 homozygote wild type groups. For 12 F_2 populations the homozygote wild

type group significantly differed from the Alpowa parent. The average of this difference was 10.2 hardness units. In order to avoid the plausible side effects of seed characteristics or environment on functional analysis of *Pin* mutants, the net intra-family grain hardness difference between homozygote mutant and homozygote wild type groups of each F_2 population was used to evaluate the impact of the mutations on grain texture (Table 11). The F_2 populations were categorized based on differences between wild type and mutant classes into bins of grain hardness difference (Figure 4). Each bin is 5 units harder or softer than its neighboring bins and includes the mutant F_2 lines whose homozygote mutant and wild type groups have the hardness difference of that bin ± 2.5. Functional analysis of the four F_2 populations with PIN nonsense mutations (PINAW71Stop, PINAQ93Stop, PINB Q20Stop and PINBW116Stop) delimited the function abolishing mutations to bins ≥ 25 (Figure 4). Almost 18.6 % (8/43) of the mutants in *Puroindolines* were located in the function abolishing bins (25 to 55), whereas ~ 62.8% (27/43) of the mutants were categorized into neutral (silent) bins (-5 to 5) with the Least Significant difference of 6.2 (Table 11 and Figure 4). Only 16.2% (7/43) of mutants were located in the intermediate bins (10-20) and one mutation (PINBP41S) located in the bin of -15 resulted in a softer texture than wild type. Although PINAV121M was categorized in bin of -5, it resulted in a significant softer texture than wild type. Duplicate lines of a neutral mutation (PINBS48N) didn't show difference in grain hardness between mutant and wild type sibling groups of either independent line. Similarly duplicate populations of a severe mutation (PINBG75D) resulted in a 34.2 increase in grain hardness in one population and 30.4 in another population with the same mutation. The grain characteristics of one *Pina*

nonsense, one *Pina* missense and two *Pinb* missense mutant lines in Table 11 are from F_3:F_4 seeds planted and harvested along with the rest of the populations which were harvested and analyzed at the F_2:F_3 seed stage (Feiz et al. 2009). The difference in grain hardness between F_3:F_4 seeds obtained from mutant and wild type sibling groups of PINAW71Stop, PINAP63S, PINBT67I and PINBS49F lines were 41.1, 33.7, 33.7 and -2.8, respectively (Table 12). In fall 2007 the difference between F_2:F_3 seeds obtained from mutant and wild type groups of the same lines was obtained 39.3, 34, 36.6 and 5, respectively (Feiz et al. 2009). These results indicate the reproducibility of our approach in functional analysis of *in planta Pin* mutants.

The difference in kernel weight between mutant and wild type groups of each F_2 population was non-significant for 76% of the populations. The mutant group of 8 of the 43 unique missense mutant populations had a significantly lower seed weight than their sibling wild type group. Two of the F_2 populations had higher kernel weight in their mutant versus wild type sibling group. Smaller seeds show more protein and harder texture than larger seeds with the same genotype. Among the lines with lower seed size in the mutant versus wild type sibling group, three were located in the intermediate grain hardness bin of 10 (Table 11).

Table 11. Kernel characteristics and *Puroindoline* allelic state of Alpowa non-mutant parent and F_2 populations derived from four crosses between Alpowa non-mutant × new *Puroindoline* mutants.

Pin allele group	*Pina-D1* alleles	*Pinb-D1* alleles	Effect	grain Hardness[b]	difference[c]	Kernel Wt[b] mg	difference[c]
F_2 progeny homozygote populations[a]							
Pina-D1G49A	*a*	*a*	PINA	33.9 ± 2.8	1.2	41.4 ± 1.3	-1.4
	D1G49A	*a*	PINAA17T	35.1 ± 4.8		40.0 ± 2.2	
Pina-D1G65A	*a*	*a*	PINA	32.0 ± 0.7	8.9*	38.2 ± 0.6	0.8
	D1G65A	*a*	PINAS22N	40.9 ± 0.7		39.0 ± 1.1	
Pina-D1G67A	*a*	*a*	PINA	35.5 ± 6.2	5.7	37.9 ± 1.7	-6.1*
	D1G67A	*a*	PINAE23K	41.2 ± 0.4		31.8 ± 0.3	
Pina-D1G70A	*a*	*a*	PINA	39.9 ± 2.9	4.6	36.8 ± 2.1	-0.5
	D1G70A	*a*	PINAV24I	44.5 ± 0.4		36.3 ± 0.5	
Pina-D1G77A	*a*	*a*	PINA	36.4 ± 2.2	-4.7	38.4 ± 0.8	-0.2
	D1G77A	*a*	PINAG26D	31.7 ± 0.1		38.2 ± 0.1	
Pina-D1C107T	*a*	*a*	PINA	26.3 ± 1.0	-1.0	37.0 ± 0.9	-1.4
	D1C107T	*a*	PINAA36V	25.3 ± 2.0		35.6 ± 0.8	
Pina-D1C119T	*a*	*a*	PINA	30.5 ± 1.9	4.3	41.1 ± 1.2	-1.1
	D1C119T	*a*	PINAP40L	34.8 ± 0.2		40.0 ± 1.1	
Pina-D1G146A	*a*	*a*	PINA	29.7 ± 5.4	8.6*	44.8 ± 2.7	-4.0*
	D1G146A	*a*	PINAR49K	38.3 ± 0.6		40.8 ± 0.3	
Pina-D1G180A	*a*	*a*	PINA	34.5 ± 4.5	3.5	40.6 ± 0.7	-0.9
	D1G180A	*a*	PINAK60K	38.0 ± 0.5		39.7 ± 0.8	
Pina-D1G181A	*a*	*a*	PINA	34.2 ± 7.6	3.2	38.9 ± 1.2	-1.0
	D1G181A	*a*	PINAD61N	37.4 ± 0.6		37.9 ± 1.6	
Pina-D1C187T	*a*	*a*	PINA	30.4 ± 0.7	33.7*	39.9 ± 0.1	2.3
	D1C187T	*a*	PINAP63S	64.1 ± 2.2		42.2 ± 1.6	

