# Molecular analysis of grain hardness genes in wheat and barley

Gurvinder Kalra
Master of Science (Plant Breeding)

This thesis is presented for the degree of Doctor of Philosophy
February 2017

**Abstract**

Grain hardness or endosperm texture is a commercially important trait that governs the end use of wheat and this variability in grain texture enables the production of different food products and eating qualities. *Puroindoline* (*Pina, Pinb*) genes control grain hardness in wheat. Variations, in either or both of *Pin* genes result in a hard texture of common wheat. In the recent years, several new *Pina* and *Pinb* alleles have been discovered in different wheat germplasm. However, the reason for difference in the grain hardness of genotypes of same *Pin* class is not known. The *Pinb-2* genes could be one of the minor genetic factors involved in influencing grain hardness. Thus, this study investigated *Pinb-2* genes of 22 Australian wheat cultivars that belonged to three different *Pin* genotype classes (*Pina-D1a/Pinb-D1a*, *Pina-D1a*/*Pinb-D1b, Pina-D1b*/*Pinb-D1a*) and varied for grain hardness as tested by Single Kernel Characterisation System (SKCS). Twenty-two worldwide landraces were also genotyped for *Pinb-2 variants* in this study. The seed samples were provided by Australian Winter Cereal Collection (AWCC; Tamworth, NSW, Australia). The grain hardness values of cultivars used in this study were determined by the SKCS, using the AACC Method 55-30, and ranged from 30.4-79.0. The grain texture of selected wheat accessions was also assessed in the broad catagories (hard or soft) by Scanning Electron Microscopy (SEM). *Pin* allele genotyping studies were undertaken by gene amplifications, cloning and/or DNA sequencing. Several findings were made that might enhance our understanding of the relationship of *Pinb-2* genes with grain texture. The *Pinb-2v3* allele was predominant compared to *Pinb-2v2*, and these two alleles did not co-exist. The *Pinb-2v2 and Pinb-2v3* were found to be allelic. Similarly, the subtypes *Pinb-2v3-1* and *Pinb-2v3-1a* (Val104Ala) did not co-exist. *Pinb-2* variants: v1, v2 or v3, v4, v5, v6 occurred together in various combinations and are likely non-allelic. The occurrence/non-occurrence of *Pinb-2* variants (v2 or v3) could not explain the difference in SKCS hardness values of same *Pin* class. The Val104Ala substitution in *Pinb-2v3* was not found to be associated with increased SKCS grain hardness values. Further, six novel *Pinb-2* subtypes of the main variant type were discovered. A preliminary investigation into the protein-protein interactions of *Pinb-2* genes using yeast- two hybrid studies revealed weak interaction of PINB-2v3 with PIN proteins. The trends in gene expression levels of *Pinb-2* variants, and any association thereof with grain hardness were also investigated using semi-quantitative reverse transcriptase PCR (sq-RTPCR) for quick analysis. The gene expression level of *Pinb-*

*2v3-1a* (Val104Ala) was found higher as compared to *Pinb-2v2* and *Pinb-2v3-1*. However, the increased gene expression level of *Pinb-2v3-1a* was not found to be associated with grain hardness values. The findings suggested that the PINB-2 variant proteins seem to have greater significance as antimicrobial proteins and make a minor contribution in affecting grain hardness. Thus, while *Pinb-2* genotyping alone might not be adequate for wheat breeders trying to improve wheat quality, it will be useful as an additional parameter, together with the main determinants (*Pina-D1* and *Pinb-D1* genes), due to their relevance to minor differences in grain texture which may also impact the end-uses of a cultivar.

Similarly, grain texture affects malting and processing quality of barley with relevance to its end-use. *Hordoindoline (Hin)* genes are the genetic determinants of barley grain texture. The role of *Hin* genes in influencing grain texture is not very clear and needs investigation. Thus, this study investigated *Hina*, *Hinb-1* and *Hinb-2* gene diversity in 14 worldwide landraces and 12 Australian barley cultivars. The seed samples were tested for SKCS and were provided by AWCC (landraces) and Dr. Joesph Pannozo (cultivars). The methodology used for determining barley endosperm hardness was the same as used for wheat. The cultivars represented a range of grain hardness (24.2-58.1 SKCS units), and the landraces ranged from 48.7- 94.4 SKCS units. Several different *Hina*, *Hinb-1* and *Hinb-2* haplotypes (novel and published) were found in cultivars and landraces. A novel *Hinb-1* allele with insertion at nucleotide position 210 and leading to frame-shift mutation was discovered in landrace L400211 from Ethiopia that had high grain hardness SKCS value. The variation in SKCS hardness values could not be directly associated with SNPs occurring in their *Hin* genes except for L400211 that showed frame shift mutation in *Hinb-1* gene. Potentially important substitutions with implications on lipid binding/antimicrobial properties were noted in cultivars/landraces. The findings suggest that *hordoindoline* sequence variation may play some role in barley grain hardness along with other factors and the selection of certain haplotypes might serve as a useful tool for barley breeders aiming to improve barley quality.

**Acknowledgements**

First and foremost, I would like to thank the Almighty for giving me the strength to complete my PhD. I am extremely grateful to my principal supervisor, Prof. Mrinal Bhave for giving me the opportunity to work in this project. Thank you for all your encouragement, patience and advice that kept me in high-spirits all the time. I am very thankful for the valuable feedback on thesis write-ups. I was also fortunate enough to have Dr. Joesph Panozzo as my co-supervisor for whom I thank for his encouragement and assistance.

I am extremely grateful for the Australian Postgraduate Award (APA) scholarship that supported this project. I would like to thank Dr. Rebecca Alfred (SUT) for her guidance with the yeast-two hybrid work. Also, I would like to thank Dr. James Wang (SUT) for his expertise in using scanning electron microscopy. Thank you to Dr. Humei Wu, Mrs Soula Mougos, Mr Chris Key, Mr Ngan Nguyen, Ms Angela McKellar for technical support. I am also grateful to former and current PhD students, Dr. Atul Kamboj, Dr. Shanthi Joseph, Dr. Abirami Ramalingam, Dr. Rue Hove, Dr. Snehal Jadhav, Dr. Avinash Karpe, Dr. Rohan Shah, Ms. Yen, Ms. Nadine, Ms. Zaynab for their support and encouragement.

To my darling husband, Satinder, thank you for being so understanding and supportive all the time. Thank you to my cheeky son, Yuvraj for making me laugh each day. Your purity and innocence were my teachers and will always be. To my mum and dad- thank you for your constant encouragement. To my late brother, Sukhmeet, for your fond memories that gave me strength throughout this journey.

**Publications arising from this work**

Journal articles

Kalra G, Panozzo J, Bhave M. (2017) Analysis of gene diversity of barley *Hordoindolines* in Australian germplasm and worldwide landraces (advanced stage draft).

Kalra G, Panozzo J, Bhave M. (2017) Investigation of *Pinb-2* gene diversity in Australian wheat cultivars and worldwide landraces (advanced stage draft).

Conference presentations

Kalra G, Panozzo J, Bhave M. (2015) Analysis of *Puroindoline b-2* variant genes in selected wheat lines. Poster presentation at the 36th Lorne Genome Conference held at Lorne, Victoria (15-17 Feb., 2015).

Kalra G, Panozzo J, Bhave M. (2016) Genetic diversity of *Hordoindoline* genes in barley and potential implications for grain texture. Oral presentation at the 17th International Biotechnology symposium and Exhibition held at Melbourne (24-27 Oct., 2016).

**Declaration**

I, Gurvinder Kalra, declare that the PhD thesis entitled 'Molecular analysis of grain hardness genes in wheat and barley' is no more than 100,000 words in length, exclusive of tables, figures, appendices, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma, and has not been previously published by another person. Except where otherwise indicated, this thesis is my own work.

Gurvinder Kalra

2017

**Abbreviations**

Standard chemical symbols, genetic notations, gene names and SI units are used without definition. Full gene names are given at their first mention.

| | |
|---|---|
| 3-AT | 3-Amino-1,2,4-triazole |
| Ade | Adenine |
| Ala (or A) | Alanine |
| Asn (or N) | Asparagine |
| Asp (or D) | Aspartic acid |
| Arg (or R) | Arginine |
| BDT | Big dye terminator |
| bp | Base pair (s) |
| CAP | Cleaved amplified polymorphic sequence |
| CDS | Coding sequence |
| cDNA | Complementary DNA |
| C-terminal | Carboxyl terminal (of a protein) |
| cv | Cultivar |
| Cys (or C) | Cysteine |
| DEL | deletion |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxyribonucleotide (any) |
| DMSO | Dimethyl sulfoxide |
| DO | Drop-out |
| EST | Expressed sequenced tag |
| g | centrifugal force |
| gDNA | Genomic DNA |
| Glu (or E) | Glutamic acid |
| Gln (or Q) | Glutamine |
| Gly (or G) | Glycine |
| GSP-1 | Grain softness protein-1 (a component of friabilin) |
| *Hin* | *Hordoindoline* genes |

| | |
|---|---|
| *Hina* | *Hordoindoline a* |
| *Hinb-1* | *Hordoindoline b-1* |
| *Hinb-2* | *Hordoindoline b-2* |
| HIN | Hordoindoline proteins |
| HINA | Hordoindoline a protein |
| HINB-1 | Hordoindoline b-1 protein |
| HINB-2 | Hordoindoline b-2 protein |
| His (or H) | Histidine |
| Ile (or I) | Isoleucine |
| INS | insertion |
| IPTG | Isopropyl β-D-1-thiogalactopyranoside |
| K | Keto; degenerate code for nucleotides T or G |
| Kb | Kilobase pairs; 1kb= 1000 nucleotide bases |
| kDa | KiloDalton |
| LB | Luria-Bertani broth or medium |
| Leu (or L) | Leucine |
| LiAc | Lithium acetate |
| Lys (or K) | Lysine |
| Met (or M) | Methionine |
| mRNA | Messenger RNA |
| MW | Molecular weight |
| N-terminal | Amino terminal (of a protein) |
| OD | Optical density |
| ONP | O-nitrophenol |
| PCR | The Polymerase chain reaction |
| PEG | Polyethylene glycol |
| Phe (or F) | Phenylalanine |
| *Pin* | *Puroindoline* genes |
| *Pina* | *Puroindoline-a* |
| *Pinb* | *Puroindoline-b* |
| PIN | *Puroindoline* proteins |
| PINA | Puroindoline-a protein (wild type) |

| | |
|---|---|
| PINB | Puroindoline-b protein (wild type) |
| *Pinb-2* | *Puroindoline-b2* genes |
| PINB-2 | Puroindoline-b2 protein |
| Pro (or P) | Proline |
| QTL | Quantitative trait locus |
| RNA | Ribonucleic acid |
| RNase A | Ribonuclease A |
| Ser (or S) | Serine |
| SAP | Shrimp alkaline phosphatase |
| SNP | Single nucleotide polymorphism |
| TAE | Tris acetate ethlenediaminetetraacetic acid buffer |
| Thr (or T) | Threonine |
| Tris | Tris (hydroxymethyl)aminomethane |
| Tyr (or Y) | Tyrosine |
| Trp (or W) | Tryptophan |
| U | Units of enzyme activity |
| UV | Ultra-violet |
| Val (or V) | Valine |
| X-gal | 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside |
| Y2H | Yeast two-hybrid |
| YPD | Yeast peptone dextrose medium |
| YSD | Yeast selective defined medium |

**Table of contents**

**1 General introduction and literature review**

**List of figures**

**List of tables**

# CHAPTER 1

# General introduction and literature review

## 1.1 General introduction

This chapter presents a review of literature regarding the genetic determinants of grain hardness in wheat and barley. Grain hardness is a commercially important trait that governs the end use of cereals. *Puroindoline* (*Pin*) genes control grain hardness in wheat and *Hordoindoline* (*Hin*) genes are the barley counterparts. The term indoline is derived from the indoline ring of tryptophan present in the unique tryptophan domain of Puroindoline and Hordoindoline proteins. A multigene family influencing wheat grain texture called *puroindoline b-2,* that has been studied less extensively as compared to *Pin genes* has been reviewed in this chapter for its impact on grain texture and other features like gene expression and antimicrobial properties. Similar aspects have been reviewed in case of *Hordoindoline* genes. Thus, the information provided in the review allows the reader to understand the importance of: a) genotyping of wheat germplasm for *Pin* genes and different *Pinb-2* variant types b) discovering new *Pin*/*Pinb-2* alleles that might influence grain texture and can be incorporated into breeding programs for developing wheat varieties with more suited end-use c) studying the gene expression of *Pinb-2* variants d) investigating the possibility of physical interaction among PIN and PINB-2 proteins e) genotyping of barley germplasm for *Hin* genes f) discovering new allelic variation for *Hin* genes with possible implications on grain texture.

## 1.2 Grain texture

Endosperm hardness, vitreosity, grain density, structural packing of the endosperm, grain appearance, and porosity are components of grain texture (reviewed in Walker and Pannozo, 2016). Endosperm hardness and vitreousness are considered two distinct characteristics of grain texture. Endosperm texture (hardness) is a trait first described in wheat by Greer and Hinton (1950). Hard grain is resistant to penetration of foreign matter or resistant to destruction and breakdown to particles while soft grain easily breaks under pressure. Hardness is a physical parameter defined as “the degree of resistance to deformation” and is mainly controlled by genetic factors (Turnbull and Rahman, 2002). Grain vitreousness is mainly determined by the environmental conditions during grain development and affects the rheological properties (related to dough formation) of the endosperm (Vejrazka et al. 2008). Grain texture is the organisation of individual grain

components in the endosperm and is significantly affected by the quantity, quality and ratio of protein and starch (reviewed in Psota et al. 2007).

### 1.3 Origin and importance of wheat as a crop

Wheat, rice, maize, barley and soybean are amongst the important world food crops. Wheat originates from the Fertile Crescent region (Near East), but is now cultivated globally. The primary use of wheat is to produce food for humans and has been staple food in many parts of the world (http://faostat.fao.org/). Wheat is the most important grain worldwide based on grain acreage and ranks third on the basis of total production volume (www.statista.com). A variety of nutritious foods are produced from wheats such as *Triticum aestivum* L. ssp. *aestivum* (bread wheat), and *Triticum turgidum* ssp. *durum* (Desf.) Husnot (pasta or durum wheat) (reviewed in Morris, 2002). Several factors contribute to the success of wheat as a major crop plant as reviewed in Shewry (2009) such as: a) wheat is adapted to a wide range of temperate environments as there is sufficient genetic diversity in bread wheat (with more than 25,000 varieties b) the crop yields are high (>10 tonnes $ha^{-1}$) with sufficient water and mineral nutrients, and effective control of pest and pathogens, c) easily harvested using mechanical combine harvesters or using traditional methods and d) long term storage of the grains before consumption. The unique dough-forming properties from wheat flour allows a wide range of food products to be processed, that include a wide range of breads, noodles, other baked products and many more (reviewed in Shewry, 2009).

#### 1.3.1 Wheat production in Australia

Most Australia's grain production is accounted by wheat. It is used to produce breads, noodles and pastas. Australia produces just three per cent of the world's wheat (about 25 million tonnes per annum) but accounts for 10-15% of the world's 100 million tonnes annual global wheat trade. Various types of wheat are produced in Australia, including Australian Prime Hard (APH), Australian Hard (AH), Australian Premium White (APW), Australian Noodle Wheat (ANW), Australian Standard White (ASW), Australian Premium Durum (ADR), and Australian Soft (ASFT) (www.aegic.org.au/australian-grain-production-a-snapshot/). The main wheat producing states of Australia are Western

Australia, New South Wales, South Australia, Victoria and Queensland. Western Australia the largest exporting state and majority of Australian wheat is sold overseas. The major export markets are in the Asian and Middle East regions and include Indonesia, Japan, South Korea, Malaysia, Vietnam and Sudan (www.agriculture.gov.au/ag-farm-food/crops/wheat).

### 1.3.2 Grain texture in wheat

Grain texture is an important trait in wheat as it determines the marketing and the technological utilization of wheat (reviewed in Morris, 2002). The texture of endosperm plays an important role in determining the commercial properties such as milling yield, flour particle size, shape and degree of starch damage (Pomeranz and Williams 1990). Soft wheat requires less energy to mill as compared to hard wheat (reviewed in Nadolska et al. 2009). In hard wheat, the starch-protein adhesion is strong that causes starch granules to fragment during grain milling, called "starch damage" and it affects the water absorption of the flour (Tranquilli et al. 1999). Starch damage also determines the amount of soluble carbohydrates available during bread dough fermentation, resulting in increased gas production and bread loaf volume (Mikulikova, 2007) and crumb structure (Dubreil et al. 1997). Soft wheats are used to make biscuits, cakes, pastries while hard wheat is suited for breads and buns and other yeast raised products (reviewed in Morris, 2002). Wheat grain texture is most commonly determined by one of three methods, the particle size index (PSI), near-infrared reflectance (NIR), and the Single Kernel Characterisation System (SKCS) (Morris, 2002). PSI and NIR are methods that rely on particle size distribution, with PSI quantifying granularity through sifting the milled material and NIR indirectly assesses particle size through optical reflectance of ground samples (Yamazaki, 1972). The NIR calibrations are developed from PSI/SKCS data set (Fox et al. 2007a). In the case of SKCS the resistance of a crushing force on single grain is used to measure grain hardness. The SKCS value is obtained by averaging over a large number of individual grains. The hardness values (within an expected range) from hard to soft corresponds with high to low values (reviewed in Fox et al. 2007a).

Grain hardness is affected by several factors and up to 40% is attributed to N management, tillage systems, pest infestations, and environmental factors (temperature,

rainfall and location) (Mikuliková, 2007). Approximately (the other) 60% of variation in the grain texture of wheat is genetically determined by the *Hardness* genes (Law et al., 1978). Wheat grain hardness is also shown to be influenced by additional factors such as such as differences in *Pin* expression (Nirmal et al. 2016), environment and biochemical factors (Dessalegn et al. 2006; Gazza et al. 2008; Tranquilli et al. 2002; Turnbull and Rahman 2002). Minor loci (1A, 1B, 6D) other than *Ha* locus and several QTLs (1AS, 1BL, 5AS, 5DL, 6B) may be involved in influencing grain hardness (reviewed in Bhave and Morris, 2008b).

### 1.3.3 Molecular basis of wheat grain hardness

The molecular basis of grain texture in wheat is associated with the *Hardness* locus on the short arm of chromosome 5D (Law et al., 1978). The discovery of a 15 kDa protein through electrophoretic separation of proteins extracted from the surface of water-washed starch granules was associated with endosperm texture (Greenwell and Schofield, 1986). This protein was present in large amounts in soft wheats while hard wheat showed small amounts and durum wheats lacked it. Greenwell and Schofield (1986) suggested that friabilin acts as a 'non-stick' surface between starch granules and the protein matrix. This starch granule associated protein was termed 'friabilin', indicating that soft wheats were more friable than hard wheats (Morrison et al. 1992). Later, friabilin was later found to be composed of multiple proteins (Morris et al. 1994; Oda and Schofield, 1997; Rahman et al. 1994). The two polypeptides puroindoline-a (PINA), puroindoline-b (PINB) (Gautier et al.1994) make up the major components and the Grain Softness Protein-1 (GSP-1) (Rahman et al. 1994) the minor component. The genes that code PINA, PINB and GSP-1 are *Puroindoline-a* (*Pina*), *Puroindolone-b* (*Pinb*) and Grain Softness Protein (*Gsp-1*), respectively. These three genes occur at the *Ha* locus on wheat chromosome 5D (reviewed in Bhave and Morris, 2008a). A second group of *Pinb* genes has been reported on the group 7 chromosomes (Wilkinson et al. 2008), designated as *Pinb-2* (Chen et al. 2010a) and is detailed in Section 1.3.6.

### 1.3.4 Biochemical properties of PIN proteins

The *Puroindoline-a* (*Pina-D1*) and *Puroindoline-b* (*Pinb-D1*) isolated from common wheat are intronless. Both genes are about 447bp long and 70.2% identical at the coding regions (Gautier et al. 1994). The puroindoline proteins have a molecular mass of 13kDa, consist of 148 amino acids and are highly basic (Gautier et al. 1994). The PINA and PINB proteins are 55% similar (Gautier et al.1994) and contain a highly conserved cysteine-rich backbone of 10 cysteine residues, forming five disulphide bonds, and unique tryptophan-rich domain (TRD) (Blochet et al.1993). The TRD in PINA consists of five Trp residues and three Trp residues in PINB. Both PINs contain a 28 to 29- residue long N-terminal cleavable peptide (Fig.1.1). These N-terminal cleavable peptides comprise the signal peptide (first 28 residues in PINA and first 29 residues in PINB) that could play a role in intracellular targeting (Gautier et al. 1994, reviewed in Bhave and Morris, 2008a).

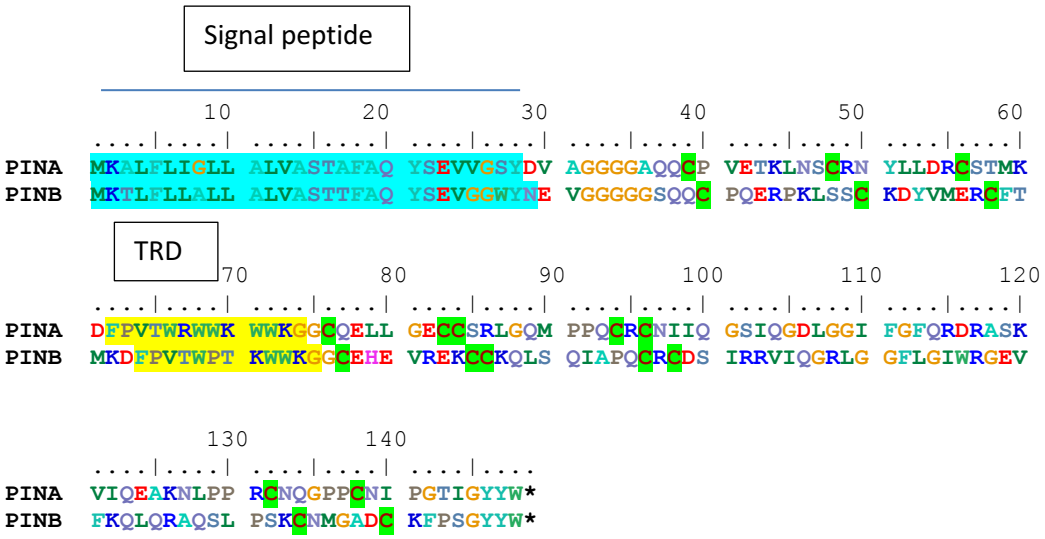



**Figure 1.1. Structure of PINA and PINB**. The signal peptide is highlighted in blue, TRD in yellow and the cysteine residues in green. The full-length PIN protein structures were first described by Gautier et al. (1994).

### 1.3.5 Relationship between *Pin* genes and grain texture

Both *Pina-D1* and *Pinb-D1* are required in their wild type (alleles *Pina-D1a* and *Pinb-D1a*) for soft grain texture, and any mutation (including deletion) in either or both genes leads to hard grain texture (reviewed in Morris, 2002). The mutations in *Pina-D1a* and *Pinb-D1a* genes that result in different soft and hard grain phenotypes in wheat have been reviewed in Bhave and Morris (2008a). Several new mutations have been reported

thereafter (http://www.shigen.nig.ac.jp/wheat/komugi/genes/symbolClassList.jsp). The mutations found in *Pina-D1a* and *Pinb-D1a* are mentioned below in Table 1.1.

**Table 1.1 *Pin* alleles in in *T. aestivum* and *Ae. tauschii***

| Allele designation | Nature of mutation[#] | Genetic source | References |
|---|---|---|---|
| *Pina-D1a* | Wild type | *T. aestivum* | Chantret et al. (2005)<br>Gautier et al. (1994)<br>Giroux and Morris (1997) |
| *Pina-D1b* | Gene deletion | *T. aestivum* | Giroux and Morris (1998) |
| *Pina-D1c* | One SNP. Arg58Gln | *Ae. tauschii* | Massa et al. (2004) |
| *Pina-D1d* | Two SNPs. Arg58Gln + one synonymous mutation | *Ae. tauschii* | Massa et al. (2004) |
| *Pina-D1e* | Two SNPs. Arg58Gln + one synonymous mutation | *Ae. tauschii* | Massa et al. (2004) |
| *Pina-D1f* | Three SNPs. Arg58Gln + two synonymous mutations | *T. aestivum* | Massa et al. (2004) |
| *Pina-D1g* | One SNP. One synonymous mutation | *Ae. tauschii* | Massa et al. (2004) |
| *Pina-D1h* | Two SNPs. Arg58Gln + one synonymous mutation | Synthetic wheat | Gedye et al. (2004) |
| *Pina-D1i* | Two SNPs. Arg58Gln +Arg21Ser | Synthetic wheat | Gedye et al. (2004) |
| *Pina-D1j* | Arg58Gln +Pro108Arg + one synonymous mutation | Synthetic wheat | Gedye et al. (2004) |
| *Pina-D1k* | Multiple deletions | *T. aestivum* | Chang et al. (2006)<br>Ikeda et al. (2005)<br>Tanaka et al. (2008)<br>Tranquilli et al. (2002) |
| *Pina-D1l* | One-base deletion, frame-shift Gln 61Lys, then a stop codon downstream | *T. aestivum* | Chen et al. (2006)<br>Gazza et al. (2005) |
| *Pina-D1m* | One SNP. Pro35Ser | *T. aestivum* | Chen et al. (2006) |
| *Pina-D1n* | One SNP. Trp43Stop | *T. aestivum* | Chen et al. (2006) |
| *Pina-D1o* | Two SNPs. Arg58Gln + one synonymous mutation | *Ae. tauschii* | Huo et al. (unpublished) |
| *Pina-D1p* | One SNP, V13E in the leader peptide. Then one-base deletion, frame-shift at Cys110Ala, then a stop codon downstream | *T. aestivum* | Chang et al. unpublished<br>McIntosh et al. (2006) |
| *Pina-D1q* | Two SNPs. Asn111Lys, Ile112Leu | *T. aestivum* | Chang et al. (2006) |
| *Pina-D1r* | Gene deletion | *T. aestivum* | Ikeda et al. (2010) |
| *Pina-D1s* | Gene deletion | *T. aestivum* | Ikeda et al. (2010) |
| *Pina-D1t* | Single SNP. Trp41stop, premature stop codon | *T. aestivum* | Ramalingam et al. (2012) |
| *Pina-D1u* | Gene deletion | *T. aestivum* | Chen et al. (2013 a) |
| *Pina-D1w* | Two SNPs. one synonymous mutation + Arg58Gln | Synthetic wheat | Ali et al. (2015) |

| | | | |
|---|---|---|---|
| *Pina-D1x* | Three SNPs. Arg58Gln + two synonymous mutations | Synthetic wheat | Ali et al. (2015) |
| *Pina-D1y* | Two SNPs. Gly53Val + Arg58Gln | Synthetic wheat | Ali et al. (2015) |
| *Pinb-D1a* | Wild type | *T. aestivum* | Gautier et al. (1994)<br>Giroux and Morris (1997) |
| *Pinb-D1b* | One SNP. Gly46Ser | *T. aestivum* | Chantret et al. (2005)<br>Giroux and Morris (1997) |
| *Pinb-D1c* | One SNP. Leu60Pro | *T. aestivum* | Lillemo and Morris (2000) |
| *Pinb-D1d* | One SNP. Trp44Arg | *T. aestivum* | Lillemo and Morris (2000) |
| *Pinb-D1e* | One SNP. Trp39 to stop codon | *T. aestivum* | Morris et al. (2001) |
| *Pinb-D1f* | One SNP. Trp44 to stop codon | *T. aestivum* | Morris et al. (2001) |
| *Pinb-D1g* | One SNP. Cys56 to stop codon | *T. aestivum* | Morris et al. (2001) |
| *Pinb-D1h* | Twenty-nine SNPs.<br>14 amino acid substitutions | *Ae. tauschii* | Massa et al. (2004) |
| *Pinb-D1i* | Thirty SNPs.<br>14 amino acid substitutions | *Ae. tauschii* | Chen et al. (2005)<br>Chantret et al. (2005)<br>Massa et al. (2004)<br>Morris et al. (2000)<br>Simeone et al. (2006) |
| *Pinb-D1j* | Nineteen SNPs.<br>9 amino acid substitutions | *Ae. tauschii* | Massa et al. (2004) |
| *Pinb-D1k* | Thirty-one SNPs.<br>14 amino acid substitutions | *Ae tauschii* | Lillemo et al. (2002) |
| *Pinb-D1l* | Lys45Glu | *T. aestivum* | Pan et al. (2004) |
| *Pinb-D1m* | Twenty-eight SNPs.<br>14 amino acid substitutions | Synthetic wheat | Gedye et al. (2004) |
| *Pinb-D1n* | Twenty-nine SNPs.<br>14 amino acid substitutions | Synthetic wheat | Gedye et al. (2004) |
| *Pinb-D1o* | Twenty-eight SNPs.<br>14 amino acid substitutions | Synthetic wheat | Gedye et al. (2004) |
| *Pinb-D1p* | One-base deletion, frame-shift at Lys42Asn, then a stop codon at 60 | *T. aestivum* | Chang et al. (2006)<br>Ikeda et al. (2005)<br>Xia et al. (2005) |
| *Pinb-D1q* | One SNP. Trp44Leu | *T. aestivum* | Chen et al. (2005) |
| *Pinb-D1r* | One-base deletion, frame-shift at Glu14Gly, then a stop codon at 48 | *T. aestivum* | Ram et al. (2005) |
| *Pinb-D1s* | One-base deletion + one SNP. Frame-shift at Glu14Gly, then a stop codon at 48 | *T. aestivum.* | Ram et al. (2005) |
| *Pinb-D1t* | One SNP. Gly47Arg | *T. aestivum* | Chen et al. (2006) |

| | | | |
|---|---|---|---|
| *Pinb-D1u* | one-base deletion, frame-shift at Glu14Ser, then a stop codon at 18 | *T. aestivum* | Chen et al. (2007) |
| *Pinb-D1v* | Two SNPs, Ala8Thr and Leu9Ile in the leader peptide | *T. aestivum* | Chang et al. (2006) |
| *Pinb-D1w* | One SNP. Ser115Ile | *T. aestivum* | Chang et al. (2006) |
| *Pinb-D1aa* | One SNP, one synonymous mutation. Then one-base deletion, frame-shift at Lys42Asn, then a stop codon at 60 | *T. aestivum* | Li et al. (2008) |
| *Pinb-D1ab* | One SNP. Gln99stop | *T. aestivum* | Tanaka et al. (2008) |
| *Pinb-D1ad* | One SNP. Val2Ala | *T. aestivum* | Kumar et al. (2015) |
| *Pinb-D1ae* | One SNP, one synonymous mutation | *T. aestivum* | Kumar et al. (2015) |
| *Pinb-D1af* | One SNP. Glu49stop | *T. aestivum* | Kumar et al. (2015) |
| *Pinb-D1ag* | One SNP. Leu95Pro | *T. aestivum* | Kumar et al. (2015) |

# mature protein sequence, SNP: Single nucleotide polymorphism

It was suggested that the mutations *Pinb-D1b* (resulting in Gly46Ser) and *Pinb-D1d* (Trp44Arg) alleles, which result in mutations within the TRD of PINB, significantly influence the interaction between PINs and anionic phospholipids (Clifton et al. 2007a; Clifton et al. 2007b). Mutations that affect the lipid binding ability of one PIN protein can influence the binding of the other and likely influence grain hardness (Clifton et al. 2007a). Transgenic studies have demonstrated the grain-softening effects of *Puroindoline* genes. Transformation of a hard red spring wheat (*Pina-D1a/Pinb-D1b*) with wild type *Pin* alleles has been shown to reduce grain hardness (Hogg et al. 2004). Transformation with *Puroindolines* have been reported to induce grain softness in crops like rice (Krishnamurthy et al. 2001) and maize (Zhang et al. 2009) lacking *Pin* genes.

### 1.3.6 *Puroindoline b-2* (*Pinb-2*) genes

Though *Pina* and *Pinb* genes are a major contributor to grain texture, the *Ha* locus does not account for the full range of grain texture variation observed in common wheat. A second set of *Pinb-2* genes, associated with a minor hardness qualitative trait locus (QTL) on all group 7 chromosomes in common wheat and durum wheat was reported by Wilkinson et al. (2008). They also identified transcripts for three variant forms of this gene (variants 1, 2 and 3) from developing grains of common and durum wheat. The putative protein sequences share around 57-60% identity with wild type PINB. All exhibit

the conservation of the characteristic 10 Cys residues as well as the TRD, but with the TRD further truncated to only two Trp residues (Fig. 1.2).

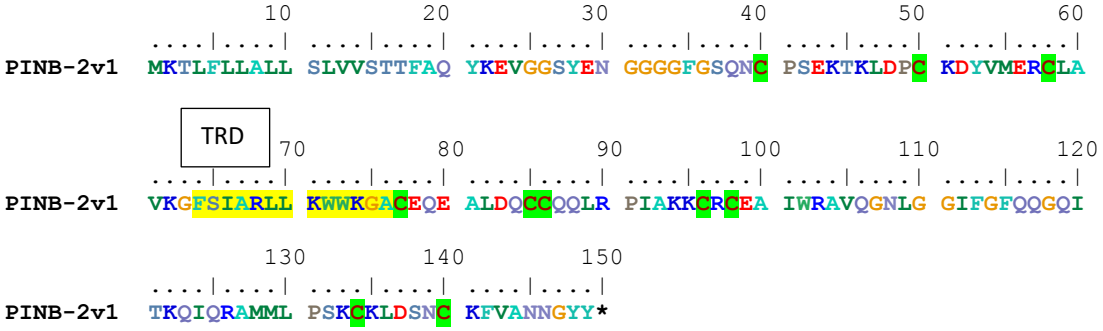



**Figure 1.2 PINB-2v1 putative protein sequence** (Wilkinson et al. 2008)
The TRD is highlighted in yellow and the cysteine residues are highlighted in green

Chen et al. (2010 a) confirmed the presence of these variants and named them *Pinb-2v*1, *Pinb-2v*2 and *Pinb-2v*3. This study also provided additional 5' and 3' flanking sequence for these variants and a new *Pinb*-2 variant, *Pinb-2v*4 was reported. They also localized *Pinb-2v1* on chromosome 7D, *Pinb-2v2* on 7BL, *Pinb-2v3* on 7B and *Pinb-2v4* on 7AL. The study proposed *Pinb-2v2* and *Pinb-2v3* to be allelic. Further remapping of *Puroindoline b-2* variants in durum wheat by Chen et al. (2011) revealed a novel variant identified at *Puroindoline b-2* locus and designated as *Pinb2-v5* in durum wheat cv. Langdon with only five SNPs difference compared to variant 4. The genetic survey of 19 durum cultivars for *Pinb-2* undertaken by Chen et al. (2011) also reported the allelic nature of *Pinb-2v2* and *Pinb-2v3* confirming the reports of Chen et al. (2010 a). Another study by Geng et al. (2012) confirmed the physical location of *Pinb-2* variant genes using 'Chinese Spring' chromosome 7 deletion lines. *Pinb-2v1*, *Pinb-2v2*, *Pinb-2v3* were found to be physically located on chromosome 7AL, 7BL, and 7DL in Chinese Spring. Further, Ramalingam et al. (2012) confirmed the five *Pinb-2* variants and reported a new group *Pinb-2v6*, showed *Pinb-2* genes comprise a multigene family. Interestingly all wild-type members within each subfamily (*Pinb-2v1* to *Pinb-2v6*) depict a conservation of the characteristics of basic pI, the structurally important 10 Cys and the functionally important TRD with at least 2 Trp in it.

### 1.3.7 Relationship of *Pinb-2* variants with grain hardness and other grain traits

Chen et al. (2010b) assessed the relationship between *Puroindoline D1* alleles and *Puroindoline b-2* variants in terms of grain hardness, other grain traits, yield components and flag-leaf size. *Pinb-2* genes were found to have greater impact on soft wheat varieties than hard. *Pinb-2v3* allele was found to have favourable association with grain yield and related traits as compared to *Pinb-2v2*. The variant 3 varieties had a potential yield advantage in their greater grain number per spike which was also reflected in greater grain weight per spike and had significantly wider flag leaves. The varieties with *Pinb-2v3* possessed about 10 % heavier kernels, 10 % more kernels per spike, as compared to varieties with *Pinb-2v2*. As mentioned earlier *Pinb-2v1*, *Pinb-2v2*, *Pinb-2v3* and *Pinb-2v4* are located on chromosomes 7DL, 7BL, 7B and 7AL in bread wheat (Chen et al. 2010a). Interestingly, the strongest QTL effects controlling grain yield, especially for grain weight have been reported on chromosomes 7 AL and 7BL (Quarrie et al. 2005). A QTL associated with grain yield has also been identified on chromosome 7D (Kuchel et al. 2007). Further, several QTLs including 7A and 7B associated with grain yield and its components were evaluated in a recombinant inbred line population (McIntyre et al., 2010).

Variant 2 and 3 found were also found to be allelic. Interestingly, Mohler et al. (2012) studied frequencies of *Pinb-D1* genes and *Pinb-2* variants and their association with quality traits. *Pinb-D1* was found to be associated with 11 quality traits and *Pinb-2* variants were found to influence semolina extraction. Further, Geng et al. (2012) studied wheat cultivars in U.S. Pacific Northwest frequencies of *Puroindoline D1* alleles and *Pinb-2* variant 2 and 3. In case of *Pinb-2* variants, variant 3 was found to be predominant. However, *Pinb-2* variants played an insignificant role in determining kernel texture (Geng et al. 2012).

In another study, Chen et al. (2013b) reported variation within *Pinb-2v3* i.e. alleles *Pinb-2v3a*, *Pinb-2v3b* and *Pinb-2v3c*. The variants *Pinb-2v3a* differing from *Pinb-2v3* by a SNP at position 6 (G to T- causing amino acid mutation Lys-Asn), which involves the signal peptide region and *Pinb-2v3b* by an SNP at 311 (T to C- showing an amino acid

change Val104Ala at the C terminal section of the deduced protein sequence while *Pinb-2v3c* contains a C325T SNP that results in a silent mutation. *Pinb-2v3* was found to be predominant genotype in wheat populations and *Pinb-2v3b* was found to be predominant amongst other variant 3 types. The cultivars with *Pinb-2v3b* were associated with increased grain hardness value of 5.5, 2.2, 1.2 SKCS units as compared to *Pinb-2v2*, *Pinb-2v3a* and *Pinb-2v3c* respectively in soft wheats. However, the difference in SKCS hardness between *Pinb-2* variants was found insignificant in hard wheat (*Pina-D1a/Pinb-D1b*) class. The substitution Val104Ala found in *Pinb-2v3b* was linked with increased grain hardness likely due to its higher gene expression levels as described in the following section.

### 1.3.8 Expression of *Pin* and *Pinb-2* genes

Based on the expression pattern in wheat seeds, the *Pins* seem to have an endosperm specific role. The *Pins* appear to be expressed during the development of wheat seed, with *Pins* possibly expressed actively during the middle stages (7-18 days post anthesis) (reviewed in Bhave and Morris, 2008b). *Pin* genes showed strict seed specific expression and no expression was found in embryo, leaf, root or shoot (Digeon et al. 1999). Both PINs were found localized in aleurone and endosperm in mature seeds by Capparelli et al. (2005). Dhatwalia et al. (2011) and Zhang et al. (2009) also established that *Puroindoline* genes show strict seed specific expression. Transcriptome analysis of the wheat seeds has detected the expression of *Pina and Pinb* genes (Gillies et al. 2012; Singh et al. 2014; Nirmal et al. 2016).

Interestingly, Wilkinson et al. (2008) noted *Pinb-2* transcripts in developing grains of bread wheat cultivar 'Hereward' for the first time and later Giroux et al. (2013) reported that *Pinb-2* variants are expressed at low levels as compared to *Pin* genes in wheat seeds. Only one study so far by Chen et al. (2013b) reported the expression of *Pinb-2* genes in seeds (different stages during filling), leaves and root tissues. Further, Chen et al. (2013b) also reported higher gene expression level of *Pinb-2v3b* relative to *Pinb-2v3a*, *Pinb-2v3c* and *Pinb-2v2*. It was suggested that the higher gene expression level of *Pin-2v3b* may be the likely reason of the increased grain hardness (see above). However, a recent study by Nirmal et al. (2016) investigated the correlation between *Pina*, *Pinb*, *Pinb-2* gene

expression and grain hardness. The study indicated that *Pinb-2* genes did not have significant influence on grain hardness on their own but a combined influence of *Pina*/*Pinb*/*Gsp-1*/*Pinb-2* explained 60 % of variation in grain hardness.

### 1.3.9 Antimicrobial properties of PIN and PINB-2 proteins

PINs influence the commercially important trait of grain hardness due to their lipid binding properties. These proteins also play biological role in seed defence against pathogens as evidenced by the antimicrobial properties of PIN proteins and peptides (Capperelli et al. 2005; Dubreil et al. 1998; Jing et al. 2003; Krishnamurthy and Giroux, 2001; Phillips et al. 2011). Chan et al. (2006) suggested the presence of the TRD, with Trp side chains is associated with antimicrobial activity due to its preference for the interfacial region of lipid bilayers. Miao et al. (2012) found antimicrobial proteins (AMPs) containing mutant forms of recombinant PINA (with an extra copy of the TRD) had significantly higher antibacterial activities than the wild-type PINA.

Ramalingam et al. (2012) showed two peptides designed on the TRD of the PINB-2 proteins displayed strong activity against bacteria and phytopathogenic fungi. The antimicrobial activity of the PINB-2 TRD peptides, PINB-2v1 and PINB-2v3 was considerably higher than PINB based peptides. These observations suggest that PINB-2 proteins play a biological role in wheat like PINA and PINB.

### 1.3.10 Possible physical interactions among PINs and PINB-2

The presence of both PINs in their functional form is required at the surface of starch for the soft phenotype (Hogg et al. 2004). Although PINA and PINB can bind to starch granules independently of each other, but the resulting grain texture is intermediate (Feiz et al. 2009b; Wanjugi et al. 2007). Caparelli et al. (2003) and Amoroso et al. (2004) indicated the expression and presence of both PINs are required for increased amount of total PINs as well as association with starch granules. PINA and PINB act cooperatively to prevent polar lipid breakdown during seed maturation and the lipid-binding property of these proteins is likely involved in development of grain softness (Kim et al. 2012).

The lipid binding properties of PINs have been associated with their functionality, the TRD being the most important region (reviewed in Bhave and Morris, 2008b, Feiz et al. 2009a). Li et al. (2014) observed low amount of starch bound PINA levels in transgenic durum wheat (overexpressing *Pina*) as compared to common wheat cultivar Chinese Spring. The study demonstrated weak association of PINA protein with starch granules in the absence of *Pinb*. Kaczmarek et al. (2015) investigated the interactions of Puroindolines with polar lipids on the surface of starch granules using fluorimetric studies. The results confirmed lipids were involved in Puroindoline-starch granule surface interactions. Thus, these studies showed the two PIN proteins act co-operatively, or in an inter-dependant manner.

Alfred et al. (2014) analysed the role of the tryptophan-rich domain (TRD), hydrophobic domain (HD), Arg39Gly substitution in PINA and *Pinb-D1c* (Leu60Pro, grain hardness associated PIN mutation) on protein-protein interactions (PPI). The TRD deletion or Arg39Gly substitution in PINA did not adversely affect its PPI, while deletion of HD resulted in significant reduction. The PINBD1-C mutant did not show any effect on PPI. The results suggested HD was essential (but not sufficient) in higher-order association of PINs.

In a recent study by Geneix et al. (2015) purified PIN proteins (PINA and PINB) from hard and soft wheat cultivars were analysed by dynamic light scattering, asymmetrical flow field fractionation and size exclusion chromatography. PINA formed small aggregates, mainly dimers in both hard and soft varieties while PINB isolated from hard varieties (PINB-D1b and PINB-D1d) assembled into large aggregates while PINB-D1a formed small aggregates. These results suggest a single amino acid substitution in PINB can affect its self- assembly and interaction with PINA.

The information on physical interactions of PINB-2 proteins is limited. Only one study so far by Ramalingam (2012) reported weak PPI of PINB-2V1 with PINA and or/PINB. *Pinb-2* variants have been linked with grain texture variations. A variant form of *Pinb-2*

variant 3 (Val104Ala substitution) has been linked with harder grain texture and its expression has been reported in seed. This suggests that PINB-2v3 might be involved in PINA/PINB co-operative binding and thus affecting grain texture.

## 1.4 Origin and domestication of barley

Barley (*Hordeum vulgare L.*) belongs to the grass family Poaceae, subfamily Pooideae and tribe Triticeae (Gaut 2002). Cultivated barley (*Hordeum vulgare*) is a diploid (HH; 2n = 2x = 14) species and has a 5.1 Giga base pair genome (The International Barley Genome Sequencing Consortium, 2012). The origins of cultivated barley date back about 10,000 years in the fertile crescent, an area of relatively abundant water in Western Asia, and near the Nile river of northeast Africa (Zohary and Hopf, 2000). Environmental and human selection of wild *H. spontaneum* is proposed to have led to the development of tough, less brittle rachis, six-rowed spikes and naked caryopsis (Salamini et al. 2002). Barley has been used for human use since 8,000 BC throughout North Africa, Southern Europe, East Asia, and North and South America. Barley was originally used mostly for human consumption but over time it has been cultivated as a feed grain and for brewing beer (Newman and Newman, 2008). Although barley is now used mainly for animal feed, malt and seed, it is still a major food in parts of Asia and North Africa, and its nutritional value is generating a renewed interest as a source for food (reviewed in Byung-Kee and Ullrich, 2008).

### 1.4.1 Importance of barley as a crop

Barley is an important cereal grain that has major contribution in feed, food and malting industry. The world barley acreage and annual production in the year 2015 is reported to be 49.8 million hectares and 145.8 million metric tonnes respectively (www.statista.com: last visited June 2016) and ranks fourth amongst cereal crops based on production. Globally, 70 % of the world barley production is utilised for animal feed and 30 % for malting purpose (FAO, 2014). The barley grain has a highly nutritive composition of carbohydrates (80%), protein (7-25%), lipids (3%), vitamins, minerals and other phytochemicals (Newman and Newman, 2008). The protein composition is variable due to environmental effects (Fox et al. 2007b). Whole grain barley is a good source of the

B-complex vitamins, especially vitamin B3 (Baik et al. 2011), and vitamin E (Kerckhoffs et al. 2002). It is also a functional food and reported to reduce blood cholesterol, assisting in prevention of cardiovascular disease, hypertension, stroke as well as diabetes (Qureshi et al. 1991; Newman and Newman, 2008; p 208). Beta-glucans, a component of the barley cell walls has been reported to lower blood cholesterol levels and is also associated with increased satiety and weight loss (reviewed in Byung-Kee and Ullrich, 2008).

### 1.4.2 Barley production in Australia

Barley is Australia's second largest grain crop (the first being wheat) and 60% of the total barley crop is exported. It meets 30% of the world's malting barley trade and 20% of the world's feed barley trade and contributes around 5% to the world's annual barley global production. The annual barley production is around 8 million tonnes/year and occupies an area of almost 4 million hectares. The majority of the barley is grown along the eastern seaboard of Australia, south-eastern Australia and western Australia (Fig. 1.3). Australia is the world's largest exporter of malting barley. Australia produces around 2.3 million tonnes of malting barley and 5.7 million tonnes of feed barley and the average Australian malting selection rate is the highest of the world's exporting nations with around 30-40% of barley national crop selected as malt. (www.barleyaustralia.com: last visited June 2016). Australian malting barley varieties have a high germination rate resulting in rapid modification of the endosperm and produce malt with a high level of extract and a range of enzyme levels to suit different style end-products.

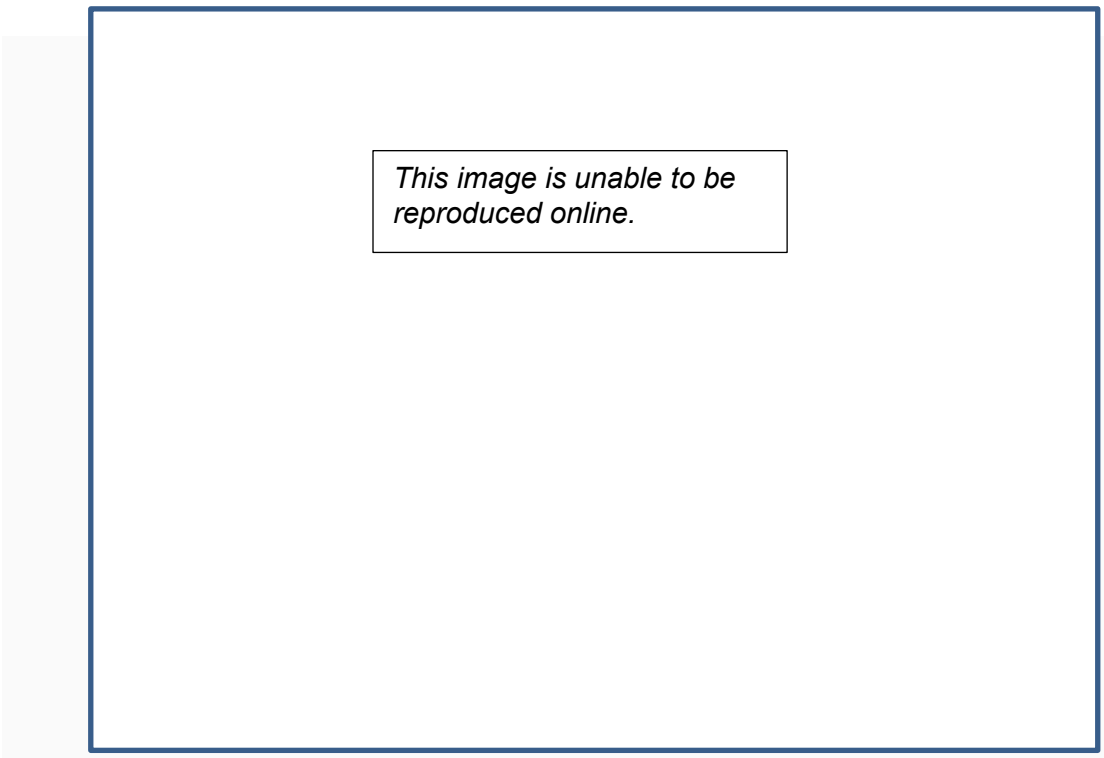



**Figure 1.3: The Australian 'barley belt'**
Source: AEGIC: Australian Export Grain Innovation Centre 2015 (www.aegic.org.au)

### 1.4.3 Grain texture of barley and its relevance to main end-uses of barley (Malt/Feed)

The terms 'mealiness' and 'steeliness' are often used to describe soft and hard endosperm characteristics of barley grain, respectively. Mealy endosperm has loosely packed starch granules in protein matrix, while steely endosperm is characterized by larger numbers of starch granules with compact embedding in the protein matrix (Palmer et al. 1989). Mealy endosperm is easily degraded by hydrolytic enzymes during malting (Swanston et al. 1995), while slower rate of water uptake during steeping and poorer distribution of enzymes during malting are characteristics of steely kernels (Chandra et al. 1999).

Barley with hard (''steely'') endosperm requires more energy to grind than soft ("mealy") barley (Camm et al. 1990). Steely barley is associated with higher protein content and lower malt extract as compared to mealy barley (Holopainen et al. 2005). The differences in steely and mealy barley for grain/malt quality parameters are listed in Table 1.2. Soft grain textured barley is reported to have better malting characteristics than hard-textured

(Allison et al. 1976). The breakdown of cell walls during malting can be quantified in barley by the friability test. Friability and hot water extract (HWE) are important selection criteria for malting and brewing. Friability measures the level of breakdown of endosperm cell wall components and protein matrix, malt varieties having much higher friability values than feed varieties. Hot water extract is also referred as malt extract and high extract levels are observed in malt varieties as compared to feed varieties. Either friability or milling energy can be used as an indicator of malt endosperm modification (reviewed in Fox et al. 2003). Greater vitreousness has been associated with poorer quality malt (Edney et al. 2002). Mealy barley contains lower concentrations of β-glucans and modifies easily in the malting process (Psota et al. 2007). The magnitude of compactness of endosperm is referred to as grain density and grains with soft endosperm have lower grain density as compared to grains with hard endosperm characteristics (Psota et al. 2007).

**Table 1.2 Grain/malt quality parameters in steely and mealy barley**

| **Parameter (unit)** | **Hard texture barley** | **Soft texture barley** | **References** |
|---|---|---|---|
| Grain hardness (PSI units) | Lower 10-15 | Higher 18-30 units | Fox et al.2003 |
| Grain hardness (SKCS units) | Higher >40 units | Lower < 30 units | Fox et al.2003 |
| Grain hardness (PS units) | Lower 1211-1290 units | Higher 1167-1318 | Fox et al. 2009 |
| Friability (%) | Lower (60.6-76.6%) | Higher (76.3-85.6%) | Fox et al. 2009 |
| Hot water extract (%) | Lower (76.7-78.9%) | Higher (77.6-78.8%) | Fox et al. 2009 |
| Malt extract (%/average) | Lower 82.7-83.0 | Higher 84.1-84.3 | Holopainen et al. 2005 |
| Protein content | Higher (11.4-12%) | Lower (9.9-10.3%) | Holopainen et al. 2005 |

Barley is classified as malting or feed according to the physical and chemical composition and processing properties (reviewed in Byung-Kee and Ullrich, 2008). Soft barley varieties are preferred for malting whereas harder endosperm is recommended for feed uses (reviewed in Bleidere and Gaile, 2012). According to the specifications of Australian malting and brewing industry, barley with a SKCS value of <30 and >40 is recommended for malting and feed uses respectively (reviewed in Fox et al. 2003). The

feed quality is influenced by the physical parameters of the grain such as grain weight, size, hardness and hull content, and chemical composition, e.g., carbohydrates, non-starch polysaccharides, fibre and protein (reviewed in Bleidere and Gaile, 2012). Digestion of barley in ruminants is reported to be dependent on the size of endosperm in ruminants (Bowman et al. 1996). *In Sacco* dry matter disappearance (ISDMD) is another indicator of feed quality, malt varieties are reported to have lower ISDMD values indicating slower disappearance in an animal's stomach, which is more desirable (Fox et al. 2009). Turuspekov et al. (2008) proposed that grain hardness was negatively correlated with DMD (dry matter digestibility), kernel diameter and kernel weight. This implies that low DMD (desirable for feed uses) is associated with harder and smaller kernels.

### 1.4.4 Determination of grain hardness in barley

Endosperm hardness is determined by a range of different methods such as milling energy, particle size index (PSI), particle size (PS), single kernel characterisation system (SKCS) and scanning electron microscopy (SEM). The PSI method is used for determining grain hardness of cereal grains. The grain samples are ground in burr mill and passed through a single 75-micron sieve. Hard barley typically has PSI values of 10-15 units and soft barley PSI values of 18-30 units (reviewed in Fox et al. 2003). A range of 18.4-27.2 PSI units (from the hardest to softest barley) was reported by Fox et al. (2007a) for barley breeding lines and commercial varieties. Particle size (PS) is determined on cracked samples by dry sieving through a series of five different size sieves. Just like PSI, low PS values are indicative of hard grain. Fox et al. (2009) reported a PS range of 1211-1290 units and 1167-1318 units for feed and malt cultivars respectively. An alternate method SKCS involves determination of resistance of crush force on a single grain, which is averaged over large numbers of grains to obtain representative data. Unlike PSI and PS, higher SKCS values are indicative of hard grain. It is noteworthy that the SKCS was initially developed to determine wheat grain hardness, and not considered suitable for barley. Iwami et al. (2005) emphasised the benefit of improving the wheat algorithm using additional information from the crush-response profile. The SKCS method using improved wheat algorithm was later shown by Fox et al. (2007a) to be reliable for measuring barley grain hardness, in a set of forty-four barley genotypes assessed for effects of genetic and environmental factors in grain hardness.

The SKCS method has since been used routinely in Dr Joseph Panozzo's laboratory (Walker et al. 2011; Walker et al. 2013) for barley, and is also reported in other literature (Galassi et al. 2012; Mohammadi et al. 2014; Takahashi et al. 2010; Turuspekov et al. 2008). Additionally, it has also been used recently for oats, using samples as small as 10 kernels (Gazza et al. 2015).

The SKCS has been widely used to measure the hardness variation within cereal mapping populations (Beecher et al. 2002; Turuspekov et al. 2008; Walker et al. 2011; Wang et al. 2008). Nagamine et al. (2009) studied the relationship between malting quality parameters and grain hardness measured by SKCS. A significant negative correlation (r=-0.48) was noted between grain hardness and malt extract. It was found that cultivars with hard grain had lower malt extract, e.g., most of the cultivars that had grain hardness value > 70 SKCS units had < 82% extract. Thus, SKCS was found to be a useful tool for selection breeding of malting barley line. Additionally, scanning electron microscopy (SEM) was used by Brennan et al. (1996) to visually examine hard and soft barley cultivars (Fig. 1.4). A high degree of starch-protein association was observed in hard barley cultivars (Fig. 1.4 B, C) and the large A-type starch granules were found to be embedded in dense layer of protein matrix, such that smaller B-type starch granules were masked by A-type granules and only A-type granules appeared to be present. In contrast, soft cultivars showed a low level of starch-protein binding (Fig. 1.4 A). Thus, SEM was found to be a useful tool for studying grain texture of barley (Brennan et al. 1996).

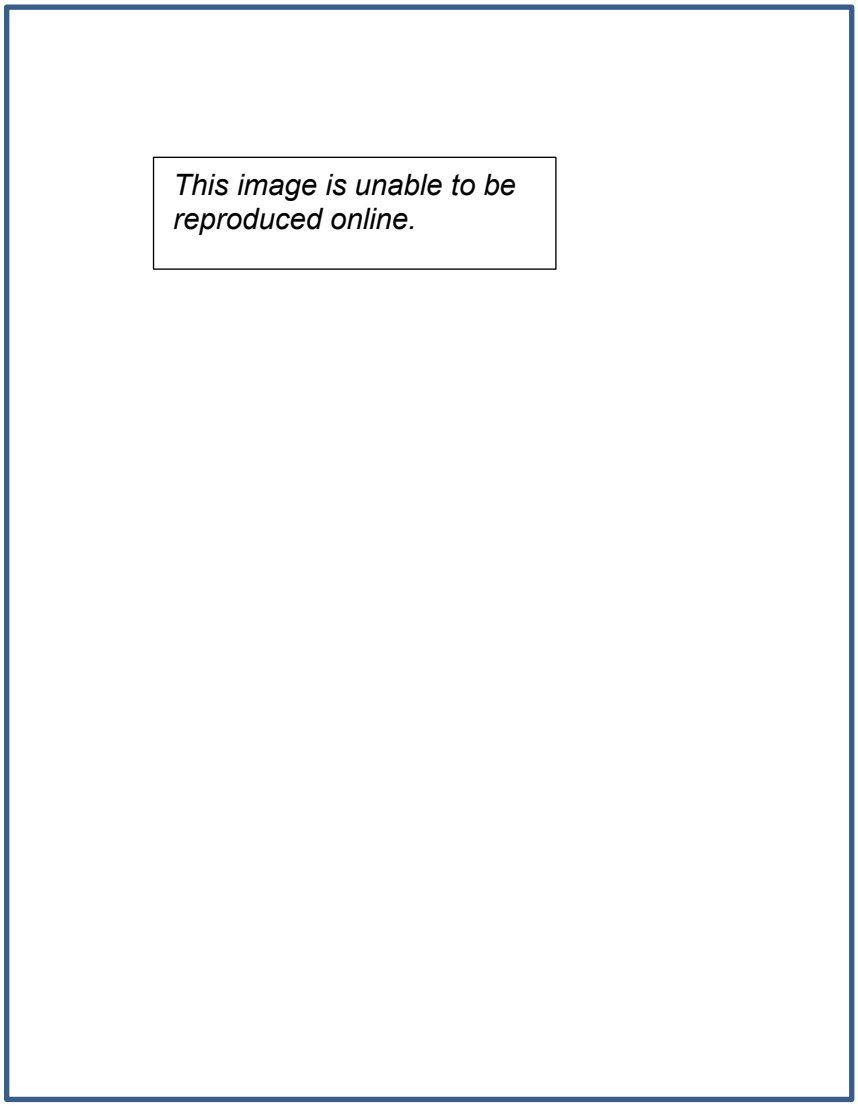



**Fig. 1.4. SEM images of soft (A) and hard (B, C) barley cultivars**
$S^a$: A type starch granule, $S^b$: B type starch granule P: protein matrix (Source: Brennan et al. 1996)

Nair et al. (2011) also used SEM to study micro structural features of hard and soft barley kernels and the flour particles. Hard-kernel lines showed continuous protein matrix and greater starch-protein adhesion than soft -kernel lines (Fig.1.5). In addition, flour particles of soft grains showed well defined individual A-type and B-type starch granules, whereas the flour particles of hard grain types showed small flour aggregates with few individual starch granules (Fig. 1.6). It was suggested that continuity of protein matrix and degree of starch – protein association may contribute to barley kernel hardness variation (Nair et al. 2011).

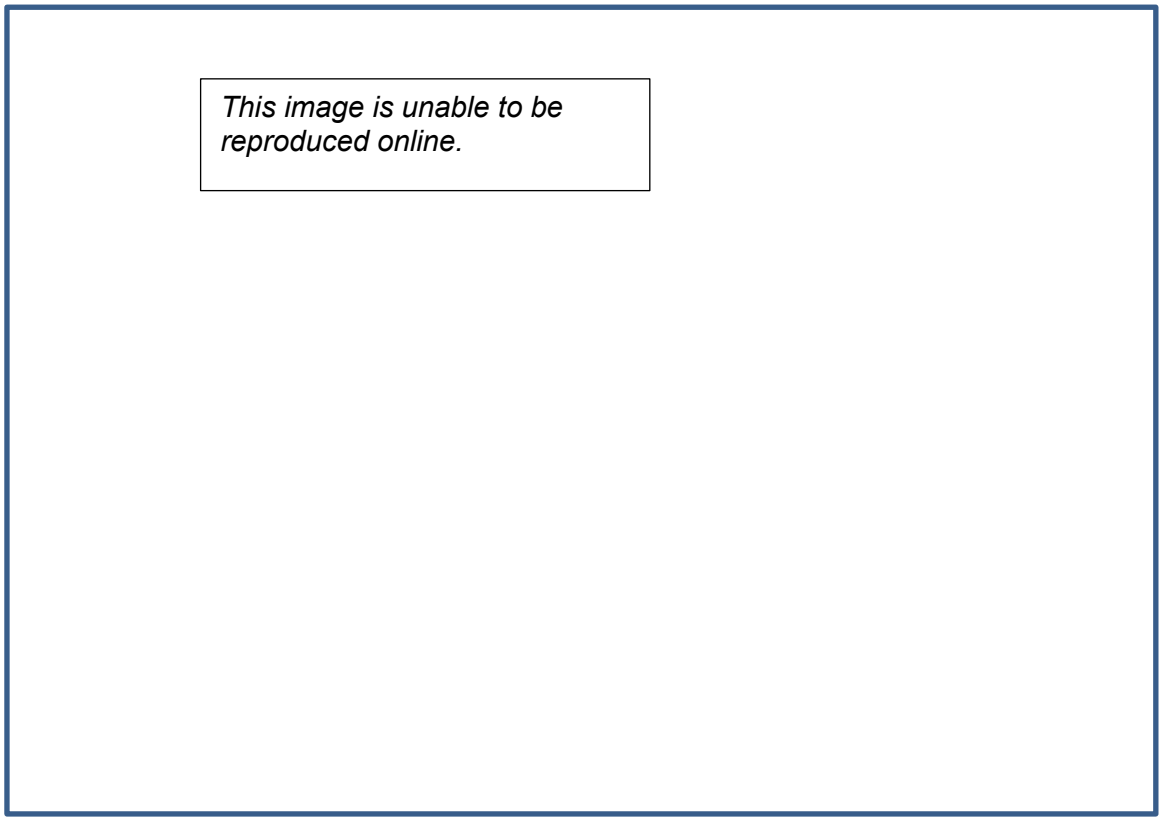



**Figure 1.5: SEM images of endosperm region of soft and hard kernel barley lines at 2000x magnification** (Source: Nair et al. 2011)

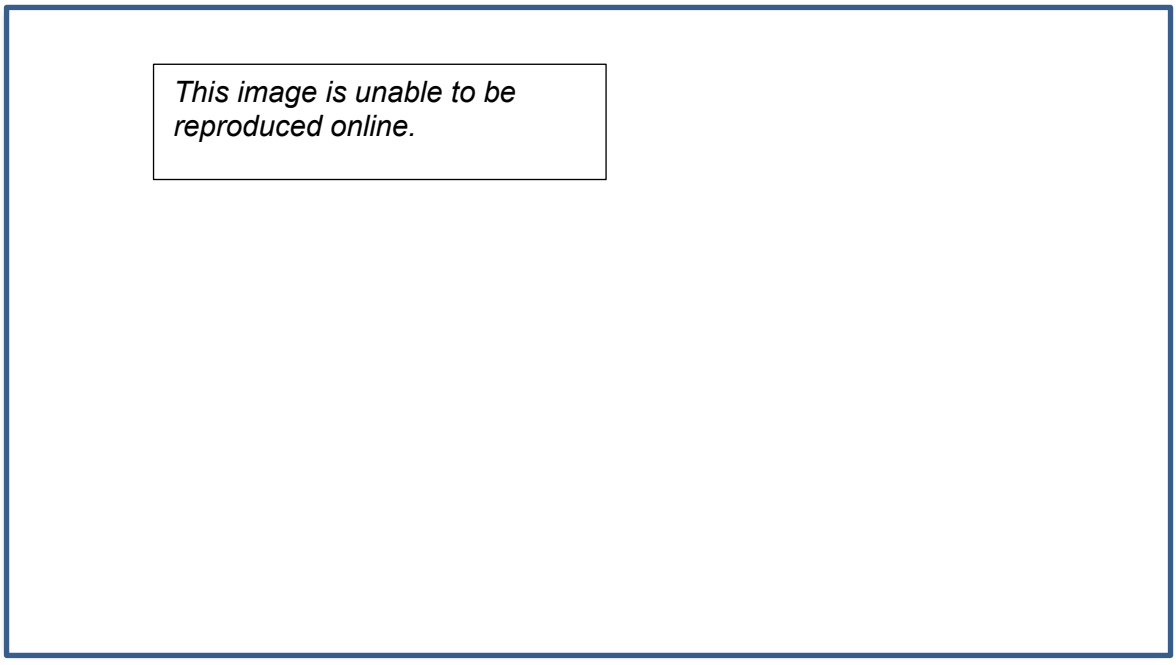



**Figure 1.6: SEM images of flour particles of soft and hard kernel barley lines at 1300x mag** (Source: Nair et al. 2011)

### 1.4.5 Occurrence of endosperm proteins in barley related to wheat friabilin

Greenwell and Schofield (1986) reported that soft wheat grains differ from hard wheat grains by the presence of a Mr 15000 protein known as friabilin which is present on the surface of water washed starch granules of soft wheat. This starch-surface associated protein was proposed to have non-stick effects preventing adhesion of starch and protein matrix. In the case of barley, Morrison et al. (1992) reported an absence of friabilin on water-washed starch granules, using caesium chloride centrifugation. However, Jagtap et al. (1993) reported the presence of friabilin in barley milled grains using a highly specific monoclonal antibody to wheat friabilin. In another study conducted by Darlington et al. (2000), it was found that starch granules of soft wheats were associated with higher amount of friabilin as compared to hard wheats, but similar levels of friabilin were noted in two types (feed and malt) of barley. It was also found that the distribution and extractability of friabilin differed in hard and soft varieties of wheat and barley. Friabilin was found on the starch granule surface in soft varieties of wheat and barley and was readily removed by repeated washing with water. In case of hard varieties, it was associated with the protein matrix and was removed by washing in case of wheat but not barley (Darlington et al. 2000). This difference was attributed to the difference in binding properties of matrix proteins in wheat and barley.

### 1.4.6 Grain hardness and related QTLs in barley

Beecher et al. (2002) reported the largest QTL associated with grain hardness on the short arm of chromosome 5H close to *Hina/Hinb/Gsp* region. QTL analysis was conducted in a doubled haploid 'Steptoe x Morex' mapping population (150 plants) segregating for *Hordoindoline* allele type and kernel hardness was measured for all the lines. The SKCS hardness values ranged from 37.2 to 76.7 units. QTLs impacting grain hardness were found on barley chromosome 1H, 4H, 5H and 7H. The grain hardness QTL on chromosome 5H accounted for 22% of the phenotypic variation in grain hardness, whereas other QTLs (1H, 4H, 7H) accounted for 9-13% variation. This chromosomal region also harbours grain texture dependent traits such as milling energy (Thomas et al. 1996), level of fine grind extract (Mather et al. 1997), and malt extract yield (Beecher et al. 2001). A number of gene regions other than 5H have been associated with quality

traits that are indirectly related to grain hardness, such as friability (Edney et al. 2002), malt tenderness (Pilen et al. 2003) and β-glucan (Psota et al. 2007).

Fox et al. (2007a) studied the genetic variation in barley grain hardness in a malting barley ('Patty x Tallon') population. Grain hardness data (PS, PSI and SKCS) was used to identify the genetic regions probably associated with grain hardness. They identified QTLs on chromosome 2H, 3H, 5H, 6H and 7H with significant logarithm of odds (LOD) values related to grain hardness. A major QTL with highest LOD score was identified on chromosome 3H. Two major QTLS were identified on 2H and it was suggested this region might coincide with a region identified for barley β-glucan as previously identified by Han et al. (1995). QTLs were identified on the distal end of both short and long arms of chromosome 5H. The region on the long arm showed significant LOD score but this was not the case for the region on the short arm associated with *Hordoindoline* genes. Walker et al. (2011) detected a significant QTL for grain hardness on chromosome 1H in a malting barley population 'Arapiles x Franklin', across three environments. Additional QTLs for endosperm hardness were detected on chromosomes 2H, 3H, 6H and 7H suggesting complex and multigenic nature of grain hardness in barley. However, no significant endosperm hardness QTL was detected in this population.

Using a malting barley mapping population (Vlamingh x Buloke), Walker et al. (2013) highlighted the complex nature of genomic regions associated with grain hardness variation, grain density and other grain parameters. Grain density QTLs were found on chromosome 2H and 6H, while endosperm hardness QTLs were found on 1H, 5H, and 7H. Interestingly, the location of *Ha* locus was not closely linked with the endosperm hardness QTLs (nearly 38.5 cM apart) in the barley mapping population. This study identified consensus QTLs (QTLs identified at more than two environments) associated with endosperm hardness, grain weight, grain width and grain yield on chromosome 5H. The regions associated with grain texture were found to coincide with QTLs for grain size, yield, flowering date and plant development genes suggesting the complex nature of barley grain hardness and related traits. Endosperm hardness and grain density were found to be significantly inter-correlated. Correlation of grain hardness and grain density

was also reported by Walker et al. (2011). Suong et al. (2016) identified QTLs for kernel hardness on chromosome 2HL and 7HL, another QTL for hot water extract on 7HL and wort-β glucan on 6HL.

### 1.4.7 Occurrence of a locus related to the wheat *Ha* locus

Wheat grain texture was linked to a single major gene on chromosome 5D (Mattern et al. 1973). Later, the *Hardness* locus was determined to be on the short arm of chromosome 5D (Law et al. 1978). Rouves et al. (1996) reported synteny between the group 5 chromosomes of wheat (the *Ha* locus being located on 5D) and chromosome 5H of barley using RFLP analysis. A set of 124 doubled haploid lines of 'Steptoe x Morex' cross was analysed to locate QTLs associated with grain hardness with RFLP probes already mapped on the homoeologous group 5 of wheat. A probe from wheat *Puroindoline* cDNA was mapped towards the distal end of chromosome 5H where the *Puroindoline* genes reside. Thus, an orthologous locus was reported on the short arm of barley chromosome 5H. Further, Beecher et al. (2001) reported a locus on short arm of chromosome 5H of barley orthologous to wheat *Ha* locus.

The *Puroindoline a* (*Pina*) and *Puroindoline b* (*Pinb*) genes are considered the main genetic determinants of kernel texture of wheat (Gautier et al. 1994). Gautier et al. (2000) first identified and designated the barley orthologs of *Puroindoline* as *Hordoindolines*, based on the latin name of its genus, *Hordeum*. Primers based on *Pin* genes were used to identify *Hordoindoline-a* (*Hina*) and *Hordoindoline-b* (*Hinb*) in barley cultivars. Beecher et al. (2001) identified a locus orthologous to wheat *Ha* on short arm of barley chromosome 5H and the *Hordoindolines* and *Gsp* were mapped there. Darlington et al. (2001) reported two tandem genes separated by 1160 nucleotides, *Hinb-1* and *Hinb-2* sequenced from barley genomic DNA (Genbank accession: AJ276143) that had 5' flanking sequences similar to *Pinb-D1*. The coding sequences of both genes were 95% identical. The *Hordoindoline-a* cDNA and protein were not found in two of the six hard varieties analysed and detected in all soft varieties. Beecher et al. (2001) found that *Hina*, *Hinb* and *Gsp* cosegregated but they did not report the two *Hinb* types (*Hinb-1* and *Hinb-2*).

The discrepancies in *Hinb* gene numbers became clear in a later study by Caldwell et al. (2004) who analysed the Morex BAC clone (Genbank accession: AY643843) depicting four tandem genes, *Hina*, *Hinb-1*, *Hinb-2* and *Gsp*. The three *Hin* genes i.e. *Hina*, *Hinb-1* and *Hinb-2* were found in same orientation, while *Gsp* in the opposite orientation. It was suggested by Caldwell et al. (2004) that all four genes are members of same family and the duplication of single ancestral gene might be their source of origin. The original duplication may have resulted in *Gsp* and one of the *Hin* genes. The divergence of *Hina*, *Hinb-1* and *Hinb-2* might be due to subsequent duplications. The *Hin* genes show high homology to wheat *Pin* genes (*Hina-Pina*: 88 %, *Hinb-1-Pinb*: 92%, *Hinb-2-Pinb*: 91%).

In another study Caldwell et al. (2006) analysed the inter- and intragenic associations across four gene loci (*Hina*, *Hinb-1*, *Hinb-2* and *Gsp*) within 215 kb of the barley *Ha* locus using haplotype-based sequence analysis. Another gene, designated as *Hinc* (Genbank accession: AB693969.1) located on chromosome 5H, has also been identified in various species of the genus *Hordeum* and is proposed to be homologue of *puroindoline b-2* variant genes in wheat (Terasawa et al. 2013).

### 1.4.8 Biochemical properties of Hordoindoline proteins

Similar to Puroindolines, Hordoindoline proteins belong to a group of cysteine-rich basic proteins and contain a tryptophan rich domain (TRD) (Gautier et al. 2000). Darlington et al. (2001) investigated the presence of Hordoindolines in barley cultivars. They reported the presence of both Hordoindoline A and Hordoindoline B proteins in the mature barley endosperm by isolating the proteins using Triton X-114 phase partitioning method of Blochet et al. (1993). Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) analysis of isolated proteins showed three major polypeptides with $M_r$ of around 1500. It was predicted that Hordoindolines like Puroindolines are synthesised as peptide signal containing precursors with a larger molecular mass than the mature proteins. Comparison of the putative PINA and HINA, PINB and HINB (HINB-1 and HINB-2) protein sequences is shown in Fig. 1.7 and Fig. 1.8 respectively. Based on the alignments, the putative TRD of HINA (FPVTWRWWTWWKG) consists of five Trp residues like PINA. Similarly, the putative TRD of HINB-1(FPLTWPTKWWKG) and HINB-2

(FPVTWPTKWWKG) consists of three Trp residues. The ten signature cysteine residues found in PINA and PINB were also found conserved in HINA and HINB pointing towards their functional similarity. Based on the alignments the signal peptide for HINA was predicted to be 28 residues long while 29 residues long for HINB-1 and HINB-2. The HINA protein consists of 150 amino acids while the HINB protein consists of 147 amino acids.

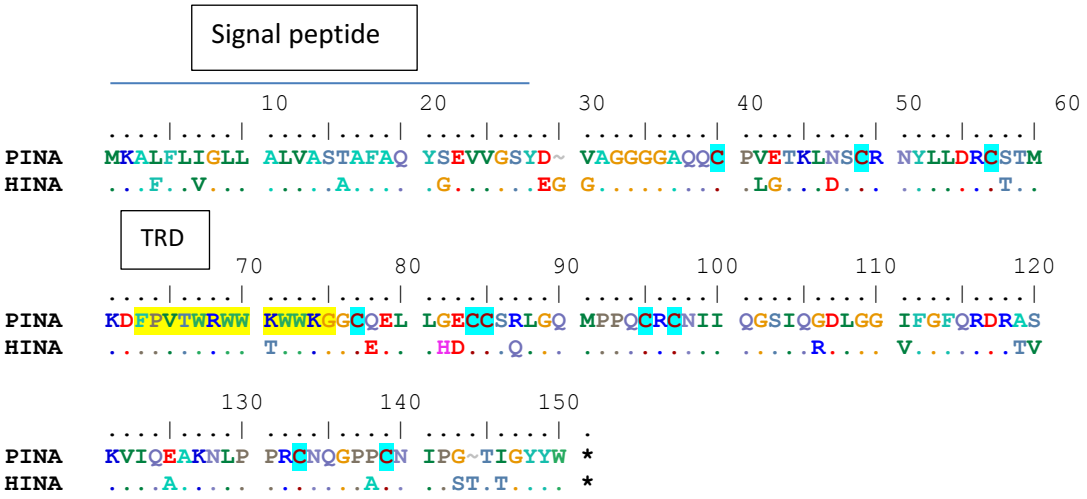



**Figure 1.7. Alignment of PINA and HINA putative protein sequences.**
Dots indicate similarity of amino acid residues. The conserved cysteine residues are highlighted in blue. The predicted TRD is highlighted in yellow. The sources of the reported putative protein sequences are: PINA (Gautier et al. 1994), HINA (Caldwell et al.2006).

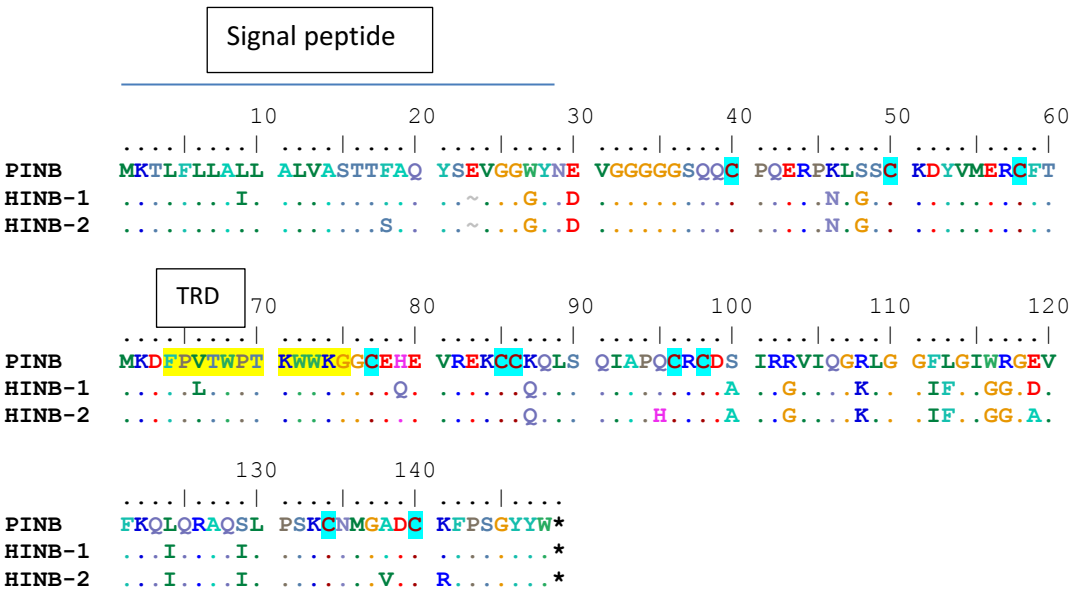



**Figure 1.8. Alignment of PINB and HINB-1, HINB-2 putative protein sequences.**
Dots indicate similarity of amino acid residues. The conserved cysteine residues are highlighted in blue. The predicted TRD is highlighted in yellow. The sources of the reported putative protein sequences are: PINA (Gautier et al. 1994), HINB1 and HINB-2 (Caldwell et al.2006).

### 1.4.9 Hordoindoline proteins and kernel texture

Galassi et al. (2012) studied Hordoindoline patterns in two-row and six-row barley cultivars grown in Italy. The Hordoindoline sequence variation was also studied by Galassi et al. (2012) and is mentioned below in section 1.14. These cultivars showed a wide range of kernel hardness values (52.4-92.3 SKCS units). The proteins accumulated on the surface of starch granules were fractionated by acidic polyacrylamide gel electrophoresis (A-PAGE) and A-PAGE x SDS-PAGE. The amount of total Hordoindolines (HINA + HINB1 + HINB2) on starch granules was comparable to Puroindolines (PINA + PINB) in soft wheat types. However, the amount of B-type Hordoindolines (HINB1 + HINB2) was found to be 50% lower than that of PINB in soft wheat types, indicating lower HINB accumulation on the surface of starch granule and thus harder texture. It was suggested that the reduced accumulation of B-type Hordoindolines on starch granules might be the reason for absence of soft textured barley cultivars, unlike wheat. Differences in kernel weight and B-type Hordoindoline levels accounted for the phenotypic variation for kernel hardness of 56 barley cultivars analysed by SKCS. Their findings indicated that some of the changes in amino acid sequences of HINA with respect to wild type PINA and HINB-1and HINB-2 with respect to wild type PINB, might lead to the lower accumulation of HINB on the surface of starch granules, thus contributing to harder texture. It is proposed that the absence of PINA in wheat leads to prevention of PINB accumulation on the surface of starch granule and thus harder texture (Corona et al. 2001; Gazza et al. 2005; Giroux and Morris 1998). Similar phenomenon can be predicted for barley but the absence of *Hina* null allele in the reports so far restricts the understanding of the relationship of HINA and grain texture.

### 1.4.10 Barley chromosome 5H plays role in reducing wheat grain hardness

Yanaka et al. (2011) demonstrated that *Hordoindoline* genes located on chromosome 5H play role reducing grain hardness in the wheat-barley chromosome addition lines. The effect of barley chromosome 5H on wheat grain hardness was studied in wheat-barley species (*H. vulgare*, *H. vulgare* ssp. *spontaneum*, and *H. chilense*) chromosome addition lines. It was found the grain hardness values of addition lines were significantly lower than the corresponding wheat parents. The grain hardness value of the three wheat-barley species addition lines was 33.3, 24.1, 18.3 SKCS units whereas the corresponding wheat

parents showed values of 39.7, 34.3 and 41.0 SKCS units respectively. Of the three barley species, the effect of *H. chilense* was greatest in reducing grain hardness. Guzman and Alvarez (2014) identified two novel alleles in *Hordeum chilense* Rom. et Schult., a wild barley species that has been used in the development of new man-made cereal (tritordeum x *Tritordeum* Ascherson et Graebner). The identified alleles (Genbank: KF717106, KF717107, KF717108) showed high identity to *Pin* genes of bread wheat which may be the likely reason for the soft texture of tritordeum.

### 1.4.11 Allelic variation in *Hin* genes and their implications on grain texture

A number of single nucleotide polymorphisms (SNPs) have been reported in *Hordoindoline* genes (*Hina*, *Hinb-1* and *Hinb-2* genes) in cultivated and wild barley (Caldwell et al. 2006; Galassi et al. 2012; Li et al. 2011; Terasawa et al. 2012; Turuspekov et al. 2008). Other *Hin* sequence variation, accessed using Genbank from unpublished sources, was also included in the comparative *Hin* alignments. In the present work, a thorough search and alignments of *Hin* sequences reported so far was undertaken to get a comprehensive picture of the genetic diversity, and any potential functional significance thereof. As mentioned earlier, the sequencing of BAC clone (Genbank accession: AY643843.1) conducted by Caldwell et al. (2004) clarified the *Hin* gene numbers and sequences; hence these sequences were used as reference sequence for all DNA and protein alignments. The existence of two *Hinb* genes (*Hinb-1* and *Hinb-2*) was not clear in earlier studies undertaken by Beecher et al. (2001) and Darlington et al. (2001) and any *Hin* sequence variation from these two studies is not included for comparison in this work. The analysis revealed 47 DNA haplotypes of *Hina*, 15 haplotypes of *Hinb-1* and 33 DNA haplotypes were found for *Hinb-2* with unique combination of SNPs inherited together. The DNA haplotypes, SNPs in the DNA sequences and amino acid substitutions in the putative protein types for *Hina*, *Hinb-1* and *Hinb-2* genes are listed in Tables 1.3, 1.4 and 1.5 respectively. The DNA and corresponding protein alignments for all three *Hin* genes are shown in Appendix I. The following sections detail the amino acid substitutions found in HINA, HINB-1 and HINB-2 sequences.

**Table 1.3 Table of all reported HINA sequences**

| Sequence | Accession | Mutation DNA | Mutation AA (in mature protein)[a] | Reference |
|---|---|---|---|---|
| HINA-1 | AY643843.1 | WT (reference sequence) | WT (reference sequence) | Caldwell et al, 2004 |
| HINA-2 | AJ249929.1 | T10C, G19A, C434T | Phe4Leu, Val7Met | Gautier et al. 2000 |
| HINA-3 | AY644142.1 | A265G, T331G | Ser89(61)Gly, Phe110(83)Val | Caldwell et al, 2006 |
| HINA-4 | GU591289.1 | A265G, T331G, INS434AAC | Ser89(61)Gly, Phe110(83)Val, INS144(116)TT | Li et al. 2011 |
| HINA-5* | DQ862190.1 | G21A, C25G, C434T | Leu9Val | Turuspekov et al. unpub |
| HINA-6* | DQ862163.1 | G21A, A314C, C434T | Gln105(77)Pro | Turuspekov et al. unpub |
| HINA-7* | DQ862148.1 | G21A, G85A, C434T | Glu29(1)Lys | Turuspekov et al. unpub |
| HINA-8* | DQ862213.1 | G21A, G85A, G232A, G305A, C434T | Glu29(1)Lys, Glu78(50)Lys, Gly102(74)Glu | Turuspekov et al. unpub |
| HINA-9 | AY644147.1 | C45T, G85A, A265G, T331G, C390T, INS434AAC | Glu29(1)Lys, Ser89(61)Gly, Phe111(83)Val, INS144(116)TT | Caldwell et al. 2006 |
| HINA-10 | GU591287.1 | A265G, A314G, T331G, INS434AAC | Ser89(61)Gly, Gln105(77)Arg, Phe110(83)Val, INS144(116)TT | Li et al. 2011 |
| HINA-11 | GU591231.1 | A265G, T331G, A361G, INS434AAC, G436A | Ser89(61)Gly, Phe110(83)Val, Lys121(93)Glu, INS144(116)TT, Gly145(117)Ser | Li et al. 2011 |
| HINA-12 | GU591234.1 | T35A, G37T, T38A, A265G, T331G, INS434AAC | Leu12Stop | Li et al. 2011 |
| HINA-13 | GU591232.1 | T187C, A265G, C268T, G325A, T331G, INS434AAC | Phe63(35)Leu, Ser89(61)Gly, Gln90(62)Stop | Li et al. 2011 |
| HINA-14 | GU591217.1 | C8T, G19A, G21A, G55A, A265G, T331G, INS434AAC, G436A | Ala3Val, Val7Ile, Ala19Thr, Ser89(61)Gly, Phe111(83)Val, INS144(116)TT, Gly145(117)Ser | Li et al. 2011 |
| HINA-15 | GU591205.1 | T187C, A265G, C268T, A314G, T331G, INS434AAC | Phe63(35)Leu, Ser89(61)Gly, Gln90(62)Stop | Li et al. 2011 |
| HINA-16 | AY644102.1 | G55A, C160T, A265G, T331G, A361G, INS434AAC | Ala19Thr, Ser89(61)Gly, Phe111(83)Val, Lys121(93)Glu, INS144(116)TT, Gly145(117)Ser Gly145(117)Ser | Caldwell et al. 2006 |
| HINA-17* | DQ862134.1 | G22A, G129A, G232T, A233T, A265G, T331G, C434T | Glu78(50)Leu, Ser89(61)Gly, Phe111(83) Val | Turuspekov et al. unpub |

| | | | | |
|---|---|---|---|---|
| HINA-18 | AY644110.1 | G55A,, C56T, C160T, A265G, T331G, A361G, INS434AAC, G436A | Ala19Met, Ser89(61)Gly, Phe111(83)Val, Lys121(93)Glu, INS144(116)TT, Gly145(117)Ser | Caldwell et al. 2006 |
| HINA-19 | GU591263.1 | DELpos92-111, A265G, T331G, A339G | DELpos33-37(5-9), Ser89(61)Gly, Phe111(83)Val | Li et al. 2011 |
| HINA-20 | GU591261.1 | DELpos92-111, A265G, T331G, A339G, A450G | DELpos33-37(5-9), Ser89(61)Gly, Phe111(83)Val, Stop151(121)W | Li et al. 2011 |
| HINA-21 | GU591262.1 | DELpos92-111, T239C, A265G, T331G, A339G | DELpos33-37(5-9), Leu80(52)His, Ser89(61)Gly, Phe111(83)Val | Li et al. 2011 |
| HINA-22 | AB611026.1 | T10C, G22A, C45G, C84T, DELpos92-111, A265G, T331G, INS434AAC | Val7Ile, Ser15Arg, DELpos33-37(5-9), Ser89(61)Gly, Met91(63)Ile, Phe111(83)Val, INS144(116)TT | Terasawa et al. 2012 |
| HINA-23 | GU591191.1 | G19A, G22A, C45G, C84T, DELpos92-111, A265G, G273A, T331G, T368C, C390T, INS434AAC | Val7Ile, Ser15Arg, DELpos33-37(5-9), Ser89(61)Gly, Met91(63)Ile, Phe111(83)Val, Ile123(95)Thr, INS144(116)TT | Li et al. 2011 |
| HINA-24 | AB605713.1 | G19A, G22A, T34C, C54T, C63A, C123T, G128A, G139A, C144A, A245G, T249G, A265G, C278A, C285T, T331G, G358T, C389G T434G, G446A | Val7Ile, Gly43(15)Glu, Asp47(19)Asn, His82(54)Arg, Asp83(55)Glu, Ser89(61)Gly, Pro93(65)Gln, Phe111(83)Val, Val120(92)Phe, Trp148(120)Stop | Terasawa et al. 2012 |
| HINA-25 | AB605716.1 | G19A, G22A, T34C, T38C, C54T, G60A, G85T, C123T, G128A, G129A, G139A, C144A, C157T, C171T, G192A, G193C, T249G, A265G, T296C, T331G, G395A | Val7Ile, Val25Ile, Gly43(15)Glu, Asp47(19)Asn, Asp83(55)Glu, Ser89(61)Gly, Ile99(71)Thr, Phe111(83)Val, Arg132(104)Lys | Terasawa et al. 2012 |
| HINA-26 | AB605714.1 | G19A, G22A, T34C, G46A, C54T, G60A, C123T, G128A, G139A, C144A, A245G, T249G, A265G, G273T, T327A, T331G, G358T, C390G, G395A, C426T, T434G, G446A | Val7Ile, Ala16Thr, Gly43(15)Glu, Asp47(19)Asn, His82(54)Arg, Asp83(55)Glu, Ser89(61)Gly, Met91(63)Leu, Phe111(83)Val, Val120(92)Phe, Arg132(104)Lys, Trp148(120)Stop | Terasawa et al. 2012 |
| HINA-27 | AB605711.1 | G19A, G22A, C28T, T34C, C54T, G60A, G101A, C123T, G128A, G139A, C144A, T187C, G192A, A245G, T249G, A265G, G303A, T331G, G337A, G358T, C390G, T434G, G446A | Val7Ile, Gly43(15)Glu, Thr44(16)Lys, Asp47(19)Asn, His82(54)Arg, Asp83(55)Glu, Ser89(61)Gly, Met91(63)Leu, Phe111(83)Val, Val120(92)Phe, Trp148(120)Stop | Terasawa et al. 2012 |
| HINA-28 | AB605712.1 | G19A, G22A, T34C, C54T, G60A, G85C, G91C, C123T, G128A, G139A, G177A, | Val7Ile, Glu29(1)Gln, Gly31(3)Arg, Gly43(15)Glu, His82(54)Arg, Asp83(55)Glu, | Terasawa et al. 2012 |

| | | | | |
|---|---|---|---|---|
| | | A245G, C246T, T249G, A265G, C277A, T331G, G337A, A357T, G358T, C390G, C426T, T434G, G446A | Ser89(61)Gly, Pro93(65)Gln, Phe111(83)Val, Val120(92)Phe, Trp148(120)Stop | |
| HINA-29 | AB605715.1 | C12G, G19A, G22A, T34C, G36A, G42A, C54T, G60A, C63T, G73A, T81C, C123T, G128A, G139A, C144A, C147T, C156T, G192A, A245G, T249G, T262C, A265G, T331G, C336T, G347A, T351C, C385T, C393G, G395A, T434G, C438T, C444T, G446A | Phe4Leu, Val7Ile, Val25Ile, Gly43(15)Glu, Asp47(19)Asn, His82(54)Arg, Asp83(55)Glu, Ser89(61)Gly, Met91(63)Leu, Phe111(83)Val, Arg116(88)His, Val120(92)Phe, Arg132(104)Lys, Gly145(117)Val, Trp148(120)Stop | Terasawa et al. 2012 |
| HINA-30 | AB605710.1 | G19A, G22A, T34C, C54T, G57A, G60A, G73A, C123T, G128A, C131A, G139A, C144A, T182C, G192A, C195G, A245G, T249G, A265G, A271T, T331G, G358T, C390G, T434G, G446A | Val7Ile, Val15Ile, Gly43(15)Glu, Thr44(16)Lys, Asp47(19)Asn, His82(54)Arg, Asp83(55)Glu, Ser89(61)Gly, Met91(63)Leu, Phe111(83)Val, Val120(92)Phe, Trp148(120)Stop | Terasawa et al. 2012 |
| HINA-31* | AB605709.1 | G19A, G22A, T34C, C54T, G60A, G73A, C123T, G128A, C131A, G139A, C144A, T182C, G192A, A245G, T249G, A265G, A271T, A293G, T331G, G357T, A363G, C390G, T432C, T434G, G446A | Val7Ile, Val25Ile, Gly43(15)Glu, Thr44(16)Lys, Asp47(19)Asn, His82(54)Arg, Asp83(55)Glu, Ser89(61)Gly, Met91(63)Leu, Phe111(83)Val, Asn98(70)Ser, Val120(92)Phe, Trp148(120)Stop | Terasawa et al. 2012 |
| HINA-32 | GU214828.1 | T10C, G19A G22A, T34C, C54T, C63A, C84T, C123T, G128A, C131A, G139A, C144A, T186C, G192A, A245G, T249G, A265G, T331G, G358T, C389G, C434T | Phe4Leu, Val7Ile, Gly43(15)Glu, Thr44(16)Lys, Asp47(19)Asn, Ser89(61)Gly, Met91(63)Leu, Phe111(83)Val, Val120(92)Phe | Gutierrez et al. unpub 2009 |
| HINA-33 | AB446469.1 | G19A, G22A, T34C, C54T, G60A, C84T, C123T, G128A, C131A, G139A, C144A, T187C, G192A, A245G, T249G, A265G, A271T, T331G, G358T, C390G, T434G, G446A | Val7Ile, Gly43(15)Glu, Thr44(16)Lys, Asp47(19)Asn, His82(54)Arg, Asp83(55)Glu, Ser89(61)Gly, Met91(63)Leu, Phe111(83)Val, Val120(92)Phe, Trp148(120)Stop | Terasawa et al. 2012 |
| HINA-34 | JX236064.1 | G19A, G22A, T34C, C54T, G60A, G84T, C95G, C123T, G128A, C131A, G139A, C144A, C160T, T182C, G192A, A245G, T249G, A265G, A271T, T331G, G358T, C390G, T434G, G446A | Val7Ile, Gly43(15)Glu, Thr44(16)Lys, Asp47(19)Asn, His82(54)Arg, Asp83(55)Glu, Ser89(61)Gly, Met91(63)Leu,, Phe111(83)Val, Val120(92)Phe, Trp148(120)Stop | Crespi et al. 2012 unpub |
| HINA-35* | DQ862178.1 | G21A, G128A, G129A, A265G, A331G, C434T | Gly43(15)Glu, Ser89(61)Gly, Phe111(83)Val | Turuspekov et al. unpub |

| | | | | |
|---|---|---|---|---|
| HINA-36* | DQ862143.1 | G22A, G128A, G232T, A233T, A265G, T331G, A339G, C434T | Gly43(15)Glu, Glu78(50)Leu, Ser89(61)Gly, Phe111(83)Val | Turuspekov et al. unpub 2006 |
| HINA-37 | JN636828.1 | G128A, C212G, T231C, G232T, A233T, A265G, T331G, A339G | Gly43(15)Glu, Thr71(43)Arg, Ser89(61)Gly, Phe111(83)Val | Galassi et al. 2012 |
| HINA-38 | JN636836.1 | G128A, G129A, C212G, T231C, G232T, A233T, A265G, T331G, A339G | Gly43(15)Glu, Thr71(43)Arg, Glu78(50)Leu, Ser89(61)Gly, Phe111(83)Val | Galassi et al. 2012 |
| HINA-39 | GU591274.1 | G22A, C49T, A114G, G128A, C212G, G232T, A233T, A265G, T331G, A339G | Gly43(15)Glu, Thr71(43)Arg, Glu78(50)Leu, Ser89(61)Gly, Phe111(83)Val | Li et al. 2011 |
| HINA-40 | GU591276.1 | G128A, C212G, T231C, G232T, A233T, A265G, C278T, T331G, A339G, C360A, A369G | Gly43(15)Glu, Thr71(43)Arg, Glu78(50)Leu, Ser89(61)Gly, Pro93(65)Leu, Phe111(83)Val, Ile123(95)Met | Li et al. 2011 |
| HINA-41* | DQ862184.1 | G22A, G128A, C212G, T231C, G232T, A233T, A265G, T331G, C434T | Gly43(15)Glu, Thr71(43)Arg, Glu78(50)Leu, Ser89(61)Gly, Phe111(83)Val | Turuspekov et al. unpub 2006 |
| HINA-42 | GU591293.1 | DELpos12-3bp, G128A, C212G, T231C, G232T, A233T, A265G, T331G, A339G, C374A | Phe5Del, Gly43(15)Glu, Thr71(43)Arg, Glu78(50)Leu, Ser89(61)Gly, Phe111(83)Val | Li et al. 2011 |
| HINA-43 | GU591283.1 | G22A, C54T, A114G, G128A, C212G, G232T, A233T, A265G, T272C, T331G, A339G, T365C | Gly43(15)Glu, Thr71(43)Arg, Glu78(50)Leu, Ser89(61)Gly, Met91(65)Thr, Phe111(83)Val, Val122(94)Ala | Li et al. 2011 |
| HINA-44* | DQ862212.1 | G22A, G128A, G129A, C212G, T231C, G232T, A233T, A265G, T331G, C434T | Gly43(15)Glu, Thr71(43)Arg, Glu78(50)Leu, Ser89(61)Gly, Phe111(83)Val | Turuspekov et al. unpub 2006 |
| HINA-45* | DQ862206.1 | G22A, C54T, A114G, G128A, C212G, G232T, A233T, A265G, T331G, A339G, C434T | Gly43(15)Glu, Thr71(43)Arg, Glu78(50)Leu, Ser89(61)Gly, Phe111(83)Val | Turuspekov et al. unpub 2006 |
| HINA-46 | AB605717.1 | G19A, G22A, T34C, C50T, C54T, C63T, G128A, G177A, G192A, C212G, C246T, T249G, C255T, A265G, T331G, A367G, G369A | Val7Ile, Ala17Val, Gly43(15)Glu, Thr71(43)Arg, Asp83(55)Glu, Ser89(61)Gly, Phe111(83)Val, Ile123(95)Val | Terasawa et al. 2012 |
| HINA-47 | AY644097.1 | C56T, G128A, C136A, C144T, C212G, T231C, G232T, A233T, A265G, T331G, A339G, INS434AAC | Ala19Val, Gly43(15)Glu, Leu46(18)Ile, Thr71(43)Arg, Glu78(80)Leu, Ser89(61)Gly, Phe111(83)Val, INS144(116)ThrThr | Caldwell et al. 2006 |

* Partial CDS, first 20bp absent from sequence,

[a]The amino acid position in the () corresponds to the site in the predicted mature protein (signal peptide predicted to be amino acid residues 1-29)

**Table 1.4 Table of all reported HINB-1 sequences**

| Seq ID | Accession | Mutation DNA | Mutation AA (in mature protein)[b] |
|---|---|---|---|
| HINB1-1 | AY643843.1 | WT (Reference seq) | WT (Reference seq) |
| HINB1-2* | DQ862246.1 | C423G | Phe141(112)Leu |
| HINB1-3* | DQ862232.1 | A366C | Gln122(94)His |
| HINB1-4 | AY644022.1 | A25C | I11Leu |
| HINB1-5 | JN636843.1 | A25C, C414G | I11Leu, Asp138(109)Glu |
| HINB1-6 | AB611029.1 | A25C, G375A | I11Leu |
| HINB1-7* | DQ862289.1 | A25C | I11Leu |
| HINB1-8 | AY644058.1 | A25C, G201C | I11Leu, Trp67(39)Cys |
| HINB1-9 | AY644025.1 | A25C, G325A | I11Leu, Gly109(81)Ser |
| HINB1-10 | AY644023.1 | A25C, A133G | I11Leu, Asn45(12)Asp |
| HINB1-11 | AY643991.1 | A25C, G329C | I11Leu, Gly110(82)Ala |
| HINB1-12 | AY643980.1 | A25C, G88A, G329C | I11Leu, Val30(2)Ile |
| HINB1-13 | AY643973.1 | A25C, G329C, C410T | I11Leu, Gly110(82)Ala |
| HINB1-14 | AY643987.1 | A25C, C56T, C176G, C225T, A231G | I11Leu, Ala19Val, Thr59(31)Arg |
| HINB1-15* | DQ862273.1 | A25C, C47T, A138G, C176G, A231G, C408G, C429T | I11Leu, Thr16I, Thr59(31)Arg |

*Partial sequence, last 3 bp missing from sequence in database

[b]The amino acid position in the () corresponds to the site in the predicted mature protein (signal peptide predicted to be amino acid residues 1-29).