Table 11-continued

Pina-D1G190A	*a*	*a*	PINA	40.0 ± 0.17	-2.0	38.8 ± 0.4	0.1
	D1G190A	*a*	PINAV64I	38 ± 3.9		38.9 ± 0.06	
Pina-D1G212A	*a*	*a*	PINA	36.7 ± 0.09	41.1*	37.4 ± 0.2	-0.1
	D1G212A	*a*	PINAW71Stop	77.8 ± 6.0		37.3 ± 1.2	
Pina-D1G223A	*a*	*a*	PINA	45.8 ± 3.6	18*	34.9 ± 1.4	0.6
	D1G223A	*a*	PINAG75S	63.8 ± 1.6		35.5 ± 0.01	
Pina-D1C259T	*a*	*a*	PINA	37.4 ± 0.7	27.9*	38.6 ± 0.1	2.9*
	D1C259T	*a*	PINAL87F	65.3 ± 1.8		41.5 ± 0.03	
Pina-D1G263A	*a*	*a*	PINA	32.8 ± 0.8	3.8	37.5 ± 0.6	-0.5
	D1G263A	*a*	PINAG88D	36.6 ± 1.3		37.0 ± 0.4	
Pina-D1C277T	*a*	*a*	PINA	39.7 ± 0.2	39.1*	41.6 ± 1.2	-2.4
	D1C277T	*a*	PINAQ93Stop	78.8 ± 0.2		39.2 ± 0.3	
Pina-D1G314A	*a*	*a*	PINA	33.9 ± 0.02	-3.1	40.6 ± 0.9	1.7
	D1G314A	*a*	PINAG105D	30.8 ± 4.1		42.3 ± 1.1	
Pina-D1G361A	*a*	*a*	PINA	37.8 ± 2.0	-7.4*	37.2 ± 0.5	0.3
	D1G361A	*a*	PINAV121M	30.4 ± 2.2		37.5 ± 0.7	
Pina-D1G370A	*a*	*a*	PINA	32.2 ± 2.7	-4.5	44.1 ± 0.4	1.0
	D1G370A	*a*	PINAE124K	27.7 ± 4.2		45.1 ± 2.3	
Pina-D1C388T	*a*	*a*	PINA	36.4 ± 1.6	-3.2	36.6 ± 0.5	2.2
	D1C388T	*a*	PINAP130S	33.2 ± 3.7		38.8 ± 2.2	
Pina-D1G392A	*a*	*a*	PINA	40.8 ± 1.9	-5.6	38.5 ± 0.6	1.9
	D1G392A	*a*	PINAR131K	35.2 ± 2.5		40.4 ± 0.3	
Pina-D1G395A	*a*	*a*	PINA	28.1 ± 4.8	39.2*	38.5 ± 3.1	2.8*
	D1G395A	*a*	PINAC132Y	67.3 ± 2.8		41.3 ± 0.3	
Pina-D1G403A	*a*	*a*	PINA	30.2 ± 0.7	4.6	43.5 ± 1.2	2.9*
	D1G403A	*a*	PINAG135S	34.8 ± 2.7		46.4 ± 0.08	
Pina-D1C410T	*a*	*a*	PINA	35.3 ± 2.9	-2.6	38.1 ± 0.1	-1.11
	D1C410T	*a*	PINAP137L	32.7 ± 1.9		37.0 ± 1.8	

Table 11-continued

Pina-D1C421T	*a*	*a*	PINA	41.0 ± 1.3	0.1	38.2 ± 1.2	-0.4
	D1C421T	*a*	PINAP141S	41.1 ± 0.3		37.8 ± 1.4	
Pinb-D1C8T	*a*	*a*	PINB	31.4 ± 0.3	-0.3	41.8 ± 0.7	-2.1
	a	*D1C8T*	PINBT3I	31.1 ± 5.4		39.7 ± 2.8	
Pinb-D1C16T	*a*	*a*	PINB	33.3 ± 0.74	10.6*	37.9 ± 1.3	-2.7
	a	*D1C16T*	PINBL6F	43.9 ± 0.01		35.2 ± 0.9	
Pinb-D1C58T	*a*	*a*	PINB	30.9 ± 1.2	33.9*	43.4 ± 1.7	-3.9*
	a	*D1C58T*	PINBQ20Stop	64.8 ± 5.4		39.5 ± 0.6	
Pinb-D1G70A	*a*	*a*	PINB	33.6 ± 0.9	0.4	40.1 ± 0.4	0.1
	a	*D1G70A*	PINBV24I	34.0 ± 0.7		40.2 ± 0.8	
Pinb-D1G77A	*a*	*a*	PINB	27.9 ± 0.5	1.9	44.3 ± 0.3	-0.2
	a	*D1G77A*	PINBG26D	29.8 ± 1.2		44.1 ± 1.4	
Pinb-D1G95A	*a*	*a*	PINB	36.5 ± 0.7	4.1	40.0 ± 0.2	-4.4*
	a	*D1G95A*	PINBG32D	40.6 ± 1.6		35.6 ± 0.0	
Pinb-D1C121T	*a*	*a*	PINB	36.8 ± 0.2	-12.6*	37.1 ± 1.4	0.71
	a	*D1C121T*	PINBP41S	24.2 ± 0.9		37.8 ± 0.2	
Pinb-D1G143A	*a*	*a*	PINB	41.7 ± 2.5	2.5	34.7 ± 1.6	-5.9*
	a	*D1G143A*	PINBS48N	44.2 ± 0.07		28.8 ± 0.6	
Pinb-D1G143A	*a*	*a*	PINB	35.8 ± 1.1	-4	38.9 ± 0.6	0.9
	a	*D1G143A*	PINBS48N	31.8 ± 0.6		39.8 ± 1.1	
Pinb-D1C146T	*a*	*a*	PINB	43.0 ± 5.1	-2.8	39.2 ± 1.6	-1
	a	*D1C146T*	PINBS49F	40.2 ± 2.3		38.2 ± 1.4	
Pinb-D1G154A	*a*	*a*	PINB	42.0 ± 0.4	-4.9	34.0 ± 0.8	-2.5
	a	*D1G154A*	PINBD52N	37.1 ± 0.0		31.5 ± 0.0	
Pinb-D1G170A	*a*	*a*	PINB	33.3 ± 1.8	-0.7	39.3 ± 0.7	1.35
	a	*D1G170A*	PINBR57Q	32.6 ± 4.5		40.7 ± 2.6	
Pinb-D1G187A	*a*	*a*	PINB	32.8 ± 0.3	27.1*	37.8 ± 0.08	-2.5
	a	*D1G187A*	PINBD63N	59.9 ± 1.2		35.3 ± 0.2	