**Table 1.5 Table of all reported HINB-2 sequences**

| Seq ID | Accession | Mutation DNA | Mutation AA (in mature protein)[b] | Reference |
|---|---|---|---|---|
| HINB2-1 | AY643843.1 | WT (Reference seq) | WT (Reference seq) | Caldwell et al. 2004 |
| HINB2-2 | AY644015.1 | C191T | Pro64(36)Leu | Caldwell et al. 2006 |
| HINB2-3 | AB611030.1 | A123G, C353A | Ala118(90)Asp | Teresawa et al. 2012 |
| HINB2-4* | DQ862369.1 | A123G | - | Turuspekov et al. unpub |
| HINB2-5* | DQ862366.1 | G419A | Arg140(111)Lys | Turuspekov et al. unpub |
| HINB2-6* | DQ862360.1 | C225T | - | Turuspekov et al. unpub |
| HINB2-7* | DQ862358.1 | C262A | Leu88(60)Met | Turuspekov et al. unpub |
| HINB2-8* | DQ862347.1 | T282A | His94(66)Gln | Turuspekov et al. unpub |
| HINB2-9* | DQ862335.1 | T410C | Val137(108)Ala | Turuspekov et al. unpub |
| HINB2-10* | DQ862313.1 | C191T | Pro64(36)Leu | Turuspekov et al. unpub |
| HINB2-11 | JN636845.1 | A123G, T410C, G419A | Val137(108)Ala, Arg140(111)Lys | Galassi et al. 2012 |
| HINB2-12* | DQ862370.1 | A123G, T410C | Val137(108)Ala | Turuspekov et al. unpub |
| HINB2-13* | DQ862346.1 | A123G, G419A | Arg140(111)Lys | Turuspekov et al. unpub |
| HINB2-14* | DQ862334.1 | A123G, C353A | Ala118(90)Asp | Turuspekov et al. unpub |
| HINB2-15* | DQ862300.1 | A123G, C262A | Leu88(60)Met | Turuspekov et al. unpub |
| HINB2-16* | DQ862299.1 | C15A, A123G | Phe5Leu | Turuspekov et al. unpub |
| HINB2-17 | JN636850.1 | A123G, C353A, T410C, G419A | Ala118(90)Asp, Val137(108)Ala, Arg140(111)Lys | Galassi et al. 2012 |
| HINB2-18* | DQ862354.1 | T93C, A123G, G210A | - | Turuspekov et al. unpub |
| HINB2-19 | AY644038.1 | A123G, C225T, T410C, G419A | Val137(108)Ala, Arg140(111)Lys | Caldwell et al. 2006 |
| HINB2-20 | AY644024.1 | A123G, C379T | Gln127(99)Stop | Cadwell et al. 2006 |
| HINB2-21 | AY644023.1 | A123G, C301T, T410C, G419A | Arg101(73)Trp, Val137(108)Ala, Arg140(111)Lys | Caldwell et al. 2006 |
| HINB2-22 | AY643989.1 | A123G, C251G, T410C, G419A | Gln,86(58)Glu, Val137(108)Ala, Arg140(111)Lys | Caldwell et al. 2006 |
| HINB2-23 | AY643987.1 | G74A, A123G, T410C, G419A | Gly)25Asp, Val137(108)Ala, Arg140(111)Lys | Caldwell et al. 2006 |
| HINB2-24 | AY643980.1 | A123G, T410C, G419A, C425A | Val137(108)Ala, Arg140(111)Lys, Pro142(113)His | Caldwell et al. 2006 |

| | | | | |
|---|---|---|---|---|
| HINB2-25 | AY644022.1 | C18T, A123G, T410C, T417C, G419A | Val137(108)Ala, Arg140(111)Lys | Caldwell et al. 2006 |
| HINB2-26 | AY644051.1 | A123G, C262T, DEL-G304 | Gly102(74)Glu, Val103(75)Stop | Caldwell et al. 2006 |
| HINB2-27* | DQ862343.1 | C16A, C51A, T93C, A123G, G237A | Leu6Ile | Turuspekov et al. unpub |
| HINB2-28* | DQ862307.1 | G57A, T93C, A99T, A123G, C414T | - | Turuspekov et al. unpub |
| HINB2-29 | AY644045.1 | C15A, A123G, G157A, G264A, T410C, G419A | Phe5Leu, Val53(25)Met, Val137(108)Ala, Arg140(111)Lys | Caldwell et al. 2006 |
| HINB2-30 | AY644020.1 | A122G, A123G, A123G, C241A, C377G, T410C, G419A | Gln41(13)Arg, Ala126(98)Gly, Val137(108)Ala, Arg140(111)Lys | Caldwell et al. 2006 |
| HINB2-31 | JN636849.1 | T10C, C51A, T93C, A123G, G21A, G237A, T282A, C353A, T410C, G419A | His94(61)Gln, Ala118(90)Asp, Val137(108)Ala, Arg140(111)Lys | Galassi et al. 2012 |
| HINB2-32 | AB611031.1 | G57A, T93C, A123G, G210A, C232A, C259G, T282A, C353A, T410C, C417C, G419A | His78(50)Asn, Gln86(58)Glu, His94(61)Gln, Ala118(90)Asp, Val137(108)Ala, Arg140(111)Lys | Teasawa et al. 2012 |
| HINB2-33* | DQ862368.1 | C16A, C51A, T93C, A123G, G210A, G237A, T282A, C353A, T410C, G419A | Leu6Ile, His94(61)Gln, Ala118(90)Asp, Val137(108)Ala, Arg140(111)Lys | Turuspekov et al. unpub |

*Partial sequence, last 3 bp missing from sequence in database

[b]The amino acid position in the () corresponds to the site in the predicted mature protein (signal peptide predicted to be amino acid residues 1-29).

### 1.4.12 Amino acid substitutions in HINA

The putative protein types of 47 DNA haplotypes are shown in Appendix I. The protein types HINA-13 and HINA-15 were identical. Similarly, HINA-38, HINA-39, HINA-41, HINA-44, HINA-45 were identical. Thus, altogether 42 HINA putative protein types were observed. The numbering of residues of PINs is traditionally from the first residue of the mature protein sequence (Giroux and Morris, 2002). However, for convenience the residue numbers in HINA, HINB-1 and HINB-2 in this study is from initiator Met, rather than from the first residue of the putative mature HIN protein after removal of the 28 (HINA), 29 (HINB-1, HINB-2) residue long predicted signal peptide based on alignment with PINA (Genbank accession: CAA49538) and PINB (CAA49537). The putative proteins showed changes at the functionally important tryptophan rich domain (TRD) along with acidic, basic, polar and hydrophobic residues. The potentially important ones are discussed below.

A mutation in the TRD region i.e. Thr71Arg was reported by Caldwell et al. (2006), Galassi et al. (2012), Li et al. (2011) and Turuspekov et al. unpublished; (Genbank: DQ862184.1, DQ862212.1, DQ862206.1). None of the studies reported any association between the amino acid substitution and grain hardness. Mutations in the TRD are of interest, as the Trp and basic residues within the TRD of PINA are proposed to be important for lipid binding and insertion into membranes (Kooijman et al. 1997; Le Guernevé et al. 1998). Mutations in the TRD also affect the starch surface lipid binding properties of PINs and hence wheat kernel texture (reviewed in Bhave and Morris, 2008). The ten signature cysteines were conserved suggesting the importance of conservation of these residues for the important role of protein folding just in case of PINs (reviewed in Pauly et al. 2013). Substitution of an acidic residue with hydrophobic residue i.e**.** Glu78Leu, near the TRD was reported by Galassi et al. (2012), Li et al. (2011) Turuspekov et al. unpublished (Genbank: DQ862143.1, DQ862213.1, DQ862134.1, DQ862212.1, DQ862206.1). Further, the substitution with a hydrophobic residue is also favourable as hydrophobic residues confer strong affinity for polar lipids located in biological membranes and starch granule surface to PINs (Douliez et al. 2000; Kooijman et al. 1997). Increased lipid binding capacity leads to increased antimicrobial activity and

also has grain softening effects. Hydrophobic residues have also been suggested to play role in protein –protein interactions (reviewed in Moreira et al. 2007).

Some mutations involve substitutions with basic residues e.g. E29K (Caldwell et al. 2006; Turuspekov et al. unpublished; Genbank: DQ862148.1, DQ862213.1,), T44K (Crespi et al. 2012 unpublished; Genbank: JX236064.1; Gutierrez et al. unpublished 2009; Genbank: GU214828.1; Terasawa et al. 2012) and Glu78Lys (Turuspekov et al. 2008). Substitutions with Arg included Gly31Arg (Terasawa et al. 2012), Thr71Arg (Galassi et al. 2012; Li et al. 2011, Turuspekov et al. unpublished; Genbank: DQ862184.1, DQ862212.1, DQ862206.1), His82Arg (Gutierrez et al. unpublished 2009; Genbank: GU214828.1), and Gly113Arg (Terasawa et al. 2012). The increased positive charge as a result of substitution with basic residues may result in stronger electrostatic forces between HINA and acidic head groups of polar lipids of bacterial membranes resulting in disruption of biological membranes and thus increased antimicrobial activity (reviewed in Pauly et al. 2013). Further, Pasupuleti et al. (2008) reported that increasing the positive charge can significantly increase the antimicrobial activity of an antimicrobial peptide (AMP). The effects of any of the above substitutions on grain texture are unknown. Comparison of mutations in the putative HINA protein types were made with hardness associated mutations reported in PINA. A substitution in PINA i.e., Arg58Gln (Gedye et al. 2004; Huo et al. unpublished; Genbank: AY608595; Massa et al. 2004) is of particular interest as glutamine was found at same position in HINA protein and was found conserved in all the 47 HINA protein types. This might offer some explanation for the comparatively hard texture of barley with no soft types as such.

### 1.4.13 Amino acid substitutions in HINB-1

The analysis revealed fifteen haplotypes of *Hinb-1* gene with their characteristic combination of SNPs (Table 1.3). As compared to *Hina*, fewer SNPs leading to changes in protein are reported in *Hinb-1*. The putative protein types of 15 DNA haplotypes are show in in Appendix I. The protein types HINB1-4, HINB1-6, HINB1-7 and HINB1-11, HINB1-13 were identical thus leading to 11 putative protein types. The putative proteins did not show any changes at TRD region. Some of the potentially important substitutions are listed below:

Substitution of hydrophobic and functionally-important residue tryptophan, with cysteine (Trp67Cys) was reported by Caldwell et al. (2006). Grain texture assessments were not involved in this study. Replacement of a tryptophan residue can have significant implications in terms of reduced lipid binding abilities. This substitution also leads to change in the number of signature cysteines from ten to eleven in the putative HINB1 protein (HINB1-8, Table 1.3). As mentioned earlier, cysteine residues are involved in important role of protein folding (reviewed in Pauly et al. 2013) and increased cysteine residues may prove beneficial for this role. A substitution with hydrophobic residue i.e. Gly110Ala (putative protein type HINB1-11) was observed in the same study (Caldwell et al. 2006). As discussed earlier increased hydrophobicity is associated with increased lipid binding capacity. Two substitutions with basic residues were found, the first one was Thr59Arg (HINB1-14) (Caldwell et al. 2006 and Turuspekov et al. unpublished; Genbank: DQ862273.1) and the other one was Gln122His (HINB1-15) (Turuspekov et al. unpublished; Genbank: DQ862232.1). An increased positive charge is favourable for HIN-lipid association. The potential effects of any of the above substitutions on grain texture are unknown. Mutations in the putative HINB-1 protein types were compared with mutations reported in PINB. Interestingly, the amino acid residues reported to be substituted in PINB (*Aegilops tauschii* and synthetic wheat) were also found substituted in all putative HINB1 protein types. These substitutions were Ser48Gly and Arg103Gly (Massa et al. 2004), Arg108Lys (Lillemo et al. 2002), Leu113Phe (Gedye et al. 2004), and Glu119Asp (Simeone et al. 2006). This might offer some explanation for harder grain texture of barley.

### 1.4.14 Amino acid substitutions in HINB-2

Thirty-three DNA haplotypes were found for *Hinb-2* with unique combinations of SNPs at different positions. These, and their resultant amino acid substitutions in the HINB-2 proteins are listed in Table 1.4. The alignments of 33 DNA haplotypes their putative HINB-2 protein types are shown in Appendix I. Some of the protein types were identical and altogether 22 putative protein types were observed. The putative proteins did not show any changes at TRD region and all the ten cysteines were conserved. Substitutions were noted at hydrophobic, basic and polar residues. The potentially important ones are mentioned below:

Caldwell et al. (2006) reported a mutation, Arg101Trp (HINB2-21) involving substitution of a basic residue with functionally important tryptophan that led to increase in number of tryptophan residues. However, this substitution would lead to significant decrease in positive charge due to replacement of basic arginine. Grain texture assessments were not involved in this study. Substitutions at hydrophobic residues were: Ala118Asp (Galassi et al. 2012; Terasawa et al. 2012; Turuspekov et al. unpublished; Genbank: DQ862334.1, DQ862368.1), P142H and A126G (Caldwell et al. 2006). The effects of these substitutions were not linked with grain texture. Two substitutions led to replacement of basic residue Histidine i.e. His94Gln (Terasawa et al. 2012 and Turuspekov et al. unpublished; Genbank: DQ862347.1, DQ862368.1) and His78Asn (Terasawa et al. 2012). The effects of these substitutions were not linked with grain texture. Caldwell et al. (2006) reported a mutation in *Hinb-2* (named *Hinb-2b*, Genbank accession: AY644051.1) with substitution Gly102Gln and one base deletion (deletion at nucleotide position 304) resulting in a null mutant due to generation of an in-frame stop codon (Val103Stop). Takahashi et al. (2010) reported that the lines with *Hinb-2* null mutation were harder as compared the lines that lacked this mutation. Higher average hardness index (HI) of 59.7 units was observed in lines with *Hinb-2b* allele as compared to lines with *Hinb-2a* allele (45.8 units). The lines containing *Hinb-2b* lacked HINB-2 protein in the seed as shown by protein studies by 2D-gel electrophoresis and N-terminal amino acid sequencing (Takahashi et al. 2010). This is the only HIN mutation reported so far that has been linked with direct increase in barley grain hardness. Mutations in the putative HINB-2 protein types were compared with mutations reported in PINB. Some substitutions reported in PINB were also found in HINB-2 (*Aegilops tauschii* and synthetic wheat) and were conserved in most of the HINB-2 protein types. These substitutions were Arg103Gly, Arg108Lys, Leu113Phe, and Glu119Asp, Leu124Ile (Massa et al. 2004; Lillemo et al. 2002; Gedye et al. 2004; Simeone et al. 2006) and might be the likely reason of harder grain texture of barley as compared to wheat.

### 1.4.15 *Hordindoline* promoters

The upstream regulatory sequences of *Pinb* genes were first identified by Digeon et al. (1999). Important features like TATA and CAAT boxes were identified in the -60 to -100 nt region. Lilemo et al. (2002) identified prolamin box (endosperm box) in both *Pina* and *Pinb* genes. The CAAT box of *Pina* was found to be inverted and other regulatory sequences i.e. NtBBF1, dyad repeats, DOF core recognition sequence (signal transduction/ tissue specificity) and Amybox 1 (response to gibberellin) were identified in both *Pin* genes. The *Pina*/*Pinb* promoters of soft wheat were found identical to *Pina*/*Pinb-D1b* promoters of hard wheat by Amoroso et al. (2004) and Simeone et al. (2006). Li et al. (2011) identified different types of regulatory elements for *Hina* gene in wild and landrace/ cultivated barleys. It was suggested that some regulatory elements of Hina gene i.e. MBS and Skn-1-motif may be highly related to grain hardness. It was also proposed that TATA box and other regulatory elements with different numbers and locations in untranslated region may affect degree of *Hina* gene expression.

In the present work, the upstream region of *Hina, Hinb-1 and Hinb-2* was accessed from barley genomic database, International Barley Sequencing Consortium (http://www.public.iastate.edu) for three cultivars i.e., Barke, Bowman and Morex. Regulatory elements were predicted by search using PLACE (http://www.dna.affrc.go.jp/PLACE/). No nucleotide variations were observed in *Hina* and *Hinb-2* genes sequences of three cultivars. However, two insertions were noted in the *Hinb-1* sequence of cv. Barke (T at -488 and A at -487) compared to that of other two cultivars. *Hina* sequence (morex_contig_413081, 710 nucleotides upstream start codon), was aligned with *Pina* sequence of *Triticum aestivum* cv. Chinese Spring [Genbank accession: DQ363911.1, Simeone et al. (2006), 849 nucleotides upstream start codon] to identify and compare important regulatory elements. Similarly, *Hinb-1* (morex_contig_9808, 1004 nucleotides upstream start codon) *and Hinb-2* (morex_contig_355704, 999 nucleotides upstream start codon) sequences of cv. Morex were aligned with *Pinb* sequence of *Triticum aestivum* variety PI 495916 [Genbank accession: AJ302100.1, Lillemo et al. (2002), 403 nucleotides upstream start codon]. *Hina/Hinb-1/Hinb-2* sequences of cv. Morex were used for locating regulatory elements (TATA box, CAAT box, Amy box, (CA)n, E box, ACGT box, P-binding site, DOF core

recognition sequence, NtBBF1, Prolamin box, AACACA box, Dyad repeats) and comparison with respective *Pin* gene sequences. Many regulatory elements were found to share common position in *Pin* and *Hin* genes indicating similarity in transcriptional regulation of these genes. These regulatory elements are shown in Fig. 1.9 (*Pina*), Fig. 1.10 (*Hina*), Fig. 1.11 (*Pinb*), Fig. 1.12 (*Hinb-1*) and Fig. 1.13 (*Hinb-2)*. The start codons of *Pin* and *Hin* sequences are highlighted in yellow. The regulatory elements are listed in Table 1.6.

> ***Pina-D1***

-849 CTTGTTAAAAAAAATTACTTCCTTGGAGACAAATACTATTTTGGAGAATAATGAAAATCATTTAGAAACA

-779 TTTGACATGAACGACATGATTCCAGAGACAATAAAAAACTTAAAAACAAACGATGGATAAGAAGATGATG

-709 TCTTGATAAATCATTACTTGTCGAGTTGTCGTGTACTACTAGTCTGTAAAATACAGTCTCTAACAAATTT

-639 GGTACAATCTGAGATCCATCTTGAACAACCTGCACAATCCTACAAGTTTAGTTTCGCAAAAGAATAACAA

-569 TATGAACCATGTGACTTTCTTGGTACGTACAAAACCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

-499 AGAGAAGGGTCACACATCATTAGATAAGGTCTATCTTCACAAAGTTGCACATACATGTACAGTACAGGAA

-429 CGACATGTATCTCAATACCACATGGTTCTAGATACTGGACGAAAAAGCAGTGGCTAGAAAGATGACGAT

-359 ATATAGATGCATTACTTATCATATACTACTACCTAGAAAAATACAATATCTAATTTCCTCTTGATCCTTC

-289 TTGAACAACCTGCACAACACTACAAGTTCAGTTTCACAAAAGCGTAAGTCTAAAGCTTTGGTACAACAAC

-219 AACTTATGGTTTATTTTGAGAAAAGGTCAGATTCAGTACACGGAACATCACATATCTCAACAACTTCCAC

-149 CAGTTTTGTGTGCTTTCAAAGTAACTTTGATTGGTATCCAGCTATACAACACACAACCGCACACAGAAAT

-79 CGTGCCACCTCAATTATAAATAAAGGTGTGGCCTCATCTCATCTATTCATCTCCACCTGCACCAAAACAC

-9 ACTGACAACATGAA

**Figure 1.9: Nucleotide sequence of *Pina* gene of cv. Chinese Spring**.
Regulatory elements are indicated: bold and wave underlined (**TATA-box**), wave underlined (CAAT box), bold double underlined (**Prolamin box**), double underlined $(CA)_n$ element, bold dash underlined: **Amy box**, dashed underlined (dyad repeats), bold in purple: **E box**, bold in blue: (**ACGT- element**), bold in green: (**P-binding site**), bold: **DOF core recognition sequence**

**>*Hina gene***

-710 AACCTCCTTCTCTAAAAAAAAAAATCATTACTTGTCTAGTTGTCGTGGACTACTAGTTTGTAAAACG**CAA**

-640 **T**CTCTAACAAATTTGGTACAGTCTGGGATCCATCTTGAACAACCTACGCAACCCTACAAGTTCAGTTTCG

-570 CAAAAGAATGA**CAAT**ATAAACCATGTGACTTCTTGGT**ACGT**A**CAAT**AACACGGAGAGAGAAGGGTCACGA

-500 TCATTAGATAAGGTCTATCTTCACAAAGTTGCACATACATCACATAAAATGACATGTCTCT**CAAT**ACCAC

-430 ATGGTTCTAGATACTGGACAAAAAAACAGTGGCTAGAAAGATGACGATACATAGATGCATTGCGTATCAT

-360 ATACTACTGCCTAGAAAAATA**CAAT**CCGTAATTTGCTCTGGATCCTTCTTGAACAACCTGCACAACACTA

-290 CAAGCCCGGTTTCACAAAATCATAAGCCTAAAGCTTTGCTACAGCAACGACTTATGGTTTATCTTGAGAA

-220 AAAGGTCTGATTCAGTACACGGGACATCACATATCTCTACAACTTCCACCAGTCCTGTGTGCTTTCAAAG

-150 TAACTTTGATTGGTATCCAGCTAAATTTGTAAGCATTAGCTGAAGCAGTGT**ACACA**ACTGCAGACAGAAA

-80 GCGTGCCACCTCATT**TATAAAT**TAAGGTGTGGCCTCATCTCATCTATTCATCTTCACCTGCACCAAAACA

-10 CATTGACAAC**ATG**AA

**Figure 1.10: Nucleotide sequence of *Hina* gene of cv. Morex.**
Regulatory elements are indicated: bold and wave underlined (**TATA-box**), bold (**CAAT box**), bold in blue: **ACGT box**, dotted underlined: Amy box, double underlined: E-box , wave underlined: DOF core recognition sequence, , bold in purple: **$(CA)_n$ element,** underlined: P binding site (AACAACC)

**> *Pinb gene***

-403 AACCACGGCTAGAAGAGGACGACATCTAGAGGCATTGCTTTTCATGTACTAATACCTTGTTA**AACACA**TT

-333 CTC**TAACAA**ATTGGTTTGGATCCTTCTTCAACAATTTCCACACACTACAAGGCCAGTTCACA**AAAG**CTT**A**

-263 **AAG**CGTGAGCAATTGGTACAAAACTAGTTGTGGTCTATCT**TGAGAAAAGG**GAACACTTAGTACACGAAAC

-193 GT**CACCTG**TCTCAACAACTTGCACCATTTCTGTTGGCTCGC**AAAG**TA**ACTTTA**TTTAGTATACCAACTTA

-123 ATTTGTGAGCATTAGCC**AAAG**C**AACACA**CAATGGTAGGCAAAAACCATGTCACTAAGCAA**TAAATAAA**GG

-53 GGAGCCTCAACCCATCTATTCATCTCCACCACCACCAAAACAACATTGAAAAC**ATG**AAGACCTTATTCCT

**Figure 1.11: Nucleotide sequence of *Pinb* gene**
Regulatory elements are indicated: bold and wave underlined (**TATA-box**), wave underlined (CAAT box), bold double underlined (**Prolamin box**), double underlined $(CA)_n$ element, bold in red: **Amy box**, bold : **AACACA box,** dashed underlined (dyad repeats), bold dash underlined: **E box**, bold dot underlined: **NtBBF1,** bold underlined: **DOF core recognition sequence**

**> *Hinb-1* gene**

-1004 TGTGGAAGTGTGTCTCTGGCGGACCT**CACGTG**ATTCGATCGATGGTTGTCTTTGATGGATCTACATGGATCTGATGTTTGT

-923 TCGTCTAATTTTGTGCGTCTTCAGGTCGGATCCTTTCAATATACTCATCTCTT**CATCTG**TGGTGGTTGTTGTTCTGTTGCA

-842 CTAGTCCTACGAGACCTTAACATGATGACTTTTGATTGTCTACTACAACAAGTTTTGCCTGGCTCCAGCG**AAAG**AGTGGCG

-761 ATGACGACGGCGCGCCTTTGGCTCGCTTCAATACTTGTAATAGTTGCTAGATGGCCTATGAATTTGAATATATTTTTTA**TT**

-680 **ATTT**CCGGTGTTCATTGTATTGTAATGATTAACGATGAATAGATCA**AAAG**TTTTCTCGTAAAAAAAACTCTACTGAAGCAG

-599 CAA**CATGTG**AGCTATCAGTTTTGCAGAATACAAGGTAACATCAATTTTACAACAAAATTAATTAGCAATAA**AAAG**GTGCAT

-518 GTACTCATTTCCCCAATATATATATATATATATTGGTAATTGAGTTTTATCTCCTTGTCCTAATAAACCCGAGCTAAGG

-437 GAGT**AAAG**ATCTGAACATTACATGATTCTAGATACAGATCAGAAAACCATGGTTGACATGATTCTAAATACAAAACAGAAA

-356 ACCACGGCTAGAAGAGGACGACATGTAGATGCATTGCTTCAAATTTGTTTGGATCCTTGTTCAACAATTTTCACAACACTA

-275 CGAGGCCAATCCACA**AAAG**CGT**AAAG**TGTCGACCTTTGGTATAACATACTTATCTTGAA**AAAG**GTTACATTTAGTACACGA

-194 AATGTCATCTATCTCAACAACTTGCACCATTTCTGTTTGCTTAT**AAAG**TA**ACTTTA**ATTAGTGCCAACTACGTTTGTAAGC

-113 ATTAGCC**AAAG**C**AACACA**CCATGGAACGCAAAAACCGTGTCACTTAGCAATAAATAAGAGGAAGTCTCAACCCATCTATTC

-32 ATCTCCACCAACACCAAAACAACGTTG**AAAG**C**ATG**AA

**Figure 1.12: Nucleotide sequence of *Hinb-1* gene** Regulatory elements are indicated: bold and wave underlined (**TATA-box**), wave underlined (CAAT box), double underlined $(CA)_n$ element, bold : **AACACA box,** dashed underlined (dyad repeats), bold dash underlined: **E box**, bold dot underlined: **NtBBF1,** bold underlined: **DOF core recognition sequence**

**>*Hinb-2* gene**

-999 ATTTGTGCTCCTTTAACAATTATAAAAATATACAGACTTGTTTAATACTTAACCGAG**AAAG**GGTTTCCCCTCGCTTTGCAT

-918 ATTTGCAACCAACATCGTGTAGAGACATTGCTAGGACGAAC**AACACA**TCAAGCCCAAAA**AAAG**AAAAATGAA**CAAATG**CCA

-837 ACAGCGTTAGCTCGAA**AAAG**AGCGGAAGATCCGCCATCGTTGCGCCTTCCGGAGATAAACCACCACAGCCCGAGGCTTCGA

-756 TCCGATGCGTGCCAAGCAGCACCTCCAAGAAGGGATGCGACGTCAACGACGCTGCTGCCGGAGCCCAACCGGAGCCGACCA

-675 GCCAAACCACCACTTAACACCATACGCTAGCGTAGTTGCGCCACCATCCATGCAGTCTCACTCCAAACACCAACACGAGGC

-594 CAAGGAGACCGGAGAGCTGCGCCAAACAGATTAATTAGCAATAA**AAAG**GTGCATGCACTCATTTCCCCAAAAAATATATAT

-513 TGGTAATTGAGTTTTCTCTCTTTGTCCTAATAAACCCGAGCTAAGGGAGT**AAAG**ATCTGAACATTACATGATTCTAGATAC

-432 AGAACAGAAAACCATGGCTGACATGATTCTAGATACA**AAAG**AGAAACCCACGGCTAGAAGAGGATGACATGTAGATGCATT

-351 GCTTTAAATTTGTTTGGATCCTTGTTCAACAATTTGCACAACACTACGTGGCCAGTTCACA**AAAG**CGT**AAAG**TGTCGACCT

-270 TTGGTACAACATACTATTTCTGGCCGTCTATCTTGAA**AAAG**GTTACATTTAGTACACGAAATGTCATCTATCTCAACAACT

-189 TGCACCATTTCTGTTTGCTTGT**AAAG**TA**ACTTTA**ATTAGTGCCAACTAAGTTTGTAAGCATTAGCC**AAAG**A**AAAG**CATTAG

-108 CC**AAAG**CAGCACACAATGGAACGCAAAAACCGTGTCACTTAGCAATAAATAATAGGAAGTCTCAACCCATCTATTCATCTT

-27 CACCAACACCAAATAAACATTG**AAAG**CATGAA

**Figure 1.13: Nucleotide sequence of *Hinb-2* gene**

Regulatory elements are indicated: bold and wave underlined (**TATA-box**), wave underlined (CAAT box), double underlined $(CA)_n$ element, bold : **AACACA box,** dashed underlined dyad repeats), bold dash underlined: **E box**, bold dot underlined: **NtBBF1,** bold underlined: **DOF core recognition sequence**

**Table 1.6: List of important regulatory elements compared in *Pin* and *Hin* genes**

| Regulatory element | Sequence | Position in *Pina* (*T. aestivum* cv. Chinese spring*) | Position in *Hina* (cv. Morex) | Position in *Pinb* (*T.aestivum* variety PI 495916**) | Position in *Hinb-1* (cv. Morex) | Position in *Hinb-2* (cv. Morex) | Function |
|---|---|---|---|---|---|---|---|
| **TATA box** | TATAAAT | -63 | -65 | -56 | Not found | Not found | Transcription regulation |
| **TATA box** | TTATTT | -209 | Not found | -143 | -682 | Not found | |
| **CAAT box** | CAAT | -69,-316,-416,-572,-605, -635,-751 | -339, -439, -528, -559,-643 | -97 | Not found | -97 | |
| **CAAT box** | CACAAT | Not found | Not found | -66, -95, -254, -302 | -66, -269, -292, -505, -534, -558, -733, -887 | -66, -95, -322, -556, -983 | |
| **Amy box** | TAACA(G/A)A | -649 | -639 | -330 | Not found | -986 | Response to gibberellin |
| **$(CA)_n$** | CAACAAC | -102, -226 | -99 | -182 | -182 | -197 | Seed specificity |
| **E box** | CANNTG | -26, -410, -562, -781 | -26,-96,-433, -549 | -191 | -596,-870,-978 | -846 | Tissue specificity, developmental control |
| **ACGT box** | ACGT | -545 | -533 | Not found | Not found | Not found | |
| **P-binding site** | AACAACC | -616 | -603 | Not found | Not found | Not found | Recognition sequence for P protein |
| **DOF core recognition sequence** | AAAG | -58,-132,-198,-238, -251, -371, -385, -459, -584 | -83,-154,-220, -261,-394,-476 | -106,-152,-264,-271 | -5,-106,-150,-216,-253, -260, -433, -528, -634, -772 | -5,-106,-118,-123,-167, -233,-283,-290,-395, -463, -550, -821, -859 -942 | Light regulation |
| **NtBBF1** | ACTTA | Not found | Not found | -146 | -146 | -161 | Tissue specific expression |
| **Prolamin box** | TG(T/A/C)AA A(A/G)(G/T) | -223 | Not found | -223 | Not found | Not found | Endosperm specificity |
| **AACACA box** | AACACA | Not found | Not found | -101,-341 | -101 | -877 | Seed specificity |
| **Dyad repeats** | $CAA(N)_{2-9}TTG$ | Not found | Not found | -18, -327 | -13,-317 | -18,-85,-325,-556 | Endosperm and aleurone specificity |

### 1.4.16 Expression of *Hin* genes

RNA from developing barley grain was shown to hybridise with oat *Puroindoline* (tryptophanin) cDNA (Tanchak et al. 1996). Darlington et al. (2001) reported the expression of *Hordoindoline b* mRNA in the starch endosperm and aleurone layer of the developing seed, but not in the embryo. The *Hordoindoline* b mRNA level was detected at 14 DAA (days after anthesis), peak at 20 DAA and declined at 30 DAA. They suggested that *Hinb* expression levels appeared to be higher in endosperm than in aleurone layer at 14 and 20 DAA based on *Hinb* transcript level using northern blot analysis. *Hordoindoline a* also showed similar expression. Beecher et al. (2001) performed RNA expression analysis on four barley cultivars (two malt-type and two feed-type). Sequences hybridising to wheat *Pina* and *Pinb* using Northern-blot analysis were detected in developing barley grain. However, no variation was detected for *Hina*/*Hinb* transcript level in the four cultivars analysed.

Lee et al. (2006) reported predominant expression of *Hordoindolines* in the aleurone layer during late kernel development, i.e., 20 days after flowering (DAA). Two cDNAs encoding *Hordoindolines* (*Hina,* Genbank accession: AY959939 and *Hinb,* Genbank accession: AY959940) were isolated and studied for molecular characteristics. Tissue specific expression of *Hin* genes was examined in grains, pericarp, stems and leaves at 14DAA. Transcripts for *Hina* were detected in grains, stem and leaves but not in the pericarp while *Hinb* gene expression was detected only in grains and not in other tissues. Northern blot analysis performed in grains from 5 DAA to 20 DAA indicated that the expression of *Hina* gene was detected at 5DAF, peaked at 8 DAA, and decreased slightly until 20 DAF, while the expression of *Hinb* began to be detected at 8 DAF, and decreased slightly until 20 DAA.

A search for Expressed Sequence Tags (ESTs) corresponding to Hordoindolines was performed using NCBI UNIGENE. The search showed two ESTS that were analysed for tissue type they were reported from and were grouped as follows; spike, seed, pericarp, pistil, stem, mixed and other. The wide range of tissues (Table 1.7) indicated the great diversity of the localisation of *Hordoindolines*. Since, these genes have been linked with antimicrobial properties their presence in various tissues is favourable from plant defense point of view. A search for Hordoindoline ESTs in Genbank revealed 12 sequences (belonging to Unigene Hv.20767) reported to be expressed in developing endosperm (Table 1.8).

**Table 1.7 Tissue types of ESTs (UNIGENE) representing barley Hordoindoline**

| No. | *Hin* gene | Number of ESTs | | | | | | | | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Spike | Seed | Pericarp | Pistil | Stem | Mixed | Other | Not specified | |
| 1 | Hv.7370 | 52 | 33 | 6 | 3 | 1 | 18 | 212 | 2 | 327 |
| 2 | Hv.20767 | 20 | 32 | | | | 1 | 69 | | 122 |

**Table 1.8 Tissue types of ESTs (Genbank) representing barley Hordoindoline**

| No. | *Hin* gene (Unigene Hv.20767) | Tissue | Development stage |
|---|---|---|---|
| 1 | CV063946.1 | endosperm | developing endosperm tissue 10, 12, 15 DAA |
| 2 | CV063888 | endosperm | developing endosperm tissue 10, 12, 15 DAA |
| 3 | CV063049.1 | endosperm | developing endosperm tissue 10, 12, 15 DAA |
| 4 | CV060605.1 | endosperm | developing endosperm tissue 10, 12, 15 DAA |
| 5 | CV060058.1 | endosperm | developing endosperm tissue 10, 12, 15 DAA |
| 6 | CV058293.1 | endosperm | developing endosperm tissue 10, 12, 15 DAA |
| 7 | CV057902.1 | endosperm | developing endosperm tissue 10, 12, 15 DAA |
| 8 | CV057160.1 | endosperm | developing endosperm tissue 10, 12, 15 DAA |
| 9 | CV055922.1 | endosperm | developing endosperm tissue 10, 12, 15 DAA |
| 10 | CV055825.1 | endosperm | developing endosperm tissue 10, 12, 15 DAA |
| 11 | CV055825.1 | endosperm | developing endosperm tissue 10, 12, 15 DAA |
| 12 | CV054101.1 | endosperm | developing endosperm tissue 10, 12, 15 DAA |

## 1.5 Conclusions

The literature establishes that *Puroindoline* genes are the major determinants of wheat grain texture. However, the reason for difference in the grain hardness of genotypes of same *Pin* class is not known. The *Pinb-2* genes could be one of the minor genetic factors involved in influencing grain hardness. *Puroindoline b-2* genes have been linked with minor grain texture variations. Similarly, allelic variation has been reported in *Hin* genes (barley orthologues of *Pins*) but lacks a very clear relationship with barley grain hardness. Hence, there is a need to further investigate the *Pinb-2* genes in wheat and *Hin* genes in barley. Such research will allow the identification of potential genetic resources for crop improvement. Thus, this study aimed to test the following hypotheses:

(i) Certain genetic variations in *Pinb-2* alleles have minor effects on wheat grain hardness

(ii) Certain genetic variations in *Hin* genes have minor effects on barley grain hardness.

## 1.6 Aims of the project

The above stated hypotheses were addressed through following specific aims:

1. To investigate the effect of any *Pinb-2* variations on grain hardness within same *Pin* genotype class in Australian wheat cultivars and worldwide landraces
2. To investigate the association of *Pinb-2v3* allelic variation with grain hardness
3. To study protein-protein interactions occurring between PINA, PINB, and PINB-2v3 that may have effect on grain hardness
4. To study the gene expression of *Pinb-2* and its association with grain hardness
5. To identify *Hin* alleles in barley landraces and Australian cultivars that can be used as potential donors in crop breeding programs to enhance the gene diversity

# CHAPTER 2

# Material and methods

# 2 Material and methods

## 2.1 Equipment and material

The instruments and apparatus used to conduct experiments are listed in Table 2.1. The commercial kits and solutions are listed in Table 2.2

**Table 2.1 List of equipment used in the project**

| Equipment | Manufacturer | Purpose |
|---|---|---|
| Plant growth cabinet with light, temperature and humidity control | Thermoline, Victoria, Australia | Growth of wheat and barley seedlings |
| MyCycler™ | Bio-Rad, California, USA | PCR |
| Electrophoresis power supply- EPS301<br>Minnie Gel Unit | General Electric (GE) Healthcare, Buckinghamshire, UK | Electrophoresis |
| Bio-Rad Chemidoc XRS documentation station | Bio-Rad, California, USA | Capturing of gel images |
| NanoDrop 2000 UV-Vis Spectrophotometer | Thermo Fisher Scientific Australia | RNA and DNA quantification |
| UV light transilluminator | UVP, USA | Visualisation of gels for gel purification |
| Finnpipette (0.5-10, 5-50, 20-200, and 100-1000 µL) | Thermo Electron, Madison, USA | Dispensing liquids |
| Sorvall RC6 | Centrifugation | Sorvall (part of Thermo Scientific, Waltham, USA) |
| Mini spin plus | | Eppendorf, Hamburg, Germany |
| Heλious ε Spectrophotometer | β-galactosidase assay $OD_{600}$ cell cultures | Thermo Scientific, Waltham, USA |
| Orbital shaker/incubator | Bacterial and fungi broth culture | Ratek, Victoria, Australia |

**Table 2.2 Commercial kits and reagents used during this project**

| Kit/material/solution | Supplier | Use |
|---|---|---|
| Restriction endonuclease, Bsrb1 | New England Biolabs (NEB), Genesearch, Australia | Identification of *Pinb-D1b* allele |
| Wizard® genomic DNA purification Kit (contains nuclei lysis solution, protein precipitation solution, DNA rehydration solution) | Promega, Alexandria, Australia | Isolation and purification of genomic DNA |
| RNase A (from bovine pancreas) | Sigma-Aldrich, St Louis, USA | Purification of genomic DNA |

| | | |
|---|---|---|
| Wizard® SV Plasmid DNA Miniprep kit (contains cell resuspension solution, cell lysis solution, neutralization solution, column wash solution, minicolumns, alkaline protease solution | Promega, Alexandria, Australia | Purification of plasmid DNA |
| Wizard® SV Gel and PCR Clean -Up kit (contains wash solution, binding solution, nuclease free water, minicolumns, collection tubes) | Promega, Alexandria, Australia | Purification of PCR products |
| pGEM-T Easy Vector System (contains pGEM-T Easy Vector*, T4 DNA ligase., 2X ligation buffer) | Promega, Alexandria, Australia | Cloning of PCR products |
| RQ1 RNase-free DNase I | Promega, Madison, USA | DNA digestion during RNA purification |
| RQ1 DNase 10× Reaction Buffer | | |
| TRIsure™ | Bioline, Alexendria, Australia | Total RNA isolation |
| RNase Inhibitor | | Inhibition of RNase activity |
| Bioscript™ Moloney Murine Leukaemia Virus Reverse Transcriptase | | Reverse transcription |
| Biomix™ 2X | | DNA amplifications by PCR |
| dNTP set | | cDNA synthesis |
| Hyperladder 1 (200-10,000 bp) | | Molecular weight markers for agarose gel electrophoresis |
| GeneRuler™ 100bp DNA ladder; | Fermentas, Waltham, USA | |
| Ethidium bromide (10 mg/mL) | Sigma | Visualisation of nucleic acids on agarose gels |
| BDT (Big Dye Terminator) v3.1 Ready Mix Applied | Applied Biosystems, Australia | DNA sequencing |

* The vector map is provided in Appendix II

## 2.2 Prepared solutions and materials

### 2.2.1 Buffers and solutions

The buffers and solutions listed in Table 2.3 were prepared using established protocols (Sambrook et al. 2001). All buffers and solutions in Table 2.3 and Table 2.4 were prepared using Milli-Q water (Millipore) unless where indicated. Sterilisation was carried out by autoclaving (121°C for 20 minutes) or filtering through 0.22 µm MF-Millipore syringe filter (Millipore, Madison, USA).

**Table 2.3 Composition of general buffers and solutions**

| Buffer/solution | Composition | Sterilization method |
|---|---|---|
| 6x loading dye | 0.25% Bromophenol blue (w/v), 0.25% Xylene Cyanol FF (w/v), 40% Sucrose (w/v) | autoclaved |
| Glacial acetic acid | 0.01% (w/v) | autoclaved |
| 70 % ethanol | 30mL ethanol, 70 mL $dH_2O$ | - |
| Chloroform/isoamyl alcohol | Chloroform: isoamyl alcohal: 24:1 (v/v) | - |
| 50 x TAE buffer (pH:8) | 40 mM Tris-acetate, 1mM EDTA (pH:8) | autoclaved |
| TB buffer | (10 mM Hepes, 15 mM CaCl2, 250 mM KCl, pH 6.7)*, then solid MnCl2 later added to final concentration of 55 mM | *autoclaved |

The solutions in Table 2.4 required for DNA sequencing were prepared according to instructions by AGRF (Australian Genome Research Facility Ltd, Melbourne, Australia (http://www.agrf.org.au/assets/files/PDF%20Documents/Guide%20to%20AGRF%20Sequencing%20Service.pdf).

**Table 2.4 Composition of buffers and solutions for sequencing**

| Buffer/solution | Composition | Sterilization method |
|---|---|---|
| 5x BDT buffer | 400 mM Tris pH 9.0, 10 mM $MgCl_2$ | autoclaved |
| 0.2 mM MgSO4 | 2µL 1M $MgSO_4$ in 10 mL of 70% ethanol | autoclaved |

Buffers and solutions required for yeast two-hybrid work were prepared per the Yeast Handbook (Clontech, CA, USA; http://www.clontech .com/xxclt_ibcGetAttachment.jsp?cItemId=17602, last accessed June, 2016) and are listed in Table 2.5.

**Table 2.5 Buffers and solutions for yeast two-hybrid work**

| Buffer/solution | Composition | Sterilization method |
|---|---|---|
| Herring testes carrier DNA | 10 mg/mL; denatured with 20 min boiling immediately prior to use | - |
| 10× Lithium Acetate (10× LiAc) | 1 M lithium acetate, pH 7.5 | autoclaved |
| 10× TE buffer | 0.1 M Tris-HCl and 10 mM EDTA, pH 7.5 | autoclaved |
| 50% PEG (polyethylene glycol) 3350 | 50% (w/v). Dissolved with gentle heating | filter sterilized |
| 1× LiAc/TE | 1 mL 10× LiAc solution, 1 mL 10× TE buffer, made up to 10 mL | - |
| 1× PEG/LiAc solution | 1 mL 10× LiAc solution, 1 mL 10× TE buffer, made up to 10 mL with sterile 50% PEG 3350 | - |
| 100% DMSO | Analytical grade dimethyl sulfoxide | - |
| 20% SDS | 20% (w/v) sodium dodecyl sulphate | - |
| Z buffer | Na2HPO4 • 7H2O 16.1 g/L, NaH2PO4 • H2O 5.50 g/L, KCl 0.75 g/L, MgSO4 • 7H2O 0.246 g/L dissolved in water, pH 7.0 | autoclaved |
| Z buffer with β-mercaptoethanol | 0.27 ml of β-mercaptoethanol added per 100 mL of Z buffer. | - |
| ONPG (o-nitrophenyl β-D-galactopyranoside) | 4 mg/mL ONPG in Z buffer, pH 7.0. | - |
| $Na_2CO_3$ | 1M | autoclaved |

### 2.2.2 Media and solutions for microbial growth

The media and solutions used for culturing bacteria were prepared according to Sambrook and Russell (2001) and listed in Table 2.6. Solutions and media required for yeast two-hybrid work were prepared as per the Yeast Handbook (Clontech, CA, USA; http://www.clontech .com/xxclt_ibcGetAttachment.jsp? cItemId=17602, last accessed June, 2016) and are listed in Table 2.7.

**Table 2.6 Solutions and media used for culturing bacteria**

| Solution | Composition | Sterilization method |
|---|---|---|
| 0.1 M IPTG | 1.2 g (Isopropylthio-β-D galactoside) in 50 ml $dH_2O$ | filter- sterilised |
| 20 mg/mL Ampicillin | 20 mg in $dH_2O$ in 1 mL | filter- sterilized |
| 50 mg/mL Kanamycin | 20 mg in dH2O* in 1 mL | filter- sterilized |
| 5% (w/v) X-gal | 5% (w/v) in DMSO (dimethyl sulfoxide) | |
| Luria Bertani broth (LB) | 10g/L tryptone, 5g/L yeast extract, 5g NaCl, 15g/L agar[#] | autoclaved |
| Super Optimal broth (SOB medium) | 0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2 , 10 mM MgSO4 | autoclaved |
| Super Optimal broth with Catabolic repressor (SOC) medium | 0.5 % yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10mM $MgSO_4$, 20mM glucose | autoclaved |
| 2×YT | 16 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl, pH 7.0 | autoclaved |

* $dH_2O$ – sterile distilled water
# agar- 15 g/L of agar was added only for preparation of LB plates and not for liquid media

**Table 2.7. Solutions and media for yeast two-hybrid work**

| Solution and media | Composition | Sterilization method |
|---|---|---|
| Dextrose | 40% w/v, dissolved with stirring and gentle heating in water* and added to YPD or YSD when required | *autoclaved |
| Yeast Peptone Dextrose (YPD) | 20 g/L Difco peptone, 10 g/L yeast extract, 20 g/L agar (for plates only), 30 mg/L adenine hemisulfate* and dextrose added to a final concentration of 2% | *autoclaved |
| Yeast Selective Defined (YSD) | 6.7 g/L of yeast nitrogen base without amino acids (BD, New Jersey, USA), 20 g agar (for plates only),* dextrose added to a final concentration of 2%, and 100 mL of the appropriate sterile 10× dropout solution (DO) per liter | *autoclaved |
| 10× Dropouts (10× DO) | prepared as media lacking Trp, Leu, Ade and His (-WLAH), or as -WL, or -L or -W. 10× dropouts contain 200 mg/L L-adenine | |

| | | |
|---|---|---|
| | hemisulfate salt, 200 mg/L L-arginine HCl, 200 mg/L L-histidine HCl monohydrate, 300 mg/L L-isoleucine, 1000 mg/L L-leucine, 300 mg/L L-lysine HCl, 200 mg/L L-methionine, 500 mg/L L-phenylalanine, 2000 mg/L L-threonine, 200 mg/L L-tryptophan, 300 mg/L L-tyrosine, 200 mg/L L-uracil, 1500 mg/L L-valine. | autoclaved |
| 3-amino-1,2,4-triazole (3-AT) | stored at a concentration of 1 M* and added to 1mM after cooling of the solid medium to ~55°C | *Filter sterilized |

### 2.2.3 Microbial strains and cloning vectors

*Escherichia coli* JM109, was used in standard cloning procedures. *Saccharomyces cereviceae* AH109 and Y187α strains were used for yeast two-hybrid experiments, and were kindly provided by Dr. Rebecca Alfred (Swinburne University of Technology, Melbourne, Australia). The vector pGEM®-T Easy (Promega) was used for cloning of PCR products. Two vectors i.e. pGBK-T7 (Clontech, Appendix III) and pGAD-T7 (Clontech, Appendix IV) were used for cloning 'bait' and 'prey' genes respectively.

### 2.2.4 Plant material and its propagation

The plant material used in this study is listed in Table 2.8, Table 2.9 (wheat) and Table 2.10, Table 2.11 (barley). The seeds of wheat (cultivars and landraces) and barley (landraces) were multiplied at, and provided by, the Australian Winter Cereal Collection (AWCC; Tamworth, NSW, Australia) from certified seed stock, and were typically of sufficient quantity (>5g) for the Single Kernel Characterisation System (SKCS) testing of grain hardness (described below). The seeds of barley cultivars were provided by Dr Joseph Panozzo (Victorian Government, Department of Economic Development, Jobs, Transport and Resources; Horsham, Victoria) and were also of sufficient quantity. These seeds were obtained from different field trials conducted at Horsham (Victoria, Australia) in 2014 (Dr Joseph Panozzo, personal communication). All wheat cultivars and landraces belong to Triticum aestivum (genome AABBDD). All barley cultivars and landraces belong to *Hordeum vulgare* (genome HH). The pedigrees of wheat and barley cultivars were obtained from literature (Park et al. 2009, Pickering et al. 2007,) (Tables 2.8, 2.10). Healthy seeds were soaked overnight on wet filter paper in petri plates. The seeds were transplanted in to trays containing soil and grown in plant growth cabinet with conditions

of 16 hours light and 8 hours dark, 24 °C temperature, 80% humidity. Leaves from single seeds were harvested after 2-3 weeks, snap frozen in liquid nitrogen and stored in -80 °C.

**Table 2.8 Australian wheat cultivars used in this project**

| S.no. | Cultivar | AWCC No. | Pedigree | Year seed increased at Tamworth |
|---|---|---|---|---|
| 1 | Amery | 25598 | Lr21-SrX/2*Shortim//3*Bodallin | 1994 |
| 2 | Batavia | 25271 | Brochis(Sib)/Banks | 1992 |
| 3 | Carnamah | 27194 | Bolsena-1-CH(RAC529)/3/(77W660)Siete-Cerros/XBVT223//AWX-011-G-48-2/XBVT221 | 2010 |
| 4 | CD87 | 91169 | Not known | 2003 |
| 5 | Chara | 30031 | CD-87/3/(BD-225)Millewa/2*Cook//TM-56 | 2010 |
| 6 | Clearfield Stilleto | 25923 | Stiletto*3//Spear/Fidel | 2011 |
| 7 | Cranbrook | 11612 | Wren,Mex//Ciano-67(Sib) /Noroeste-66/3/Zambezi | 2014 |
| 8 | Cunningham | 25046 | 3-Ag-3/4*Condor//Cook | 2013 |
| 9 | Diaz | 23326 | Combination-III/3*Oxley//3*Cook | 2014 |
| 10 | Egret | 16037 | Not known | 1974 |
| 11 | Falcon | 90252 | Bencubbin//Dundee/Gular/3/Gular | 1972 |
| 12 | Halbred | 99124 | Scimitar/Kenya C6042//Bobin/3/Insignia-49 | 1983 |
| 13 | Hartog | 21533 | Vicam 71//Ciano-67(Sib)/Siete-Cerros-66/3/ Kalyansona/Bluebird | 2014 |
| 14 | Janz | 24794 | 3-AG-3/4*Condor//Cook | 2013 |
| 15 | Katepwa | 24314 | Neepawa*6/RL-2938/3/Neepawa*6//CI-8154/2*Frocor | 2010 |
| 16 | Kukri | 29472 | ((DRP((FNK58xN10B/Gb55)NAI60)/(TOB-CNO 'S' x TOB8156/CALxBb-CNO)/2/MDN/6*RAC177 | 2011 |
| 17 | Machete | 23038 | MEC-3/2*Gabo(RAC-177)//Madden | 2014 |
| 18 | Ouyen | 25571 | Takari/TM-56//Cocamba | 2014 |
| 19 | Rosella | 25923 | Farro-Lunga/Heron//2*Condor/3/Quarrion (Sib) | 2011 |
| 20 | Sunco | 23455 | SUN9E-27*4/3Ag14//WW15/3/3*Cook | 2014 |
| 21 | Sunstar | 22660 | Hartog*4//Cook*5/VPM1 | 2013 |
| 22 | Tasman | 25557 | Torres/3/Gaboto/Siete-Cerros-66//Bluebird/Ciano-67 | 1994 |

**Table 2.9 Worldwide wheat landraces used in this project**

| S.no. | Landrace Aus number (from AWCC) | Origin | Year seed increased at Tamworth |
|---|---|---|---|
| 1 | L28266 | Abyssinia | 1997 |
| 2 | L27212 | Afghanistan | 1996 |
| 3 | L27287 | Afghanistan | 1996 |
| 4 | L27216 | Afghanistan | 2001 |
| 5 | L27295 | Algeria | 1996 |
| 6 | L27356 | Burma | 1996 |
| 7 | L27439 | China | 1996 |
| 8 | L27441 | China | 1996 |
| 9 | L27442 | China | 1996 |
| 10 | L27440 | China | 1996 |
| 11 | L27414 | China | 1996 |
| 12 | L27415 | China | 1996 |
| 13 | L27416 | China | 1996 |
| 14 | L28076 | India | 1997 |
| 15 | L28112 | Iraq | 1997 |
| 16 | L27892 | Leon (Spain) | 1997 |
| 17 | L27975 | Rumania | 1997 |
| 18 | L28013 | Turkestan | 1997 |

**Table 2.10 Australian barley cultivars used in this project**

| S.no. | Cultivar (Malt/Feed) | Pedigree | Year of field trials at Horsham |
|---|---|---|---|
| 1 | Arapiles (Malt) | Noyep/Proctor/ICI3576/Union/4/Kenia/3/ Research/2/Noyep/Proctor/5/Domen | 2014 |
| 2 | Arivat (Feed) | | 2014 |
| 3 | Barque (Feed) | Triumph/Galleon | 2014 |
| 4 | Baudin (Malt) | Stirling/Franklin | 2014 |
| 5 | Buloke (Malt) | Franklin/VB9104//VB9104 | 2014 |
| 6 | Capstan (Feed) | Waveney/WI2875//Chariot/Chebec | 2014 |
| 7 | Franklin (Malt) | Shannon/Triumph | 2014 |
| 8 | Hamelin (Malt) | Stirling/Harrington | 2014 |
| 9 | Keel (Feed) | CPI18197/Clipper//WI2645 | 2014 |
| 10 | Sloop SA (Malt) | CCN6-3/Sloop3 | 2014 |
| 11 | Stirling (Malt) | Dampier/Prior/Ymer/3/Piroline | 2014 |
| 12 | Vlamingh (Malt) | WABAR0570/TR118 | 2014 |

**Table 2.11 Worldwide barley landraces used in this project**

| S.no. | Landrace Aus number (from AWCC) | Origin | Year seed increased at Tamworth |
|---|---|---|---|
| 1 | L400097 | Ethiopia | 1987 |
| 2 | L400211 | Ethiopia | 1984 |
| 3 | L400619 | North Korea | 1984 |
| 4 | L402433 | Afghanistan | 1982 |
| 5 | L402448 | Iran | 1982 |
| 6 | L403089 | Russia | 1982 |
| 7 | L403156 | China | 1984 |
| 8 | L411144 | Morocco | 2008 |
| 9 | L411145 | Morocco | 2008 |
| 10 | L411152 | Morocco | 2008 |
| 11 | L411859 | Nepal | 2010 |
| 12 | L411813 | Kyrgystan | 2011 |
| 13 | L411890 | Nepal | 2010 |
| 14 | L491320 | Japan | 2004 |

## 2.3 General molecular methods

### 2.3.1 Genomic DNA isolation and quantification

DNA isolation was conducted from leaves of single seedlings using the Wizard genomic DNA purification kit (http://www.promega.com/resources/protocols/technical-manuals/0/wizard-genomic%20dna-purification-kit-protocol/). About 40 mg of leaf tissue, kept frozen in liquid nitrogen, was crushed in 1.5 mL microcentrifuge tube using a pestle. The fine powder was mixed with 600 µL of nuclei lysis solution and incubated at 65° C for 15 minutes. RNAse A solution (3 µl) was added and incubated at 37 °C for 30 min. Then 200 µL of protein precipitation solution was added to the lysate, followed by centrifugation at maximum 14.5 x 1000 r.p.m for 5 minutes. The supernatant was mixed with 600uL isopropanol and centrifuged using same conditions. The DNA pellet was washed with 600 µL of 70% ethanol and dissolved in 100 µL of DNA rehydration solution and stored at -20°C. Single seed genomic DNA extraction was also performed for some of the landraces and cultivars, using a method modified from Chao and Somers (http://maswheat.ucdavis.edu/protocols/general_protocols/DNA_extraction_003.htm) and Mohammadi et al. (2012), kindly provided by Dr Rebecca Alfred (Swinburne University of Technology, Victoria, Australia). An aliquot of 5 µL was electrophoresed on 1% gel to determine the DNA quality by looking for any degradation. The quantity of

DNA was also estimated using spectrophotometer. The absorbance recorded at 260nm and 280 nm. An absorbance unit of 1 at $A_{260}$ corresponds to 50 µg/mL of double stranded DNA, and DNA to protein absorbance ratio ($A_{260}/A_{280}$) of~ 1.8 indicated high purity (Sambrook and Russell, 2001; pp 5.4-5.17).

### 2.3.2 Agarose gel electrophoresis of DNA and RNA samples

Agarose gel electrophoresis was used to quantify DNA samples and determine the size of amplified genes. DNA quality was also checked by examining the sample integrity and looking for any smear if present (as indicative of degraded DNA). Generally, 5 µL of DNA and 4 µL of RNA samples were mixed with 1 µL of 6× loading dye. The typical gel strengths were 1.0- 1.5 % (w/v), prepared in 1X TAE buffer. 0.5 µg/mL of ethidium bromide was added to the gel solutions before setting the gels, to visualise the DNA over UV light. The gels were typically run at 100V for 40-60 minutes. A quantitative DNA molecular weight marker, Hyperladder I (Bioline) or GeneRuler™ 100bp DNA ladder (Fermentas) was used for determining the size of bands. The concentration of DNA (ng) is known for each band/fragment of marker. The intensity of bands obtained from unknown DNA samples was compared with the intensity of marker fragments containing known amount of DNA for a rough estimate of DNA quantity. Gel documentation system was used to capture gel images.

## 2.4 The Polymerase Chain Reaction (PCR)

### 2.4.1 Design and synthesis of primers

The primers used to amplify *Pina-D1* (Genbank: DQ363911) and *Pinb-D1* (Genbank: DQ363913) genes were designed as reported by Gautier et al. (1994). *Pinb-2* primers (Genbank: AM944731, AM944733) were based on sequences reported by Wilkinson et al. (2008). These primers amplify full length genes i.e. the coding sequence and the start and stop codons are shown in Table 2.12. Two pairs of *Pinb-2* primers were used. The D1 primer pair was used to amplify *Pinb-2* variants 1, 3, 5 whereas D2 primer pair was used to amplify *Pinb-2* variants 2, 3, 4, 6. *Pinb-2* variant 1 (v1), v2, v3 and v4 specific primers, were designed as per Chen et al. (2010 a). Each variant has a unique nucleotide;

hence the forward/reverse primers are designed to start or end at that site. Vector specific primers i.e., T7 and SP6 were used for cloning *Pinb-2* genes. These primers are listed in Table 2.12. The primers used to amplify *Hina* (Genbank: AY644174), *Hinb-1* (Genbank: AY644090) and *Hinb-2* (Genbank: AY644090) genes were designed as reported by Takahashi et al. (2010). These primers amplify full length genes i.e. the full coding sequence and are listed in Table 2.13. For directional cloning into the yeast two-hybrid expression vectors pGBK-T7(bait) and pGAD-T7 (prey) (described below), the primers were designed to amplify the gene sections encoding mature proteins (i.e., without the putative signal peptides) in the correct reading frame. The restriction site for EcoRI was incorporated at the 5' end of the forward primers and BamHI at the 5' end of reverse primers (Table 2.14). Tri-adenine spacers were incorporated into the 5`end of all the primer sequences to enhance restriction enzyme activity. The primer sequence for housekeeping gene actin were designed to span introns as per Genbank: KC775780.1 (wheat actin CDS sequence) and the gene specific primers were designed as per Chen et al. (2013). These primers were used for testing expression patterns of *Pinb-2* genes using semi-quantitative reverse transcriptase PCR (sqRT-PCR) and are listed in Table 2.15. The primers were synthesized commercially (Invitrogen, Australia) and supplied as dried pellets. Primers were resuspended in MilliQ water to a concentration of 100 µM (stock solution) and diluted to final concentration of 10 µM for working primers.

### 2.4.2 Semi-quantitative reverse transcriptase-PCR (sqRT-PCR)

Gene expression analysis was conducted using this technique. Three biological replicates, i.e., RNAs from three plants were used to compare the changes in expression of *Pinb-2* genes. The actin primers were used as housekeeping genes (Suprunova et al. 2004). The total reaction volume was made to 50 µL containing 1 µL of synthesised first strand cDNA as template, 25 µL Biomix (Bioline), and 1 µL (of the 10 µM working solution) each of the forward and reverse primer of a pair. Each gene was amplified for 20, 25, 30, 35 cycles to determine the cycle number where the PCR is still in the exponential phase. The sqRT-PCR (5 µL) products were electrophoresed on agarose gels. The intensity of bands was recorded using the ChemiDoc$^{TM}$ XRS+ Documentation Station which uses Image lab v 5.1 (Bio-Rad). The relative expression of the *Pinb-2*

variant genes was determined by comparing band intensities to those of actin using the following formula (adapted from Jang et al.2004)

$$\text{Relative gene expression} = \frac{\text{Band intensity of } \textit{Pinb-2} \text{ variant}}{\text{Band intensity of } \textit{actin}}$$

**Table 2.12: Primer pairs used for amplification of *Pina*, *Pinb*, *Pinb-2* genes and cloning of *Pinb-2* genes**

| Primer | Sequence[a] | Annealing temperature[b] | Expected product and size[c] |
|---|---|---|---|
| Pina-D1F | 5' **ATG**AAGGCCCTCTTCCTCATAGG 3' | 54°C | *Pina-D1*; 447 bp (Genbank: DQ363911) |
| Pina-D1R | 5' **TCA**CCAGTAATAGCCAATAGTGC 3' | | |
| Pinb-D1F | 5' **ATG**AAGACCTTATTCCTCCTA 3' | 54°C | *Pinb-D1*;447 bp (Genbank: DQ363913) |
| Pinb-D1R | 5' **TCA**CCAGTAATAGCCACTAGGGAA 3' | | |
| Pinb-2DF | 5' **ATG**AAGACCTTATTCCTCCTAGCTC 3' | 54°C | *Pinb-2* alleles; 453 bp (Genbank: AM944731) |
| Pinb-2D1R | 5' **TCA**CTAGTAATAGCCATTA**K**TAGCGACA 3' | | |
| Pinb-2DF | As above | As above | *Pinb-2* alleles; 453 bp (Genbank: AM944733) |
| Pinb-2D2R | 5' **TCA**GTAGTAATAGCCATTAGTA**K**GGACG 3' | 54°C | |
| Pinb2v1F | 5'GGTTCTCAAAACTGCCCAT 3' | 57°C | *Pinb-2v1*; 319 bp (Genbank: GQ496616) |
| Pinb2v1R | 5' ACTTGCAGTTGGAATCCAG 3' | | |
| Pinb2v2F | 5'CTTGTAGTGAGCACAACCTTTGCA3' | 60°C | *Pinb-2v2*; 401 bp (Genbank: GQ496617) |
| Pinb2v2R | 5'GTATGGACGAACTTGCAGCTGGAG3' | | |
| Pinb2v3F | 5'GAGCACAACCTTTGCGCAATG3' | 60°C | *Pinb-2v3*; 398 bp (Genbank: GQ496618) |
| Pinb2v3R | 5'CATTAGTAGGGACGAACTTGCAGCTA3' | | |
| Pinb2v4F | 5' CCTTTCTCTTGTAGTGAGCACAACCA 3' | 65°C | *Pinb-2v4*; 403bp (Genbank: GQ496619) |
| Pinb2v4R | 5' GACGAACTTGCAGTTGGAATCCAA | | |
| T7 | 5' GTAATACGACTCACTATAGGGC 3' | 50 °C | Vector-based primers for amplifications of inserts in clones in pGEM-T Easy and for sequencing |
| SP6 | 5' TATTTAGGTGACACTATAGAAT 3' | | |

[a]**K** is G or T in the degenerate reverse primers Pinb-2D1R and Pinb-2D2R. The nucleotides shown in bold represent the start and stop codons
[b]An annealing temperature about 5ºC below the lowest Tm (primer melting temperature) of the pair of primers was used
[c]The primers used to amplify *Pina-D1* and *Pinb-D1* genes were based on the sequences reported by Gautier et al. (1994) and Pinb-2 primers were based on sequences reported by Wilkinson et al. (2008)

**Table 2.13: Primer pairs used for amplification of *Hina, Hinb-1* and *Hinb-2* genes**

| Primer | Sequence (5'-3') * | Annealing temperature | Expected amplicon size** |
|---|---|---|---|
| Hina F | GTGTACACAACTGCAGACAGAAAGC | 60°C | *Hina*; 631 bp (Genbank: (AAV49987.1) |
| Hina R | ATTATTCCAAGACCACTTTTATTTGTC | | |
| Hinb-1 F | CAACACCAAAACAACG | 54°C | *Hinb-1*; 535 bp (Genbank: (AAV49986.1) |
| Hinb-1 R | GACCTCATTGATTTGTC | | |
| Hinb-2 F | ACCAACACCAAATAAACA | 54°C | *Hinb-2*; 576 bp (Genbank: (AAV49985.1) |
| Hinb-2 R | CCAATATACAAGCGGAATTTTATTC | | |

*Primers designed as reported by Takahashi et al. (2010), and these sequence areas are identical for all three genes. **The sequences reported by Caldwell et al. (2004) used as reference sequences.

**Table 2.14 Primer pairs used for cloning and sequencing in yeast vectors**

| Primer name | Sequence (5'-3') * | Annealing Temp °C | Size of genes coding for mature protein sequence | Applications |
|---|---|---|---|---|
| PINA-Y2HF | AAA**GAATTC**GATGTTGCTG | 54 | 363 bp (coding for DVAGGG…….*) | Amplification of gene section encoding mature PINA |
| PINA-Y2HR | AAA**GGATCC**TCACCAGTA | | | |
| PINB-Y2HF | AAA**GAATTC**GAAGTTGGCG | 54 | 360bp (coding for EVGGGG…….*) | Amplification of mature PINB |
| PINB-Y2HR | AAA**GGATCC**TCACCAGTA | | | |
| PINB-2v3-1 Y2HF | AAA**GAATTC**AATGGTGGTGGA | 54 | 395 bp (coding for ENGGGG…….*) | Amplification of gene section encoding mature PINB-2v3-1 |
| PINB-2v3-1 Y2HR | AAA**GGATCC**TCAGTAGTAATAG | | | |
| 3'AD pGADT7 R | AGATGGTGCACGATGCACAG | 55 | | Amplification and sequencing from 3'AD cloning vectors |

*The incorporated restriction enzyme sites are shown in bold underlined

**Table 2.15 Primers used for sqRT-PCR**

| S.no. | Primer | Sequence | Annealing temperature | Expected product size |
|---|---|---|---|---|
| 1 | ActinF | 5'-TGAACCCAAAAGC CAACAGAG-3' | 54°C | 147 bp |
| 2 | ActinR | 5'-CACCATCACCAGA GTCGAGAAC-3' | | |
| 3 | Pinb2v-RT F | 5'-GATGTGAGGCCA TTTGGAGG-3' | 54°C | 146 bp |
| 4 | Pinb2v-RT R | 5'-AGTAKGGACGAA CTTGCAGC-3' | | |

### 2.4.3 Typical PCR conditions

The total reaction volume was made to 50 µL containing 200 ng of genomic DNA (gDNA) template, 25 µL Biomix (Bioline), and 1 µL (of the 10 µM working solution) each of the forward and reverse primer of a pair. The reactions were carried out in a thermal cycler. All amplifications involved initial denaturation of the template at 94 °C for 5 minutes. This was followed by 35 cycles of denaturation at 94 °C (45s), annealing at annealing temperature (45s), extension at 72 °C (1 min), and a final extension at 72 °C (5 min). Five µL aliquots of the PCR products were mixed with 1µL of 6x loading dye and subjected to agarose gel electrophoresis as described above.

Amplification of mature protein sections of inserts with restriction sites incorporated in the primers (Table 2.14) was carried out using following PCR conditions: (i) initial denaturation at 94 °C for 3-5 min (ii) 10 cycles of denaturation (94 °C) for 40 s, annealing at 40 °C (temperature estimated excluding the AAA+ restriction site sequence introduced in the primers; see Table 2.3) for 40 s, and extension (72 °C) for 40 s. (iii) another 20 cycles of denaturation (94 °C) for 40 s, annealing at 54 °C (temperature estimated including the overhangs) for 40 s, and extension (72 °C) for 40 s. (iv) a final extension at 72 °C for 5 min.

### 2.4.4 Restriction enzyme digestion of *Pinb-D1* PCR products

Restriction enzyme digestion was carried out on PCR products to screen for known mutations and detect Cleaved Amplified Polymorphic Sequences (CAPs) sites. The enzyme Bsrb1 (GGCGAG) was used to detect *PinbD1b* mutation containing GGC to AGC codon change (Giroux and Morris, 1997). This SNP (G to A) causes gain of restriction site in the *Pinb-D1b* sequence at 223bp. The wild type *Pinb-D1* has one restriction site whereas the mutated *Pinb-D1b* has two restriction sites. The digestion of PCR products containing non-*Pinb-D1b* alleles result in 318 bp and 129 bp fragments and products with the *Pinb-D1b* mutation result in 223bp, 129bp, 95bp fragments (Figure 2.1). The reaction volume was made to 15 µL consisting of 10 µL *Pinb* PCR product, 5U of restriction enzyme and 1.5 µL of buffer. The reaction was incubated at 37 °C for 1 hour, followed by agarose gel electrophoresis to look for fragments.

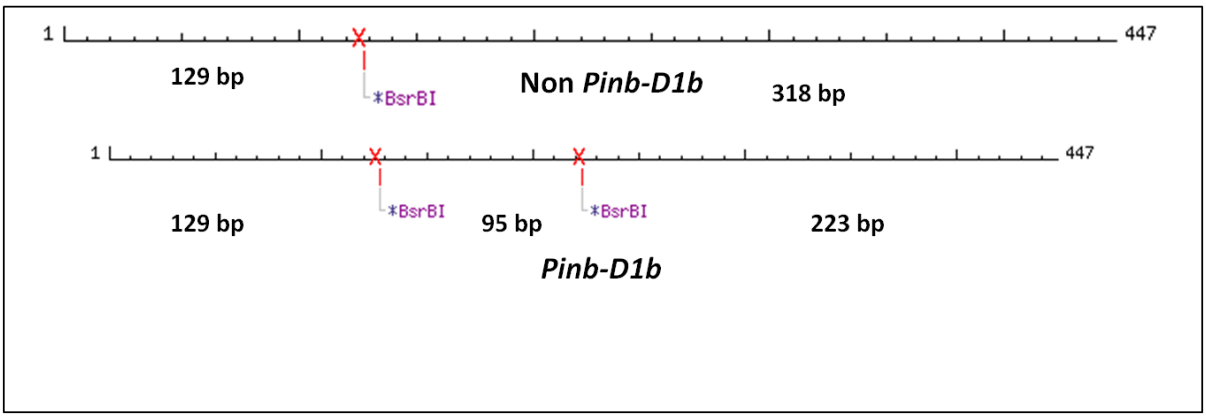



**Fig. 2.1 Bsrb1 restriction sites in *Pinb-D1***

### 2.4.5 Purification of PCR products

The PCR products were purified using Wizard® SV Gel and PCR Clean -Up kit to remove excess primers and PCR reagents. Briefly, an equal volume of membrane binding solution (from the kit) was added to the PCR amplification and the tube centrifuged at 16,000 x g for 1 minute. The flowthrough was discarded and 700 µL of membrane wash solution was added and centrifuged as above. Lastly, washing was repeated with 500 µL wash solution. DNA was eluted in 50 µL MilliQ water and 5 µL aliquot was electrophoresed in 1% agarose gels.

## 2.5 Cloning of inserts into pGEM-T Easy vector

The pGEM-T Easy Vector system (Promega, Australia) was used to clone any PCR products of interest, especially where the PCR would have generated mixed products form different homeo-alleles or gene duplicates. Direct sequencing of *Pinb-2* PCR products with either primer pair indicated in Table 2.12 would have been inappropriate as these are expected to contain multiple sequences (Chen et al. 2010a; Wilkinson et al. 2008). Ligation reactions (final volume: 10 µL) contained purified PCR product (~23 ng), pGEM-T Easy vector with T-overhangs (1 µL: 50 ng), 2x Rapid Ligation Buffer and T4 DNA ligase 1 µL (3 units/µL). The reactions were incubated, held at 4°C overnight. The amount of PCR product (insert) to be included in a ligation reaction was calculated using the following equation (Promega):

$$\text{Insert (ng)} = \frac{\text{vector (ng) x size of insert (kb)}}{\text{Size of vector (kb)}} \text{ x insert: vector molar ratio}$$

Thus, for a 0.45 kb *Pinb-2* insert, with insert: vector ratio of 3:1, and 50 ng of vector, approximately 23 ng of insert was required.

## 2.6 Transformation of chemically competent *E. coli* JM109 cells

Chemically competent *E. coli* JM109 cells were prepared per Inoue et al. (1990), with minor modifications. The frozen stock culture was streaked onto a non-selective LB agar plate and incubated at 37°C overnight. A single colony was inoculated into 10mL LB medium and incubated overnight at 37 °C on a shaker (180 rpm). This culture was used to inoculate 500mL of SOB media. The inoculation was incubated at 18 °C with shaking until the $OD_{600}$ reached 0.4. The culture was held on ice for 10 min and the cells were transferred into 50mL tubes and centrifuged at 2, 500 × g for 15 at 4 °C. The cells were resuspended in 80 mL ice-cold TB buffer and held on ice for 10 min and centrifuged as above. The cells were washed again in 20 mL TB buffer and DMSO was added with gentle mixing to a final concentration of 7 % (v/v). The cells were again held on ice and dispensed in 0.1mL aliquots. These aliquots were then snap frozen in liquid nitrogen and stored at -80 °C. The ligation reaction mixes were transformed into *E. coli* JM109 cells.

A tube containing 100 µL of competent JM109 cells was thawed briefly on ice and the ligation mixture was added. The cell suspension was held on ice for 20 minutes, followed by heat shock at 42°C for 90 seconds and again held on ice for 2 minutes. 500 µL of SOC medium was added to the cells and subjected to incubation at 37°C with shaking at 200 rpm for 1.5 hours. This was followed by plating of 100 µL of culture onto LB plates containing Amp/IPTG/X-Gal for blue-white screening and incubated at 37°C overnight. The pGEM-T Easy plasmid contains the *lac*Z gene, encoding the α-peptide of β-galactosidase that can be detected using X-gal (blue colour). An internal multiple cloning site (MCS) is present within the *lac*Z gene (Appendix II) where the DNA of interest can be inserted that consequently disrupts the function of gene and no functional peptide is produced and thus no blue colour. Blue colonies therefore contain a vector with an uninterrupted *lacZ* (therefore no insert), while white colonies indicate the presence of an insert in *lacZα* which disrupts its function.

## 2.7 Plasmid DNA isolations

For plasmid purifications, about 18-20 white colonies (with recombinant vectors) were picked from each plate and each inoculated into a tube containing 4 mL of LB medium with 20 mg ampicillin per mL and incubated overnight at 37°C with shaking (200 rpm). Plasmid DNA isolations were conducted using the Wizard® Plus SV Minipreps DNA purification system (Promega, Australia). The cultures were pelleted by centrifugation at 14.5 x 1000 rpm for 5 minutes. 250 µL of resuspension solution was added to the pellet and the cells resuspended by vortexing. This was followed by addition of 250 µL of cell lysis solution and 10 µL of alkaline protease solution (provided with the plasmid isolation kit) and incubation for 5 minutes at room temperature. Neutralization solution (350 µL) was added and the lysate was centrifuged at 14.5 x 1000 rpm for 10 minutes at room temperature. The clear supernatant was transferred to spin column assembly and centrifuged at 14.5 x 1000 rpm for 1 minute. The spin column containing the bound plasmid DNA was washed with 750 µL of wash solution. Second washing was done with 250µL of wash solution. The plasmid DNA was eluted into 50 µL of sterile MilliQ water.

### 2.8 Screening of clones for inserts

Screening for inserts was carried out by PCR amplification of plasmid DNAs using vector specific, insert-flanking primers, T7 and SP6 (Table 2.12). The expected sizes of amplification products of the different genes are given in Table 2.12. The PCR products were electrophoresed to determine the presence of *Pinb-2* insert.

### 2.9 DNA sequencing reactions

The inserts were amplified from clones of interest by using vector specific primers (T7 and Sp6) and the PCR products were purified using purification kit as described in section 2.4.4. The purified product was eluted into 50 uL of sterile MilliQ water and quantified using agarose gel electrophoresis. The sequencing reactions were undertaken with ABI BigDye Terminator (BDT) version 3.1 (Applied Biosystems; California, USA) according to the instructions provided by Australian Genome Research Facility (AGRF) (Melbourne, Australia). The reaction mix typically included 10-50 ng PCR product or 500 ng plasmid DNA, 3.2 pmole sequencing primer (T7 or SP6), 0.5 µL of DNA sequencing dye (Big Dye v3.1) and 2 µL 5x BDT buffer, the final volume made to 10 µL with sterile MilliQ water. In case of *Hin* genes, the purified *Hin*a, *Hinb-1* and *Hinb-2* PCR products were directly sequenced without cloning using gene specific primers as mentioned in Table 2.13. The cycling conditions were: initial denaturation at 94°C for 5 minutes followed by 30 cycles of denaturation (94 °C, 10 seconds), annealing (50°C, 5 seconds) and primer extension (60°C, 4 minutes). The sequencing reactions were mixed with 75 µL 0.2mM $MgSO_4$ and incubated for 15 minutes. This was followed by centrifugation for 15 minutes at 14.5 x 1000 g. The supernatant was discarded and the pellet was air-dried and submitted to AGRF (Melbourne) for capillary separation using a 3730x1 DNA analyser (Applied Biosystems).

### 2.10 Bioinformatics methods

The raw DNA sequence data was analysed using Bioedit v7.2.3 (http://www.mbio.ncsu.du/BioEdit/bioedit.html). The DNA sequences were translated by using the key function 'CTL G'. The alignments of DNA and protein sequences with reported sequences were done using sequence alignment editor program Bioedit v 7.2.3. The alignments were displayed using 'plot identities' to first sequence with a dot, to

identify any single nucleotide polymorphisms (SNPs) and amino acid variations. The % identity of *Pinb-2* sequences to published *Pinb-2* variant group sequences were calculated by making use of the 'Sequence identity matrix icon' in Bioedit. The signal peptides for PINB-2, HINA, HINB-1 and HINB-2 were determined based on the alignment with PINA-D1 (Genbank: CAA49538) and PINB-D1(Genbank: CAA49537) using the Bioedit program. The isoelectric point (pI) of the putative PINB-2, HINA, HINB-1 and HINB-2 mature sequences were determined using the 'compute pI/MW Tool' at the Expert Protein Analysis system (ExPAsy) site (http://au/expasy.org/tools/pi_tool.html).

### 2.11 Grain texture and protein content determination

The grain texture of wheat and barley samples, obtained as detailed in section 2.2.4, was determined using SKCS (Chen et al. 2010b; Geng et al. 2012; Galassi et al. 2012; Walker et al. 2013). The SKCS values were determined using the AACC Method (55-30) in the laboratory of Dr Joseph Panozzo (Department of Economic Development, Jobs, Transport and Resources, Horsham, Victoria) through cumulative testing using 50 seeds. The samples were tested in duplicate. Two commercial barley varieties of known SKCS hardness values, i.e., Hindmarsh (SKCS; 65 units) and Barque (SKCS; 35 units) were included as internal controls.

Scanning Electron Microscopy (SEM) was used on progeny lines and landraces which were available only in limited quantities, as per Chen et al. (2005) and Ramalingam (2012) in order to visualise the degree of association between starch granules and protein matrix, to attempt to classify them in broad categories of texture (hard or soft). The cryo-fractured seed was dissected longitudinally and fixed onto glass slides using double-sided adhesive tape, exposing the starchy endosperm while mounting the sample. All samples were mounted on pin-type aluminium SEM mounts with double-sided conducting carbon tape and then coated in Dynavac CS300 coating unit with carbon and gold to achieve better conductivity of the seed surface. Images were captured at 2000× and 3000× magnification on ZEISS supra 10 VP field emission scanning electron microscope (Carl Zeiss Microscopy, NY, USA).

The grain protein content was determined in the laboratory of Dr. Panozzo using the Dumas combustion method (AACC 46-30.01).

## 2.12 Methods specific to yeast two-hybrid work

### 2.12.1 Principles of the Yeast two-hybrid- protein -protein interaction system

The interactions between proteins expressed in 'bait' constructs and 'prey' constructs in yeast were identified and quantitated using the yeast two-hybrid system (Chein et al. 1991). In a yeast two-hybrid assay, when the bait (pGBK-T7) and prey (pGAD-T7) fusion protein interact, they bring the DNA-BD (binding domain: bait fusion protein) and AD (activation domain: prey fusion protein) into proximity to activate transcription of the reporter genes (ADE2, HIS3 and *LacZ*) (Fig. 2.2). When the bait and prey proteins interact ADE2 and HIS3 expression allow the cells to grow on yeast synthetic defined (YSD) medium deficient in Ade and His, allowing for identification of interactions. The activity of the β-galactosidase enzyme (encoded by *LacZ*), allows for the strength of the interaction to be quantified. The steps involved in yeast-two hybrid work are described below

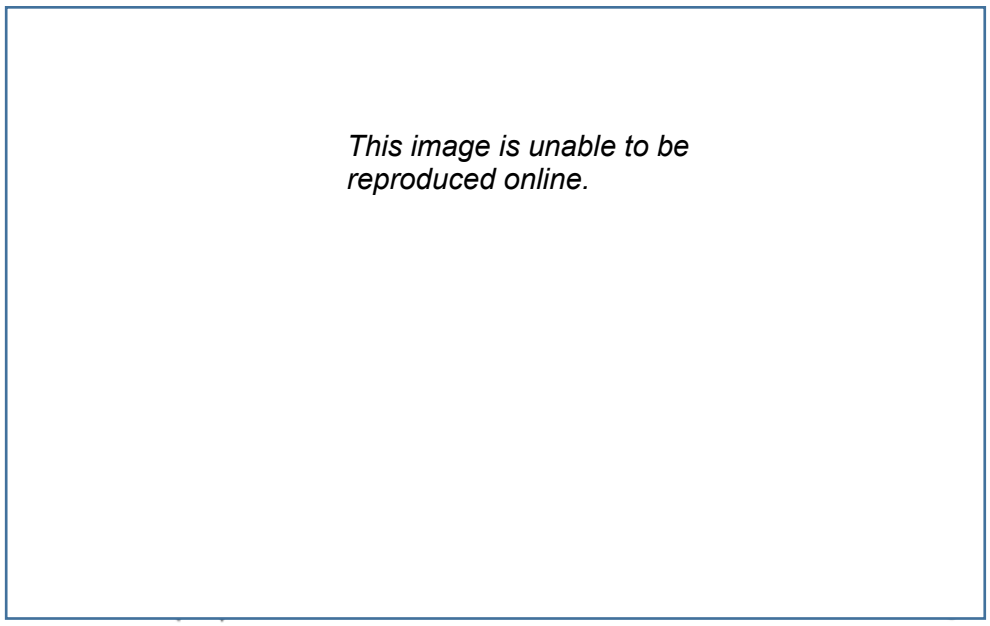



**Figure 2.2 Principle of the yeast two-hybrid system** (Promega, Protein interaction guide, (http://www.promega.com/resources/product-guides-and-selectors/protein-interaction-guide/)

### 2.12.2 Isolation of PINA and PINB constructs made in bait and prey vectors

The pGADT7 and pGBK clones containing gene sections encoding the putative mature PINA and PINB proteins were previously constructed in our group directionally by inserting the appropriate gene sections at the EcoRI and BamHI sites of the vector (Ramalingam, 2012) and used for testing certain Y2H interactions (Alfred et al.2014). The bait and prey constructs in i.e., pGADPINA, pGADPINB, pGBKPINA, pGBKPINB, empty vectors (pGBKev, pGADev) were grown from their respective frozen glycerol stock (*E.coli* cells) on LB plates with appropriate selection antibiotic (ampicillin for pGADT7 and kanamycin for pGBKT7). Plasmid isolations were undertaken for the bait and prey vectors as described in Section 2.7. The extracted plasmids were quantified using spectrophotometer and subjected to agarose gel electrophoresis to determine the quality.

### 2.12.3 Cloning, transformation and DNA sequencing of pGADPINB-2v3-1

*Directional cloning of Pinb-2v3-1 into yeast vector*

*Pinb-2*v3-1 insert (obtained from pGEM-T Easy clone of wheat cultivar Sunco) was cloned into prey vector pGADT7 in this study. The cloning involved vector and insert preparation followed by ligation. Five micrograms of DNA of the purified recombinant clone pGADPINA obtained as above, were digested with 25 U of each restriction enzyme (BamH1 and EcoR1double digest) in a suitable 1× restriction enzyme buffer (Buffer E, Promega) in a volume of 100 µL for 3 hr at 37°C. To prevent religation of linearized plasmid SAP was added (10 units SAP/µg vector in a final volume of 30-50µl) and incubated at 37°C for 15 min followed by heat inactivation (65 °C for 15 min). The digested DNA was loaded onto an agarose gel and electrophoresed for gel purification. The DNA band was excised and purified using the Wizard® SV Gel and PCR Clean -Up kit and eluted in sterile MilliQ water (30 µL). The extracted vector was quantified using spectrophotometer. The gene section corresponding to the mature PINB-2v3-1 protein was amplified from the pGEM® -T Easy clones of full length gene (as described in Section 2.5) using primers that have restriction sites incorporated into them (Table 2.14). The PCR product was column-purified and double digested as above with 10U of each enzyme

(BamH1 and EcoR1), in a volume of 50 µL (~3 hr) at 37°C. The digested PCR products were column- purified and quantified using spectrophotometer.

To determine the appropriate amount of PCR product (insert) required for ligation into the yeast vectors, the equation mentioned earlier in Section 2.5 was used. For example, for a 0.395 kb insert (*Pinb-2v3*), with insert to vector ratio of 3:1, and 50ng of 8 Kb vector (pGAD), ~8 ng of insert was required. Thus, ligation reactions involved 50 ng of vector DNA, to ~8 ng insert, 1 × Rapid Ligation buffer, 1 U of T4 DNA ligase (1µL) and ATP (to a final concentration of 1mM). The reactions were carried out in a final volume of 10-15 µL and incubated at 4°C overnight. The ligation was then transformed into chemically competent *E.coli* JM109 cells prepared as per Inoue et al. (1990) (Section 2.6). Transformants containing pGADT7 clones were spread on LB plates containing 100 µg/mL ampicillin. A volume of 100 µL was inoculated and spread on each plate and the plates were incubated for 16 hr at 37 ºC. Since the recombinant colonies could not be distinguished visually from the non-recombinant colonies (with colour selection), an initial screening with colony PCR was applied. Gene specific primers (Table 2.12) were used for colony PCR. The PCR reaction consisted of 2× Biomix, 100 ng of each of the forward and reverse primer, and some cells of a colony touched with the help of a sterile loop, the total volume made up to 50µL with sterile MilliQ water. The PCR products were subjected to agarose gel electrophoresis. For this, the remaining half of a colony of the PCR -positive clone was grown overnight for purification of plasmids at 37°C. The pelleted cells were purified using the Wizard® Plus SV Minipreps DNA Purification System per manufacturer's instructions (Promega). These purified clones were used for sequencing the inserts. Sequencing of the recombinant pGADPINB-2v3-1 vector was carried out using the vector specific primer 3'BD reverse (Table 2.14).

### 2.12.4 Yeast transformation and selection

The recombinant plasmids i.e, pGADPINA, pGADPINB, pGBKPINA, pGBKPINB, made previously (Ramalingam, 2012), pGADPINB-2v3-1 described above and the empty vectors (pGBKev, pGADev) were then transformed into yeast cells by the lithium acetate method as per the yeast protocols handbook (Clontech, CA, USA;

http://www.clontech.com/xxclt_ibcGetAttachment.jsp?cItemId=17602, last accessed June, 2016). The bait (pGBK) constructs were transformed into yeast strain AH109 and the prey (pGAD) constructs were transformed into yeast strain Y187α. Single colonies of *S. cereviceae* strains AH109 and Y187α were separately inoculated into 50 mL YPD media, and grown at 30°C for 16-18 hr with shaking (250 rpm) to an $OD_{600}$ > 1.5. 30 mL of this culture was used to inoculate 300 mL YPD in a 2L baffled flask and incubated at 30°C with shaking (230 rpm) to an $OD_{600}$ = 0.4-0.6 (4-6 h). The cultures were transferred into 50 mL centrifuge tubes and were gently centrifuged (1,000 ×g for 5 min). The cell pellets were washed in TE buffer and pooled into 1 tube (final volume 25-30 mL) and centrifuged again. The pellet was resuspended in 1.5 mL freshly prepared 1xTE/1xLiAC and held on ice for 2 min. For the yeast transformations, 100 µL of the above cell suspension, 0.1 µg plasmid DNA and 0.1 mg denatured carrier DNA (10 mg/mL Herring testes DNA), were mixed by vortexing. This was followed by addition of 600 µL PEG/LiAC solution and incubation at 30°C for 30 min with shaking (200 rpm). 70 µL DMSO was added gently and the cells heat-shocked at 42°C for 15 min and then held on ice for 2min. The cells were centrifuged at 14,100 x g for 5 sec and the supernatant discarded. The cells were resuspended in 100 µL sterile MilliQ water and spread on YSD selection medium deficient in Trp, (bait constructs) and medium deficient in Leu (prey constructs), and incubated at 30°C for 3-5 days. Bait plasmid DNA (pGBKT7 clones) was transformed into freshly prepared competent cells of *S. cereviceae* AH109 and the prey plasmid DNA was transformed into *S. cereviceae* Y187α.

### 2.12.5 Yeast mating

Different pairs of haploid cultures of *S. cerevisiae* AH109 (carrying the bait vector with various desired inserts) and S. cerevisiae Y187α (carrying the prey vector with various desired inserts) were mated overnight and plated on YSD-WL media (double drop out) as described in the yeast protocols handbook (Clontech, CA, USA; http://www.clontech.com/xxclt_ibcGetAttachment.jsp?cItemId=17602, last accessed June, 2016). The matings included empty vector controls to rule out the possibility of false positives. One large colony of each yeast type of the desired pair was picked from freshly transformed plates and inoculated into microcentrifuge tube containing 0.5 mL YPD medium. The suspension incubated overnight at 30°C with shaking (200 rpm). The

mated cultures were spread on plates of double drop out media (YSD-WL) and were incubated at 30 °C for 3-5 days to allow diploid cells to form visible colonies.

### 2.12.6 Screening of protein-protein interactions

The screening of protein- protein interactions were undertaken a as described in the yeast protocols handbook (Clontech, CA, USA; http://www.clontech.com/xxclt_ibcGetAttachment.jsp?cItemId=17602, last accessed June, 2016). After 3-5 days, a single colony was picked from each YSD -W-L plates, inoculated into 1mL sterile MilliQ water and vortex-mixed to disperse any clumps. This was followed by replica-spotting of 5 µL of this suspension on plates containing selective media lacking Trp, Leu, Ade and His (YSD -W-L-A-H), for selecting yeast cells expressing HIS3 and ADE2 reporter genes activated during any bait-prey interactions. The cultures were also spotted on duplicate plates containing YSD -W-L-A-H and 1 mM 3-amino-1,2,4-triazole (3-AT). This was done to assess the strength of interaction between the expressed proteins. The plates were incubated at 30 °C for 3-5 days.

### 2.12.7 Quantification of interaction strength by β-galactosidase assay

The detection of ONP ortho-Nitrophenyl (ONP) in the reaction: ortho-Nitrophenyl-β-galactoside (ONPG) → ONP + galactose by spectrophotometry (Clontech, CA, USA; http://www.clontech.com/xxclt_ibcGetAttachment.jsp?cItemId=17602) was used to measure the β-galactosidase activity of the reporter gene (*lacZ*) in mated yeast cells . Five mL YSD (-WL) media was inoculated with a single colony of each diploid culture and incubated at 30°C overnight. The overnight culture was diluted to $OD_{600}$ = 0.2 and grown at 30 °C with vigorous shaking (230 rpm) to an $OD_{600}$ of 0.6-0.8 (approximately 4-6 h). This was followed by centrifugation of the culture (1.5 mL) at 14,100 × g for 30 sec. The supernatant was removed and the cell pellets were washed in 1.5 mL Z buffer. The washed cell pellets were centrifuged and resuspended in 300 µL of Z buffer. A fresh microcentrifuge tube containing 100 µL of the cell suspension and freeze/thawed three times to lyse cells. To the lysed cells, 0.7 mL of Z buffer + β-mercaptoethanol and 160 µL of 4 mg/mL ONPG were added and the reactions were incubated at 30°C until a

yellow colour developed (3 h). 0.4 mL of 1 M $Na_2CO_3$ was used to stop the reaction. The tubes were centrifuged at 14,000 x g for 10 min and the supernatant was transferred to a clean cuvette and $OD_{420}$ measured using a spectrophotometer. The following formula was used to calculate enzyme activity:

β-galactosidase units = 1000 × $OD_{420}$/ (t ×V×$OD_{600}$), t = time (min) elapsed after incubation, V = 0.1 mL × concentration factor (Miller, 1972).

## 2.13 Methods specific to gene expression analysis

### 2.13.1 Total RNA extraction and quantification

RNA was extracted from leaf and root tissues of wheat cultivars harvested after 21 days of germination. These tissues were snap frozen in liquid nitrogen and stored at -80$^0$C for RNA extraction. RNA was extracted using TRIsure reagent per the manufacturer's protocol (Bioline, Australia). About 100 mg of crushed tissue was mixed with 1ml of TRIsure and incubated at room temperature (RT) for five minutes. To this suspension, 200 µL of 24:1 chloroform: isoamyl alcohol was added and the tube was shaken vigorously for 15 seconds and then incubated at RT for 3 minutes. This was followed by centrifugation at 4$^0$C at 14,000 xg for 15 minutes and the supernatant was transferred to a sterile microcentrifuge tube containing 500 µL of isopropanol. The tube was held at RT for 10 minutes, followed by centrifugation at 4$^0$C at 14,000 xg for 10 minutes. Finally, the RNA pellet was washed with 1ml of 75% ethanol (made with DEPC-treated water) and centrifuged at 4$^0$C at 4,000 xg for 5 minutes. The pellet was air-dried, re-suspended in 40 µL DEPC-treated water and incubated for 10 minutes at 60°C. The quality of the total RNA extracted was assessed by running 4 µL of RNA on an agarose gel. The quantity of RNA was estimated using spectrophotometer. An absorbance unit of 1 at $A_{260}$ corresponds to 40 µg/mL of single stranded RNA, and RNA to protein absorbance ratio ($A_{260}/A_{280}$) of~1.8 indicated high purity (Sambrook and Russell, 2001).

### 2.13.2 DNase treatment of extracted total RNA

The RNA preparations were treated with DNase to get rid of genomic DNA contamination. Total RNA was treated with RNase-free DNaseI (Promega) according to the supplier's instructions (http://au.promega.com/?origUrl=http%3a%2f%2fwww.promega.com%2f). The total RNA extracted above (36 µL) was incubated was incubated with 10 U of RQ1 RNase-free DNase I (Promega), 5 µL of 10× Reaction Buffer (Promega) and 2 U of RNase Inhibitor (Bioline) in a 50 µL volume at 37 ºC for 30 minutes. The precipitation of RNA was performed using the LiCl precipitation method (http://www.ambion.com/techlib/tb/tb_160.html). The DNase I mix was combined with 20 µL of 10 M LiCl made in DEPC-treated water and then diluted to 80 µL with DEPC-treated water to a final concentration of 2.5 M LiCl. This mixture was kept at -20 ºC for 30 minutes followed by centrifugation at 14,500 rpm for 15 minutes at 4ºC. The RNA pellet was washed twice with 75% ethanol made in DEPC-treated water. The RNA was centrifuged as above and the pellet was dried in air for 10 minutes. 40 µL DEPC-treated water was used to resuspend the air-dried RNA pellet. The lack of gDNA in the RNA preparations was checked using agarose gel electrophoresis. The concentration of RNA preparations were assessed using spectrophotometer and stored at -80 °C for further analysis.

### 2.13.3 cDNA synthesis

The Bioscript MMLV reverse transcriptase (Bioline, Australia) was used to synthesise first strand complementary DNA (cDNA) according the supplier's instructions. This was done by incubation of 1 or 2 µg of purified total RNA with 0.5 µg/µL oligo d(T)18 primer (Invitrogen) at 70 ºC for 5 minutes in a volume made up to 12 µL with DEPC-treated water. To this reaction mixture, 1 µL dNTPs (10 mM each), 10 U RNase inhibitor (Bioline), 4.0 µL 0f 5× reaction buffer (Bioline), 2.5 µL of DEPC-treated water and 50 U Bioscript (Bioline) were added and incubated at 37 ºC for 1 hour. The reaction was stopped by incubation at 70 ºC for 10 minutes. A minus-reverse transcriptase control (-RT) was included which means reverse transcription containing all the RT PCR reagents except the reverse transcriptase. The lack of genomic DNA contamination in the cDNA

was assessed by PCR using 1 µL of the cDNA preparations and the intron-spanning actin primers and comparing the size of PCR product from cDNA and gDNA (Table 2.13, Fig. 6.8). For the semi-quantitative reverse transcriptase, PCR (sqRT-PCR) 1 µL cDNA was used as template and amplified using gene specific primer (Pinb2v-RT) and reference gene primer (actin) listed in Table 2.15.