Table 11-continued

Pinb-D1C200T	*a*	*a*	PINB	43.3 ± 0.2	33.2*	34.7 ± 0.9	-1.8
	a	*D1C200T*	PINBT67I	76.5 ± 0.8		32.9 ± 0.6	
Pinb-D1G224A	*a*	*a*	PINB	37.5 ± 3.4	34.2*	38.4 ± 0.7	-5.5*
	a	*D1G224A*	PINBG75D	71.7 ± 0.3		32.9 ± 0.5	
Pinb-D1G224A	*a*	*a*	PINB	45.1 ± 3.3	30.4*	38.3 ± 0.8	-5*
	a	*D1G224A*	PINBG75D	75.5 ± 2.8		33.3 ± 2.5	
Pinb-D1G238A	*a*	*a*	PINB	23.8 ± 3.1	53.9*	42.7 ± 1.1	-5.2*
	a	*D1G238A*	PINBE80K	77.7 ± 2.9		37.5 ± 0.9	
Pinb-D1G254A	*a*	*a*	PINB	32.9 ± 1.7	37*	45.0 ± 0.4	-2.3
	a	*D1G254A*	PINBC85Y	69.9 ± 1.9		42.7 ± 0.4	
Pinb-D1C325T	*a*	*a*	PINB	35.3 ± 1.2	6.3*	33.3 ± 1.7	1.6
	a	*D1C325T*	PINBL109F	41.6 ± 3.2		34.9 ± 2.0	
Pinb-D1G329A	*a*	*a*	PINB	37.2 ± 0.4	8.7*	36.7 ± 1.7	-3.7*
	a	*D1G329A*	PINBG110D	45.9 ± 5.0		33.0 ± 1.0	
Pinb-D1G332A	*a*	*a*	PINB	32.0 ± 3.8	10.5*	39.7 ± 0.1	0.4
	a	*D1G332A*	PINBG111D	42.5 ± 0.7		40.1 ± 1.1	
Pinb-D1G348A	*a*	*a*	PINB	36.1 ± 0.4	36.8*	41.5 ± 0.7	0.4
	a	*D1G348A*	PINBW116Stop	72.9 ± 0.9		41.9 ± 1.1	
Pinb-D1G355A	*a*	*a*	PINB	35.6 ± 2.5	3.3	36.9 ± 0.8	1.4
	a	*D1G355A*	PINBE119K	38.9 ± 1.4		38.3 ± 0.5	
Pinb-D1G377A	*a*	*a*	PINB	42.2 ± 3.1	10.34*	36.5 ± 0.1	-4.2*
	a	*D1G377A*	PINBR126K	52.54 ± 4.9		32.3 ± 0.3	
LSD (0.05)					6.2		2.8
Alpowa parent	*a*	*a*		31.5 ± 2.7		39.6 ± 0.7	

* Denotes significance at 0.05.

[a] Two F_2 groups from each mutant population, one homozygote mutant and one homozygote wild type group, and one group of Alpowa non mutant parent were used to determine kernel characteristics. One –hundred and fifty F_2: F_3 seeds were bulked from each one of the two groups of homozygote mutant plants and homozygote wild-type plants with each group composed of 5 plants.

[b] Mean of kernel characteristics averaged over two sample of 50 seeds each group. Kernel characteristics were determined using the SKCS.

[c] Difference in kernel characteristics between two groups of each family, homozygote mutant minus homozygote wild type.

Table 12. Comparing grain hardness of homozygote mutant and homozygote wild type sibling groups of *Pin* mutant populations between two generations

Pin allele group[a]	*Pina-D1* alleles	*Pinb-D1* alleles	Effect	F_2:F_3 seed hardness	Difference[b]	F_3:F_4 seed hardness	Difference[b]
Pina-D1G212A	*a*	*a*	PINA	36.7 ± 0.09	41.1	32.8 ± 4.2	39.3
	D1G212A	*a*	PINAW71Stop	77.8 ± 6.0		72.1 ± 5.4	
Pina-D1C187T	*a*	*a*	PINA	30.4 ± 0.7	33.7	26.3 ± 7.1	34
	D1C187T	*a*	PINAP63S	64.1 ± 2.2		60.3 ± 7.6	
Pinb-D1C200T	*a*	*a*	PINB	43.3 ± 0.2	33.2	35.0 ± 3.6	36.6
	a	*D1C200T*	PINBT67I	76.5 ± 0.8		71.6 ± 3.7	
Pinb-D1C146T	*a*	*a*	PINB	43.0 ± 5.1	-2.8	29.3 ± 1.7	5
	a	*D1C146T*	PINBS49F	40.2 ± 2.3		34.3 ± 3.7	

[a] Two F_2 groups from each mutant family, one homozygote mutant and one homozygote wild type group planted in 2007 (Feiz et al. 2009), and two F_3 groups from the same mutant families planted in 2008 (Table 11) were used to determine grain hardness.

[b] Difference in kernel characteristics between two groups of each family, homozygote mutant and homozygote wild type.

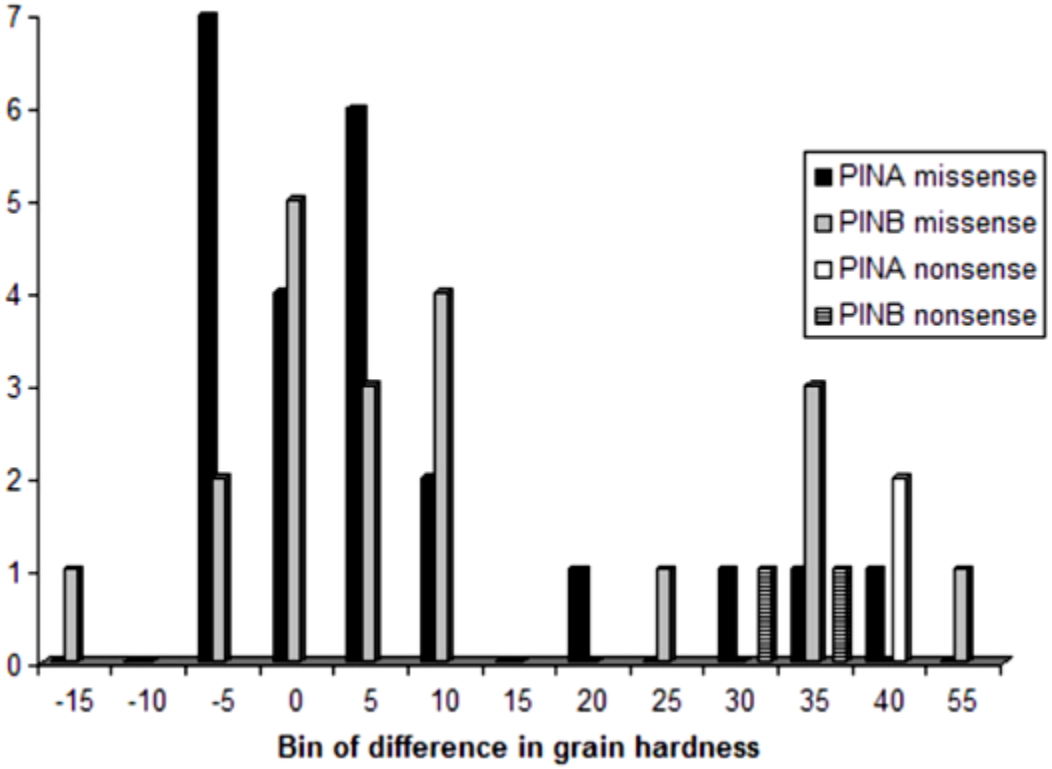

Figure 4. Classification of F_2 families with missense mutation in one of the puroindoline proteins into bins of grain hardness difference between their homozygote mutant and homozygote wild type groups. Two nonsense mutations for each PIN were used to define the function-abolishing bins.

Probability that a Missense Mutation
in PINA or PINB Affects Grain Hardness

We used the formula recently published by Gue et al. (2004) and used by Georgelis et al. (2007) to determine the probability that a random missense mutation abolishes gene function in bacterial clones expressing a target protein. This probability is termed the X-factor (X_f) and it can be calculated via:

$S = f1 (1 - X_f)$ where in our case study S is the fraction of functional proteins, f1 is the fraction of lines having one missense mutations and X_f is the probability that a random missense mutation in PINA or PINB changes protein function in conferring grain texture. Since we determined the impact of some of the missense mutations on grain texture, we obtained S via: $S = f1 \times M/N$ where M is the number of functional lines which fall into a certain threshold of protein functionality for either one of the PINs and N is the number of unique missense mutations which were used in this analysis. So we simply calculated

$X_f = 1 - M/N$

The threshold of protein functionality in this study is defined as the threshold of the changes in grain texture determining function of *Pin*s (Table 13). The M for each threshold is the number of mutant lines which the difference in grain hardness between their homozygote mutant and wild type F_2 groups do not exceed that threshold. The X_f value percentage for PINA and PINB was 26.08% and 49.99%, respectively (Table 13). By decreasing the stringency of classification of functionality of the proteins, X_f changes to reach to 0. However it stays higher for PINB than PINA. Via SIFT 48% (11/23) of the PINA and 55% (11/20) of the PINB missense mutations used for the functional analysis

were predicted to have severe impact on the function of these proteins (Table 9). SIFT uses the amino acid conservation of the proteins in the protein family database to predict the deleterious mutations. The less percentage of SIFT predicted deleterious mutations for PINA than PINB indicates that lower X_f value obtained for PINA may be due to the introduction of less conservative changes into PINA than PINB.

All of the PIN missense mutations which resulted in an increase of >20 in grain hardness, were among the amino acids which are shared between the two PIN proteins (Figure 5).

In total from the 12 PINA substitutions at the shared amino acids between two PINs, 7 were neutral, 4 were mild, and 1 was severe substitutions. From the 14 PINB substitutions at the shared amino acids between two PINs, 5 were neutral, 3 were mild and 5 were severe substitutions. PINBP41S which resulted in a significant decrease of hardness (gain of function) was one of the PINB shared amino-acids with PINA.

Table 13: X_f of PINA and PINB using different stringencies to classify mutants as nonfunctional

Threshold of the change in function	X_f (PINA)	X_f (PINB)
0±5	26.08	49.99
0±10	17.39	29.99
0±20	13.04	24.99
0±35	4.34	4.99
0±55	0	0

The F_2 families are classified according to the difference in grain hardness between their homozygote mutant and homozygote wild type groups. X_f is the probability that a missense mutation in either one of PINs increases grain hardness.

Identification of Important Amino Acid Residues and Protein Regions Involved in Puroindoline Grain-texture-determining Function

Figure 5 indicates the distribution of the total EMS-induced mutations within PINA and PINB. The position of the missense mutations used for functional analysis of puroindolines has been marked on the proteins.

The Tryptophan rich domain of PINB has been predicted to be in a coiled loop which joins the first two α –helices of PIN (Figure 5). PINAC132Y and PINBC85Y resulted in the deleterious mutations due to the loss of conserved disulfide bridges of PINA and PINB, respectively. Aside from these two severe mutations, 1 of 2 function-abolishing mutations in PINA and 3 of 4 severe mutations in PINB happened to be in the Trp loop. PINAL87F is a deleterious mutation which happened to be in the predicted helical region close to the Trp loop. Although both Leu and Phe are hydrophobic amino acids, the aromaticity of the Phe is likely the causal agent of protein destabilization in this position. None of the neutral or mild mutations occurred in Trp loop of PINB. The substitution of glycine 75 to aspartic acid in two of the F_2 mutant lines was categorized among the most function abolishing mutations. The fact that the substitution of this glycine in a turn of the polypeptide by serine is the most common mutation among natural hard wheats all over the world is an indication that this glycine is located in the active site of PINB. In this study the substitution of glycine by aspartic acid is one of the most common mutations over the entire sequence of PINs. The frequency of unique glycine to aspartic changes which used for functional analysis was 5 for PINB and 3 for PINA. However the only functional abolishing change happened where glycine 75 was

substituted by aspartic acid. Substitution of the same glycine in PINA with serine although is not categorized among the function abolishing mutations, resulted in an increase of 18 in grain hardness (Table 11). All this evidence lends support to this hypothesis that the non-polar nature of this glycine in the vicinity of the tryptophan rich region plays an important role in the function of both puroindolines and that this glycine interacts within a different environment than other PIN's glycines. Due to its covalently bound side chain, proline has a unique role in determining the local conformation of a protein (MacArthur and Thornton 1991). The substitution of proline 63 by serine in tryptophan rich loop of PINA resulted in the loss of PIN function. The same substitution in other regions of both proteins, P130S and P41S had no effect and a gain of function effect on their proteins, respectively. Two substitutions of proline by leucine in PINB (P137L and P40L) had no effect on grain hardness. Similar to glycine substitutions, proline substitutions showed a regional effect pattern. Substitution of aspartic acid 61 by asparagine had no effect on grain hardness (PINAD61N), whereas the same substitution of the same conserved amino acid (PINBD63N) resulted in an increase of 27 in grain hardness and was categorized among the function-abolishing mutations (Figure 5).