# CHAPTER 3

## Study of sequence diversity of *Pinb-2* genes in Australian wheat cultivars

### 3.1 Abstract

Wheat is one of the most important cereal crops worldwide and the variability in its grain texture (hardness) enables production of different food products and eating qualities. The *puroindoline* genes (*Pina-D1* and *Pinb-D1)* located on chromosome 5D of common wheat (Triticum aestivum L) are the main genetic determinants of grain hardness, certain mutations leading to hard texture. Members of the gene copy called *Puroindoline b-2* (*Pinb-2),* present on all group 7 chromosomes, also have minor contributions to texture. Its alleles *Pinb-2 variant 2* (*Pinb-2v2*) *and Pinb-2v3* are considered true alleles, with *Pinb-2v3* associated with harder textures, and the subtype *Pinb-2v3b* (Val104Ala) has also been linked with harder grains compared to other *Pinb-2v3* and *Pinb-2v2* types in soft wheats. In this study, 22 Australian wheat cultivars from the three well-reported *Pin* genotype classes but showing a range of SKCS grain hardness values, were genotyped for alleles *Pinb-2v1*, *Pinb-2v2*, *Pinb-2v3* (two subtypes) and *Pinb-2v4.* The SKCS grain hardness values of cultivars used in this study ranged from 30.4-79.0 and were predicted using the AACC Method 55-30. This investigation aimed to understand the contribution of *Pinb-2* genes to grain texture variation within genotypes of same *Pin* class. Two parental lines from the (Tasman x Sunco population) and 13 randomly selected doubled-haploid (DH) progeny lines of their genetic cross, were also analysed for their *Pinb-2v3* subtypes. Further, the full *Pinb-2* genotype diversity (study of all six variant alleles) was investigated in four cultivars. Studies were undertaken using gene amplifications, cloning and/or DNA sequencing. Based on these results, the grain texture of seeds of selected lines was examined by scanning electron microscopy (SEM). The results confirmed the reported multi-genic nature of *Pinb-2* and *Pinb-2v1* to *Pinb-2v6* groups and indicated notable sequence diversity. Some of the single nucleotide polymorphisms (SNP) leading to amino acid substitutions in the deduced proteins, at the functionally important tryptophan- rich domain (TRD) and/or certain basic and hydrophobic residues, which may have effects on lipid binding properties and thus antimicrobial properties/grain texture. The *Pinb-2v3* allele was predominant (detected in 18/22 cultivars) compared to *Pinb-2v2*, and these two alleles did not co-exist, supporting these being true alleles. Similarly, the subtypes *Pinb-2v3-1* and *Pinb-2v3-1a* (Val104Ala) did not co-exist, confirming their allelic nature. Interestingly, v1, v2 or v3, v4, v5 or v6 occurred together in various combinations and are likely non-allelic. Interestingly, occurrence/non-occurrence of *Pinb-2* variants (v2 or v3) could not explain the difference in SKCS

hardness values of same *Pin* class (*PinaD1-a/Pinb-D1a*, *PinaD1a/Pinb-D1b*, *Pina-D1b/Pinb-D1a*) genotypes. *Pinb-2v3-1a* (Val104Ala) was not found in soft cultivars and was found in five *Pina-D1a/Pinb-D1b cultivars* which had 52.39 to 78.53 SKCS grain hardness values and three *Pina-D1b/Pinb-D1a* cultivars which had SKCS grain hardness values of 56.04 to 70.37. The Val104Ala substitution in variant 3 was detected in Tasman and not Sunco, and seven of their 13 DH progeny lines. However, the SEM analysis of these seven lines did not show relation with harder grain texture. Thus, it is suggested that Val104Ala substitution in variant 3 could not be associated with increased grain hardness and *Pinb-2* variants do not exert a notable effect on kernel texture.

## 3.2 Introduction

Grain hardness or texture in wheat is determined by the *Pin* genes, *Pina-D1* and *Pinb-D1*, located at the *Ha* locus on chromosome 5D of common wheat. As detailed in Chapter 1 (Section 1.3.6), another gene copy, called *Pinb-2*, reported to occur on all homeologous copies of chromosome 7, has been linked with minor contributions to grain texture. The *Pinb-2* genes were first mapped to chromosome 7AL and linked to a minor quantitative trait loci (QTL) for hardness (Wilkinson et al. 2008). Chapter 1 details the literature reports regarding the following:

a) identification, sequences and chromosomal localisation of the five *Pinb-2* gene variants (Wilkinson et al. 2008, Chen et al. 2010a, Chen et al.2011, Geng et al. 2012)
b) identification of a sixth variant (*Pinb-2v6*) in our laboratory (Ramalingam et al. 2012); v6 not mapped
c) the allelic nature of *Pinb-2v2* and *Pinb-2v3* (Chen et al.2010a) and the association of *Pinb-2v3* to increased grain yield or harder texture (Chen et al. 2010 b)
d) the association of *Pinb-2v3* with higher SKCS hardness index (27.2) as compared to *Pinb-2v2* (22.2) in soft wheats (*Pina-D1a/Pinb-D1a*) whereas no significant difference in hard *Pin* genotypes (*Pina-D1a/Pinb-D1b*, *Pina-D1a/Pinb-D1p*)
e) variation within *Pinb-2v3* i.e. alleles *Pinb-2v3a*, *Pinb-2v3b* and *Pinb-2v3c*;
f) association of *Pinb-2v3b* with increased grain hardness value of 5.5, 2.2, 1.2 SKCS units as compared to *Pinb-2v2*, *Pinb-2v3a* and *Pinb-2v3c* respectively
g) anti-microbial nature of PINB-2 peptides, like PIN peptides (Phillips et al.2011, Ramalingam ,2012)

In summary, *Pinb-2v2* and *Pinb-2v3* show 94.3 % identity at DNA level. The overall sequence identity between variant 2,3 with the *Pinb-D1a* protein was reported to be 57% and 60% respectively. The first tryptophan residue at position 39 of the mature *Pinb* protein is replaced by arginine in variant 2 and tyrosine in variant 3 and only two tryptophans are present (Wilkinson et al. 2008). Further work (Chen et al. 2013) indicated that *Pinb-2v3b* differs from *Pinb-2v3a* by a SNP at position 311(T to C, amino acid change V104A of full length protein) while *Pinb-2v3c* has a C325T that results in silent mutation. Cultivars with *Pinb-2v3b* have a higher SKCS hardness index i.e., they are harder (27.7) as compared to *Pinb-2v2* (22.2), *Pinb-2v3a* (25.5) and *Pinb-2v3c* (26.5) in soft wheats (*Pina-D1a/Pinb-D1a*). However, the difference in SKCS hardness between *Pinb-2* variants was found insignificant in hard wheat (*Pina-D1a/Pinb-D1b*) class (Chen et al. 2013). Thus, all *Pinb-2* genes need investigations of genetic diversity for applications in both grain hardness and biotic defence areas. In particular, *Pinb-2v3* is an important variant that needs to be investigated within lines of common wheat of the two main classes (soft and hard), and in durum wheat (null for *Pina* and *Pinb*, very hard). Thus, 22 in-bred commercial varieties and an in-bred population were tested for genetic diversity. This study involved the following:

(i) Screening of 22 Australian wheat cultivars, selected on the basis of their reported *Pina/Pinb* genotypes (falling into three classes) but with variable SKCS grain hardness values within each class (listed Table 3.1), for *Pinb-2* variation. The cultivars were tested to confirm their *Pina/Pinb* genotypes, followed by analysis of presence/absence and sequences of *Pinb-v2*, and/or *Pinb-2v3a/ Pinb-2v3b/ Pinb-2v3c* alleles. This was undertaken using sequence-specific primers for the reported *Pinb-2* variant alleles.

(ii) The parents (Tasman x Sunco) of a genetic cross studied previously for chromosomal regions associated with grain hardness (Osborne et al. 2001) and 13 randomly selected doubled-haploid (DH) progeny lines of this cross were genotyped as above. They were also analysed by scanning electron microscopy (SEM) to investigate if the mutations had any implications on grain texture.

(iii) The full *Pinb-2* genotype (all variant alleles) of four of the 22 Australian cultivars i.e. Tasman, Sunco, Egret, Sunstar that are parents of genetic crosses studied previously by Osborne et al. (2001) for grain hardness was investigated for the extent of diversity in this multi-gene family.

## 3.3 Results

### 3.3.1 Selection of Australian wheat cultivars

Twenty-Two Australian wheat cultivars with known *Pina*/*Pinb* genotypes but exhibiting a range of SKCS grain hardness values within each class (Table 3.1) were chosen for investigation of their *Pinb-2* alleles. The *Pina/Pinb* genotypes of these cultivars had been determined earlier (Pickering et al. 2007) and the grain hardness of these cultivars was determined in this study by using the SKCS. The seed samples provided by AWCC (as mentioned in Section 2.2.4) were used for determination of SKCS grain hardness. The seed samples were multiplied by AWCC under unknown field conditions (Chapter 2, Table 2.8). As these samples were not grown in the same field trial, an accurate comparison of grain hardness values is not possible. The SKCS grain hardness values are not used as absolute values for comparison among cultivars. These values have been used for comparison in a broader context keeping in mind the environment induced variation. Higher SKCS value is indicative of harder grain texture. *Pina-D1a/Pinb-D1a* genotype (wild type for both *Pina* and *Pinb*) are considered soft, *Pina-D1a/Pinb-D1b* as moderately hard and *Pina-D1b/Pinb-D1a* as hard (Giroux and Morris, 1998). The three *Pin* classes that were investigated were *Pina-D1a/Pinb-D1a* (four cultivars), *Pina-D1a/Pinb-D1b* (twelve cultivars) *and Pina-D1b/Pinb-D1a* (six cultivars). These cultivars showed a range of SKCS grain hardness: *Pina-D1a/Pinb-D1a:*30.47-35.42 units, *Pina-D1a/Pinb-D1b:* 51.43-78.53, *Pina-D1b/Pinb-D1a*: 56.04-78.95. It is well known that the genotypes with identical *Pin* alleles (wild-type or mutant) vary for grain hardness, but the molecular basis of this variation is not known. These cultivars were chosen with the aim to investigate if *Pinb-2* genes contribute to the variation in grain hardness within the same *Pin* class, and to see if the *Pin* genotype has any association with the *Pinb-2* diversity.

### 3.3.2 Confirmation of the reported *Ha* locus genotypes

The twenty-two Australian wheat cultivars selected as above (Table 3.1) were tested to confirm their reported *puroindoline* alleles (Turnbull et al. 2000, Osborne et al. 2001, Cane et al. 2004, Pickering et al. 2007). The entire coding sequences of *Pina-D1* and *Pinb-D1* genes were amplified from 200 ng of genomic DNA (as described in Section 2.4.3) using the primers listed in Table 2.12, designed based on the sequences (Gautier et al. 1994; Genbank: DQ363911 and DQ363913, respectively). The expected product

length for both *Pina-D1* and *Pinb-D1* is 447 base pairs (bp) (Gautier et al. 1994) (Fig. 3.1 and Fig. 3.2 respectively). *Pina* amplicons of approximately 450 bp were obtained for all samples except six cultivars (i.e. Carnamah, Cranbrook, Hartog, Kukri, Machete, Ouyen), indicating the latter six had *Pina-D1b* (null) allele and the other 16 likely had the wild-type allele *Pina-D1a* (examples shown in Fig. 3.1). Sequencing of PCR products of ten of these (Amery, Clearfield Stilleto, Diaz, Falcon, Halbred, Janz, Katepwa, Rosella, Tasman) has confirmed the *Pina-D1a* allele (Pickering et al. 2007), and the other six (Batavia, Chara, Cunningham, Egret, Sunco, Sunstar) are likely also wild type for *Pina* based on their pedigrees and intra-breeding in the Australian wheat programs.

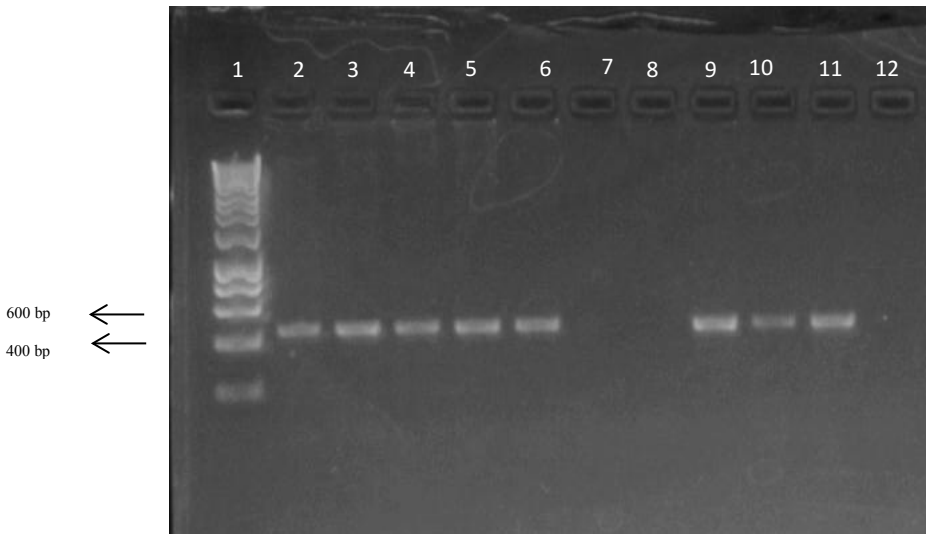



**Figure 3.1. Example of *Pina* PCR products of common wheat cultivars.**
Lane1: Hyperladder 1; lane 2: Tasman, lane 3: Sunco, lane 4: CD87, lane 5: Katepwa, lane 6: Janz, lane 7: Kukri, lane 8: Cranbrook, lane 9: Halbred, lane 10: Egret, lane 11: Sunstar, lane 12: negative control (no template).

Amplifications from 200 ng of genomic DNA (as described in Section 2.4.3) with the Pinb-D1 primers also led to a 450 bp product (calculated size 447 bp) in all cultivars (examples shown in Fig. 3.2). The results for *Pina*/*Pinb* genotypes in wheat cultivars are presented in Table 3.1.

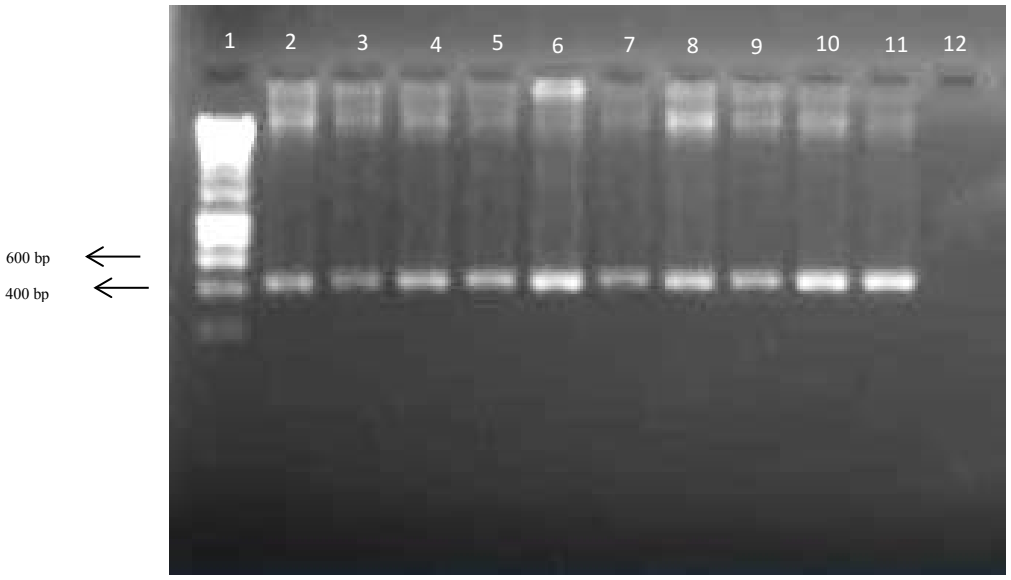



**Figure 3.2. Example of *Pinb* PCR products of common wheat cultivars.**
Lane1: Hyperladder 1; lane 2: Tasman, lane 3: Sunco, lane 4: CD87, lane 5: Katepwa, lane 6: Janz, lane 7: Kukri, lane 8: Cranbrook, lane 9: Halbred, lane 10: Egret, lane 11: Sunstar, lane 12: negative control (no template).

### 3.3.3 Confirmation of *Pinb-D1b* allele using Cleaved Amplified Polymorphic Sequence (CAPS) analysis

Cleavage of above *Pinb-D1* purified PCR products with the enzyme Bsrb1 was used to determine the presence of the hardness-associated *Pinb-D1b* allele, as this mutation (G to A SNP at position 223 of the coding sequence), results in a gain of another restriction site in the *Pinb-D1b* sequence. This leads to RFLPs giving 129 bp and 318 bp and bands in a wheat sample with WT *Pinb-D1a* allele, and 129 bp, 95 bp, and 223 bp bands in wheat sample with *Pinb-D1b* mutation as explained in Section 2.4.4 (Fig. 3.3). *Pinb-D1b* was present in 12 (Amery, Batavia, Chara, Clearfield Stilleto, Cunningham, Diaz, Halbred, Janz, Katepwa, Sunco, Sunstar, Tasman) of 22 cultivars.

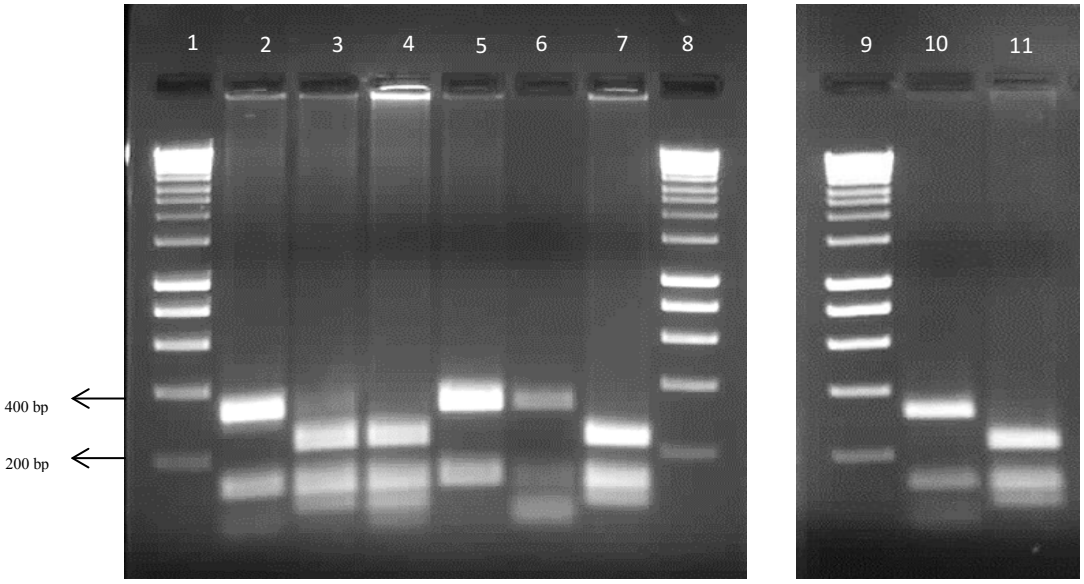



**Figure 3.3. Example of BsrB1 digested PCR products of *Pinb***
Lane 1: Hyperladder 1, lane 2: CD87, lane 3: Katepwa, lane 4: Janz, lane 5: Kukri, lane 6: Cranbrook, lane 7: Halbred, lane 8: Hyperladder 1, lane 9: Hyperladder 1, lane 10: Egret, lane 11: Sunstar

### 3.3.4 Amplification of *Pinb-2* genes in Australian wheat cultivars

The amplification of *Pinb-2* genes in wheat cultivars (round I PCR) was performed using three degenerate primers (Table 2.12) as pair D1 (Pinb-2F/ Pinb-2D1R) or D2 (Pinb-2F/Pinb-2D2R) using 200 ng of genomic DNA as template. These primers were designed as per Wilkinson et al. (2008). The Pinb-2D1 primer pair was used to amplify *Pinb-2* variants 1 (450bp), 3 (453bp), 5(450 bp) whereas D2 primer pair was used to amplify *Pinb-2* variants 2 (453 bp), 3(453 bp), 4 (450 bp), 6 (450 bp). Amplicons of approximately 450 base pairs (bp) were obtained for all the cultivars for both Pinb-2 D1 and Pinb-2 D2 primer pairs. An example of *Pinb-2* PCR amplifications in wheat cultivars is shown in Fig. 3.4. The results for *Pinb-2* gene amplifications in wheat cultivars are presented in Table 3.1.

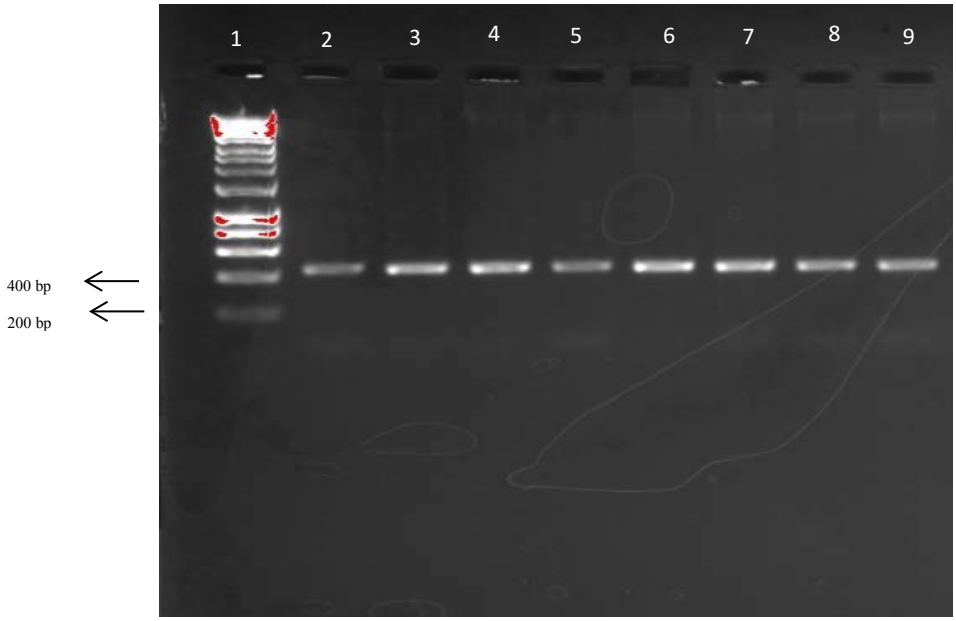



**Figure 3.4 Example of the amplification *Pinb-2 PCR* products**
Lane 1: Hyperladder I, lanes 2-5: *Pinb-2* PCR products generated using primer pair D1; Lanes 6 – 9: *Pinb-2* PCR products generated using primer pair D2.

**Table 3.1: PCR amplification of *Pina, Pinb, Pinb-2* genes and SKCS value of Australian wheat cultivars**

| S.no | Cultivar | AWCC no. | Class | Grain hardness (SKCS) | *Pina*/*Pinb* genotype* | PinaD1 PCR amplification | PinbD1 PCR amplification | *Pinb-D1b* | *Pinb-2D1* PCR amplification | *Pinb-2D2* PCR amplification |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Amery | 25598 | Hard | 67.24 | *Pina-D1a*/*Pinb-D1b* | ✓ | ✓ | ✓ | ✓ | ✓ |
| 2 | Batavia | 25271 | Hard | 78.53 | *Pina-D1a*/*Pinb-D1b* | ✓ | ✓ | ✓ | ✓ | ✓ |
| 3 | Carnamah | 27194 | Hard | 63.22 | *Pina-D1b*/*Pinb-D1a* | × | ✓ | × | ✓ | ✓ |
| 4 | CD87 | 91169 | Soft | 70.27 | *Pina-D1a*/*Pinb-D1a* | ✓ | ✓ | × | ✓ | ✓ |
| 5 | Chara | 30031 | Hard | 71.62 | *Pina-D1a*/*Pinb-D1b* | ✓ | ✓ | ✓ | ✓ | ✓ |
| 6 | Clearfield Stilleto | 25923 | Hard | 55.41 | *Pina-D1a*/*Pinb-D1b* | ✓ | ✓ | ✓ | ✓ | ✓ |
| 7 | Cranbrook | 11612 | Hard | 56.04 | *Pina-D1b*/*Pinb-D1a* | × | ✓ | | ✓ | ✓ |
| 8 | Cunningham | 25046 | Hard | 51.43 | *Pina-D1a*/*Pinb-D1b* | ✓ | ✓ | ✓ | ✓ | ✓ |
| 9 | Diaz | 23326 | Hard | 77.03 | *Pina-D1a*/*Pinb-D1b* | ✓ | ✓ | ✓ | ✓ | ✓ |
| 10 | Egret | 16037 | Soft | 30.87 | *Pina-D1a*/*Pinb-D1a* | ✓ | ✓ | × | ✓ | ✓ |
| 11 | Falcon | 90252 | Soft | 35.42 | *Pina-D1a*/*Pinb-D1a* | ✓ | ✓ | × | ✓ | ✓ |
| 12 | Halbred | 99124 | Hard | 69.88 | *Pina-D1a*/*Pinb-D1b* | ✓ | ✓ | ✓ | ✓ | ✓ |
| 13 | Hartog | 21533 | Hard | 78.95 | *Pina-D1b*/*Pinb-D1a* | × | ✓ | × | ✓ | ✓ |
| 14 | Janz | 24794 | Hard | 68.77 | *Pina-D1a*/*Pinb-D1b* | ✓ | ✓ | ✓ | ✓ | ✓ |
| 15 | Katepwa | 24314 | Hard | 75.24 | *Pina-D1a*/*Pinb-D1b* | ✓ | ✓ | ✓ | ✓ | ✓ |
| 16 | Kukri | 29472 | Hard | 67.65 | *Pina-D1b*/*Pinb-D1a* | × | ✓ | × | ✓ | ✓ |
| 17 | Machete | 23038 | Hard | 70.37 | *Pina-D1b*/*Pinb-D1a* | × | ✓ | × | ✓ | ✓ |
| 18 | Ouyen | 25571 | Hard | 71.03 | *Pina-D1b*/*Pinb-D1a* | × | ✓ | × | ✓ | ✓ |
| 19 | Rosella | 25923 | Soft | 30.47 | *Pina-D1a*/*Pinb-D1a* | ✓ | ✓ | × | ✓ | ✓ |
| 20 | Sunco | 23455 | Hard | 71.98 | *Pina-D1a*/*Pinb-D1b* | ✓ | ✓ | ✓ | ✓ | ✓ |
| 21 | Sunstar | 22660 | Hard | 75.13 | *Pina-D1a*/*Pinb-D1b* | ✓ | ✓ | ✓ | ✓ | ✓ |
| 22 | Tasman | 25557 | Hard | 52.39 | *Pina-D1a*/*Pinb-D1b* | ✓ | ✓ | ✓ | ✓ | ✓ |

* The *Pina/Pinb* genotype of the cultivars was determined earlier (Pickering et al. 2007) and confirmed in this study

### 3.3.5 Development of nested PCR using allele specific primers for amplifications of variants v1-v4

*Pinb-2* variant 1 (v1), v2, v3 and v4 specific primers, were designed as per Chen et al. (2010 a) (listed in Table 2.12). These primers were then used with a second-round PCR using the purified Pinb-2D1 and Pinb-2D2 PCR products (see above) as templates. The expected product sizes are 319 bp for v1, 401 bp for v2,398 bp for v3 and 403 bp for v4. Pinb-2 D1 PCR product was used as template for second round PCR with variant 1 (v1) specific primers while Pinb-2 D2 PCR product with v2 and v4 specific primers. Positive and negative controls were included in the PCR reactions. The allele specific primers were confirmed for their specificity using the v1, v2, v3 and v4 specific clones obtained by cloning and sequencing of four cultivars: Tasman, Sunco, Egret, Sunstar as described in Section 3.3.8. The PCR product obtained using Pinb-2v1 primers and v1 specific clone as template was approximately 319 bp as expected and no amplification was obtained for other variant types used as negative controls (Fig.3.5). Similarly, the specificity of Pinb-2v2, Pinb-2v3 and Pinb-2v4 primers was established as detailed in Section 3.3.6. The results confirmed the specificity of allele specific primers. Hence these primers were used genotyping all 22 cultivars (second round PCRs) and the variant specific clones were used as controls for presence/absence of PCR products.

### 3.3.6 Genotyping of wheat cultivars into variant 1, variant 2, variant 3 and variant 4 types using *Pinb-2* variant specific primers

Genotyping of 22 cultivars was performed using *Pinb-2* variant 1 (v1), variant 2 (v2), variant 3 (v3) and variant 4(v4) specific primers. These primers were used for second round PCR of purified Pinb-2D1/Pinb-2D2 PCR products. As mentioned earlier in Section 3.3.4 the Pinb-2D1 primer pair was used to amplify v1, v3 whereas D2 primer pair was used to amplify v2, v3, v4. Thus, purified Pinb-2D1 PCR product was used as template for amplification with Pinb-2v1and Pinb-2v3 primers while purified Pinb-2D2 PCR product was used as template for amplification with Pinb-2v2 and Pinb-2v4 primers in the second-round PCR. All cultivars showed positive amplification for v1 in the second-round PCR of purified Pinb-2D1 PCR product (template), yielding 319 bp product as expected (Table 2.12) (Fig. 3.5), suggesting the presence of v1 in these. The background 450 bp bands observed in Fig. 3.5 are the template (round 1 PCR product with Pinb-2D1 primers). All cultivars also yielded the expected 403 bp product for v4

PCR of purified Pinb-2D2 PCR product (Fig. 3.6), suggesting the presence of v4 also. The above-mentioned v1 clone was used as positive control in the PCR for v1, and showed an amplicon of the expected size (319 bp), and its insert was not amplified with v4 specific primers. Similarly, a v4 clone showed the expected amplicon (403 bp) in v4 PCR and none in the v1 PCR. This confirmed the specificity of the primers used.

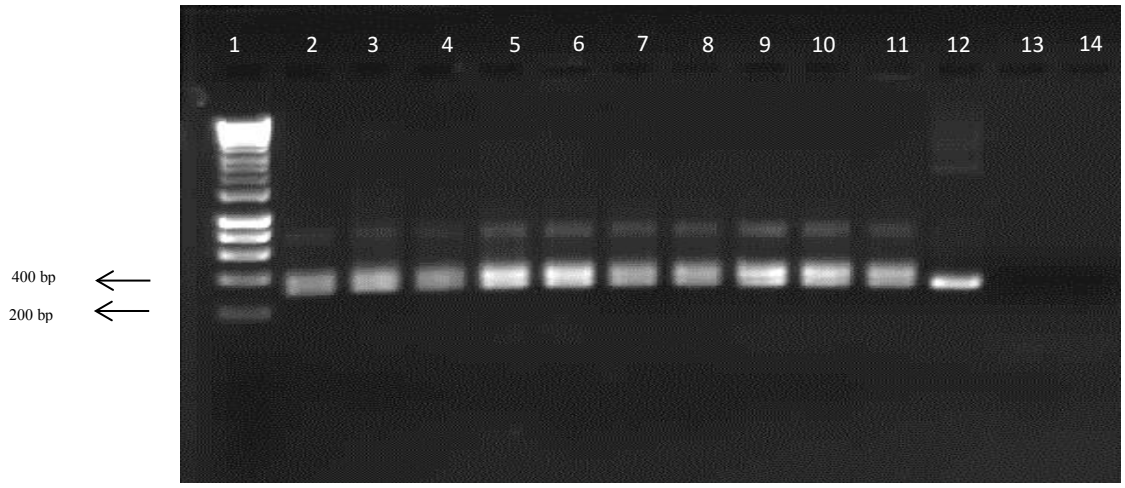



**Figure 3.5. Example of *Pinb-2v1* PCR products of wheat cultivars**.
Lane1: Hyperladder 1, lane 2: Amery, lane 3: Batavia, lane 4: Carnamah, lane 5: Chara, lane 6: Cunningham, lane 7: Diaz, lane 8: Falcon soft, lane 9: Hartog, lane 10: Machete, lane 11: Ouyen, lane 12: positive control 1(variant 1 specific clone), lane 13: negative control 1 (variant 4 specific clone), lane 14: negative control 2 (no template)

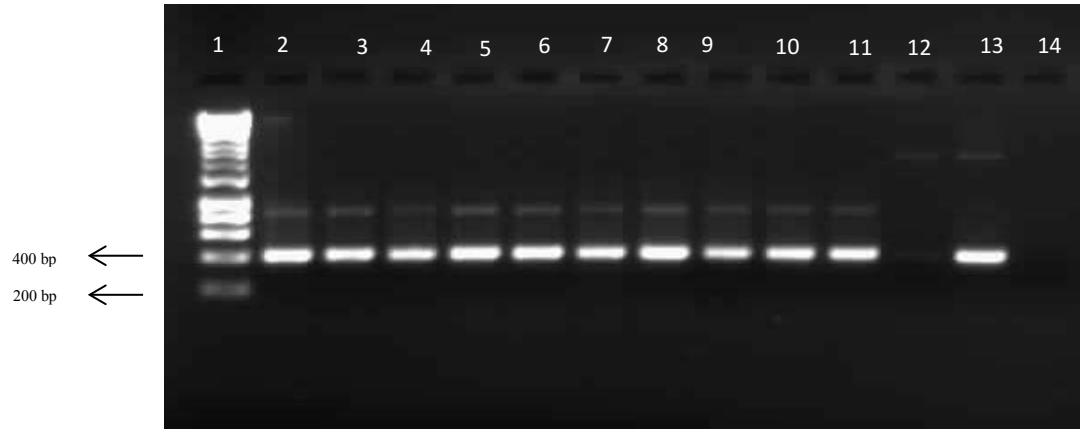



**Figure 3.6. *Pinb-2v4* PCR products of wheat cultivars**.
Lane1: Hyperladder 1, lane 2: Amery, lane 3: Batavia, lane 4: Carnamah, lane 5: Chara, lane 6: Cunningham, lane 7: Diaz, lane 8: Falcon soft, lane 9: Hartog, lane 10: Machete, lane 11: Ouyen, lane 12: negative control 1(variant 1 specific clone), lane 13: positive control (variant 4 specific clone), lane 14: negative control 2 (no template)

A 401bp PCR product was observed in v2 PCR products of purified Pinb-2D2 PCR product used as template (Fig. 3.7). Variant 2 PCR amplification was observed in four cultivars (Amery, Halbred, Kukri, Ouyen). An example of 401bp v2 PCR product observed in cultivar Amery (lane 2, Fig.3.7) and Ouyen (lane 11) is shown in Fig. 3.7. It is worth mentioning that the background bands observed in lanes 3 to 10 of Fig. 3.7 are the unamplified template (purified Pinb-2D2 PCR product). In case of v3 PCR reactions, successful amplifications of purified Pinb-2D1 PCR product (template), yielding an expected PCR product of 398 bp (Fig. 3.8) was observed in all except the four cultivars that had shown positive amplification for v2, i.e. Amery, Halbred, Kukri and Ouyen. The results suggested that v2 and v3 are not present together in any cultivar. A v2 specific clone was used as positive control in v2 PCR and negative control in v3 PCR. Similarly, variant 3 specific clone was used as positive control in v3 PCR and as negative control in v2 PCR, and all controls gave the expected results. The *Pinb-2* variant genotypes of cultivars are shown in Table 3.3.

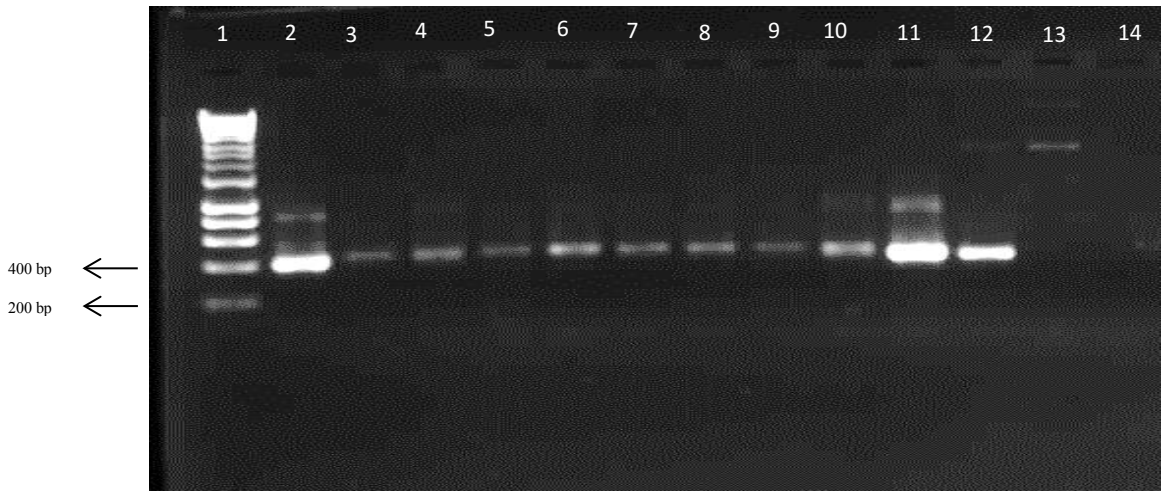



**Figure 3.7 Example of *Pinb-2v2* PCR products of wheat cultivars**.
Lane1: Hyperladder 1, lane 2: Amery, lane 3: Batavia, lane 4: Carnamah, lane 5: Chara, lane 6: Cunningham, lane 7: Diaz, lane 8: Falcon soft, lane 9: Hartog, lane 10: Machete, lane 11: Ouyen, lane 12: positive control (variant 2 specific clone), lane 13: negative control 1 (variant 3 specific clone), lane 14: negative control 2 (no template)

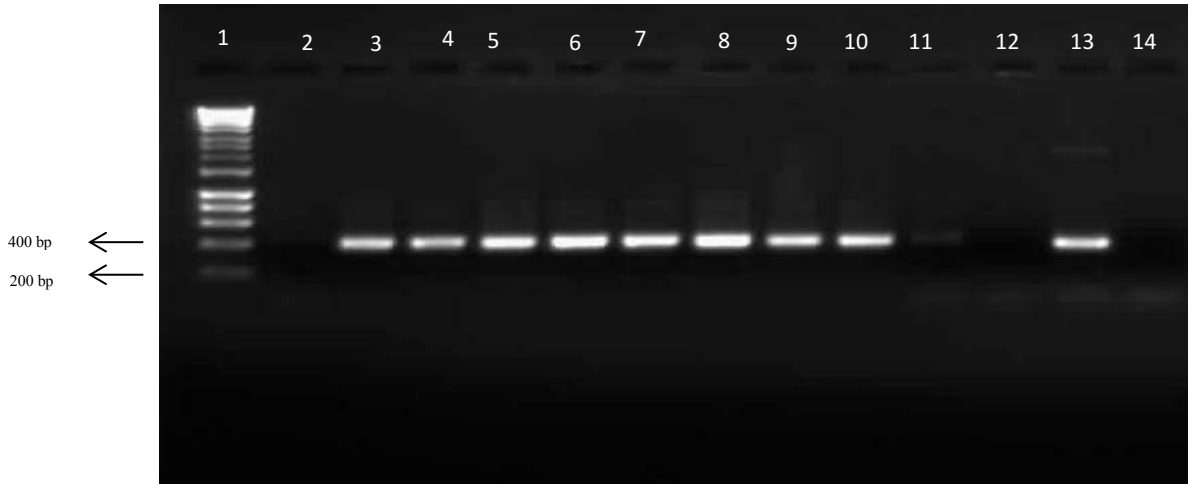
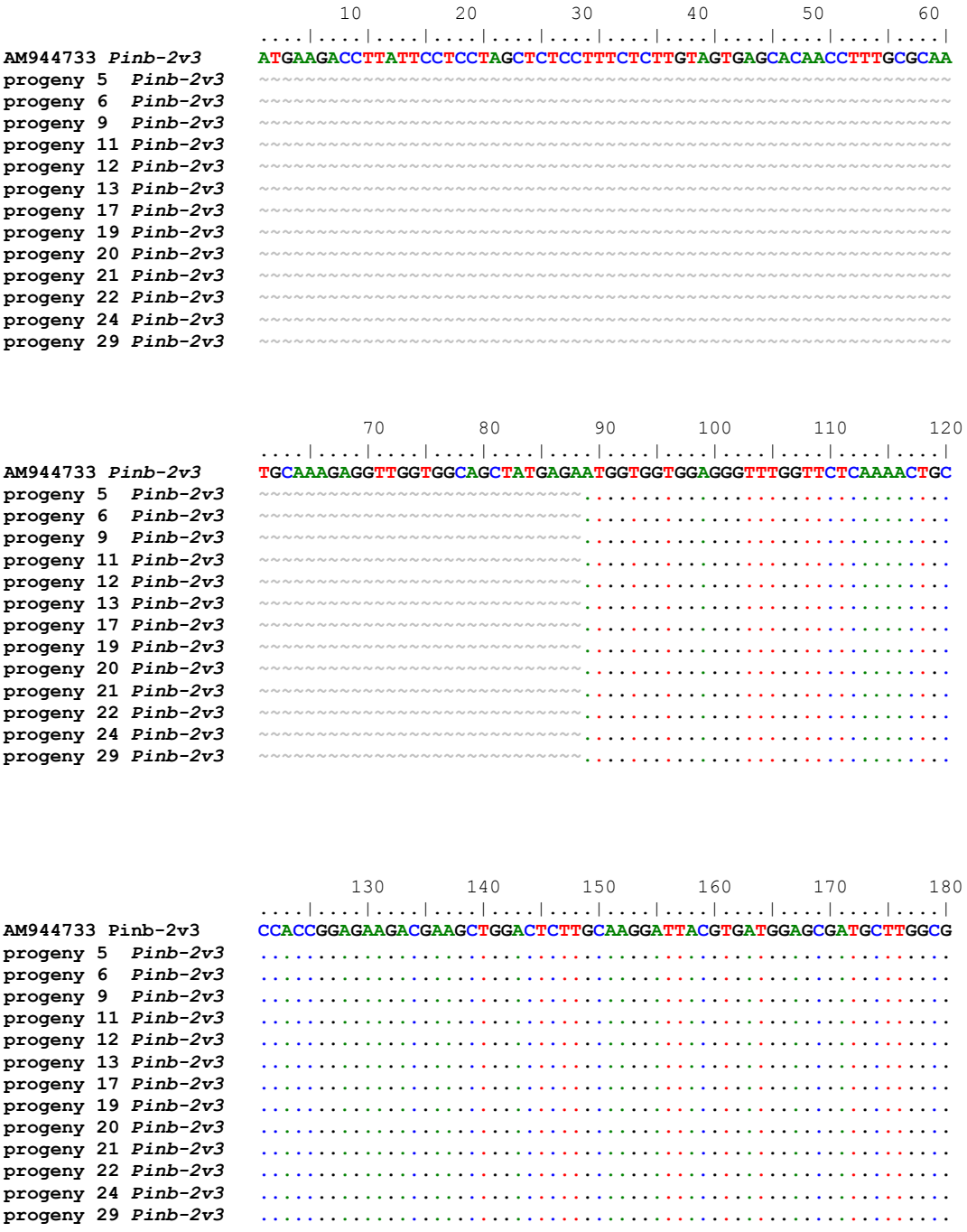


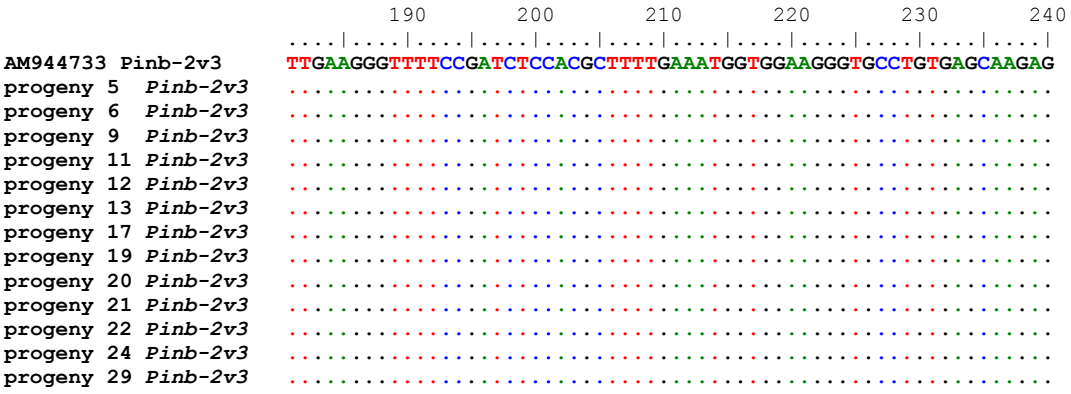

**Figure 3.8. Example of *Pinb-2v3* PCR products of wheat cultivars**.
Lane1: Hyperladder 1, lane 2: Amery, lane 3: Batavia, lane 4: Carnamah, lane 5: Chara, lane 6: Cunningham, lane 7: Diaz, lane 8: Falcon soft, lane 9: Hartog, lane 10: Machete, lane 11: Ouyen, lane 12: negative control 1(variant 2 specific clone), lane 13: positive control (variant 3 specific clone), lane 14: negative control (no template)

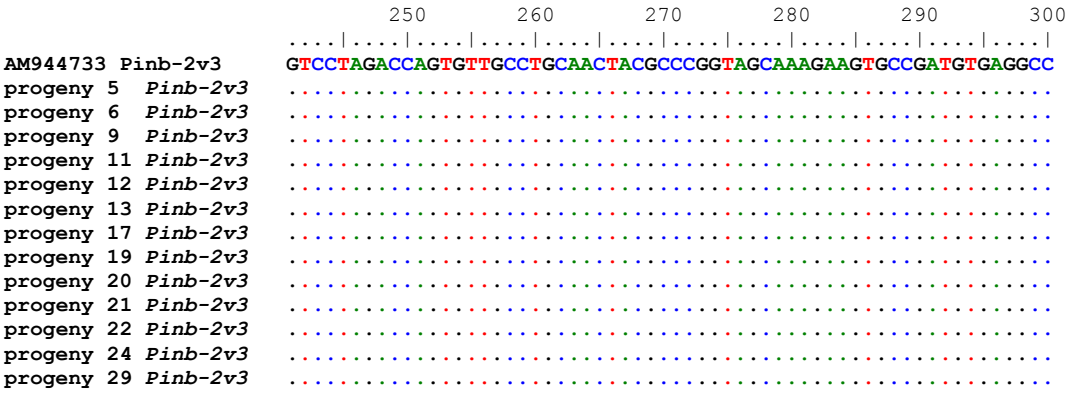
### 3.3.7 Detection of Val104Ala substitution in variant 3 associated with harder grain texture

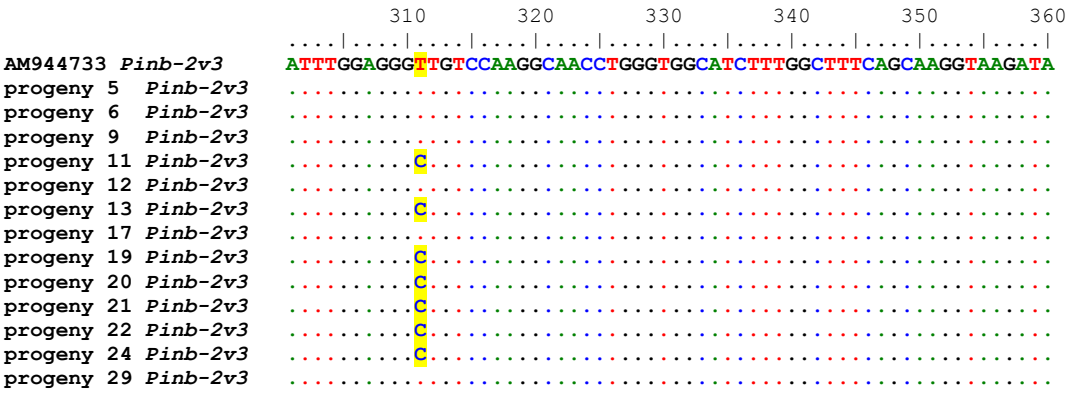
All the cultivars possessing variant 3 were investigated for the presence of T311C SNP resulting in Val104Ala substitution. This was undertaken by sequencing the purified *Pinb-2v3* PCR products. The PCR products were directly sequenced using forward primer. The DNA sequences obtained were aligned with reference sequence, *Pinb-2v3-1* (Genbank*:*AM944733). Eight of 18 cultivars showed a T311C SNP that results in Val104Ala substitution when compared with the reference sequence. The eight cultivars were: Batavia, Chara, Cranbrook Hartog, Janz, Katepwa, Machete, Tasman. In addition to above cultivars 13 doubled- haploid progeny lines of a cross Tasman x Sunco were also investigated for the presence of Val104Ala substitution. The doubled-haploid progeny lines from 'Tasman x Sunco' cross have been used earlier to construct genetic maps depicting the markers associated with grain hardness (Osborne et al.2001). Val104Ala was detected in one of the parents i.e. cv. Tasman. Therefore, the (Tasman x Sunco) progenies are expected to show segregation for this mutation. The doubled-haploid lines provide a more controlled means of examining the segregation of *Pinb-2v3* alleles. The 13 progeny lines were chosen at random. Firstly, the progenies were investigated for their *puroindoline* alleles. The *Pina-D1* and *Pinb-D1* genes were amplified using the primers listed in Table 2.12, design based on Gautier et al. (1994). The PCR amplifications indicated the presence of *Pina-D1* and *Pinb-D1*in all 13 progeny

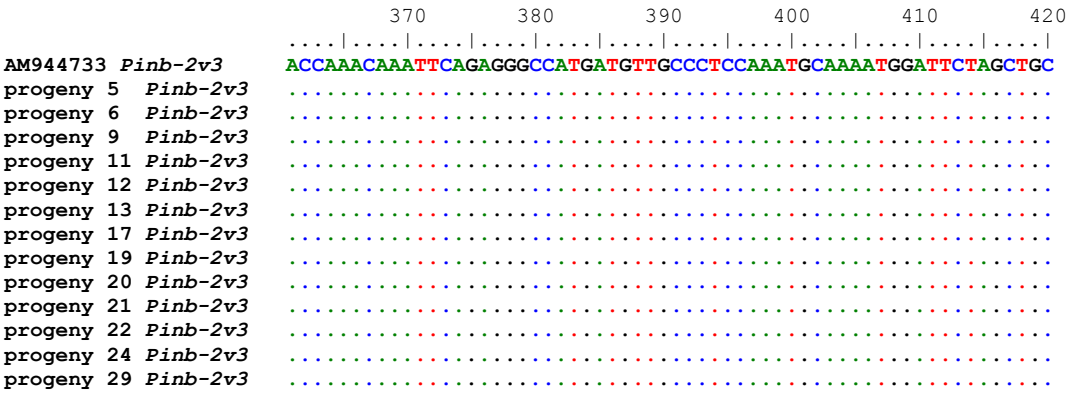
lines. Both the parents Tasman and Sunco contained *Pinb-D1b* allele, thus the progenies were also assumed to have *Pinb-D1b* allele. The sequencing of *Pinb-2v3* PCR products of the progenies revealed that seven of 13 progenies contained T311C SNP indicating the presence of Val104Ala substitution. The results are presented in Fig. 3.9 and Table 3.2.