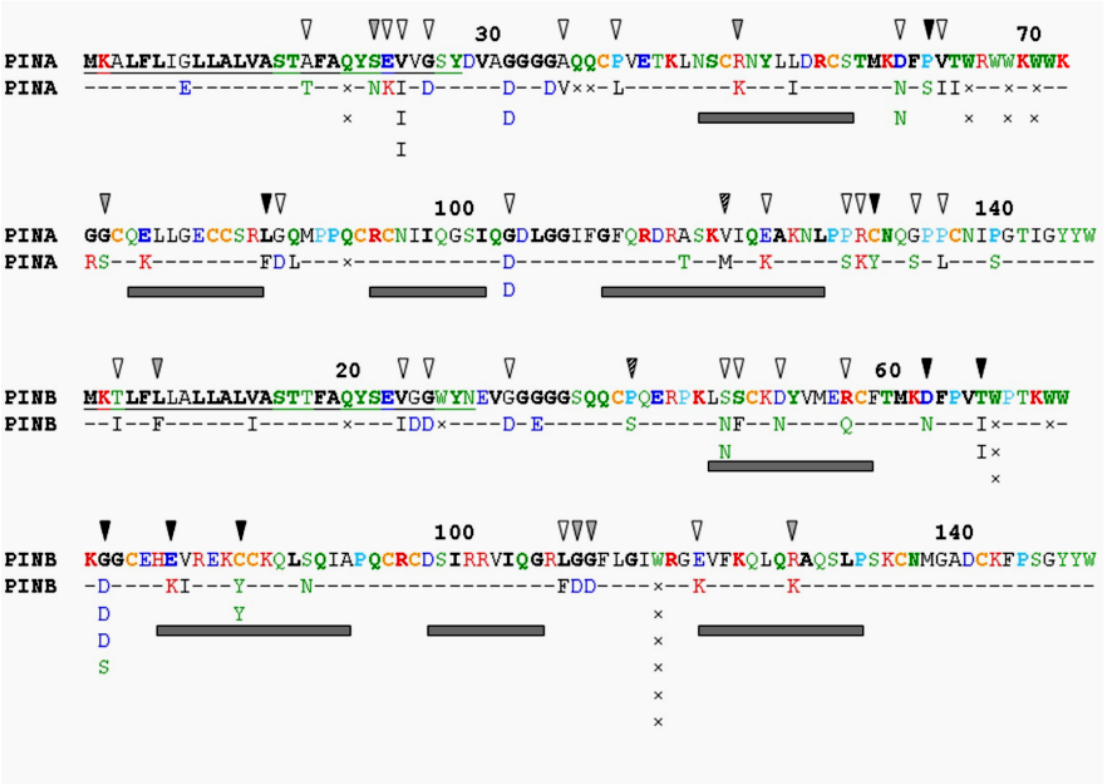

Figure 5. The position of the induced mutations within the coding region and signal peptide (underlined) of *Pina* and *Pinb*. Bold letters indicate the shared amino acids between PINA and PINB. The amino acids are colored as black, nonpolar; green, uncharged polar; red, basic; blue, acidic; turquoise, proline; orange, cysteine. Open triangles show the neutral (silent) amino acid substitution of PINs which are located in the hardness difference bins of -5 to 5 (Figure 4). Solid triangles indicate the function-abolishing amino acid substitutions in either one of the PINs (hardness difference bins of 25 or more). Gray triangles shows the position of substitutions which resulted in intermediate texture (bins of 10-20) and dashed triangle show the position of amino acid substitution in PINA and PINB with gain of function. Gray rectangles show the predicted position of α-helices in PIN proteins (Bihan et al. 1996).

Discussion

PINs are unique among plant proteins because of their tryptophan-rich hydrophobic domain. Two different functions have been proposed for PINs. First they constitute the molecular basis of grain endosperm texture, which is a primary determinant of end-use quality of wheat (Giroux and Morris 1997, 1998, Morris and Rose 1996). Second, *in vitro* antimicrobial activity of wheat puroindolines against several bacterial and fungal pathogens were demonstrated through several independent studies by Dubreil et al. (1998), Krishnamurthy et al. (2001) and Capparelli et al. (2005).

In order to improve understanding of PIN function as well as improving wheat end use quality, we created an *in planta* source of EMS-induced random mutations. We further screened it to find the mutations in *Pina* and *Pinb* genes. Via sequencing over 1,496 kb of DNA, our mutation rate was about 1 per 11.5 kb DNA (Table 10) which is about 13 times higher than that reported for *A. thaliana* (Greene et al. 2003), 22 times higher than *O. sativa* (Till et al. 2007), 77 times higher than *H. vulgare* (Caldwell et al. 2004) and two times higher than previously observed in hexaploid wheat (Slade et al. 2005).

Crossing individuals with random single amino acid mutations in either one of PINs into the wild type parent allowed us to perform functional analysis of the mutations via F_2 isolines obtained from each mutation. In terms of their impact on grain texture, most of the missense mutations were categorized into two groups; a group which had no impact on grain texture and the function-abolishing group giving hard grain texture (Figure 4). Apart from disulfide-bridge-diminishing mutations which happened on the conserved cysteines, all of the other function abolishing mutations were centered on the

tryptophan-rich domain of the proteins which strongly support the hypothesis that this region is the active site of PINs (Giroux et al. 2000). Additional evidence to support this hypothesis is that most of the *Pina* and *Pinb* natural missense alleles found in various germ plasm surveys of hard textured hexaploid bread wheat are centered on this domain (reviewed in Bhave and Morris 2008). Functional analysis of three of these natural mutations in *Pinb* via creation of near isogenic lines in Alpowa background proved that all three mutations (G75S, L89P and W73R) confer hard texture (Morris and King, 2008).

We estimated the probability that a missense mutation in PINA or PINB affects grain hardness and the results showed that this probability was twice as high for PINB compared to PINA (Table 13). To explain the significantly higher X_f value of PINB, we hypothesized that PINB grain texture determining function is more vulnerable to missense changes than PINA. The supporting evidence is that PINBD63N abolishes the PIN function in conferring kernel softness, whereas PINAD61N which is a shared amino acid in a conserved protein region between two proteins has no effect on grain texture. The fact that the same substitution of a conserved amino acid close to the tryptophan rich domain of the protein is neutral in one PIN and is function-abolishing in another PIN indicates the structural-functional specificity of these two proteins (Figure 5). Transgenic manipulation demonstrated that while both PINA and PINB levels limit grain softness, PINB is a greater limiting factor than PINA (Swan et al. 2006).

It has been hypothesized that selective pressure results in the evolution of proteins that are more tolerant of change than their homologues proteins meaning that their X_f are expected to be less than X factors of their homologous proteins (Gue et al. 2004). The specificity of PINA to adapt a different function than determining grain texture has been shown by other studies. Via analysis of the complete coding sequences from the *Triticum*/*Aegilops* taxa, rye and barley, Massa and Morris (2006) showed that *Pina*, *Pinb* and *GSP-1* genes have distinct rates of sequence evolution. They concluded that adaptive forces operated only at *Pina* locus resulting in strong positive selection ($\omega>4$) at this locus consistent with its role as a plant defense gene. A value of >1 for ω (the ratio of natural non-synonymous mutations per non-synonymous sites to silent mutations per silent sites) is the indicator of positive or adaptive selection. Massa and Morris (2006) statistically proved the presence of amino acid positions along the *Pina* sequence with a high probability of having been fixed by natural selection. In conclusion their results showed the possibility that *Pina* has evolved adaptively in response to plant pathogens to enhance fitness. Via transgenic lines overexpressing PINA, PINB or both PINs, Krishnamurthy et al. (2001) showed a significantly higher percentage of disease control in rice transgenic lines expressing PINA than the lines expressing PINB. In their study the percentage of disease control was not significantly different between lines expressing PINA and lines expressing both PINs.

Trp and Arg are the characteristic functional amino acids of antimicrobial peptides (Chan et al. 2006). Via site-directed mutagenesis in an antimicrobial peptide (bacteriocins), Fimland et al. (2002) proved that the three Trp are the functional residues

of these peptides. Jing et al. (2003) studied the mode of action of a 13-amino-acid peptide encompassing the Trp-rich domain of PINA. Their results indicated that a partially helical amphipathic structure and cation-π interactions between positively charged Lys and Arg residues and Trp aromatic indole ring allow the peptide to penetrate more deeply into vesicles resembling bacterial membranes than eukaryotic membranes. In this interaction, basic residues interact with the bacterial negative outer leaflet, and Trp resides in the membrane interface. The specificity for bacterial membranes is the result of the bacterial membrane outer leaflet being heavily populated with negatively charged phospholipid headgroups whereas in plants and animals, the dominant membrane outer leaflets lipids have no net charge (Zasloff, 2002). Further functional analysis of PIN new alleles via disease resistance tests will likely reveal the regions of PINA and PINB critical for conferring disease resistance. Via wheat standard end quality tests on the F_4:F_5 seeds derived from the F_2 populations developed from *Pin* missense mutations, the usefulness of the new alleles will be tested in improving wheat end use quality. Hard wheats with reduced grain hardness may improve milling yield as suggested by the results of Martin et al. (2001) or improve starch extractability (Feiz et al. 2008). Soft wheats with reduced grain hardness may improve break flour yield (Hogg et al. 2005; Wanjugi et al. 2007b). Soft wheats with increased grain hardness would likely have increased particle size but be more readily sifted (Campbell et al. 2007). In summary some of these new alleles may improve end use properties of soft and hard wheats.

LITERATURE CITATION

AACC International (2000) Approved Methods of the American Association of Cereal Chemists, 10th Ed. Methods 26-50, 44-15A, and 76-13. The Association: St. Paul, MN.

Anderson P (1995) Mutagenesis Methods. Cell Biol. 48:31-58.

Anjum FM, Walker, CE (1991) Review on the significance of starch and protein to wheat kernel hardness. J Sci Food Agric 56: 1-13.

Bagge M, Xia X, Lübberstedt, T (2007) Functional markers in wheat. Curr Opin Plant Biol 102:211-6.

Beecher B, Bettge A, Smidansky E, Giroux MJ (2002) Expression of wild-type pinB sequence in transgenic wheat complements a hard phenotype. Theor Appl Genet 105:870-877.

Bergthaller,W Lindhauer, MG (2004) Proposal for rapid evaluation of the separation potential of concentrated wheat flour/water slurries. Cereal Technol 8:239-245.

Bhave M, Morris CF, (2008) Molecular genetics of puroindolines and related genes: allelic diversity in wheat and other grasses. Plant Mol Biol 66:205–219.

Bihan TL, Blochet JE, Désormeaux A, Marion D, Pézolet M (1996) Determination of the secondary structure and conformation of puroindolines by Infrared and Raman spectroscopy. Biochemistry 35:2712 -12722.

Blochet JE, Chevalier C, Forest E, Pebay-Peyroula E, Gautier MF, 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 Lett 329:336-40.

Caldwell DG, McCallum N, Shaw P, Muehlbauer GJ, Marshall DF, Waugh RA (2004) A structured mutant population for forward and reverse genetics in barley (*Hordeum vulgare* L). Plant J 40:143-150.

Campbell KG, Bergman CJ, Gualberto DG, Anderson JA, Giroux MJ, Hareland G, Fulcher G, Sorrels ME, Finney PL (1999) Quantitative trait loci associated with kernel traits in a soft hard wheat cross. Crop Sci 39:11841195.

Campbell JB, Martin JM, Crutcher F, Meyer FD, Clark DR, and Giroux MJ (2007) Effects on soft wheat (Triticum aestivum L.) quality of increased puroindoline dosage. Cereal Chemistry 84(1):80-87.

Cane, K., Spackman, M., and Eagles, H.A. 2004. Puroindoline genes and their effects on grain quality traits in southern Australian wheat cultivars. Aust. J. Agric. Res. 55:89 – 95.

Capparelli R, Amoroso MG, Palumbo D, Iannaccone M, Faleri C, Cresti M (2005). Two plant puroindolines colocalize in wheat seed and in vitro synergistically fight against pathogens. Plant Mol. Biol. 58: 857-867

Chan DI, Elmar JP, Vogel HJ (2006) Tryptophan- and arginine-rich antimicrobial peptides: Structures and mechanisms of action. Biochim. Biophys. Acta 1758:1184-1202

Chang C, Zhang H, Xu J, Li W, Liu G, You M, Li B (2006) Identification of allelic variations of puroindoline genes controlling grain hardness in wheat using a modified denaturing PAGE. Euphytica 152: 225-234.

Chen F, He ZH, Xia XC, Lillemo M, Morris C (2005) A new puroindoline b mutation present in Chinese winter wheat cultivar Jingdong 11. J Cereal Sci 42**:**267–269.

Chen F, He ZH, Xia XC, Xia LQ, Zhang XY, Lillemo M, Morris CF (2006) Molecular and biochemical characterization of Puroindoline a and b alleles in Chinese landraces and historical cultivars. Theor Appl Genet 112:400-409.

Colbert T, Till BJ, Tompa R, Reynolds S, Steine MN, Yeung AT, McCallum CM, Comai l, Henikoff S (2001) High-throughput screening for induced point mutations. Plant Physiol 126:480-484.

Cox TS, (1998) Deepening the wheat gene pool. J Crop Production 1:1-25.

Czuchajowska Z, Pomeranz Y (1993) Protein concentration and prime starch from wheat flours. Cereal Chem 70:701-706.

Dubreil L, Gaborit, T, Bouchet B, Gallant DJ, Broekaert WF, 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-LTP) of Triticum aestivum seeds. Relationships with their in vitro antifungal properties. Plant Sci. 138**:** 121-135

Feiz L, Martin JM, Giroux MJ (2009) Creation and functional analysis of new *Puroindoline* alleles in *Triticum aestivum*. Theor Appl Genet DOI 10.1007/s00122-008-0893-1

Feiz L, Martin JM, Giroux MJ (2008) The relationship between wheat (Triticum aestivum L.) grain hardness and wet-milling quality. Cereal Chem 85(1):44-50.

Fimland G, Eijsink VGH, Nissen-Meyer J (2002) Mutational analysis of the role of tryptophan residues in an antimicrobial peptide. Biochemistry 41 (30):9508 - 9515.