190 200 210 220 230 240
AM944733 Pinb-2v3 TTGAAGGGTTTTCCGATCTCCACGCTTTTGAAATGGTGGAAGGGTGCCTGTGAGCAAGAG
progeny 5 Pinb-2v3
progeny 6 Pinb-2v3
progeny 9 Pinb-2v3
progeny 11 Pinb-2v3
progeny 12 Pinb-2v3
progeny 13 Pinb-2v3
progeny 17 Pinb-2v3
progeny 19 Pinb-2v3
progeny 20 Pinb-2v3
progeny 21 Pinb-2v3
progeny 22 Pinb-2v3
progeny 24 Pinb-2v3
progeny 29 Pinb-2v3


250 260 270 280 290 300
AM944733 Pinb-2v3 GTCCTAGACCAGTGTTGCCTGCAACTACGCCCGGTAGCAAAGAAGTGCCGATGTGAGGCC
progeny 5 Pinb-2v3
progeny 6 Pinb-2v3
progeny 9 Pinb-2v3
progeny 11 Pinb-2v3
progeny 12 Pinb-2v3
progeny 13 Pinb-2v3
progeny 17 Pinb-2v3
progeny 19 Pinb-2v3
progeny 20 Pinb-2v3
progeny 21 Pinb-2v3
progeny 22 Pinb-2v3
progeny 24 Pinb-2v3
progeny 29 Pinb-2v3


310 320 330 340 350 360
AM944733 Pinb-2v3 ATTTGGAGGGTTGTCCAAGGCAACCTGGGTGGCATCTTTGGCTTTCAGCAAGGTAAGATA
progeny 5 Pinb-2v3
progeny 6 Pinb-2v3
progeny 9 Pinb-2v3
progeny 11 Pinb-2v3 C
progeny 12 Pinb-2v3
progeny 13 Pinb-2v3 C
progeny 17 Pinb-2v3
progeny 19 Pinb-2v3 C
progeny 20 Pinb-2v3 C
progeny 21 Pinb-2v3 C
progeny 22 Pinb-2v3 C
progeny 24 Pinb-2v3 C
progeny 29 Pinb-2v3


370 380 390 400 410 420
AM944733 Pinb-2v3 ACCAAACAAATTCAGAGGGCCATGATGTTGCCCTCCAAATGCAAAATGGATTCTAGCTGC
progeny 5 Pinb-2v3
progeny 6 Pinb-2v3
progeny 9 Pinb-2v3
progeny 11 Pinb-2v3
progeny 12 Pinb-2v3
progeny 13 Pinb-2v3
progeny 17 Pinb-2v3
progeny 19 Pinb-2v3
progeny 20 Pinb-2v3
progeny 21 Pinb-2v3
progeny 22 Pinb-2v3
progeny 24 Pinb-2v3
progeny 29 Pinb-2v3

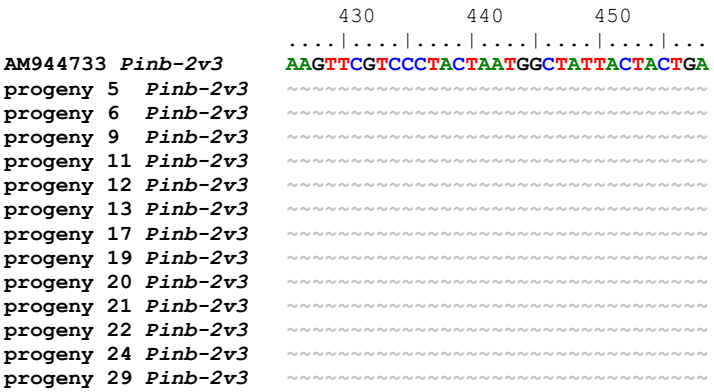



**Figure 3.9 Example of *Pinb-2v3* DNA sequence alignment showing haplotypes of (Tasman x Sunco) progeny lines.**

Dots indicate nucleotides identical to the top sequence in that particular group and˜indicates gap.

**Table 3.2: Status of Val104Ala substitution in (Tasman x Sunco) progeny lines**

| Progeny no. | Val104Ala* |
|---|---|
| 5 | × |
| 6 | × |
| 9 | × |
| 11 | ✓ |
| 12 | × |
| 13 | ✓ |
| 17 | × |
| 19 | ✓ |
| 20 | ✓ |
| 21 | ✓ |
| 22 | ✓ |
| 24 | ✓ |
| 29 | × |

*The presence of Val104Ala in Variant 3is indicated by ✓ and absence by ×

**Table 3.3: *Pinb-2* variant genotypes of Australian wheat cultivars**

| S.no | Cultivar | *Pina*/*Pinb* genotype | SKCS | V1 | V2 | V3 | Val104 Ala* | V4 |
|---|---|---|---|---|---|---|---|---|
| 1 | Amery | *Pina-D1a*/*Pinb-D1b* | 67.24 | ✓ | ✓ | × | - | ✓ |
| 2 | Batavia | *Pina-D1a*/*Pinb-D1b* | 78.53 | ✓ | × | ✓ | ✓ | ✓ |
| 3 | Carnamah | *Pina-D1b*/*Pinb-D1a* | 63.22 | ✓ | × | ✓ | × | ✓ |
| 4 | CD87 | *Pina-D1a*/*Pinb-D1a* | 70.27 | ✓ | × | ✓ | × | ✓ |
| 5 | Chara | *Pina-D1a*/*Pinb-D1b* | 71.62 | ✓ | × | ✓ | ✓ | ✓ |
| 6 | Clearfield Stilleto | *Pina-D1a*/*Pinb-D1b* | 55.41 | ✓ | × | ✓ | × | ✓ |
| 7 | Cranbrook | *Pina-D1b*/*Pinb-D1a* | 56.04 | ✓ | × | ✓ | ✓ | ✓ |
| 8 | Cunningham | *Pina-D1a*/*Pinb-D1b* | 51.43 | ✓ | × | ✓ | × | ✓ |
| 9 | Diaz | *Pina-D1a*/*Pinb-D1b* | 77.03 | ✓ | × | ✓ | × | ✓ |
| 10 | Egret | *Pina-D1a*/*Pinb-D1a* | 30.87 | ✓ | × | ✓ | × | ✓ |
| 11 | Falcon | *Pina-D1a*/*Pinb-D1a* | 35.42 | ✓ | × | ✓ | × | ✓ |
| 12 | Halbred | *Pina-D1a*/*Pinb-D1b* | 69.88 | ✓ | ✓ | × | × | ✓ |
| 13 | Hartog | *Pina-D1b*/*Pinb-D1a* | 78.95 | ✓ | × | ✓ | ✓ | ✓ |
| 14 | Janz | *Pina-D1a*/*Pinb-D1b* | 68.77 | ✓ | × | ✓ | ✓ | ✓ |
| 15 | Katepwa | *Pina-D1a*/*Pinb-D1b* | 75.24 | ✓ | × | ✓ | ✓ | ✓ |
| 16 | Kukri | *Pina-D1b*/*Pinb-D1a* | 67.65 | ✓ | ✓ | × | × | ✓ |
| 17 | Machete | *Pina-D1b*/*Pinb-D1a* | 70.37 | ✓ | × | ✓ | ✓ | ✓ |
| 18 | Ouyen | *Pina-D1b*/*Pinb-D1a* | 71.03 | ✓ | ✓ | × | - | ✓ |
| 19 | Rosella | *Pina-D1a*/*Pinb-D1a* | 30.47 | ✓ | × | ✓ | × | ✓ |
| 20 | Sunco | *Pina-D1a*/*Pinb-D1b* | 71.98 | ✓ | × | ✓ | × | ✓ |
| 21 | Sunstar | *Pina-D1a*/*Pinb-D1b* | 75.13 | ✓ | × | ✓ | × | ✓ |
| 22 | Tasman | *Pina-D1a*/*Pinb-D1b* | 52.39 | ✓ | × | ✓ | ✓ | ✓ |

*The presence of Val104Ala in Variant 3 is indicated by ✓ and absence by ×

### 3.3.8 Study of *Pinb-2* haplotypes in selected wheat cultivars

Genetic variation in *Pinb-2* haplotypes of four Australian cultivars was investigated by using the methods of gene cloning and sequencing. The aim of this study was to analyse genetic variation in *Pinb-2* haplotypes between wheat cultivars in relation to *Pin-D1* genotype/ *Pinb-2v3* genotype. The four cultivars chosen for analysis were Tasman, Sunco, Egret, and Sunstar. Of these four cultivars, only Egret is considered soft and rest are hard grain (Osborne et al. 2001). Three cultivars i.e. Tasman, Sunco, Sunstar had *Pina-D1a/Pinb-D1b* genotype while the soft cultivar, Egret had *Pina-D1a/Pinb-D1a* genotype. The Val104Ala substitution (associated with hardness) was present in one cultivar i.e. Tasman and the rest had the *Pinb-2v3-1* allele. As discussed earlier, 13 doubled- haploid progenies of Tasman x Sunco cross were investigated for the presence of Val104Ala substitution in variant 3. Thus, it was necessary to find out if there was any other *Pinb-2* genetic variation between the two parents apart from the T311C SNP

(Val104Ala) with possible implication on texture. The two remaining cultivars, Egret and Sunstar are also the parents of genetic mapping populations (Osborne et al.2001). Though both have *Pinb-2v3-1* allele but Egret (*Pina-D1a*/ *Pinb-D1a*) is soft and Sunstar is hard (*Pina-D1a*/ *Pinb-D1b*). The discovery of any potentially important genetic variation in *Pinb-2* genes in the parental cultivars might prove useful as it can be introgressed into breeding programs for grain-texture variation.

### 3.3.8 a) Amplification of *Pina/ Pinb/Pinb-2* genes and confirmation of *Pin* genotype

Two primer pairs (Table 2.12) designed on the wild-type *Pin* genes, PINA-D1F/PINA-D1R and PINB-D1F/PINB-D1R, were used to amplify the *Pina-D1* and *Pinb-D1* genes. All four cultivars showed *Pina-D1* and *Pinb-D1* gene amplifications as mentioned in Section 3.3.2 (Table 3.1). The *Pinb-D1* PCR products were subjected to digestion with the restriction enzyme BsrBI, to detect the presence of *Pinb-D1b.* Three cultivars i.e. Tasman, Sunco and Sunstar showed the presence of *Pinb-D1b* (Section 3.3.3, Table 3.1). The amplification of *Pinb-2* genes in wheat cultivars was undertaken using three degenerate primers (Table 2.12) as pair D1 (Pinb-2F/ Pinb-2D1R) or D2 (Pinb-2F/Pinb-2D2R). Positive amplification was obtained for both D1 and D2 primer pairs (Section 3.3.1, Table 3.2). The results are summarised in Table 3.4.

**Table 3.4 Wheat samples used in *Pinb-2* gene analysis**

| Cultivar | *Pina-D1a* allele | *Pinb-D1a* allele | Hard/Soft |
|---|---|---|---|
| Tasman | a | b | Hard |
| Sunco | a | b | Hard |
| Egret | a | a | Soft |
| Sunstar | a | b | Hard |

### 3.3.8 b) Cloning and sequencing of *Pinb-2* genes

All *Pinb-2* PCR products obtained using the D1 and D2 primer pairs (example shown in Fig. 3.4) were cloned into pGEM-T Easy vectors (Promega, Australia) using *E. coli* JM109 as the host prior to sequencing as these PCR products are likely to contain a mixture of haplotypes. The primer pair D1 is expected to amplify *Pinb-2v1*, *Pinb-2v3* and *Pinb-2v5* groups, while the pair D2 will amplify the members of the *Pinb-2v2*, *Pinb-2v3*, *Pinb-2v4* and *Pinb-2v6* groups (Ramalingam et al. 2012). The *Pinb-2* genes likely exist

as a multigene family in wheat (Chen et al. 2010a), thus cloning was necessary for the analysis of individual haplotypes.

Twenty colonies were used for plasmid preparations (Fig. 3.10). Good quality plasmids were obtained for all plasmid preparations. Screening for recombinant clones was carried out by PCR amplification of plasmids using vector specific primers i.e., T7 and SP6 (Table 2.12). The expected size of the amplicon was ~630bp (~450bp inset + 70bp and 98bp flanking regions; Fig. 3.11). Up to 15 recombinant clones were sequenced for each cultivar in the first round. DNA sequencing reactions were performed using sequencing primer i.e., T7. Only the sequences with clear and unambiguous chromatograms were used for analysis. A second round of PCR, gene cloning and sequencing was conducted to confirm the mutations found in round I PCR and sequencing and reduce any PCR-induced errors.

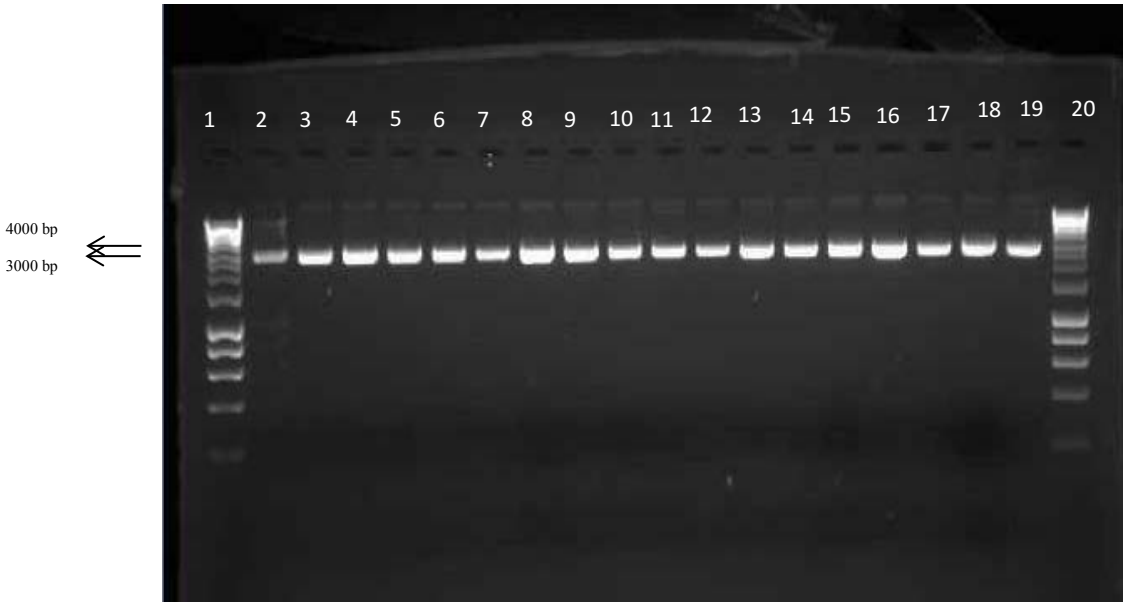



**Figure 3.10. Example of the plasmid preparations containing *Pinb-2* clones**
Lane 1, 20: Hyperladder 1, lanes 2 to 19: plasmids preps 1-18.

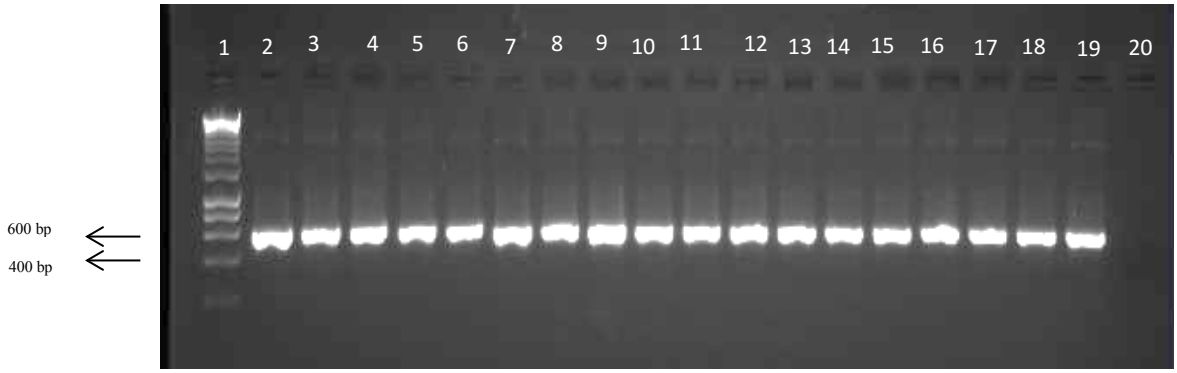
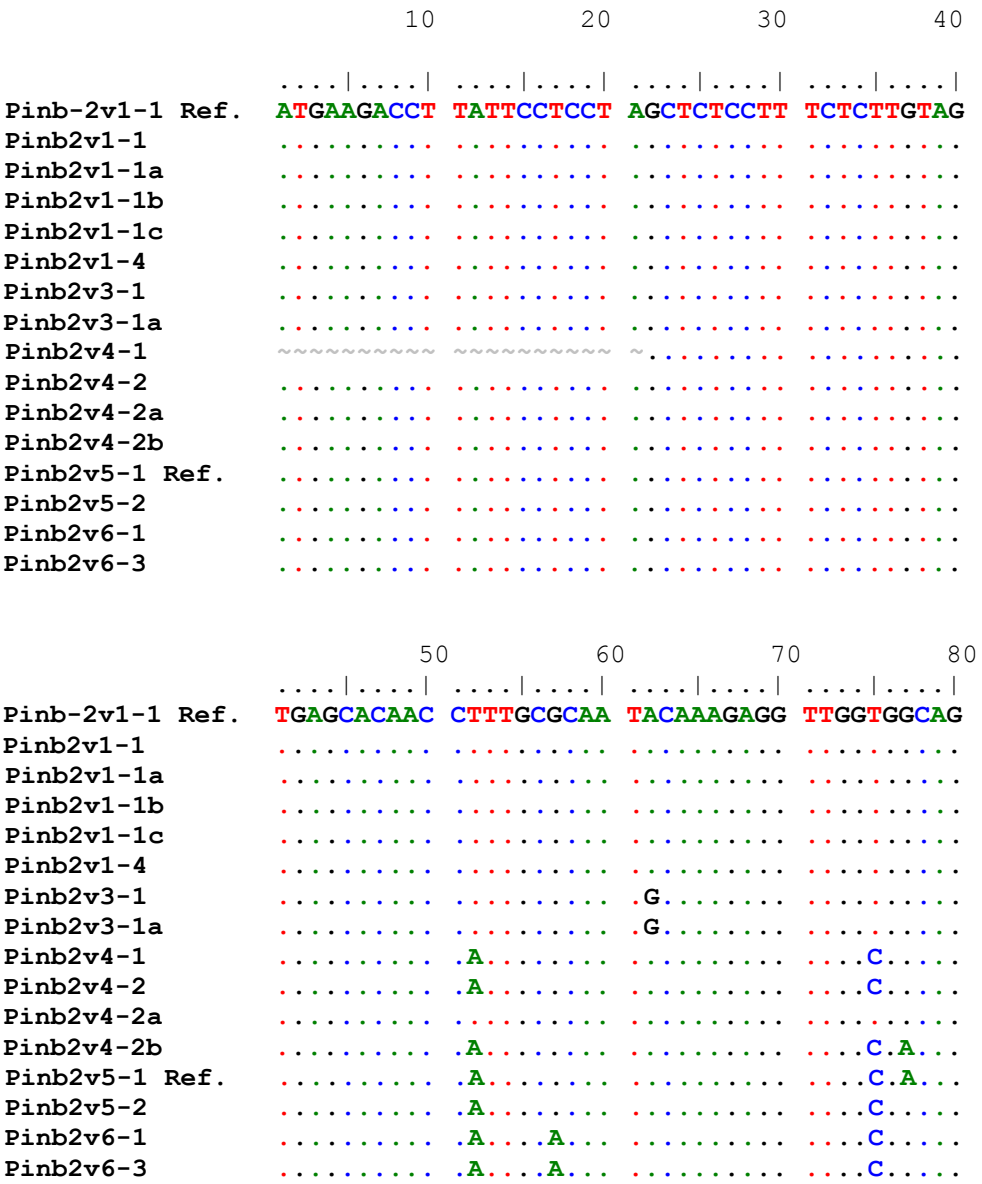



**Figure 3.11. Example of PCR products (using T7, SP6 primers) to screen for *Pinb-2* inserts**. Lane 1: Hyperladder1, lanes 2-19: amplified inserts from plasmids containing *Pinb-2* clones, lane 20: negative control (no template).

### 3.3.8 c) *Pinb-2* haplotypes in wheat cultivars

The sequences were aligned manually with the reported *Pinb-2v1* (AM944731, GQ496616), *Pinb-2v2* (AM944732, GQ496617), *Pinb-2v3* (AM944733, GQ496618) (AM944731-AM944733: Wilkinson et al. 2008; GQ496616-496618: Chen et al. 2010a), *Pinb-2v4* (GQ496619, Chen et al. 2010a), *Pinb-2v5* (HM245236*;* Chen et al. 2011), and *Pinb-2v6* (JN585979*;* Ramalingam et al. 2012) using Bioedit. *Pinb-2v1* was used as a reference sequence for alignment of sequences and inferences were drawn based on unique/shared regions with other sequences. The reported *Pinb-2* genes such as *Pinb-2v1, Pinb-2v2, Pinb-2v3, Pinb-2v4v4* and *Pinb-2v5* haplotypes was redesignated *Pinb-2v1-1, Pinb-2v2-1, Pinb-2v3-1*, *Pinb-2v4-1* and *Pinb-2v5-1*. The sequences that were 100% identical to the reported variant types were assigned the same haplotype, whereas those that differed were designated for readability as different subtypes of the main haplotypes e.g. *Pinb-2v1-1a*, *Pinb-2v1-1b* etc. Thus, a total of 14 haplotypes were obtained (Fig. 3.12, Table 3.5). Eight haplotypes have been reported earlier (Genbank accessions provided in Table 3.5) and six are novel (*Pinb-2v1-1a*, *Pinb-2v1-b, Pinb-2v1-c*, *Pinb-2v3-1a*, *Pinb-2v4-2a*, *Pinb-2v4-2b*). Four of 14 haplotypes were shared among cultivars. These haplotypes were *Pinb-2v1-1*, *Pinb-2v1-4*, *Pinb-2v3-1*, and *Pinb-2v4-2*. Sequence subtypes of the main haplotypes were found for cultivar Egret (*Pinb2v1*: *Pinb2v1-1a*, *Pinb2v1-1b, Pinb2v1-1c*), Sunco (*Pinb-2v4: Pinb2v4-2, Pinb2v4-2a),* Tasman (*Pinb-2v3-1a*) and Egret (*Pinb-2v4-2b*). These subtypes have not been reported earlier. The haplotype *Pinb-2v3-1a* has the T311C SNP (Val104Ala) associated with grain hardness in an earlier study by Chen et al.2013. It was observed that the first 51 nucleotides of the *Pinb-2* genes were highly conserved while the 3' gene sections showed numerous SNPs

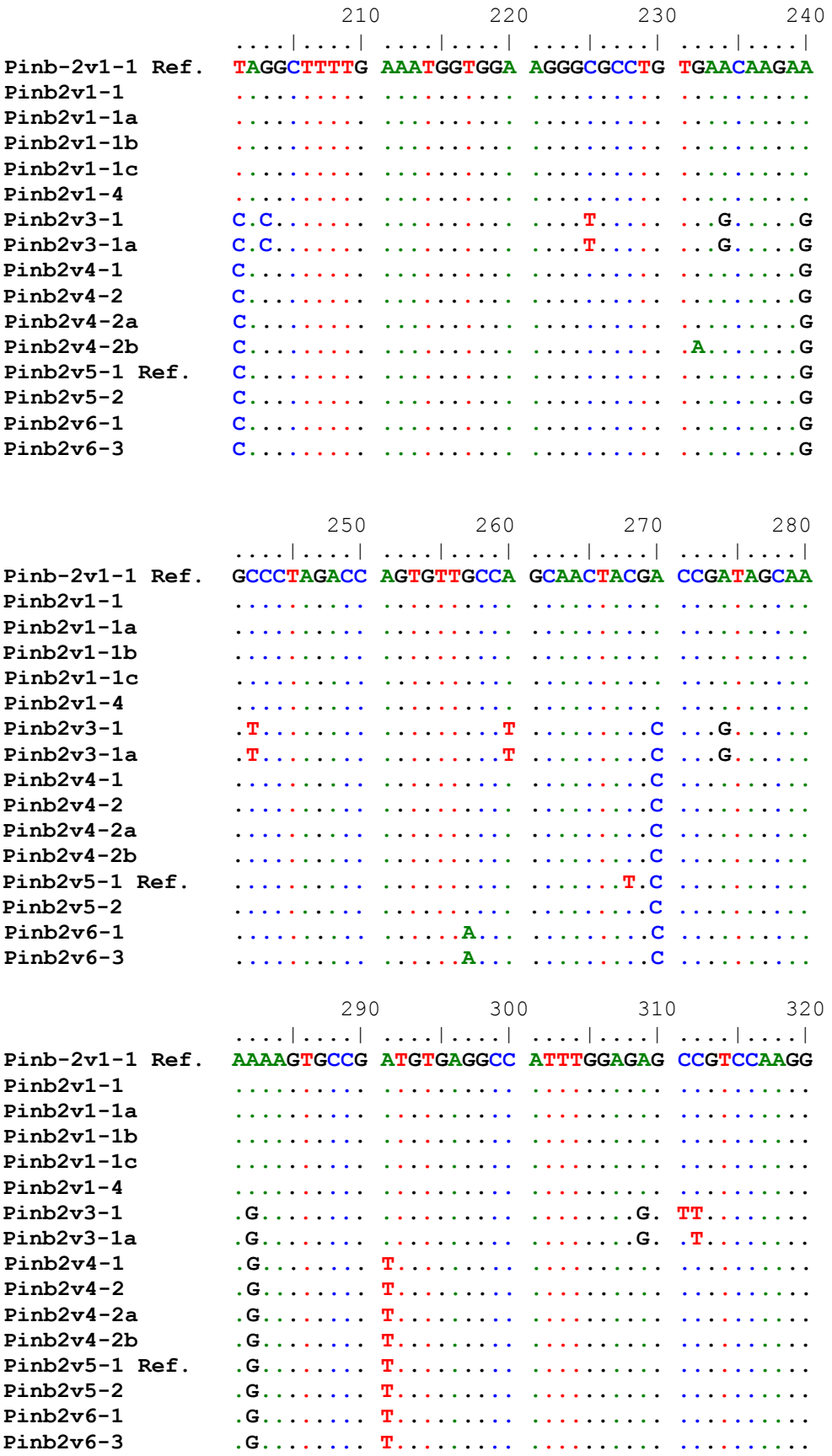
(Fig. 3.12). Multiple variations were observed in most of the sequences, unlike single point mutations found in *Pina-D1* or *Pinb-D1* (reviewed in Bhave and Morris, 2008a). Thus, the genes formed haplotypes, each having a 'characteristic combination of SNPs that are inherited together as a single gene' (Schwartz et al. 2004). Previously such variations in *Pinb-2* haplotypes were noted by Ramalingam et al. (2012).



```
                          90         100         110         120
                    ....|....| ....|....| ....|....| ....|....|
Pinb-2v1-1 Ref.     CTATGAGAAC GGTGGCGGAG GGTTTGGTTC TCAAAACTGC
Pinb2v1-1           .......... .......... .......... ..........
Pinb2v1-1a          .......... .......... .......... ..........
Pinb2v1-1b          .......... .......... .......... ..........
Pinb2v1-1c          .......... .......... .......... ..........
Pinb2v1-4           .......... .......... .......... ..........
Pinb2v3-1           .........T .....T.... .......... ..........
Pinb2v3b            .........T .....T.... .......... ..........
Pinb2v4-1           .......... .......... .......... ..........
Pinb2v4-2           .......... .......... .......... ..........
Pinb2v4-2a          .......... .......... .......... ..........
Pinb2v4-2b          .......... .......... .......... ..........
Pinb2v5-1 Ref.      .......... ....A..... .......... ..........
Pinb2v5-2           .......... .......... .......... ..........
Pinb2v6-1           .......... .......... .......... ..........
Pinb2v6-3           .......... .......... .......... ..........

                         130         140         150         160
                    ....|....| ....|....| ....|....| ....|....|
Pinb-2v1-1 Ref.     CCATCGGAGA AGACGAAGCT AGACCCTTGC AAGGATTATG
Pinb2v1-1           .......... .......... .......... ..........
Pinb2v1-1a          .......... .......... .......... ..........
Pinb2v1-1b          .......... .......... .......... ..........
Pinb2v1-1c          .......... .......... .......... ..........
Pinb2v1-4           .......... .......... .......... ..........
Pinb2v3-1           ...C...... .......... G...T..... .........C.
Pinb2v3-1a          ...C...... .......... G...T..... .........C.
Pinb2v4-1           ...C...... .......... G...T..... ..........
Pinb2v4-2           ...C...... .......... G...T..... ..........
Pinb2v4-2a          ...C...... .......... G...T..... ..........
Pinb2v4-2b          ...C...... .......... G...T..... ..........
Pinb2v5-1 Ref.      ...C...... .......... G...T..... ..........
Pinb2v5-2           ...C...... .......... G...T..... ..........
Pinb2v6-1           ...C...... .......... G...G..... ..........
Pinb2v6-3           ...C...... .......... G...G..... ..........

                         170         180         190         200
                    ....|....| ....|....| ....|....| ....|....|
Pinb-2v1-1 Ref.     TGATGGAGCG GTGCTTGGCT GTGAAGGGTT TTTCGATCGC
Pinb2v1-1           .......... .......... .......... ..........
Pinb2v1-1a          .......... .......... .......... ..........
Pinb2v1-1b          .......... .......... .......... ..........
Pinb2v1-1c          .......... .......... .......... ..........
Pinb2v1-4           .......... .......... .......... ..........
Pinb2v3-1           .......... A........G T.......... ..C.....T.
Pinb2v3-1a          .......... A........G T.......... ..C.....T.
Pinb2v4-1           .........T. .........G .......... ....A.....
Pinb2v4-2           .........T. .........G .......... ....A.....
Pinb2v4-2a          .........T. .........G .......... ....A.....
Pinb2v4-2b          .......... .........TG .......... ....A.....
Pinb2v5-1 Ref.      .......... .........G .......... ....A.....
Pinb2v5-2           .......... .........G .......... ....A.C.A.
Pinb2v6-1           .......... .........G .......... ....A.....
Pinb2v6-3           .......... .........G .......... ....A.....
```


210 220 230 240
Pinb-2v1-1 Ref. TAGGCTTTTG AAATGGTGGA AGGGCGCCTG TGAACAAGAA
Pinb2v1-1
Pinb2v1-1a
Pinb2v1-1b
Pinb2v1-1c
Pinb2v1-4
Pinb2v3-1 C.C.....T.....G.....G
Pinb2v3-1a C.C.....T.....G.....G
Pinb2v4-1 C.....G
Pinb2v4-2 C.....G
Pinb2v4-2a C.....G
Pinb2v4-2b C.....A.....G
Pinb2v5-1 Ref. C.....G
Pinb2v5-2 C.....G
Pinb2v6-1 C.....G
Pinb2v6-3 C.....G
250 260 270 280
Pinb-2v1-1 Ref. GCCCTAGACC AGTGTTGCCA GCAACTACGA CCGATAGCAA
Pinb2v1-1
Pinb2v1-1a
Pinb2v1-1b
Pinb2v1-1c
Pinb2v1-4
Pinb2v3-1 .T.....T.....C...G.....
Pinb2v3-1a .T.....T.....C...G.....
Pinb2v4-1 .....C.....
Pinb2v4-2 .....C.....
Pinb2v4-2a .....C.....
Pinb2v4-2b .....C.....
Pinb2v5-1 Ref. .....T.C.....
Pinb2v5-2 .....C.....
Pinb2v6-1 .....A.....C.....
Pinb2v6-3 .....A.....C.....
290 300 310 320
Pinb-2v1-1 Ref. AAAAGTGCCG ATGTGAGGCC ATTTGGAGAG CCGTCCAAGG
Pinb2v1-1
Pinb2v1-1a
Pinb2v1-1b
Pinb2v1-1c
Pinb2v1-4
Pinb2v3-1 .G.....G. TT.....
Pinb2v3-1a .G.....G. .T.....
Pinb2v4-1 .G.....T.....
Pinb2v4-2 .G.....T.....
Pinb2v4-2a .G.....T.....
Pinb2v4-2b .G.....T.....
Pinb2v5-1 Ref. .G.....T.....
Pinb2v5-2 .G.....T.....
Pinb2v6-1 .G.....T.....
Pinb2v6-3 .G.....T.....

```
                          330         340         350         360
                   ....|....| ....|....| ....|....| ....|....|
Pinb-2v1-1 Ref.    GAACCTTGGT GGCATCTTTG GTTTTCAGCA AGGCCAGATA
Pinb2v1-1          .......... .......... .......... ..........
Pinb2v1-1a         .......... .......... .......... ..........
Pinb2v1-1b         .......... .......... .......... ..........
Pinb2v1-1c         .......... .......... .......... ..........
Pinb2v1-4          .......... .......... .......... ..........
Pinb2v3-1          C.....G... .......... .C........ ...TA.....
Pinb2v3-1a         C.....G... .......... .C........ ...TA.....
Pinb2v4-1          C......... .......... .......... ..........
Pinb2v4-2          C......... .......... .......... ..........
Pinb2v4-2a         C......... .......... .......... ..........
Pinb2v4-2b         C......... .......... .......... ..........
Pinb2v5-1 Ref.     C......... .......... .......... ..........
Pinb2v5-2          C......... .......... .......... ..........
Pinb2v6-1          C......... .......... .......... ..........
Pinb2v6-3          C......... .......... .......... ..........

                              370         380         390         400
                   ....|....| ....|....| ....|....| ....|....|
Pinb-2v1-1 Ref.    ACCAAACAAA TTCAGAGGGC CATGATGTTG CCCTCCAAAT
Pinb2v1-1          .......... .......... .......... ..........
Pinb2v1-1a         .......... .......... .......... ..........
Pinb2v1-1b         .......... .......... .......... ..G.......
Pinb2v1-1c         .......... .......... .......... ..........
Pinb2v1-4          .......... .......... .......... ..........
Pinb2v3-1          .......... .......... .......... ..........
Pinb2v3-1a         .......... .......... .......... ..........
Pinb2v4-1          .......... .......... .......... ..G.......
Pinb2v4-2          .......... .......... .......... ..G.......
Pinb2v4-2a         .......... .......... .......... ..........
Pinb2v4-2b         .......... .......... .......... ..G.......
Pinb2v5-1 Ref.     .......... .......... .......... ..G.......
Pinb2v5-2          .......... .......... .......... ..G.......
Pinb2v6-1          .......... .......... .......... .TG.......
Pinb2v6-3          .......... .......... .......... .TG.......

                   410         420         430         440
                   ....|....| ....|....| ....|....| ....|....|
Pinb-2v1-1 Ref.    GCAAACTGGA TTCCAACTGC AAGTTTGTCG CTAATAATGG
Pinb2v1-1          .......... .......... .......... ..........
Pinb2v1-1a         .......... .......... .....C...C ...C......
Pinb2v1-1b         .....T.... .......... .....C...C ...C......
Pinb2v1-1c         .......... .....G.... .....C...C ...C......
Pinb2v1-4          .......... .......... .......... ...C......
Pinb2v3-1          .....A.... ...T.G.... .....C...C ...C......
Pinb2v3-1a         .....A.... ...T.G.... .....C...C ...C......
Pinb2v4-1          .....T.... .......... .....C.... ...C......
Pinb2v4-2          .....T.... .......... .......... ...C......
Pinb2v4-2a         .......... .......... .......... ...C......
Pinb2v4-2b         .....T.... .......... .......... ...C......
Pinb2v5-1 Ref.     .....T.... .......... .....C.... ..........
Pinb2v5-2          .....T.... ........T. .......... ...C......
Pinb2v6-1          .....T.... .......... .......... ...C......
Pinb2v6-3          .....T.... .......... .....C...C ...C......
```

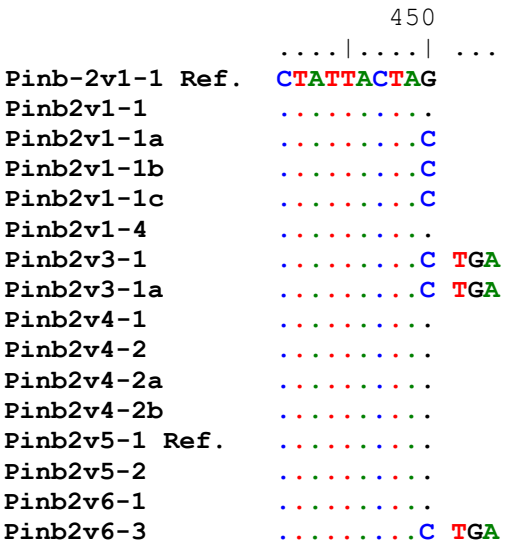



**Figure 3.12. *Pinb-2* haplotypes of wheat cultivars**
Dots indicate nucleotides identical to the top sequence in that particular group and˜indicates gap.

**Table 3.5: *Pinb-2* haplotypes and deduced protein types of selected wheat cultivars**

| DNA haplotype | Deduced protein type | Cultivar | *SNP(s) | Characteristic amino acid(s) | pI | Hydrophobic residues |
|---|---|---|---|---|---|---|
| *Pinb2v1-1* (GQ496616) | PINB2v1 | Tasman, Sunco, Sunstar | **T124,** A141**, C145,** A240, G321 | Ser42, Pro49 | 9.08 | 52 |
| *Pinb2v1-1a* | PINB2v1-1a | Egret | T426C, **G430C, A434C, G450C** | Ala144Pro, Asn145Thr, stop codon Tyr | 9.06 | 52 |
| *Pinb2v1-1b* | PINB2v1-1a | Egret | C393G, C406T, T426C, **G430C**, **A434C**, **G450C** | Ala144Pro, Asn145Thr, stop codon Tyr | 9.06 | 52 |
| *Pinb2v1-1c* | PINB2v1-1b | Egret | **A416G**, T426C, **G430C, A434C, G450C** | Asn139, Ala144Pro, Asn145Thr, stop codon Tyr | 9.08 | 52 |
| *Pinb2v1-4* (JN626913) | PINB2v1-4 | Sunco, Egret, Sunstar | **A434C** | Asn145Thr | 9.08 | 52 |
| *Pinb2v3-1* (GQ496618) | PINB2v3-1 | Sunco, Egret, Sunstar | **G62,** T90, T96, C124, G141, T145, C159, A171, G180, **T181**, **C193**, **T199**, **C201**, **C203**, T225, G234, G240, **T242**, G252, **T260**, C270, **G274**, G282, G309, T311, T312, C321, G327, C342, T354, **A355**, **A406,** T414, **G416**, C426, **C430**, **C434, C450** | 21Cys (signal peptide) 61Leu, 65Pro, 67Ser (caused by two SNPs), 68Thr, 81Val, 87Leu, 92Val, 119Lys, 136Met, 139Ser, 144Pro, 145Thr, 150Tyr | 9.05 | 52 |
| *Pinb-2v3-1a* | PINB2v3-1a | Tasman | **T311C** | Val104Ala | 9.05 | 52 |
| *Pinb2v4-1* (GQ496619) | PINB2v4-1 | Sunstar | **A52,** C75, **C124**, G141, **T145, T169**, G180, A195, C201, G240, C270, G282, T291, C321, G393, T406, C426, C434 | 18Ile, 42Pro, 49Ser, 57Trp | 8.96 | 53 |

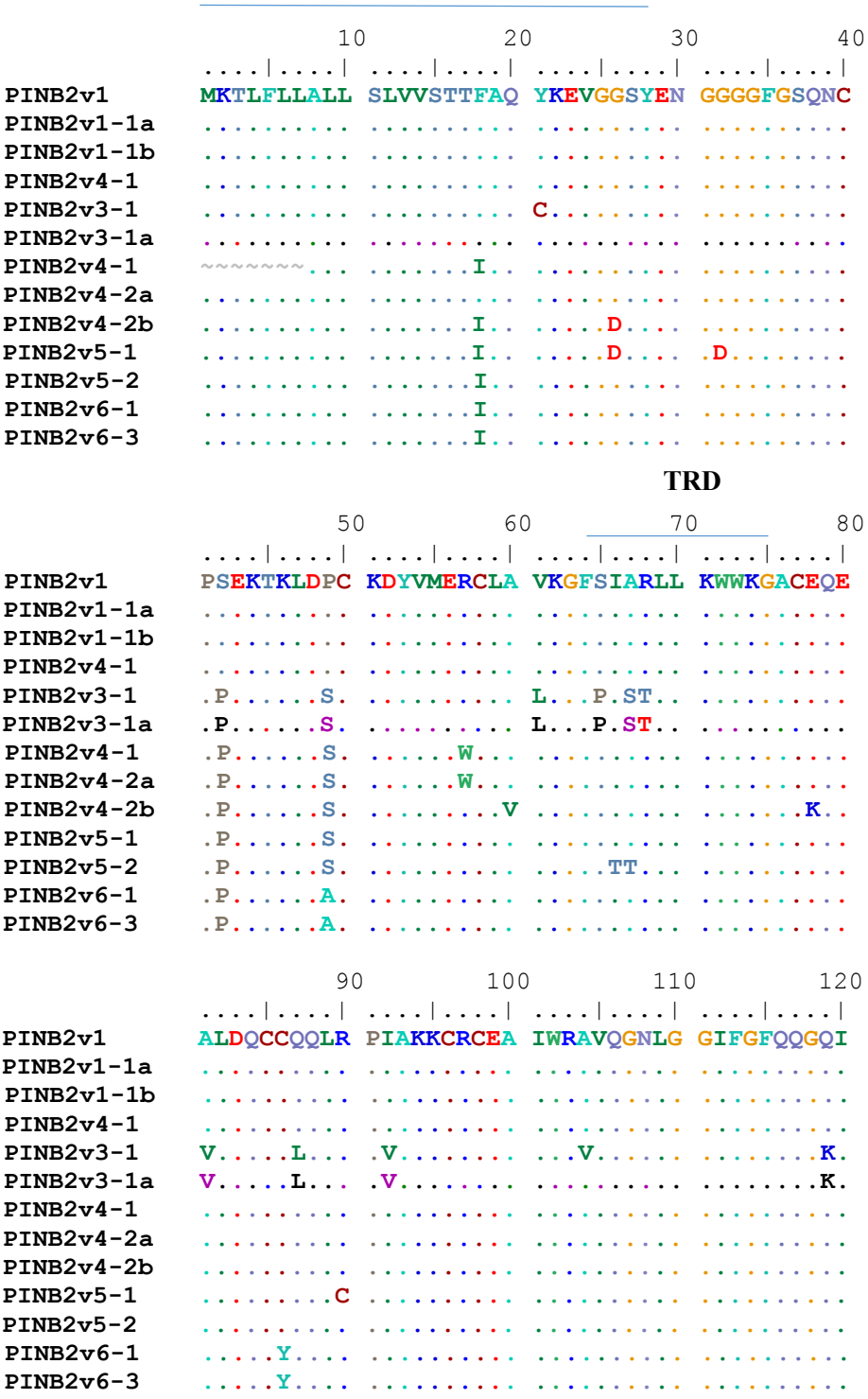
| DNA haplotype | Deduced protein type | Cultivar | *SNP(s) | Characteristic amino acid(s) | pI | Hydrophobic residues |
|---|---|---|---|---|---|---|
| *Pinb2v4-2* (JN585978) | PINB2v4-1 | Egret, Sunco | C426T | | 8.96 | 53 |
| *Pinb-2v4-2a* | PINB2v4-2a | Sunco | **A52T,** C75T, G393C, T406C, C426T | Ile18Phe | 8.96 | 53 |
| *Pinb2v4-2b* | PINB2v4-2b | Egret | **G77A, T169C, C179T, G232A** | Gly26Asp, Trp57Arg, Ala60Val, Glu78Lys | 9.27 | 52 |
| *Pinb2v5-1* (JQ619934) | PINB2v5-1 | Not found | **G77A**, T180G, T201C, **G226A**, A270G, A282G, C346T, **A434C** | Gly26Asp, Ala76Thr, Gln116stop codon | 8.79 | 53 |
| *Pinb2v5-2* (JN626928) | PINB2v5-2 | Sunstar | **A77G**, **A95G**, **T197C**, **G199A**, **T268C**, **G419T**, C426T, **A434C** | Asp26Gly, Asp32Gly, Ile66Thr, Ala67Thr, Cys90Arg, Cys140Phe, Asn145Thr | 9.14 | 50 |
| *Pinb2v6-1* (JN585979) | PINB2v6-1 | Sunstar | **A52,** A57, C75, **C124**, G141, **G145**, G180, A195, C201, G240, **A257**, C270, G282, T291, C321, T392, **G393**, T406, T426, C434 | 18Ile, 42Pro, 49Ala, 86Tyr, 131Leu | 9.13 | 54 |
| *Pinb2v6-3* (JQ619936) | PINB2v6-3 | Sunstar | T426C, **G430C**, G450C | Ala144Pro, stop codon Tyr | 9.11 | 54 |

*SNPs shown in bold result in amino acid substitution and the rest are silent mutations

### 3.3.8 d) Sequence diversity in the deduced PINB-2 proteins

The *Pinb-2* haplotype sequences (Fig. 3.12, Table 3.5) were translated into their deduced amino acid sequences leading to 13 deduced protein types. (Fig. 3.13, Table 3.5). Two *Pinb-2v1* haplotypes, *Pinb-2v1-1a* and *Pinb-2v1-1b* led to the same protein types. Similarly, two *Pinb-2v4* haplotypes (*Pinb-2v4-1* and *Pinb-v4-2*) led to similar deduced protein types. The signal peptide was located from residues 1 to 28 based on alignment with the *Pinb-2v1* reference sequence (GQ496616) indicating the mature protein starts at residue 29 (ENGGGG). The counting of the amino acid residues was applied from the initiator Met, rather than the mature protein sequence (ENGGGG). It was observed that the deduced PINB-2 variant protein types had characteristic amino acids specific to the particular group. Ser42 and Pro49 were shared by the PINB-2v1 group, while Ala49, Tyr86 and Leu131 were characteristic of PINB-2v6 group. The residues LLKWWK were highly conserved within the TRD, highlighting its functional importance. A conserved region FPISTLLKWWKG at the TRD was noticed in the members of the PINB-2v3 with Thr replacing the characteristic Arg noted in all other groups. The ten characteristic cysteine residues forming disulphide bonds in PINs (Le Bihan et al. 1996), were completely conserved except a Cys86Tyr mutation found in the PINB-2v6 members. The predicted pI of the deduced mature proteins ranged from 8.79 to 9.27 (Table 3.5). The highest pI (9.27) was noted in PINB2v4-2b, probably due to Trp57Arg substitution. However, the other PINB2v4 types i.e. PINB2v4-1 and PINB2v4-2a depicted lowest pI (8.96) that may be explained due to the presence of Trp57 instead of Arg. The PINB2v3-1a deduced protein type found in cultivar Tasman (associated with hardness mutation Val104Ala) depicted pI value of 9.05. The deduced protein type PINB2v5-2 showed an increased pI value (9.14) probably due to the Cys90Arg substitution. The PINB-2v4 and PINB-2v6 members (and the published PINB2-v5-1) had 1-2 additional hydrophobic residues compared to other deduced proteins whereas the PINB2-v5-2 protein type had two fewer hydrophobic residues.

**Signal peptide**



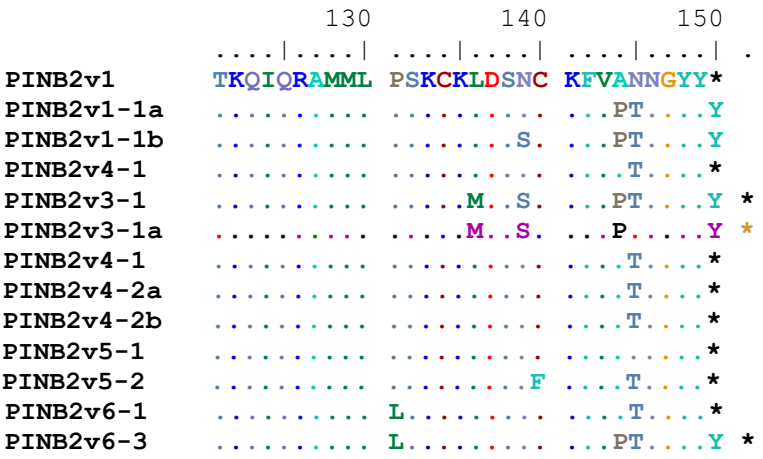
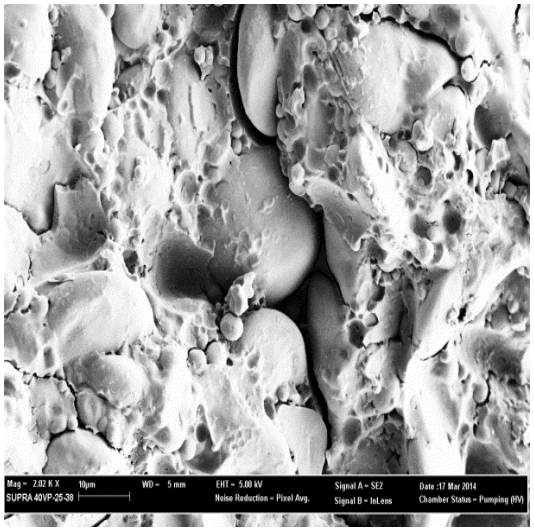



**Figure 3.13. Deduced *Pinb-2* protein types of wheat cultivars**
Dots indicate nucleotides identical to the top sequence in that particular group and˜indicates gap.

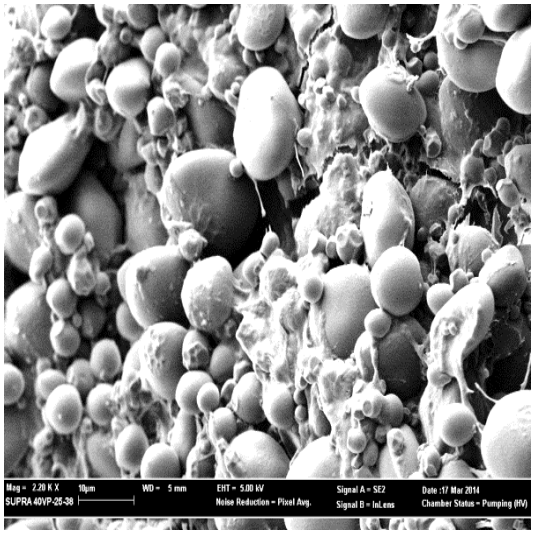
### 3.3.9 Investigation of grain hardness of selected wheats using Scanning electron microscopy (SEM)

Scanning electron microscopy was used to look for differences in grain texture by investigating the amount of starch granules bound to the protein matrix. Due to limited seed sample SEM was adopted to analyse grain hardness of progeny lines and other grain hardness determination techniques such as PSI, SKCS were not used. SEM is a direct method for analysing grain texture by determining the extent of binding between starch granules and protein matrix (Chen et al. 2005). SEM has been used to distinguish grain texture in wheat cultivars and landraces previously by Qamar et al. (2014) and Ramalingam (2012). The seed samples provided by AWCC were used for SEM. The sample preparation for SEM (Section 2.11) was performed with consistency but the variations in seed size, type of seed fracture cannot be ignored. Parent cultivars, Tasman, Sunco and their progenies were analysed by SEM. Figures 3.14 A and B are SEM images of cultivars Tasman and Sunco, respectively. Both Tasman (52.39 SKCS units) and Sunco (71.98 SKCS units) are hard cultivars with *Pina-D1a/Pinb-D1b* genotype. The Val104Ala substitution in variant 3 was found in Tasman but not Sunco. Greater number of A type and B-type granules were seen in Sunco (Fig. 3.14 A) as compared to Tasman (Fig. 3.14 B). In case of progeny lines, Figures 3.14 C, D, E, F are SEM images of progenies having Val104Ala allele. These lines depicted starch granules embedded with dense protein matrix as compared to starch granules in progeny line lacking Val104Ala

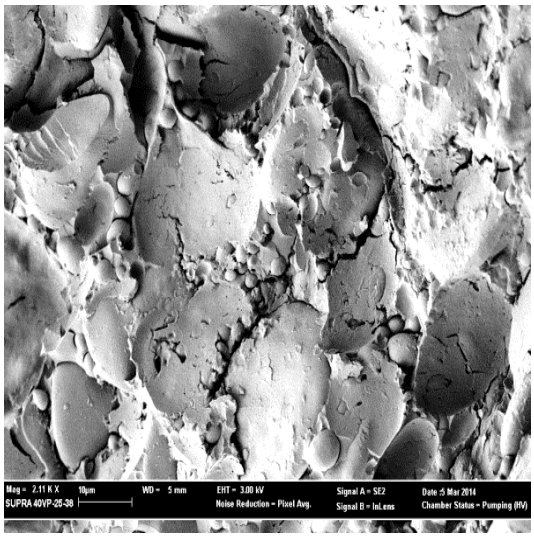
substitution (Fig. 3.14 G). It was difficult to draw a conclusion regarding the grain hardness of progeny lines with and without Val104Ala substitution.



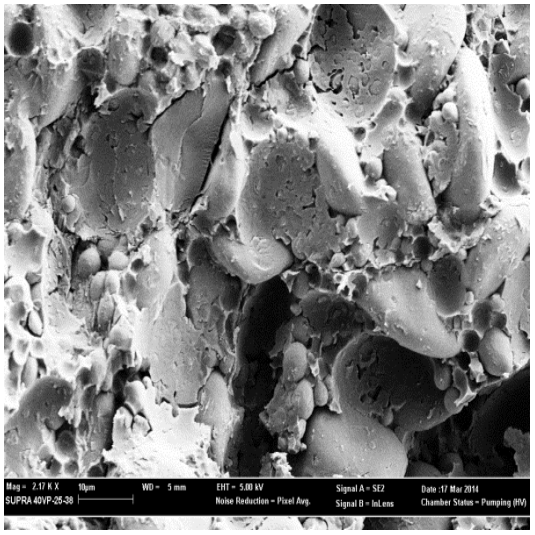
**A**



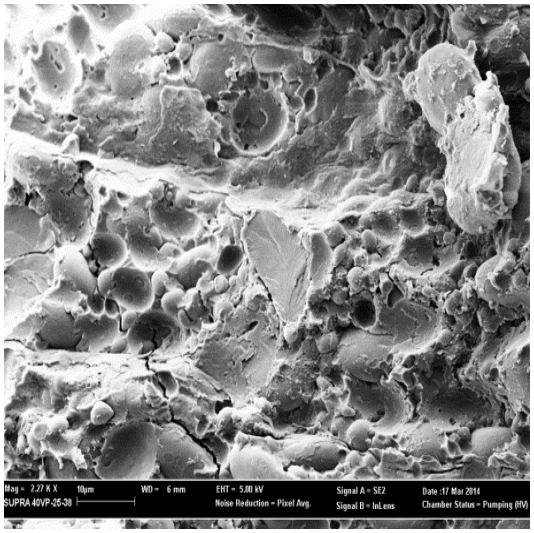
**B**



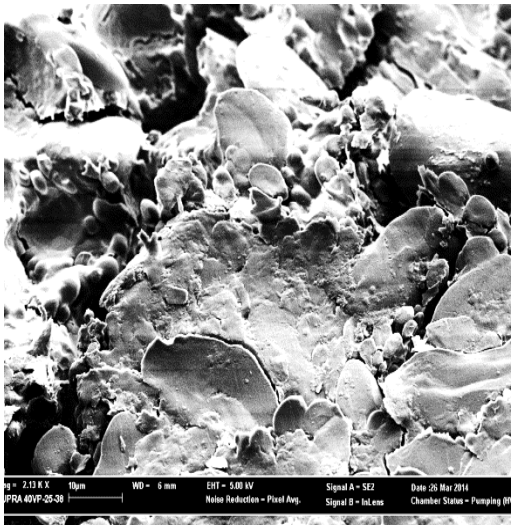
**C**



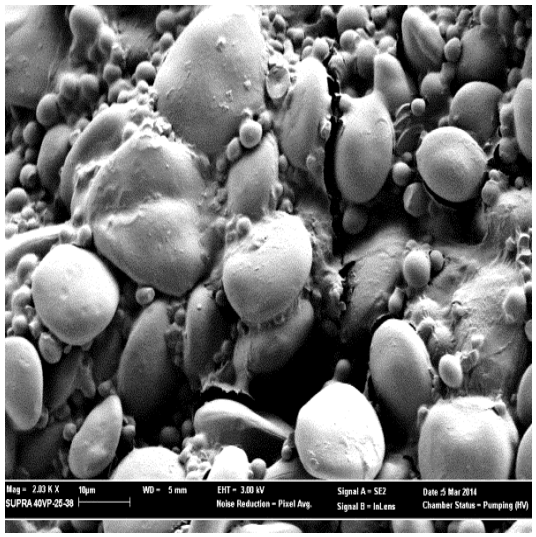
**D**




**E**




**F**




**G**

**Figure 3.14. SEM images of wheat samples.** Magnification: X 2000, A: Tasman, B: Sunco, C: Progeny line, D: Progeny line, E: Progeny line, F: Progeny line, G: Progeny line

### 3.3.10 Wheat grain protein content and its possible relationship with grain hardness

The grain protein of wheat cultivars was determined using the Dumas combustion method described in Section 2.11 with an objective to determine a possible correlation with grain hardness. It was observed that in general increasing protein resulted in harder grain (Table 3.6). The SKCS grain hardness showed a range of 30.5-78.9. A mean of 63.1±14.9 was observed for SKCS. The grain protein depicted a range of 11.0-15.9 and mean value of 13.0 ± 1.43 respectively (Table 3.7). Despite these observations the correlation between these two traits was not statistically significant ($r = 0.22$, $p < 0.01$).

**Table 3.6 Grain protein content and hardness in wheat cultivars**

| S no. | Cultivar | SKCS | GP (%) |
|---|---|---|---|
| 1 | Amery | 67.24 | 11.1 |
| 2 | Batavia | 78.53 | 15.3 |
| 3 | Carnamah | 63.22 | 14.5 |
| 4 | CD87 | 70.27 | 15.9 |
| 5 | Chara | 71.62 | 13.8 |
| 6 | Clearfield Stilleto | 55.41 | 13.4 |
| 7 | Cranbrook | 56.04 | 12.9 |
| 8 | Cunningham | 51.43 | 13.3 |
| 9 | Diaz | 77.03 | 12.2 |
| 10 | Egret | 30.87 | 13.9 |
| 11 | Falcon | 35.42 | 11.1 |
| 12 | Halbred | 69.88 | 13.2 |
| 13 | Hartog | 78.95 | 11.8 |
| 14 | Janz | 68.77 | 12.8 |
| 15 | Katepwa | 75.24 | 14.5 |
| 16 | Kukri | 67.65 | 14.0 |
| 17 | Machete | 70.37 | 14.5 |
| 18 | Ouyen | 71.03 | 11.8 |
| 19 | Rosella | 30.47 | 11.7 |
| 20 | Sunco | 71.98 | 11.0 |
| 21 | Sunstar | 75.13 | 12.1 |
| 22 | Tasman | 52.39 | 11.3 |

**Table 3.7 Descriptive statistics of grain hardness compared to grain protein**

| Parameter | Range | Mean | SD |
|---|---|---|---|
| Grain hardness | 30.47-78.95 | 63.13 | 14.86 |
| Grain protein content | 11.0-15.9 | 13.00 | 1.43 |

## 3.4 Discussion

Allele diversity of *puroindoline* genes is always sought after because of its implications on the grain hardness variation. Australian wheat cultivars possess limited grain hardness diversity (Pickering and Bhave 2007), therefore it is important to look for allele variations that might affect grain hardness. A variant form of *Pinb-2v3* i.e. *Pinb-2v3b* (Val104Ala) has been reported earlier by Chen et al. (2013) and linked with higher SKCS hardness indices than other *Pinb-2* variants. Thus, it is important to look for the presence of such variation in Australian cultivars. In this study, Australian wheat cultivars were screened for *Pinb-2v3* sequence diversity. The *Pinb-2* diversity analysis confirmed the six variant types and provided further evidence for the multigenic nature of the *Pinb-2* gene family. The present work involved molecular analysis of *Pinb-2* gene diversity in selected Australian wheat cultivars. A number of polymorphisms with likely implications on the structure and function of the deduced proteins were identified. These observations suggest the role of these genes in moderating grain texture.

### 3.4.1 Allelic and non-allelic nature of *Pinb-2* variants

The genotyping of 22 Australian wheat cultivars using the *Pinb-2* variant 2 and variant 3 specific primers identified the allelic nature of variant 2 and variant 3. None of the cultivars showed the simultaneous presence of variant 2 and variant 3. The cultivars either possessed variant 2 or variant 3 genotypes. These observations suggest the allelic nature of variant 2 and variant 3. Further, the physical location of variant 2 and variant 3 on chromosome 7B reported by Chen et al. (2010 a) supports the above suggestion. The allelic nature of variant 2 and variant 3 has been reported earlier by Chen et al. (2010 a) and Geng et al. (2013). Variant 1 and variant 4 were found in all the cultivars analysed indicating their non-allelic nature in all cultivars. Interestingly, the gene cloning and sequencing studies indicated the presence of variant 1, variant 3, variant 4, variant 5, variant 6 in a single cultivar Sunstar. Thus, the results suggested variant1/ variant 3/ variant 4/ variant 5/ variant 6 to be non-allelic. In this study, it was found that the *Pinb-2v3-1a* (Val104Ala) allele is present in the Australian wheat cv. Tasman and absent in cv. Sunco that had *Pinb-2v3-1*. Further, the doubled- haploid progeny lines of these cultivars indicated segregation for *Pinb-2v3*, with seven of 13 progenies showing *Pinb-2v3-1a* and the rest showing the other parental allele *Pinb-2v3-1*, thus confirming that these are truly

allelic also, and not homeoalleles. Hence the breeding programs using these lines (and others with this mutation) may wish to consider this observation.

### 3.4.2 Occurrence *Pinb-2* variant types in wheat cultivars and relationship with grain texture

Variant 3 was found in 18 of 22 cultivars indicating that variant 3 is predominant as compared to variant 2. Variant 1/ variant 3/variant 4 was found to be the most commonly occurring genotype in cultivars. Of the two variant 3 (v3-1 and v3-1a) types, v3-1a (Val104Ala) was found to be present in eight cultivars suggesting that though Australian wheat germplasm lacks diversity for *puroindoline* alleles there exists significant *Pinb-2* diversity particularly in context to *Pinb-2v3-1a,* associated with harder grain texture. *Pinb-2v3-1a* was either found in *PinaD1-a*/ *Pinb-D1b* or *PinaD1-b*/*Pinb-D1a* genotypes and not detected in wild-type (*Pina-D1a/ Pinb-D1a*) genotypes. The occurrence of *Pinb-2v3-1a* in cultivars with wild type *Pin* alleles would give the opportunity to relate to the SKCS grain hardness values excluding the hardening effect of *Pina-D1b* or *Pinb-D1b* allele. Previously, variant 3 has been associated with higher SKCS grain hardness values in soft wheats (*Pina-D1a*/ *Pinb-D1a*) and no such association was reported in hard wheats i.e. *Pina-D1b*/ *Pinb-D1a* or *Pina-D1a*/*Pinb-D1b* (Chen et al. 2010b, Chen et al. 2013). Occurrence/non-occurrence of *Pinb-2* variants (v3 or v3-1a) could not explain the difference in SKCS hardness values of same *Pin* class (*PinaD1-a/Pinb-D1a*, *PinaD1a/Pinb-D1b*, *Pina-D1b/Pinb-D1a*) genotypes. Thus, Val104Ala substitution in variant 3 could not be associated with increased SKCS grain hardness values as reported in an earlier study by Chen et al. (2013) Further, no pattern was observed in relation to the SKCS grain hardness values and the *Pinb-2v2*/*Pinb-2v3* alleles, in case of all three *Pin* classes and *Pinb-2* variants likely do not exert notable effect on kernel texture. Similar findings were reported by Geng et al. (2013) who genotyped 388 wheat cultivars and advanced breeding lines of U.S. germplasm and found no association between *Pinb-2* variants and SKCS grain hardness.

### 3.4.3 Point mutations in PINB-2 protein and their possible implications on membrane binding abilities

The Trp residues in the TRD of PINs appear to be most relevant in context of PIN functionality. The TRD has been shown to be important in interactions with lipids and the antimicrobial abilities of PIN peptides (reviewed in Bhave and Morris, 2008b; Jing et al. 2003; Kooijman et al. 1997). It has also been shown that the TRD of PINs directly binds to the starch granule surface (Wall et al. 2010). Secondly, the membrane activity of Trps is influenced by the vicinity to Arg as Arg provides cationic charges as well as H-bonding for interaction with the anionic bacterial membranes (Chan et al. 2006). Increased hydrophobicity is favourable for membrane binding functionality (Lee et al. 2011) Hydrophobic residues are also involved in weak, non-specific interactions with other proteins and/or lipids. Thus Trps, increased positive charge and increased hydrophobicity may lead to increased lipid binding capacities of PINs that might have biological (plant defence) and technological (grain hardness) implications. The presence of favourable residues (Trps, Arg and hydrophobic residues) in multiple copies in PINB-2 variants might enhance the unique lipid binding capacities of PINs thus relevant to antimicrobial properties. The antimicrobial activity of PINB-2 based peptides has been reported earlier by Ramalingam et al. (2012). It was noted that the Trp57 (instead of Arg57) found only in PINB-2v4 members may reduce their pI but improve hydrophobicity. It has been noted from this study that the PINB-2v4 and PINB-2v6 members have 1-2 additional hydrophobic residues compared to other groups that might favour PINB-2 and PIN interactions. The deduced protein type PINB2v4-2b (cultivar Egret) depicting highest pI (9.27) probably due to W57R substitution could be potentially useful in terms of increased antimicrobial properties.

### 3.4.4 Relationship between occurrence of Val104Ala substitution and grain hardness based on SEM analysis of selected wheat cultivars

It is important to study the association of allelic variation with grain hardness due to its significant effect on grain quality attributes. Therefore, the parental cultivars Tasman, Sunco and their progenies were analysed by SEM to study textural differences. Based on the SEM images of lines with and without Val104Ala substitution, this substitution could not be associated with grain hardness. However, it is important to mention that the

environmental variation in grain texture cannot be quantified using SEM analysis since the grain texture of a single seed from a specific environment is analysed. Environmental factors have been reported to have minor to moderate effects on grain texture as mentioned in Section 1.3.2 and SEM analysis has the above-mentioned limitation. Thus, further studies are required to understand the association between *Pinb-2* variants and grain hardness and any identification of potentially useful *Pinb-2* variation in the doubled-haploid progeny lines allows for early screening of segregating generations. This can prove as a useful plant breeding tool allowing breeders to develop varieties with desired grain hardness.

# CHAPTER 4

# Study of diversity of *Pinb-2* genes in worldwide wheat landraces

## 4.1 Abstract

As mentioned in previous chapter (Chapter 3) variation in *Pin* genes affects the degree of grain hardness of wheat, thus determining its end use. However, the role of *Pinb-2* gene variation in affecting wheat grain hardness is a relatively new area with report of one variant type *Pinb-2v3* associated with favourable grain yield and related traits. The allele *Pinb-2v3b* is considered potentially useful as it is associated with increased grain hardness values. Discovery of sources of useful gene variations with implications on grain hardness might lead to their selection and incorporation into breeding programs. There is always a quest for identifying sources of useful *Pin* gene variation. Landraces offer a unique resource for searching *Pin* gene variation. Previous studies have found that the *Pin* alleles are extremely diverse among the Chinese wheat germplasm. Therefore, in this study *Pinb-2* gene diversity was investigated in 18 overseas hexaploid wheat landraces from different centres of wheat cultivation ie., China, Burma, Afghanistan, Iraq, India, Asia, Africa, Europe and Ethiopia. These landraces possessed different background *Pina/Pinb* genotypes (*Pina-D1a/Pinb-D1a*, *Pina-D1a/Pinb-D1b*, *Pina-D1b/Pinb-D1a*, *Pina-D1a/Pinb-D1f*). The landraces were genotyped for *puroindoline* alleles and *Pinb-2* variant types. The study also involved screening of landraces for the presence of Val104Ala substitution in variant 3 associated with harder texture. The results indicated that variant1/variant 3/ variant 4 was most common genotype. Variant 2 and variant 3 did not co-exist and hence likely allelic. Variant 3 was found in 17 out of 18 landraces indicating it to be predominant genotype. *Pinb-2v3-1a* (Val104Ala) was detected in eight of 17 landraces. The occurrence of *Pinb-2v3-1a* was noted in landraces of different geographic origins and particularly China was identified an important source of this favourable allele. *Pinb-2v3-1a* (Val104Ala) was observed in all four *Pina/Pinb* genotypes (*Pina-D1a*/*Pinb-D1f*, *Pina-D1a*/*Pinb-D1a*, *Pina-D1a/Pinb-D1b*, *Pina-D1b*/*Pinb-D1a).* Three landraces were identified in this study that had combination of grain hardness *Pin* genes and *Pinb-2v3-1a*. These landraces might offer a prospective genetic source for enhancing the Australian wheat germplasm.

## 4.2 Introduction

As detailed in Chapter 1, *Pinb-2* genes have been associated with minor wheat grain texture variation. The Val104Ala substitution in *Pinb-2v3b* in particular has been associated with increased grain hardness values and is a potentially important allele with implications on grain texture. Thus, there is a quest for identifying genetic sources of useful *Pinb-2* variants that might serve to enhance the limited *Pin* gene diversity of Australian cultivars. *Pinb-2* gene diversity of Australian wheat cultivars has been discussed in Chapter 3. This chapter investigates *Pinb-2* gene diversity in 18 overseas hexaploid wheat landraces with different background *Pina/Pinb* genotypes (*Pina-D1a/Pinb-D1a*, *Pina-D1a/Pinb-D1b*, *Pina-D1b/Pinb-D1a*, *Pina-D1a/Pinb-D1f*). Landraces serve as unique resources that may prove invaluable in the search for new alleles for any trait whereas commercial cultivars show limited genetic diversity due to inbreeding depression. These landraces offer a potential genetic source especially those with combination of mutant *Pin* genes (e.g. *Pina-D1b*/ *Pinb-D1b*) and *Pinb-2v3b* allele. The landraces chosen for investigation were from different centres of wheat cultivation ie. China, Burma, Afghanistan, Iraq, India, Asia, Africa, Europe and Ethiopia (Pickering et al. 2007). China has been identified as important centre for *Pin* gene diversity in several studies (Chen et al. 2006, Li et al. 2008a, Wang et al. 2008, Ikeda et al. 2010, Chen et al. 2011). A novel *Pinb-2* variant (*Pinb-2v6*) was reported in Chinese winter wheat cultivar Yunong 202 by Chen et al. (2013). Thus, it is worthwhile to investigate Chinese landraces for *Pinb-2v3* allelic variation. This study included seven Chinese wheat landraces. Thus, this investigation explored *Pinb-2* variation in worldwide landraces with the aim to identify potentially useful *Pinb-2* variants, the effect of background *Pin* genotype on *Pinb-2* variation, potentially important regions associated with occurrence of *Pinb-2v3* allelic variation. The methods (detailed in Chapter 2) involved amplification of *Pina* and *Pinb* genes, identification of *Pinb-D1b* using CAPS analysis and genotyping into *Pinb-2* variants (v1, v2, v3 and v4) using variant specific primers.

### 4.3.1 Selection of worldwide wheat landraces with different background *Pina/Pinb* genotype

Eighteen landraces of common wheat from different centres of wheat cultivation ie., China, Burma, Afghanistan, Iraq, India, Asia, Africa, Europe and Ethiopia were chosen to investigate their *puroindoline* alleles. The *Pin* genotype of the landraces was *Pina-D1a/Pinb-D1a* (10 landraces), *Pina-D1a/Pinb-D1b* (four landraces) and *Pina-D1b/Pinb-D1a* (three landraces), *Pina-D1a/Pinb-D1f* (one landrace). As mentioned earlier in Section 3.3.1 (Chapter 3), the cultivars with *Pina-D1a/Pinb-D1a* genotype are considered soft, *Pina-D1a/Pinb-D1b* as moderately hard and *Pina-D1b/Pinb-D1a* as hard. Thus, the germplasm consisted of soft, hard and hard classes. These landraces were chosen with the aim to investigate the *Pinb-2* gene diversity in wheat samples of different geographic origin, different *Pin* genotype and grain hardness class. The *Pin* genotype and grain hardness class of landraces is mentioned in Table 4.1.

## 4. 4 Results

### 4.4.1 Genomic DNA extraction and identification/confirmation of the *Ha* locus genotypes

The wheat landraces used in this work are listed in Chapter 2 (Table 2.9). Genomic DNA was extracted from single plant i.e. plant from single seed using the Wizard DNA Purification Kit (Promega) as described in Section 2.3.1. The *Pina-D1* and *Pinb-D1* genes were amplified using the primers listed in Table 2.12 design based on Gautier et al. (1994). Amplicons of approximately 450 base pairs (bp) were obtained for all except three landraces (China Sh8, Leon 4, Rumania 10) for *Pina-D1* amplifications, indicating the presence of *Pina-D1*. The results indicated that the three landraces (China Sh8, Leon 4, Rumania 10) that did not show any *Pina-D1* amplification had *Pina-D1b* (null) allele and the rest 15 landraces probably had *Pina-D1a*. All the landraces also showed a 450 bp product in *Pinb-D1* amplifications, indicating that these landraces are probably wild type for *Pinb-D1*. An example of *Pina-D1* and *Pinb-D1* amplifications is shown in Fig. 4.1 and Fig. 4.2 respectively. No products were amplified in the negative controls (no DNA template) for both *Pina-D1* and *Pinb-D1*. *Pina*/*Pinb* allele designations of 11 landraces (Abyssinia AV12.4, Afghanistan 49, Afghanistan 51, Afghanistan 77, Algeria 37, Burma 7, India 223, Iraq 55, Leon 4, Rumania 10, Turkestan W84525) were known prior to start

of our studies (Pickering et al.2007). The results confirmed the reported *Pina/Pinb* genotypes of these 11 landraces and identified *Pina/Pinb* alleles of seven landraces (China Sh6, China Sh8, China Sh9, China Sh75, China Sh154, China Sh164, China Sh166) that had not been reported earlier. The results are presented in Table 4.1.

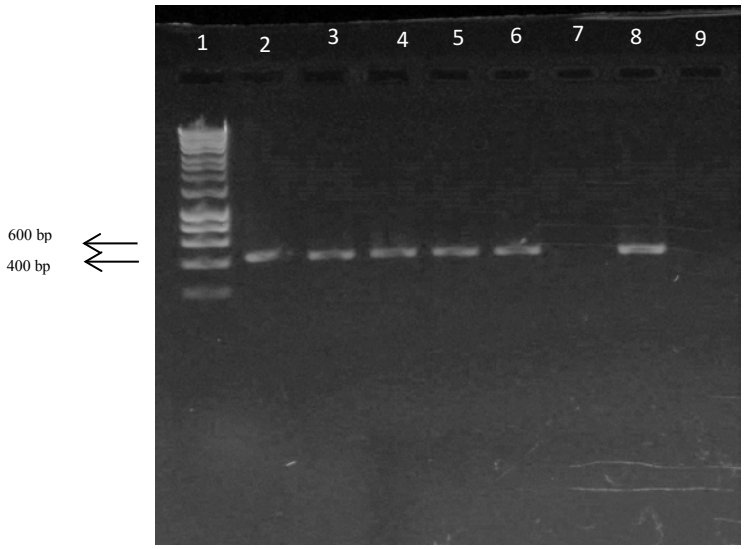



**Figure 4.1. Example of *Pina* PCR products of wheat landraces.**
Lane1: Hyperladder 1; lane 2: Afghanistan 49, lane 3: Afghanistan 51, lane 4: Algeria 37, lane 5: Burma 7, lane 6: China Sh 6 , lane 7: China Sh 8 , lane 8: China Sh 9 , lane 9: negative control (no template).

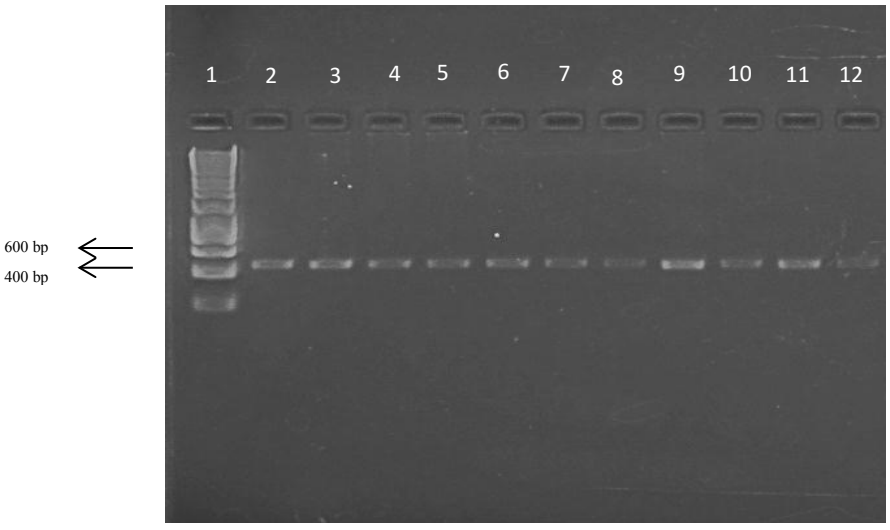



**Figure 4.2. Example of *Pinb* PCR products of wheat landraces**.
Lane 1: Hyperladder 1, lane 2: Abyssinia AV12.4, lane 3: Afghanistan 49, lane 4: Afghanistan 51, lane 5: Afghanistan 77, lane 6: Algeria 37, lane 7: Burma 7, lane 8: China Sh6, lane 9: China Sh8, lane 10: China Sh9, lane 11: China Sh75, lane 12: China Sh154

### 4.4.2 Identification/confirmation of *Pinb-D1b* allele using CAPS analysis

Cleavage of PCR-amplified *Pinb-D1* genes with the enzyme Bsrb1 was used to determine the presence of *Pinb-D1b* mutation as detailed in Chapter 3, Section 3.3.3. *Pinb-D1b* was present in four (Afghanistan 51, Algeria 37, India 223, Turkestan W84525) out of 18 landraces. As mentioned in Section 3.3.3, digestion of purified Pinb-D1 PCR products of a wheat sample with *Pinb-D1b* mutation leads to 129 bp, 95 bp, and 223 bp bands (Fig. 4.3). *Pinb* allele designations of 11 landraces (Abyssinia AV12.4, Afghanistan 49, Afghanistan 51, Afghanistan 77, Algeria 37, Burma 7, India 223, Iraq 55, Leon 4, Rumania 10, Turkestan W84525) were known prior to start of our studies (Pickering et al.2007). Another mutation *Pinb-D1f* (Trp44 to stop codon, Morris et al. 2001a) was present in landrace Abyssinia AV12.4 (Pickering et al.2007). The results confirmed the reported *Pinb-D1b* alleles of these 11 landraces and identified *Pinb* alleles of seven landraces (China Sh6, China Sh8, China Sh9, China Sh75, China Sh154, China Sh164, China Sh166) that had not been reported earlier. The results are presented in Table 4.1.

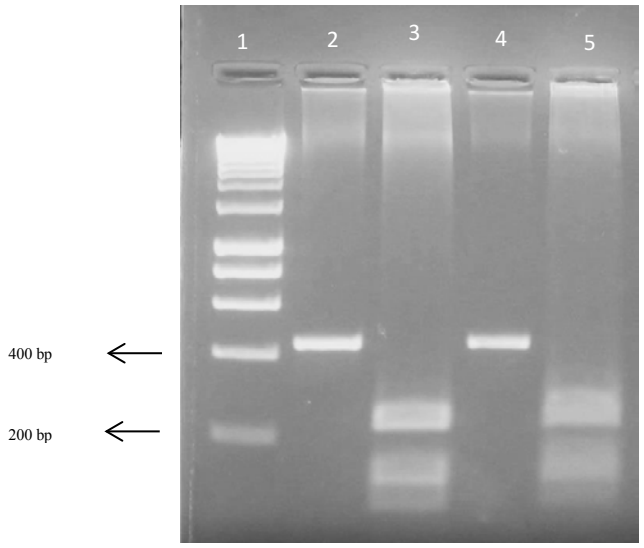



**Figure 4.3. Example of BsrB1 digested *Pinb* PCR products of wheat landraces**
Lane 1: Hyperladder 1, lane 2: undigested *Pinb* PCR product of Afghanistan 51, lane 3: BsrB1 digested PCR product of Afghanistan 51, lane 4: *Pinb* PCR product of Algeria 37, lane 5: BsrB1 digested PCR product of Algeria 37

### 4.4.3 Amplification of *Pinb-2* genes in wheat landraces

The amplification of *Pinb-2* genes in wheat landraces (round I PCR) was undertaken using three degenerate primers (Table 2.12) as pair D1 (Pinb-2F/ Pinb-2D1R) or D2 (Pinb-2F/Pinb-2D2R). These primers were designed as per Wilkinson et al. (2008). The Pinb-2D1 primer pair was used to amplify *Pinb-2* variants 1, 3, 5 whereas Pinb-2D2 primer pair was used to amplify *Pinb-2* variants 2, 3, 4, 6. Amplicons of approximately 450 base pairs (bp) were obtained as expected for all the landraces for both Pinb-2D1 and Pinb-2D2 primer pairs. The results for *Pinb-2* gene amplifications in wheat landraces are presented in Table 4.1.

**Table 4.1: PCR amplification of *Pina*, *Pinb* and *Pinb-2* genes in wheat landraces**

| S.no | Landrace | AWCC no. | *Pina*/*Pinb* genotype | Class | *Pina-D1* | *Pinb-D1* | *Pinb-D1b* | *Pinb-2*D1 | *Pinb-2*D2 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Abyssinia AV12.4 | 28266 | *Pina-D1a*/*Pinb-D1f* | - | ✓ | ✓ | × | ✓ | ✓ |
| 2 | Afghanistan 49 | 27212 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | ✓ | × | ✓ | ✓ |
| 3 | Afghanistan 51 | 27287 | *Pina-D1a*/*Pinb-D1b* | Hard | ✓ | ✓ | ✓ | ✓ | ✓ |
| 4 | Afghanistan 77 | 27216 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | ✓ | × | ✓ | ✓ |
| 5 | Algeria 37 | 27295 | *Pina-D1a*/*Pinb-D1b* | Hard | ✓ | ✓ | ✓ | ✓ | ✓ |
| 6 | Burma 7 | 27356 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | ✓ | × | ✓ | ✓ |
| 7 | China Sh6 | 27439 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | ✓ | × | ✓ | ✓ |
| 8 | China Sh8 | 27441 | *Pina-D1b*/*Pinb-D1a* | Hard | × | ✓ | × | ✓ | ✓ |
| 9 | China Sh9 | 27442 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | ✓ | × | ✓ | ✓ |
| 10 | China Sh75 | 27440 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | ✓ | × | ✓ | ✓ |
| 11 | China Sh154 | 27414 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | ✓ | × | ✓ | ✓ |
| 12 | China Sh164 | 27415 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | ✓ | × | ✓ | ✓ |
| 13 | China Sh166 | 27416 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | ✓ | × | ✓ | ✓ |
| 14 | India 223 | 28076 | *Pina-D1a*/*Pinb-D1b* | Hard | ✓ | ✓ | ✓ | ✓ | ✓ |
| 15 | Iraq 55 | 28112 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | ✓ | × | ✓ | ✓ |
| 16 | Leon 4 (Spain) | 27892 | *Pina-D1b*/*Pinb-D1a* | Hard | × | ✓ | × | ✓ | ✓ |
| 17 | Rumania 10 | 27975 | *Pina-D1b*/*Pinb-D1a* | Hard | × | ✓ | × | ✓ | ✓ |
| 18 | Turkestan W84525 | 28013 | *Pina-D1a*/*Pinb-D1b* | Hard | ✓ | ✓ | ✓ | ✓ | ✓ |

### 4.4.4 Genotyping of landraces into variant 1, variant 2, variant 3 and variant 4 types using *Pinb-2* variant specific primers

Genotyping of 18 landraces was done using *Pinb-2* variant 1 (v1), variant 2 (v2), variant 3 (v3) and variant 4(v4) specific primers. These primers were used for second round PCR of purified Pinb-2D1/Pinb-2D2 PCR products. The primers were designed as per Chen et al. (2010 a) and their specificity, choice of template, use of negative/positive controls have been described earlier in Chapter 3, Section 3.3.5. All the landraces showed positive amplification for variant 1 yielding approximately 300 bp product (expected size 319 bp, Fig. 4.4). This suggested the presence of variant 1 in these landraces. Similarly, all the landraces showed positive amplification for variant 4 and yielded approximately 400 bp product (expected size 403 bp, Fig. 4.5). This suggested the presence of variant 4 in the landraces. No amplification was observed in the negative controls.

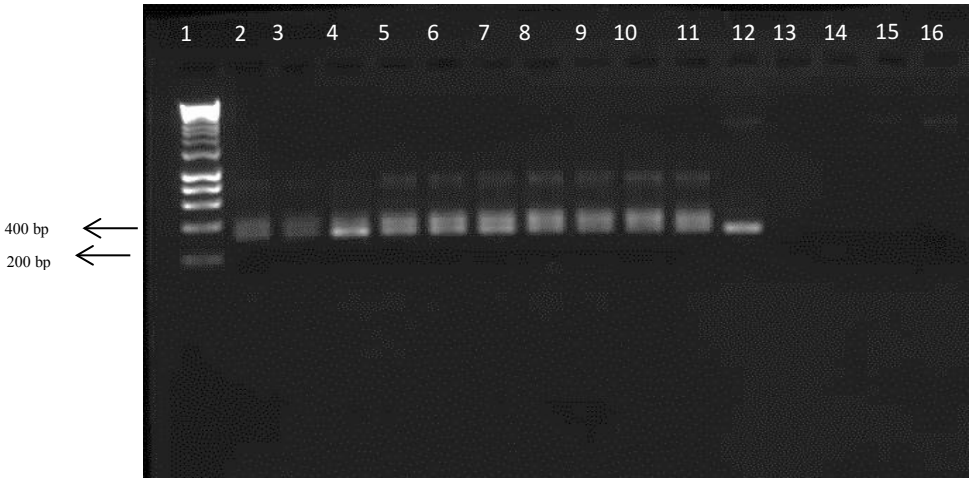



**Figure 4.4. Example of *Pinb-2v1* PCR products of wheat landraces**.
Lane1: Hyperladder 1, lane 2: China Sh 154, lane 3: China Sh 164, lane 4: China Sh 166, lane 5: China Sh 6, lane 6: China Sh 75, lane 7: China Sh 8, lane 8: China Sh 9, lane 9: Burma 7, lane 10: Iraq 55, lane 11: Afghanistan 51, lane 12: positive control (variant 1 specific clone), lane 13: negative control 1 (variant 2 specific clone), lane 14: negative control 2 (variant 3 specific clone), lane 15: negative control 3 (variant 4 specific clone), lane 16: negative control 4 (no template)

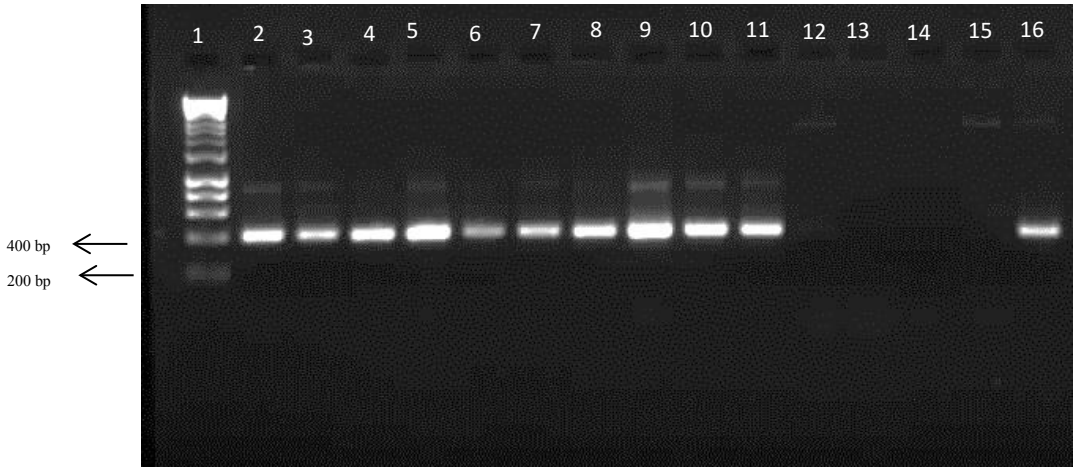



**Figure 4.5. *Pinb-2v4* PCR products of wheat landraces**.
Lane1: Hyperladder 1, lane 2: China Sh 154, lane 3: China Sh 164, lane 4: China Sh 166, lane 5: China Sh 6, lane 6: China Sh 75, lane 7: China Sh 8, lane 8: China Sh 9, lane 9: Burma 7, lane 10: Iraq 55, lane 11: Afghanistan 51, lane 12: negative control 1(variant 1 specific clone), lane 13: negative control 2 (no template), lane 14: negative control 3 (variant 2 specific clone), lane 15: negative control 4 (variant 3 specific clone), lane 16: positive control (variant 4 specific clone)

Variant 2 PCR amplification was observed in only one landrace (Iraq 55) depicting an approximately 400 bp PCR product (expected size: 401 bp, Fig. 4.6). The purified v2 PCR product was sequenced to confirm its variant type and the DNA sequence obtained was *Pinb-2v2* type. Interestingly, all the landraces except Iraq 55 gave positive amplification for variant 3 specific primers, depicting an approximately 400bp PCR product (expected size: 398 bp, Fig. 4.7). The landrace Iraq 55 that showed positive amplification for variant 2 did not show any amplification with variant 3 specific primers. The results suggest the presence of variant 3 in all landraces except Iraq 55 and absence of variant 2 in all landraces except Iraq 55. As explained in the case of v2 PCR in case of cultivars (Chapter 3, Section 3.3.6) the background bands observed in lanes 2-9, lane 11, lane 13 of Fig. 4.6 is the unamplified template (*Pinb-2* PCR product). The *Pinb-2* variant genotypes of landraces are shown in Table 4.2.

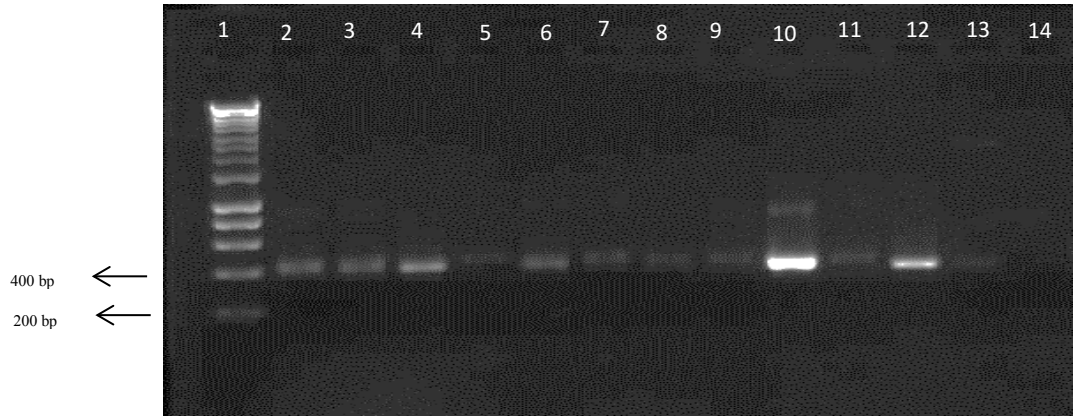



**Figure 4.6. *Pinb-2v2* PCR products of wheat landraces**.
Lane1: Hyperladder 1, lane 2: China Sh 154, lane 3: China Sh 164, lane 4: China Sh 166, lane 5: China Sh 6, lane 6: China Sh 75, lane 7: China Sh 8, lane 8: China Sh 9, lane 9: Burma 7, lane 10: Iraq 55, lane 11: Afghanistan 51, lane 12: positive control (variant 2 specific clone), lane 13: negative control 1 (variant 3 specific clone), lane 14: negative control 2 (no template)

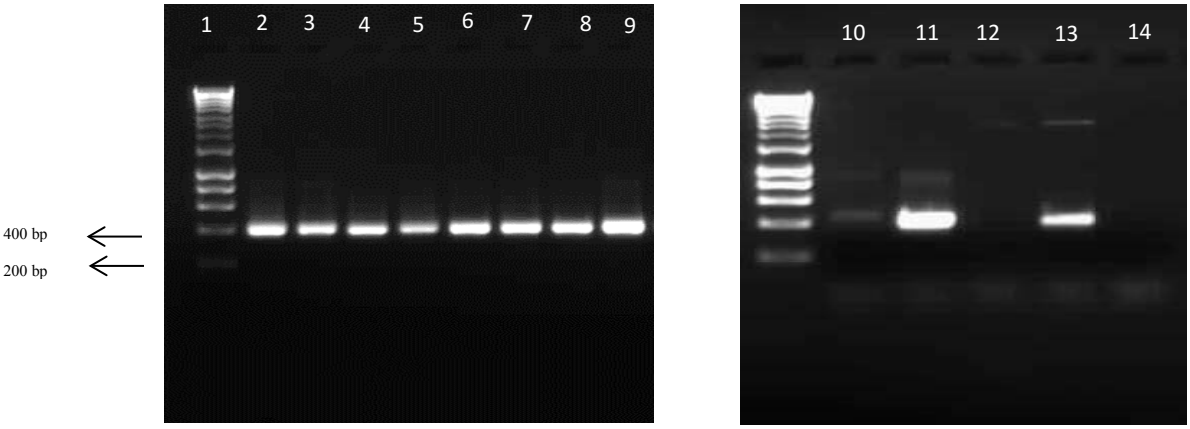
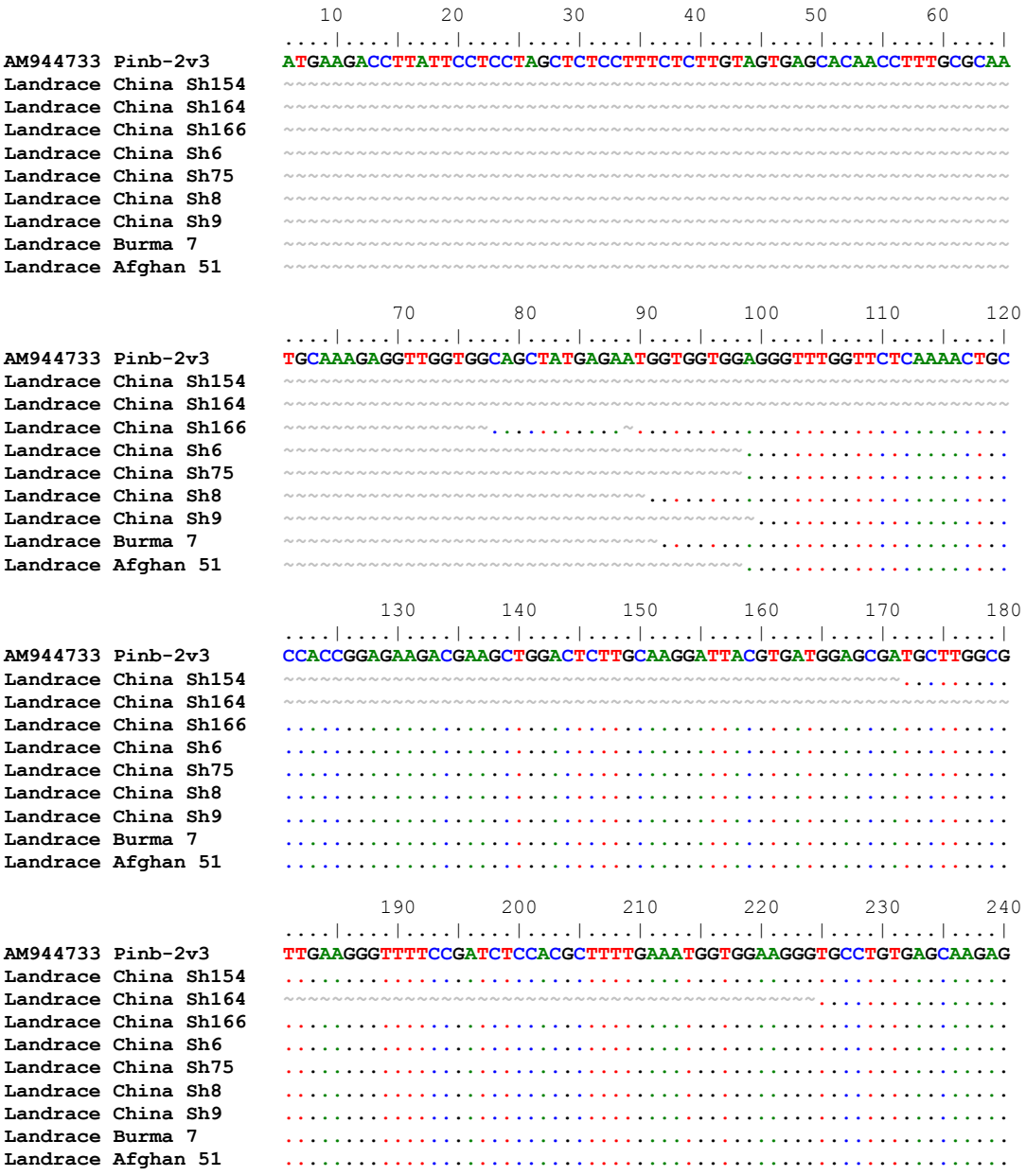



**Figure 4.7. *Pinb-2v3* PCR products of wheat landraces**.
Lane1: Hyperladder 1, lane 2: China Sh 154, lane 3: China Sh 164, lane 4: China Sh 166, lane 5: China Sh 6, lane 6: China Sh 75, lane 7: China Sh 8, lane 8: China Sh 9, lane 9: Burma 7, lane 10: Iraq 55, lane 11: Afghanistan 51, lane 12: negative control (variant 2 specific clone), lane 13: positive control 1 (variant 3 specific clone), lane 14: negative control 2 (no template)

### 4.4.5 Detection of Val104Ala substitution in variant 3 associated with harder grain texture in wheat landraces

All the landraces possessing variant 3 were investigated for the presence of T311C SNP resulting in Val104Ala substitution. This was carried out by sequencing the purified *Pinb-2v3* PCR products The PCR products were directly sequenced using forward primer. The DNA sequences obtained were aligned with reference sequence, *Pinb-2v3* (Genbank:AM944733). Eight of 17 landraces showed a T311C SNP when compared with the reference sequence (Fig.4.8). These landraces were: Abyssinia AV12.4, Afghanistan 49, Algeria 37, Burma 7, China Sh6, China Sh8, China Sh9, China Sh 166. The results are presented in Table 4.2.