Finn RD, Tate J, Mistry J, Coggill PC, Sammut JS, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL, Bateman A (2008) The pfam protein families database. Nuc Acid Res 36:281-288.

Gautier MF, Aleman ME, Guirao A, Marion D, Joudrier P (1994) Triticum aestivum Puroindolines, two basic cystine-rich seed proteins: cDNA sequence analysis and developmental gene expression. Plant Mol Biol 25:43-57.

Gedye KR, Morris CF, Bettge AD (2004) Determination and evaluation of the sequence and textural effects of the Puroindoline a and Puroindoline b genes in a population of synthetic hexaploid wheat. Theor Appl Genet 109:1597-603.

Georgelis N, Braun EL, Shaw JR, Hannah LC (2007) The two AGPase subunits evolve at different rates in angiosperms, yet they are equally sensitive to activity-altering amino acid changes when expressed in bacteria. The Plant Cell 19:1458-1472.

Gill BS, Raupp WJ (1987) Direct genetic transfers from *Aegilops squarrosa* L to hexaploid wheat. Crop Sci 27: 445–450.

Giroux MJ, Morris CF (1997) A glycine to serine change in Puroindoline b is associated with wheat grain hardness and low levels of starch-surface friabilin. Theor Appl Genet 95:857-864.

Giroux MJ, Morris CF (1998) Wheat grain hardness results from highly conserved mutations in the friabilin components Puroindoline a and b. Proc Natl Acad Sci USA 95:6262-6266.

Giroux MJ, Talbert L, Habernicht DK, Lanning S, Hemphill A, Martin JM (2000) Association of puroindoline sequence type and grain hardness in hard red spring wheat. Crop Sci 40:370–374.

Greenblatt GA, Bettge AD, Morris CF (1995) The relationship among endosperm texture, friabilin occurrence, and bound polar lipids on wheat starch. Cereal Chem 72:172 -176.

Greene EA, Codomo CA, Taylor NE, Henikoff JG, Till BJ, Reynolds SH, Enns LC, Burtner C, Johnson JE, Odden AR, Comai L, Henikoff S (2003) Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. Genetics 164:731-740.

Greenwell P, Schofield JD (1986) A starch granule protein associated with endosperm softness in wheat. Cereal Chem 63:379-380.

Greffeuille V, Abecassis J, Rousset M, Oury FX, Faye A, L'helgouac'h CB, Lullien-Pellerin V (2006) Grain characterization and milling behavior of near-isogenic lines differing by hardness. Theor Appl Genet 114:1-12.

Gue HH, Choe J, Loeb LA (2004) protein tolerance to random amino acid change. PNAS 101:9205-9210.

Hogg AC, Beecher B, Martin JM, Meyer F, Talbert L, Lanning S, Giroux MJ. (2005) Hard wheat milling and bread baking traits affected by the seed-specific overexpression of Puroindolines. Crop Sci 45(3): p. 871-878.

Hogg AC, Sripo T, Beecher B, Martin, JM, Giroux MJ. (2004) Wheat Puroindolines interact to form friabilin and control wheat grain hardness. Theor Appl Genet 108:1089-1097.

Jander G, Baerson SR, Hudak JA, Gonzalez KA, Gruys KJ, Last RL (2003) Ethylmethanesulfonate saturation mutagenesis in Arabidopsis to determine frequency of herbicide resistance. Plant Physiol 131:139–146.

Jing W, Demcoe AR, Vogel HJ (2003) Conformation of a bactericidal domain of puroindoline a: structure and mechanism of action of a 13-residue antimicrobial peptide. J of Bacteriol 185:4938-4947.

Kooijman M, Orsel R, Hessing M, Hamer RJ, Bekkers ACAPA (1997) Spectroscopic characterization of the lipid-binding properties of wheat puroindolines. J Cereal Sci 26:145–159.

Krishnamurthy K, Giroux MJ (2001) Expression of wheat puroindoline genes in transgenic rice confers grain softness. Nature Biotech 19:162-166.

Krishnamurthy K, Balconi C, Sherwood JE, Giroux MJ. (2001) Increased tolerance to fungal diseases of rice plants transformed with puroindoline genes. Molec Plant-Microbe Interact 14:1255-1260.

Lanning SP, Talbert LE, McNeal FH, Alexander WL, McGuire CF, Bowman H, Carlson G, Jackson G, Eckhoff J, Kushnak G (1992) Registration of 'Hi-Line' wheat. Crop Sci 32: 283-284.

Law CN, Young CF, Brown JWS, Snape JW, Worland JW (1978) The study of grain protein control in wheat using whole chromosome substitution lines In: Seed protein improvement by nuclear techniques. International Atomic Energy Agency, Vienna, Austria, PP 483-502.

Letang C, Samson MF, Lasserre TM, Chaurand M, Abecassis J (2002) Production of starch with very low protein content from soft and hard wheat flours by jet milling and air classification. Cereal Chem 79:535–543.

Lillemo M, Morris CF (2000) A leucine to proline mutation in *Puroindoline b* is frequently present in hard wheats from Northern Europe. Theor Appl Genet 100:1100–1107.

LMC. (2002) Evaluation of the community policy for starch and starch products. Online at http://www.ienica.net usefulreports/starch.pdf. LMC International: New York.

MacArthur MW, Thornton JM (1991). Influence of proline residues on protein conformation. J Mol Biol 218:397-412.

Marion D, Gautier MF, Joudrier P, Ptak M, Pezolet M, Forest E, Clark DC Broekaert W (1994) Structure and function of wheat lipid binding proteins Pages 175–180 in: S Martino al Cimino (ed) Wheat kernel proteins — Molecular and functional aspects: Universita degli studi della tuscia Viterbo, Italy.

Martin C R, Rousser R, Brabec D L (1993) Development of a single-kernel wheat characterization system. Trans ASAE 36:1399-1404.

Martin JM, Frohberg RC, Morris CF, Talbert LE, Giroux MJ (2001) Milling and bread baking traits associated with puroindoline sequence type in hard red spring wheat. Crop Sci 41:228-234.

Martin JM, Meyer FD, Morris CF, Giroux MJ (2007) Pilot scale milling characteristics of transgenic isolines of a hard wheat over-expressing puroindolines. Crop Sci 47 (2):497-5040.

Martin JM, Meyer FD, Smidansky ED, Wanjugi H, Blechl AE, Giroux MJ (2006) Complementation of the pina (null) allele with the wild type Pina sequence restores a soft phenotype in transgenic wheat. Theor Appl Genet 113:1563-1570.

Martin JM, Sherman JD, Lanning SP, Talbert LE, Giroux MJ (2008) Effect of variation in amylose content and puroindoline composition on bread quality in a hard spring wheat population. Cereal Chem 85:266-269.

Massa AN**,** Morris CF (2006) Molecular evolution of the puroindoline-a, puroindoline-b, and grain softness protein-1 genes in the tribe *Triticeae*. J Mol Evol 63:526-536

Massa AN, Morris CF, Gill BS (2004) Sequence diversity of puroindoline-a, puroindoline-b, and the grain softness protein genes in *Aegilops tauschii* cross. Crop Sci 44:1808-1816.

Mattern PJ, Morris R, Schmidt JW, Johnson, VA (1973) Location of genes for kernel properties in the wheat variety 'Cheyenne' using chromosome substitution lines In: Proceedings of the 4th international wheat genetics symposium. Columbia, MO, pp 703-707.

McIntosh RA, Devos KM, Dubcovsky J, Rogers WJ, Morris CF, Appels R, Somers DJ, Anderson OA (2007) Catalogue of gene symbols for wheat: 2007 supplement. http://www.grain.jouy.inra.fr/ggpages/awn/53Textfile/WGC.html.

Morris CF, Allan RE **(**2001) Registration of hard and soft near-isogenic lines of hexaploid wheat genetic stocks. Crop Sci 41:935-936.

Morris CF, Greenblatt GA, Bettge AD, Malkawi HI (1994) Isolation and characterization of multiple forms of friabilin. J Cereal Sci 21:167-174.

Morris CF, King GE (2008) Registration of hard kernel puroindoline allele near-isogenic line hexaploid wheat genetic stocks. Journal of Plant registrations 2:67-68.

Morris CF, King, GE, Allan RE, Simeone MC (2001a) Identification and characterization of near-isogenic hard and soft hexaploid wheats. Crop Sci 41:211-217.

Morris CF, Lillemo M, Simeone MC, Giroux MJ, Babb SL, Kidwell KK (2001b) Prevalence of Puroindoline grain hardness genotypes among historically significant North American spring and winter wheats. Crop Sci 41:218-228.

Morris C F, Rose SP (1996) Wheat. In: R J Henry and P S Kettlewell (ed) Cereal grain quality. Chapman and Hall, London, pp 3-54.

Morrison WR, Law CN, Wylie LJ, Conventry AM, Seeking J (1989) The effect of group 5 chromosomes on the free polar lipids and breadmaking quality of wheat. J Cereal Sci 9:41-52.

Ng PC, Henikoff S (2003) SIFT: predicting amino acid changes that affect protein function. Nuc Acid Res 31:3812-3814.

Pan Z, Song W, Meng F, Xu L, Liu B, Zhu j (2004) Characterization of genes encoding wheat grain hardness from Chinese cultivar GaoCheng 8901. Cereal Chem 81: 287-289.

Parris N, Moreau RA, Johnston DB, Singh V, Dickey LC (2006) Protein distribution in commercial wet- and dry-milled corn germ. J Agric Food Chem 54:4868-4872.

Ram S, Boyko E, Giroux MJ, Gill BS (2002) Null mutation in Puroindoline A is prevalent in Indian wheats: Puroindoline genes are located in the distal part of short arm of 5D chromosome. J Plant Biochem Biotech 11:1-5.

Riede CR, Anderson JA (1996) Linkage of RFLP markers to an aluminum tolerance gene in wheat. Crop Sci 36:905–909.

SAS Institute, Inc (2004) SAS/STAT 91 User's guide SAS Institute Inc, Cary, NC

Sayaslan A (2006) Characterization of volatile compounds of starches extracted with aqueous solutions. Int J Food Properties 9:463–473.

Sayaslan A, Seib PA, Chung, OK (2006) Wet-milling properties of waxy wheat flours by two laboratory methods. J Food Engineer 72:167-178.

Simeone MC, Gedye KR, Mason-Gamer R, Gill BS, Morris CF (2006) Conserved regulatory elements identified from a comparative puroindoline gene sequence survey of *Triticum* and *Aegilops* diploid taxa. J Cereal Sci 44:21-33.

Slade AJ, Fuerstenberg SI, Loeffler D, Steine MN, Facciotti D (2005) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. Nat Biotechnol 23:75-81.

Swan CG, Bowman JGP, Martin JM, Giroux MJ (2006) Increased puroindoline levels slow ruminal digestion of wheat (Triticum aestivum L.) starch by cattle. J Anim Sci 84:641-650.

Swan CG, Meyer FD, Hogg AC, Martin JM, Giroux MJ. (2006) Puroindoline B limits binding of Puroindoline A to starch and grain softness. Crop Sci 46(4):1656-1665.

Symes K J (1965) The inheritance of grain hardness in wheat as measured by particle size index. Aust J Agric Res 16:113–123.

Tester RF, Patel T, Harding SE (2006) Damaged starch characterisation by ultracentrifugation. Carbohydr Res 341:130-137.

Till BJ, Cooper J, Tai TH, Colowit P, Greene EA, Henikoff S, Comai L (2007) Discovery of chemically induced mutations in rice by TILLING. BMC Plant Biol 11:7-19.

Thorgeirsson TE, Russel CJ, King DS, Shin YK (1996) Direct determination of the membrane affinities of individual amino acids. Biochemistry 35:1803-1809.

USDA. 2007. World agricultural supply and demand estimates. Published on line at http://usda.mannlib.cornell.edu/. USDA: Washington, DC.

Van Der Borght A, Goesaert H, Veraverbeke WS, Delcour JA (2005) Fractionation of wheat and wheat flour into starch and gluten: overview of the main processes and the factors involved. J Cereal Sci 41:221–237.

Vignaux N, Fox SR, Johnson LA (2006) A 10-g laboratory wet-milling procedure for maize and comparison with larger scale laboratory procedures. Cereal Chem 83:482-490.

Wanjugi HW, Hogg AC, Martin JM, Giroux MJ (2007a) The role of puroindoline A and B individually and in combination on grain hardness and starch association. Crop Sci 47:67-76.

Wanjugi, HW, Martin, JM, Giroux, MJ (2007b) Influence of Puroindolines A and B individually and in combination on wheat milling and bread traits. Cereal Chemistry 54(6):540-547.

Warburton ML, Crossa J, Franco J, Kazi M, Trethowan R, Rajaram S, Pfeiffer W, Zhang P, Dreisigacker S, van Ginkel M (2006) Bringing wild relatives back into the family: recovering genetic diversity in CIMMYT improved wheat germplasm. Euphytica 149:289-301.

Zasloff M (2002) Antimicrobial peptides of multicellular organisms. Nature 415:389–395

Zehr BE, Eckhoff SR, Nyquist WE, Keeling PL (1996) Heritability of product fractions from wet milling and related properties of maize grain. Crop Sci 36:1159-1165.

Zhu JK, Liu J, Xiong L (1998) Genetic analysis of salt tolerance in Arabidopsis: evidence for a critical role of potassium nutrition. Plant Cell 10: 1181-1192.