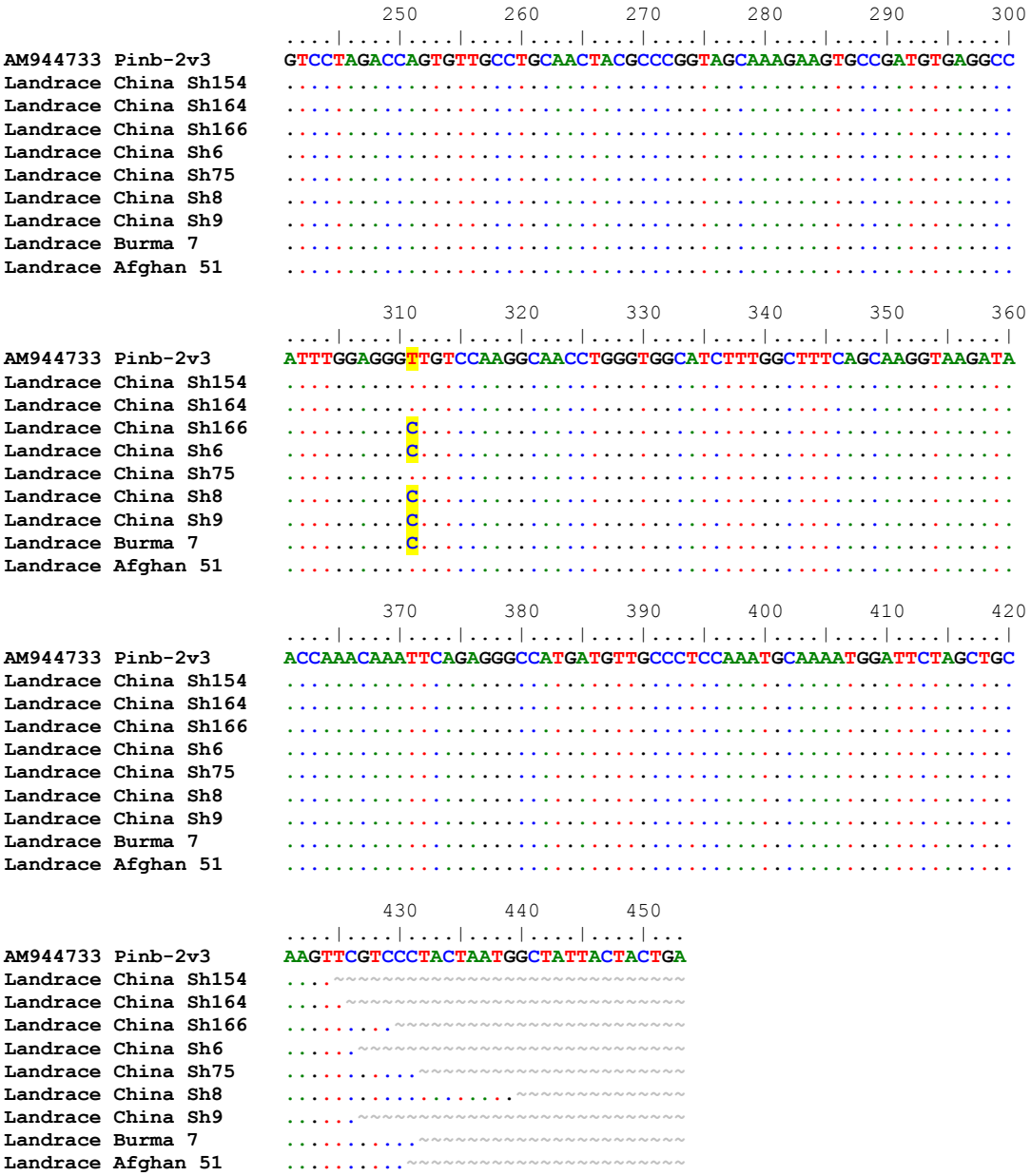



**Figure 4.8. Example of *Pinb-2v3* haplotypes of wheat landraces.**
Dots indicate nucleotides identical to the top sequence in that particular group and˜ indicates gap

**Table 4.2: *Pinb-2* variant genotypes of wheat landraces**

| S.no. | Landrace | *Pina*/*Pinb* genotype | Class | V1 | V2 | V3 | Val104 Ala* | V4 |
|---|---|---|---|---|---|---|---|---|
| 1 | Abyssinia AV12.4 | *Pina-D1a*/*Pinb-D1f* | Hard | ✓ | × | ✓ | ✓ | ✓ |
| 2 | Algeria 37 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | × | ✓ | ✓ | ✓ |
| 3 | Afghanistan 49 | *Pina-D1a*/*Pinb-D1b* | Hard | ✓ | × | ✓ | ✓ | ✓ |
| 4 | Afghanistan 51 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | × | ✓ | × | ✓ |
| 5 | Afghanistan 77 | *Pina-D1a*/*Pinb-D1b* | Hard | ✓ | × | ✓ | × | ✓ |
| 6 | Burma 7 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | × | ✓ | ✓ | ✓ |
| 7 | China Sh6 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | × | ✓ | ✓ | ✓ |
| 8 | China Sh8 | *Pina-D1b*/*Pinb-D1a* | Hard | ✓ | × | ✓ | ✓ | ✓ |
| 9 | China Sh9 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | × | ✓ | ✓ | ✓ |
| 10 | China Sh75 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | × | ✓ | × | ✓ |
| 11 | China Sh154 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | × | ✓ | × | ✓ |
| 12 | China Sh164 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | × | ✓ | × | ✓ |
| 13 | China Sh166 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | × | ✓ | ✓ | ✓ |
| 14 | India 223 | *Pina-D1a*/*Pinb-D1b* | Hard | ✓ | × | ✓ | × | ✓ |
| 15 | Iraq 55 | *Pina-D1a*/*Pinb-D1a* | Soft | ✓ | ✓ | × | × | ✓ |
| 16 | Leon 4 (Spain) | *Pina-D1b*/*Pinb-D1a* | Hard | ✓ | × | ✓ | × | ✓ |
| 17 | Rumania 10 | *Pina-D1b*/*Pinb-D1a* | Hard | ✓ | × | ✓ | × | ✓ |
| 18 | Turkestan W84525 | *Pina-D1a*/*Pinb-D1b* | Hard | ✓ | × | ✓ | × | ✓ |

*The presence of Val104Ala in Variant 3 is indicated by ✓ and absence by ×

### 4.4.6 Investigation of grain hardness of selected wheats using Scanning electron microscopy (SEM)

The differences in grain texture of selected landraces were investigated using SEM. Due to limited quantity of the seed, SEM method was adopted to investigate grain hardness of landraces and other grain hardness determination techniques such as PSI, SKCS were not used. The seed samples were provided by AWCC. The effect of the associated mutation (Val104Ala) on grain hardness was investigated in landraces with *Pina-D1a/Pinb-D1a* genotype (soft wheats). This would provide the opportunity to assess the effect of the *Pinb-2v3* mutation (Val104Ala) in the absence of any mutation in *Pina* or *Pinb* genes. Four landraces (*Pina-D1a/Pinb-D1a* genotype) were investigated using SEM. Two landraces (China Sh 154, China Sh 164) did not contain the Val104Ala substitution and the other two landraces (China Sh 6, China Sh 9) depicted the hardness associated mutation. In case of all four these landraces China Sh 154 (Fig. 4.9 A), China Sh 156 (Fig. 4.9 B), China Sh 6 (Fig. 4.9 C) and China Sh 9 (Fig. 4.9 D) a dense layer of

protein matrix covering the A and B-type starch granules was observed. However, the landrace China Sh 9 (Fig. 4.9 D) with the Val104Ala substitution depicted a comparatively tighter binding of starch granules and protein matrix, making the B-type starch granules less visible.

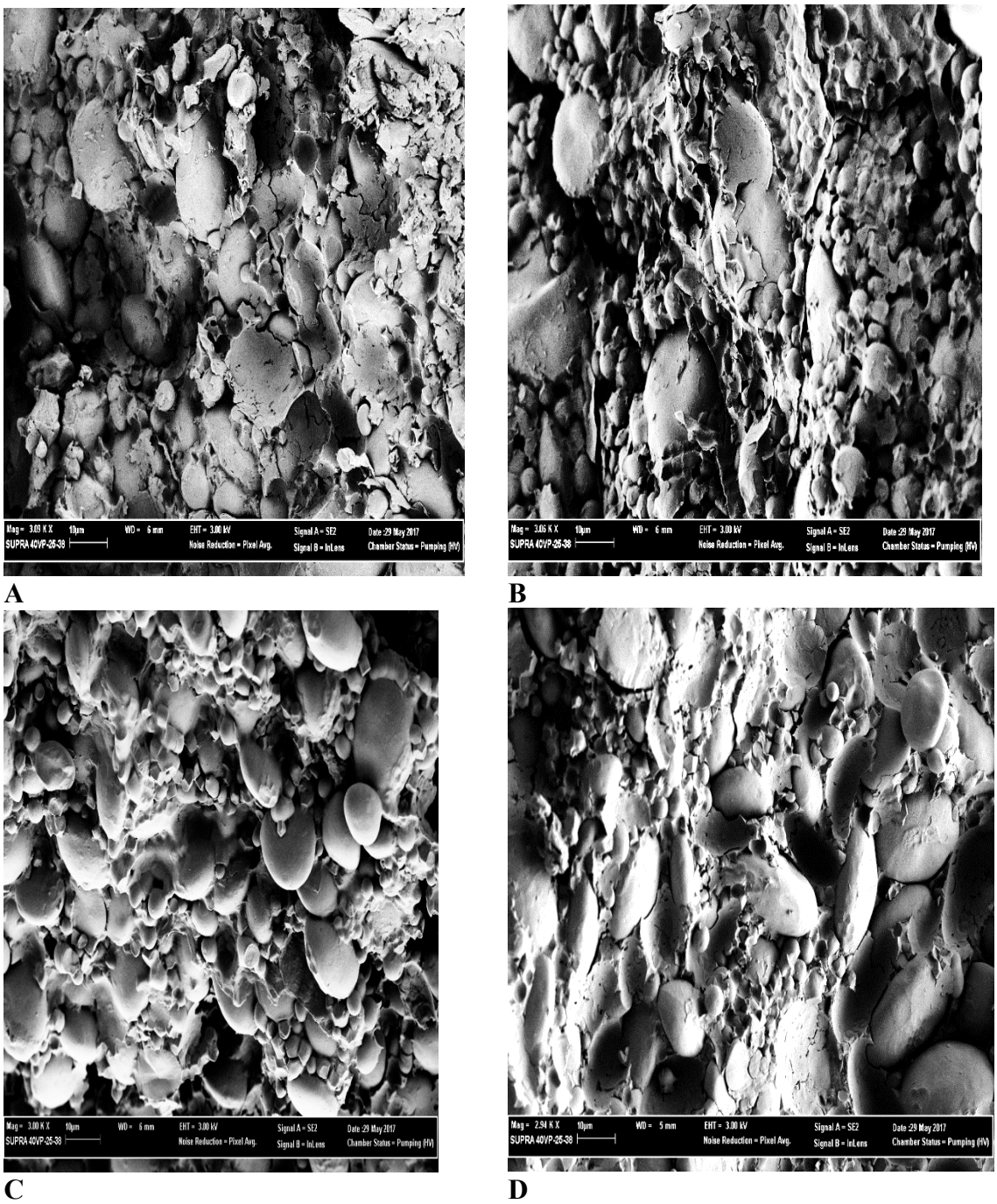



**Figure 4.9. SEM images of wheat samples.** Magnification: X 2000, A: Tasman, B: Sunco, C: Progeny line, D: Progeny line, E: Progeny line, F: Progeny line, G: Progeny line

## 4.5 Discussion

In this chapter, an investigation was conducted on 18 bread wheat landraces from centres of early wheat cultivation in search of *Pinb-2* variation. *Pina/Pinb* genotyping confirmed the *Pin* alleles in eleven landraces that had been reported earlier (Pickering et al. 2007) and identified new *Pin* variation in seven Chinese landraces. *Pina-D1a/Pinb-D1a* genotype was present in six out of seven landraces and one landrace i.e. China Sh8 had *Pinb-D1a/ Pinb-D1b* genotype. All 18 landraces were analysed for *Pinb-2* variants for the first time. This was conducted using variant specific primers as done for cultivars in Chapter 3. These primers were designed as per Chen et al. (2010 a) and were found to be appropriate for specifically amplifying variant 1/ variant 2/ variant 3/ variant 4. The approach of nested PCR (use of purified Pinb-2 D1 or Pinb-2 D2 PCR products) and inclusion of positive and negative controls in PCR reactions helped in achieving the desired results. This study made use of these primers for genotyping wheat cultivars and landraces and provides further evidence for the specificity of these primers. As in case of cultivars, the *Pinb-2* analysis in landraces confirmed the multigenic nature of these genes since multiple variant types were detected in all landraces.

The *Pinb-2* variant genotyping of landraces indicated the allelic nature of variant 2 and variant 3, as found in wheat cultivars. Variant 2 and variant 3 were not found together in any landrace, whereas variant 1 and variant 4 coexisted, suggesting the likely non-allelic nature of the latter. Similar findings were observed in case of wheat cultivars (Chapter 4). These results provide further evidence to the likely allelic nature of variant 2 and variant 3 reported earlier by Chen et al. (2010 a) and Geng et al. (2013). Interestingly, variant 3 was found in 17 out of 18 landraces suggesting predominant nature of variant 3 as compared to variant 2 as observed in case of wheat cultivars. Variant 1/ variant 3/variant 4 was found to be the most commonly occurring genotype in cultivars and landraces. The grain hardness associated mutation (Val104Ala) was observed in eight landraces: Abyssinia AV 12.4, Algeria 37, Afghanistan 49, Burma 7, China Sh6, China Sh8, China Sh9, China Sh166. The occurrence of this mutation was noted in landraces of different geographic origins. As mentioned earlier in Section 4.2, China has been reported to be a hot spot for *Pin* gene diversity. The occurrence of this substitution in four out of seven landraces suggests that Chinese landraces can be an important source of this favourable allele.

As far as *Pin* genotype is concerned, Val104Ala substitution in *Pinb-2v3* was observed in all four *Pina/Pinb* genotypes (*Pina-D1a*/*Pinb-D1f*, *Pina-D1a*/*Pinb-D1a*, *Pina-D1a/Pinb-D1b*, *Pina-D1b*/*Pinb-D1a).* This is an interesting observation since in case of cultivars this substitution was found only in hard wheat genotypes (*Pina-D1a/Pinb-D1b*, *Pina-D1b/Pinb-D1a*) and not in the soft wheat class (*Pina-D1a/ Pinb-D1a*). These observations suggest that *Pin* genotype does not affect the sequence diversity of *Pinb-2v3* and landraces were found to be a more diverse genetic source of *Pinb-2v3* genetic variation as compared to cultivars. The grain texture analysis conducted using SEM provided some evidence regarding the likely association of Val104Ala substitution with increased grain hardness. Further, three landraces were identified in this study that had combination of hardness related *Pin* and *Pinb-2v3* alleles. These landraces might have cumulative grain hardening effects of *Pin* and *Pinb-2* genes and offer a potential genetic source for enhancing the Australian wheat germplasm.

# CHAPTER 5

# A scoping study investigating the potential functionality of *Pinb-2* genes

## 5.1 Abstract

The yeast two-hybrid system has been used successfully in previous work leading to evidence of *in vivo* PINA and PINB protein-protein interactions. In the present study, the association between PINA, PINB and PINB-2V3 was tested using this system, to determine if V3 proteins can also interact with PINA and/or PINB. Results based on the selective media correlated well with the β-galactosidase assays. The PINA-PINA and the PINA(bait)–PINB (prey) interactions were the strongest. PINB-2V3-PINB and PINB-2V3-PINA interactions were found to be very weak. The study demonstrates that though the interaction of PINB-2V3 with PIN proteins is weak, PINB-2V3 may play a minor role in the overall PIN-PIN interactions that are functionally crucial in affecting grain hardness. This study also investigated the expression of *Pinb-2* genes (in leaf and root tissues) and its association with grain hardness. *Pinb-2* gene expression was detected in these tissues indicating that PINB-2 proteins are not seed specific. The gene expression was not found to be associated with grain hardness.

## 5.2 Introduction

As reported in Chapter 1 (Section 1.3.8) several gene and protein expression (Capparelli et al. 2003; Amoroso et al. 2004), transgenic studies (Li et al.2014) and fluorimetric studies (Kaczmarek et al.2015) have shown that the two PIN proteins act co-operatively, or in an inter-dependant manner. PINA and PINB act cooperatively to prevent polar lipid breakdown during seed maturation and the lipid-binding property of these proteins is likely involved in development of grain softness (Kim et al. 2012). *Pin* genes are undoubtedly involved in influencing grain hardness and the *Pinb-2* genes have been linked with minor grain texture variations (Section 1.3.7). Previous work in our laboratory has provided evidence of *in vivo* friabilin interactions (Ramalingam, 2012; Alfred et al. 2014). Ramalingam (2012) reported strong PINA(bait)-PINA (prey) and the PINA(bait)–PINB (prey) interactions and very weak interaction of PINB-2V1 with PINA and PINB. Alfred et al. (2014) analysed the role of the tryptophan-rich domain (TRD), hydrophobic domain (HD), Arg39Gly substitution in PINA and *Pinb-D1c* (grain hardness associated PIN mutation) on PPI (Section 1.3.10). The lipid binding properties of PINs have been associated with their functionality, the TRD being the most important region (reviewed in Bhave and Morris, 2008b, Feiz et al. 2009b). The Gly46Ser and Trp44Arg mutations

in the TRD of PINB have been associated with reduced lipid binding (Clifton et al. 2007). However, hardness associated mutations outside TRD like *Pinb-D1c* (Leu60Pro) suggest that regions outside TRD might be associated with lipid binding properties. As per a deduced tertiary structure of PINA by Alfred et al. (2014) the TRD (Phe34-Gly46) is located as an extended loop between α helix 1and 2 and the HD (Ile75-Phe85) is located as a smaller loop between α helix 3 and 4. Alfred et al. (2014) observed no effect on PPI for Leu60Pro, located in an alpha-helical section between the TRD and HD of PINB. Similarly, Arg39Gly substitution in PINA located in the TRD did not affect the PPI. They also suggested that the HD was found to be essential (but not sufficient) in higher-order association of PINs. In a recent study by Geneix et al. (2015) purified PIN proteins (PINA and PINB) from hard and soft wheat cultivars were analysed by dynamic light scattering, asymmetrical flow field fractionation and size exclusion chromatography. PINA formed small aggregates, mainly dimers in both hard and soft varieties while PINB isolated from hard varieties (PINB-D1b and PINB-D1d) assembled into large aggregates while PINB-D1a formed small aggregates. These results suggest that a single amino acid substitution in PINB can affect its self- assembly and interaction with PINA. Though, the evidence for PINA and PINB interactions is gathering but not much is known about the physical interactions of PINB-2 proteins. Only one study by Ramalingam (2012) reported weak PPI of PINB-2V1 with PINA/PINB. *Pinb-2* variants have been linked with grain texture variations. A variant form of *Pinb-2* variant 3 that has Val104Ala substitution has been linked with harder grain texture and its expression has been reported in seed. This suggests that PINB-2v3 might be involved in PINA/PINB co-operative binding and thus affecting grain texture. Though, PINB-2 proteins have two Trp residues in the TRD these might retain certain residues/regions required for starch granule association or PINA and/or PINB may co-operate with PINB-2's as well. Thus, this study queries the physical interaction of PINB-2V3 with PINA and PINB proteins.

It is well established that *Puroindoline* genes show strict seed specific expression (Zhang et al.2009, Dhatwalia et al. 2011). However, Wilkinson et al. (2008) reported low level of *Pinb-2* expression in developing grains of bread wheat cultivar 'Hereward' for the first time. Interestingly, a study by Chen et al. (2013b) reported the expression of these genes not only in seeds, different stages during grain filling, seedlings, roots and leaf tissues in bread wheat cultivars using real-time PCR. Thus, it is likely that expression of *Pinb-2* genes is not seed specific and it is worthwhile to investigate the expression of *Pinb-2*

genes. This study investigated the gene expression of *Pinb-2* genes in leaf and root tissues (harvested 21 days after germination) using semi quantitative reverse transcriptase PCR (sq-RTPCR) for quick analysis. Given the above background, the aims of this study were i) to investigate whether PINA, PINB and PINB-2V3 can physically interact with one another ii) to investigate if the *Pinb-2* genes express in different plant tissues iii) to investigate if the *Pinb-2* variants: *Pinb-2v2*, *Pinb-2v3-1*, *Pinb-2v3-1a* (Val104Ala substitution) identified in Chapter 3 express differentially and its association with grain hardness. The first aim was addressed using the yeast two-hybrid approach by the methodologies (Clontech, CA, USA; http://www.clontech.com/xxclt_ibcGetAttachment.jsp?cItemId=17602) explained in Chapter 2 (Section 2.12). The gene expression (second aim) was investigated using (sq-RTPCR).

## 5.3 Results

### 5.3.1 Confirmation of inserts in previously constructed clones

The gene sections of wild-type *Pina-D1a* (Genbank DQ363911) and *Pinb-D1a* (Genbank DQ363913) encoding mature protein-encoding sections had been cloned previously in our laboratory in both bait (pGBKT7) and prey (pGADT7) vectors and were kindly donated by Dr. Ramalingam (personal communication). The clone pGAD PINB-2v3 was constructed in this study. The presence of inserts in pGBKPINA, pGBKPINB, pGADPINA, pGADPINB was confirmed by PCR amplifications using insert-specific primers listed in Table 2.12. Bands at approximate positions of 363 bp for *Pina-D1a*, and 360 bp for *Pinb-D1a* were observed in bait and prey vectors (Fig. 5.1), as expected.

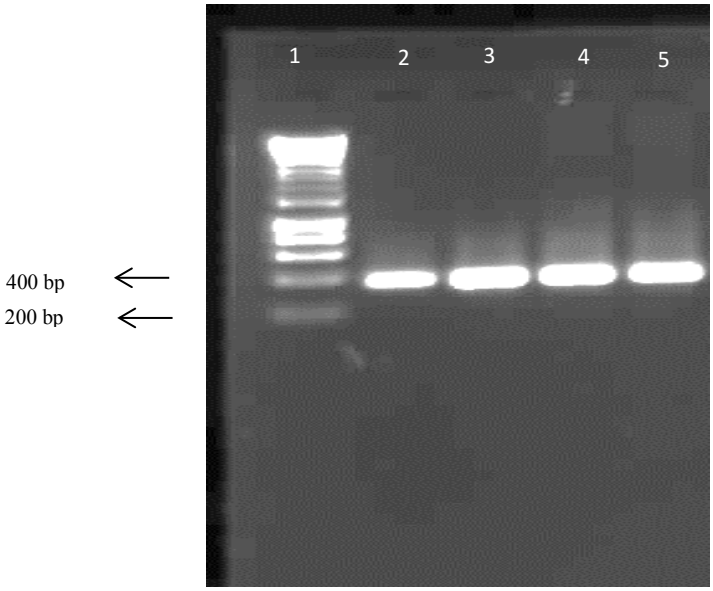



**Figure 5.1 Amplification of inserts in PINA and PINB Y2H constructs.**
PCR products from clones in bait (pGBKT7) and prey (pGADT7) vectors using insert-specific primers (Table 2.12). Lane 1: Hyperladder 1, lane 2: PINA-bait, lane 3: PINB-bait, lane 4: PINA-prey; lane 5: PINB-prey

### 5.3.2 Cloning of mature gene section of *Pinb-2v3-1*into pGADT7

The 395 bp of gene section corresponding to the mature protein was amplified from the pGEM® -T Easy clones of full length gene using primers that have restriction sites incorporated into them (Table 2.14). *Pinb-2v3* pGEM-T Easy clone (*Pinb-2v3*-1) of cultivar Sunco was obtained by gene cloning and sequencing as described in Chapter 3 (Section 3.3.8, Table 3.5) for *Pinb-2* gene diversity studies. The 395 bp PCR products of mature protein-encoding section that include the restriction sites (EcoR1 in the 5' section and BamH1 in the 3' section) are shown in Figure 5.2, Table 5.1.

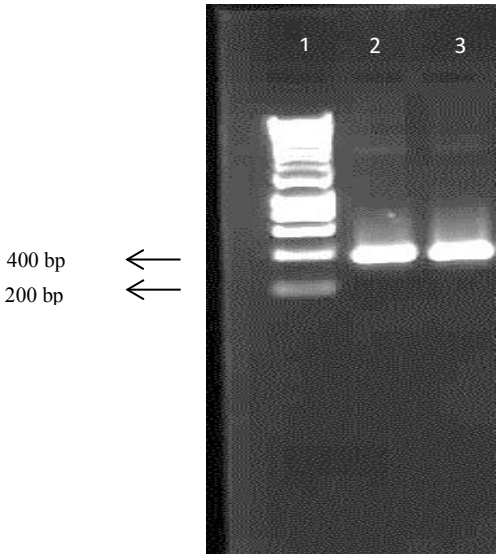



**Figure 5.2. PCR amplification of the mature protein encoding section of the *Pinb-2v3-1*gene with primers containing introduced restriction sites.** Lane 1: PGEM- T Easy *Pinb-2v3-1* clone, lane 2: PGEM- T Easy *Pinb-2v3-1* clone

**Table 5.1 Gene cloned into pGADT7 vector**

| Wheat cultivar | Gene | Genbank | Size of full length gene | Mature protein section of gene |
|---|---|---|---|---|
| Sunco | *Pinb-2v3-1* | AM944733, GQ496618 | 453 bp | 395 bp |

The vector pGADT7 with *Pina* insert was subjected to double digestion with BamH1 and EcoR1. Double digestion of the vector would release the insert (a band around 400bp) which would provide an indication that the digestion was complete (Fig. 5.3) The appropriate amount of PCR product (*Pinb-2v3-1*insert) to be included in a ligation reaction is based on the equation described in section 2.5.

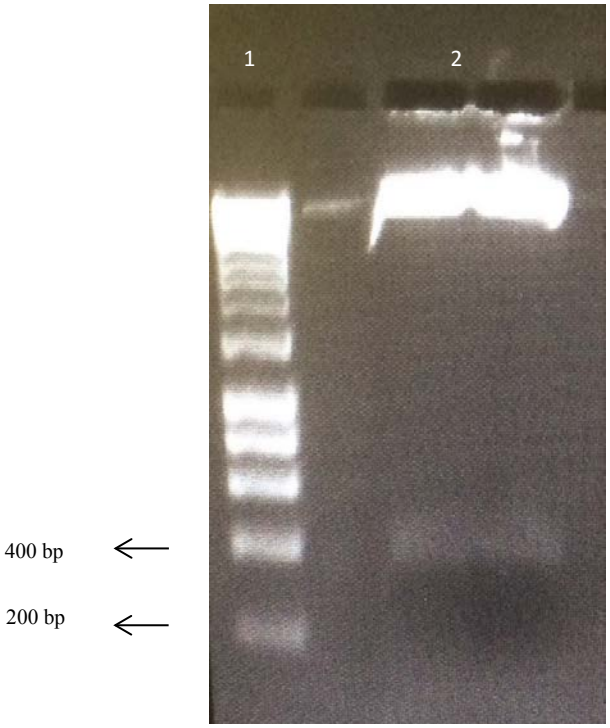



**Figure 5.3. Double digested yeast vector PGADT7**
Lane1: Hyperladder 1, lane 2: pGADT7 vector with a mature protein-encoding section of *Pina*, double-digested with EcoR1 & BamH1

The ligated insert (mature protein section of gene *Pinb-2v3-1*) and vector DNA (pGADT7) were transformed into *E.coli* JM109 cells. Upon transformation, colonies were obtained and the cloned products were grown on appropriate selective media. The pGADT7 clones were selected on LB plates with ampicillin. Screening of pGAD-*Pinb-2v3-1* constructs (designated as pGADv3-1) was carried out using colony PCR (Section 2.11.3, Fig. 5.4) using insert specific forward and reverse primers (Table 2.12). Based on this technique, the colony PCR product size expected for pGAD-*Pinb-2v3-1* construct is 395 bp (Fig. 5.4). Positive colonies (clones with inserts) were grown in LB media containing ampicillin and sequenced.

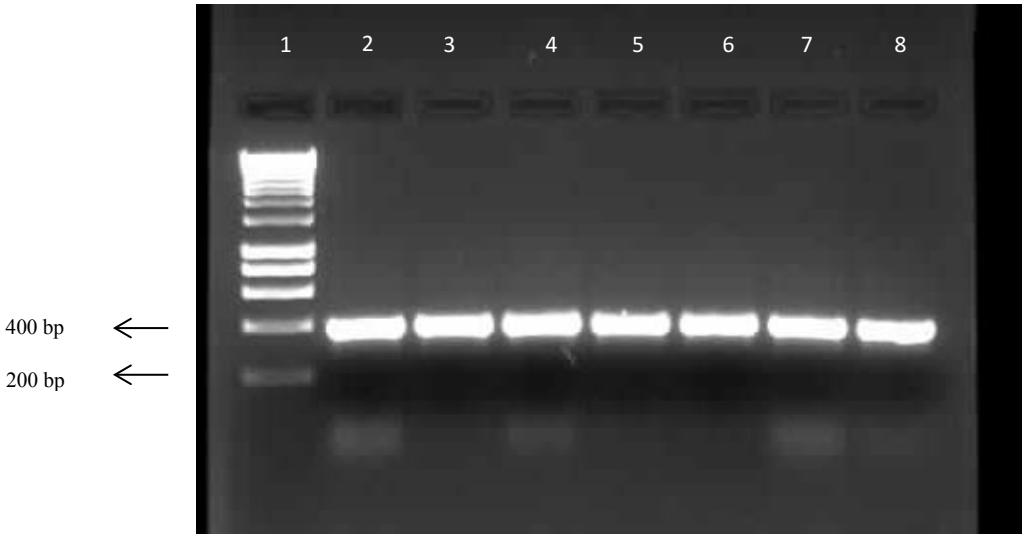



**Figure 5.4. Colony PCR for preliminary selection of clones**.
Lane 1: Hyperladder 1, lanes 2-8: PCR products of pGAD*Pinb-2v3-1* amplified using insert specific primers multiple clones of Sunco which selected?

The bait constructs (pGBKPINA, pGBKPINB), prey constructs (pGADPINA, pGADPINB, pGADv3-1) and empty vectors (pGADev, pGBKev) were transformed into yeast. pGADv3-1 is shown in the above gel. Yeast strains AH109 and Y187α were grown on YPD plates to have viable cells for transformation. The preparation of chemically competent yeast and transformation was performed per the lithium acetate method as described in Section 2.11.4. An example of plates containing prey vector pGBKPINB transformed into yeast strain AH109, is shown in Fig. 5.5 below.

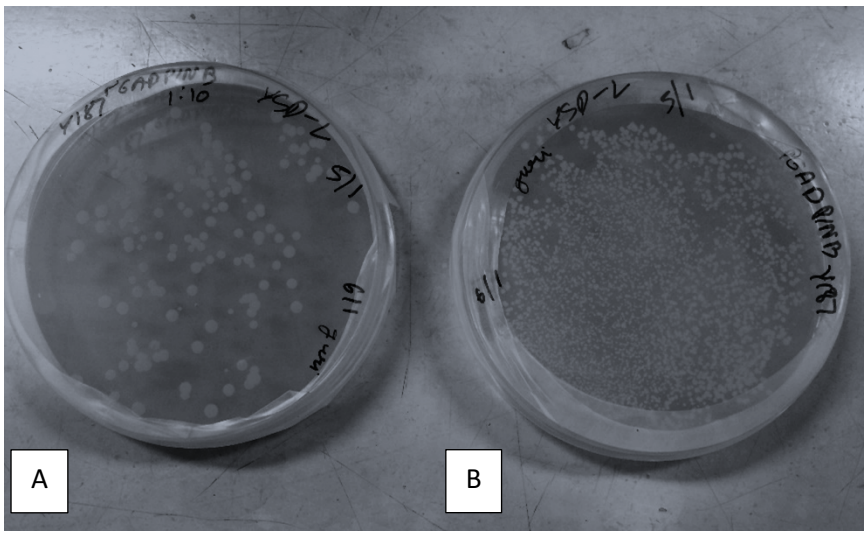



**Figure 5.5: Prey vector pGADPINB transformed into yeast strain Y187α**
A: 1:10 dilution of transformed yeast cells, B: neat

Yeast strains containing bait and prey vectors were mated in all pairwise combinations (Table 5.3) in pairwise fashion as described in Section 2.11.5. The matings included empty vector controls to rule out the possibility of false positives. All the mated pairs showed growth on YSD -W-L media indicating the presence of bait and prey plasmids. The interactions of prey and bait constructs were tested using plate assays as described in Section 2.11.6. The mated cultures were spotted onto replica selection plates containing selective media lacking Trp, Leu, Ade and His (YSD -W-L-A-H), that select for yeast expressing HIS3 and ADE2 reporter genes activated during bait-prey interactions. In addition, the cultures were spotted on duplicate plates containing YSD -W-L-A-H and 1 mM 3-amino-1,2,4-triazole (3-AT) to assess the strength of interaction between the proteins expressed by the bait and prey vectors. An example of plate-screening of yeast two-hybrid interactions listed in Table 5.3 is shown in Fig. 5.6. The PINA(prey/bait)-PINB (prey-bait) interactions were tested to establish controls for PINB-2v3-1(prey)-PINA/PINB (bait) interactions and comparison with previous data (Ramalingam, 2012, Alfred et al. 2014). Growth of yeast cells on YSD (W-L-A-H) media indicates interaction between the tested proteins expressed by the bait and prey plasmids and growth of yeast cells on YSD (W- L-A-H+1mM 3AT) indicates the strength of interaction between the expressed proteins (Yeast Protocols Handbook, Clontech,CA,USA; http://www.clontech.com/xxclt_ibcGetAttachment.jsp?cItemId=17602)).

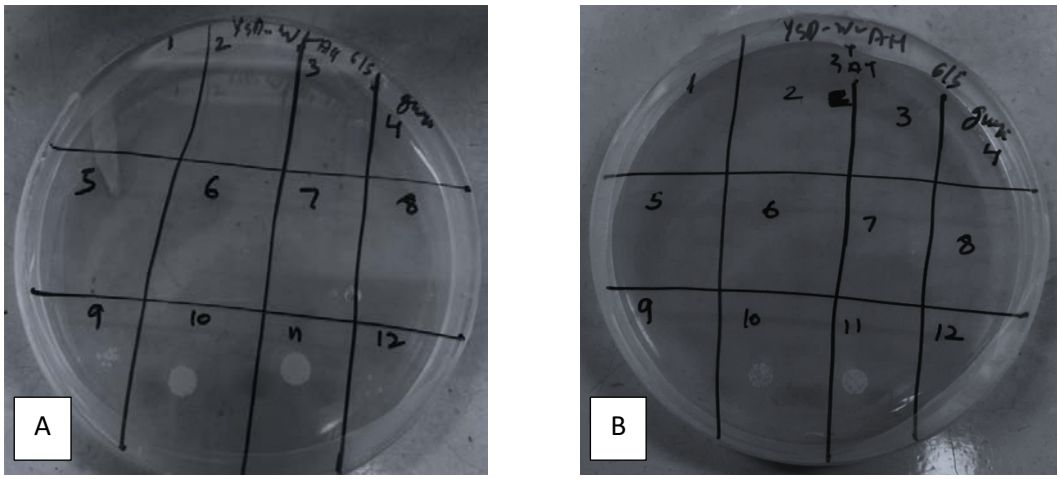



**Figure 5.6: Yeast two hybrid interactions of PIN proteins tested on selective plates**
A: (YSD-W-L-A-H), B: (YSD-W-L-A-H + 1mM 3AT) Details of each spot are given in Table 5.3.

Mated yeast cultures with empty vector control as one of the vector, did not show any growth on the YSD (W-L-A-H) suggesting that single expressed protein (PINA/PINB/PINB-2v3-1) could not induce the reporter gene activity on their own. Thus, cell growth in other mated cultures is indicative of protein-protein interactions. In addition, no growth was noticed for other two (PINB-2v3-1prey x PINA bait), (PINB-2v3-1prey x PINB bait) matings suggesting that these interactions were very low. These plate assays were scored between 0 (no growth observed) to 3 (strong growth observed) for cultures spotted on the interaction plates YSD (W-L-A-H) and YSD (W-L-A-H+1mM 3AT) as shown in Table 5.2. In case of YSD-W-L-A-H plates, the PINA-PINA, PINB (prey)-PINA (bait) interactions, had the highest score of 3. Next in line were the PINA (bait) - PINB (prey), with a score of 2. The PINB-PINB interactions were the weakest on these plates which were scored 1. PINB-2v3-1prey x PINA bait and PINB-2v3-1prey x PINB bait interactions were scored 0 (no growth on YSD-W-L-A-H). Only two interactions ie. PINA prey- PINA and PINB prey-PINA showed growth on YSD-WLAH + 3-AT plates and were scored 2 and 1 respectively. All other interactions were revealed to be very weak to negligible and were given a score of 0.

**Table 5.2 Analysis of PINA, PINB, PINB-2v3-1 interactions based on plate assays**

| **S. no.** | **Prey vector** | **Bait vector** | YSD (W-L) | YSD (W-L-A-H) | YSD (W-L-A-H + 1mM3AT) |
|---|---|---|---|---|---|
| 1 | pGADev | pGBKev | √ | 0 | 0 |
| 2 | pGADv3-1 | pGBK ev | √ | 0 | 0 |
| 3 | pGADv3-1 | pGBK PINA | √ | 0 | 0 |
| 4 | pGADv3-1 | pGBK PINB | √ | 0 | 0 |
| 5 | pGADev | pGBK PINA | √ | 0 | 0 |
| 6 | pGADev | pGBK PINB | √ | 0 | 0 |
| 7 | pGAD PINA | pGBKev | √ | 0 | 0 |
| 8 | pGAD PINB | pGBKev | √ | 0 | 0 |
| 9 | pGAD PINA | pGBK PINB | √ | 2 | 0 |
| 10 | pGAD PINB | pGBK PINA | √ | 3 | 1 |
| 11 | pGAD PINA | pGBK PINA | √ | 3 | 2 |
| 12 | pGAD PINB | pGBK PINB | √ | 1 | 0 |

(ev): empty vector, (√): growth on YSD(-W-L) media. The growth of spotted cultures was scored between 0 to 3. Based on Figure 5.7.

The β-galactosidase activity of the reporter gene (*lac Z*) for the mated yeast cultures involving PINA and PINB and PINB-2V3-1 in various bait or prey vector expression combinations were measured (tested in triplicates). The β-galactosidase activity of the reporter genes for these interactions, (Fig. 5.6), revealed a similar pattern of results (Table 5.4) as that of the plate assays. The interactions with empty vector controls were ≤ 1 Miller Unit. The relevant empty vector interaction values were subtracted from the interaction results as shown in Table 5.3. The PINB-2V3-1prey: PINB interaction value (-0.23 MU) was found to be statistically insignificant and hence not shown in Fig. 5.7. Maximum activity was observed for the PINA in the bait orientation, PINA prey: PINA bait (13.31 MU), PINA prey: PINB bait (7.61 MU). Reduced activity was observed when PINA was present in the prey orientation i.e. PINB prey: PINA bait (5.56 MU), indicating the importance of protein orientation for the interaction. Next strongest were the PINB prey-PINB bait interactions (3.47 MU). Very little β-galactosidase activity was observed in PINB-2V3-1prey-PINA bait (0.87 MU) and the least in PINB-2V3-1prey: PINB bait (-0.23 MU). A similar pattern of PINA-PINB interactions was noted in previous work carried out in our laboratory. Strongest β-galactosidase activity was observed when PINA was present in bait orientation followed by PINA prey-PINB bait and PINB bait- PINB prey showed least activity (Ramalingam, 2012, Alfred et al. 2014).

**Table 5.3 Results of PINA, PINB and PINB-2V3-1 interactions**

| S.no. | Prey vector | Bait vector | YSD-WLAH[a] | YSD-WLAH +3AT[a] | β-galactosidase activity (MU) | Controls | Average of controls | β-galactosidase activity (MU) (corrected)[b] |
|---|---|---|---|---|---|---|---|---|
| 1 | pGADev | pGBKev | - | - | 0.43 ± 0.03 | NA | | |
| 2 | pGADv3-1 | pGBK ev | - | - | 0.14 ± 0.02 | NA | | |
| 3 | pGADv3-1 | pGBK PINA | - | - | 1.33 ± 0.15 | 1,2,5 | 0.46 | 0.87 ± 0.15 |
| 4 | pGADv3-1 | pGBK PINB | - | - | 0.06 ± 0.01 | 1,2,6 | 0.29 | -0.23 ± 0.01 |
| 5 | pGADev | pGBK PINA | - | - | 0.81 ± 0.04 | NA | | |
| 6 | pGADev | pGBK PINB | - | - | 0.3±0.04 | NA | | |
| 7 | pGAD PINA | pGBKev | - | - | 0.55 ± 0.02 | NA | | |
| 8 | pGAD PINB | pGBKev | - | - | 0.41±0.04 | NA | | |
| 9 | pGAD PINA | pGBK PINB | ++ | - | 5.98 ± 0.52 | 1,6,7 | 0.42 | 5.56 ± 0.52 |
| 10 | pGAD PINB | pGBK PINA | +++ | + | 8.16 ± 0.41 | 1,5,8 | 0.55 | 7.61 ± 0.41 |
| 11 | pGAD PINA | pGBK PINA | ++++ | ++ | 13.9 ± 1.14 | 1,5,7 | 0.59 | 13.31 ± 1.14 |
| 12 | pGAD PINB | pGBK PINB | + | - | 3.85 ± 0.43 | 1,6,8 | 0.38 | 3.47 ± 0.43 |

[a]Growth is scored as no growth (-), very weak (+), weak (++), medium (+++), strong (++++) [b]β-galactosidase activity is the mean of triplicate experiments corrected for average of the relevant empty vector negative control values

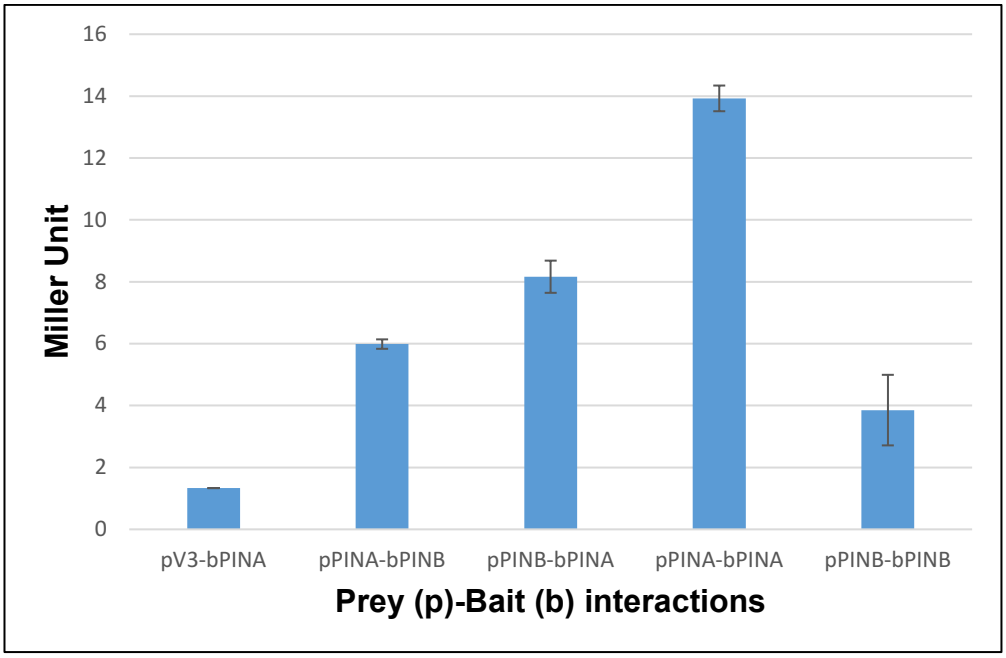



**Figure 5.7. β-galactosidase activities of PINA, PINB and PINB-2V3-1 interactions.**
b: bait (pGBKT7), p: prey (pGADT7)

### 5.4. Gene expression of *Pinb-2* variants

RNA was extracted from leaf and root tissues of three wheat cultivars (three replicates each) using TRIsure (Bioline) as described in Section 2.12. The three cultivars chosen for analysis represented different *Pinb-2* variant type. These cultivars were: Halbred (*Pinb-2v2*), Sunco (*Pinb-2v3-1*) and Tasman (*Pinb-2v3-1a*). The SKCS grain hardness value of these cultivars was 70.27, 71.98, 52.39 units respectively. All the cultivars had *Pina-D1a/Pinb-D1b* as background *Pin* genotype. cDNA was synthesized from these cultivars (Section 2.12) and used for gene expression analysis using semi-quantitative reverse-transcriptase PCR. The gene specific primers were designed as per Chen et al. (2013b) and are listed in Table 2.15. Relative gene expression was calculated using *actin* as housekeeping control that exhibit relatively constant expression (Supronova et al. 2004). Each cDNA preparation was tested for quality (lack of gDNA contamination) by amplification using intron-flanking actin primers (Section 2.4.1, Table 2.15). gDNA was used as a control in the in the PCR, different product size in cDNA and gDNA indicated the lack of genomic DNA contamination in the cDNA preparation. Thus, indicating the PCR results will be specific for gene expression. Another control that was used to rule out gDNA contamination was the -RT control (Section 2.12.3) which showed no amplification in the RT PCR. The cDNA was used for further analysis if the PCR resulted

in a single band at approximately 147 bp (expected in cDNA) and no 440 bp product (which would be expected in gDNA) as seen in Fig. 5.8.

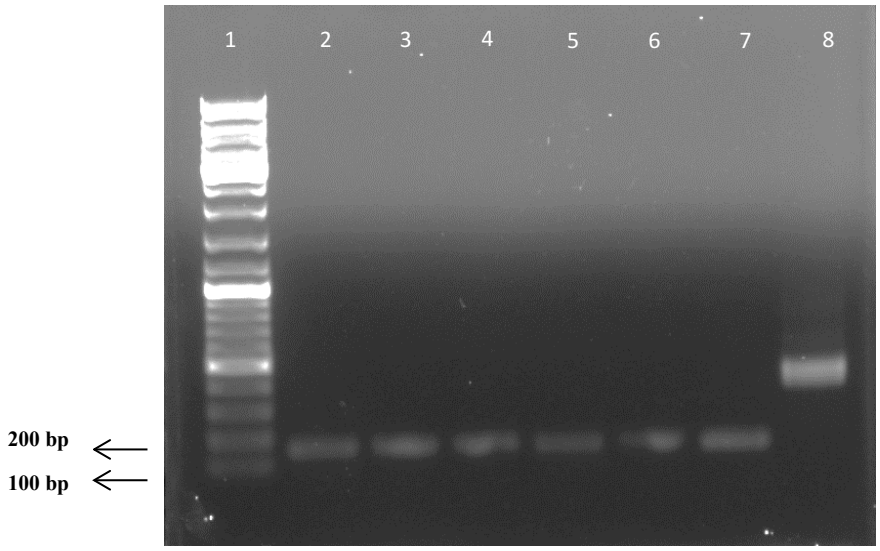



**Figure 5.8: Example of a gel image showing quality cDNA**
Lane 1: Generuler TM 100 bp ladder: lanes 2-7: Actin PCR products from different cDNA preparations, lane 8: PCR product from wheat gDNA

The cDNA was used for sq-RTPCR of the three genes, whose success was noted by the amplification of bands at the expected size of 146 bp (Table 2.15) when electrophoresed on an agarose gel. The bands were then quantitated using ChemiDoc$^{TM}$ XRS+ Documentation Station and Image lab software (Bio-Rad). The relative expression of the *Pinb-2* variant genes was determined using the formula mentioned in Section 2.4.2. When the gene expression level of *Pinb-2* variants was compared with each other, the expression of *Pinb-2v3-1a* was found to be higher than that of *Pinb-2v2* and *Pinb-2v3-1* in leaf and root tissues (Table 5.4). The expression of *Pinb-2* variants was found higher in leaf tissue as compared to root. These observations are in line with an earlier study conducted by Chen et al. (2013b). Based on the results of quantitative real-time PCR it was reported that the relative expression of *Pinb-2v3b* (Val104Ala) was higher than *Pinb-2v2* and *Pinb-2v3*. However, Chen et al. (2013b) suggested that the increased gene expression might be associated with the increased grain hardness value of *Pinb-2v3b* cultivars as compared to *Pinb-2v2* and *Pinb-2v3a*. It is important to mention here that *Pinb-2v3a* is identical to *Pinb-2v3-1* except a G6T SNP in the signal peptide and *Pinb-2v3b* (Val104 Ala) is identical to *Pinb-2v3-1a* (Val104 Ala) except a G6T SNP in the signal peptide. The substitution (Val 104 Ala) has been associated with increased grain hardness (Chen et al.2013b). In this study, the greater expression level of *Pinb-2v3-1a* could not be

associated with increase in grain hardness since the SKCS hardness value of *Pinb-2v3-1a* cultivar (52.39 units) was found to be lower than *Pinb-2v3-1* (71.98 units) *and Pinb-2v2* (70.27 units) cultivars.

**Table 5.4 Relative gene expression of *Pinb-2* variants**

| Gene | Relative expression Leaves | Relative expression Roots |
|---|---|---|
| *Pinb-2v2* | -1.49 ± 0.12 | -1.56 ± 0.07 |
| *Pinb-2v3-1* | -1.56 ± 0.17 | -1.75 ± 0.14 |
| *Pinb-2v3-1a* | -1.35 ± 0.10 | -1.44 ± 0.07 |

## 5.5 Discussion

It is well established that PINA and PINB bind to lipids and their association with starch granules is mediated by the co-operation between the two proteins (reviewed in Bhave and Morris, 2008b) and the biochemical basis for soft grain texture is the presence of both PINA and PINB in their native form (Morris, 2002). A tight interaction between puroindolines and lipids at the starch granule surface was shown by Pauly et al. (2014). Transgenic studies have also demonstrated the simultaneous requirement of both native proteins to confer soft phenotype (Hogg et al. 2004; Martin et al. 2006). Previous work based on the use of yeast-two hybrid system has provided evidence of *in-vivo* friabilin interactions (Ramalingam, 2012, Alfred et al. 2014) There is sufficient information available on the PINA/PINB protein interactions. However, not much is known about the PINB-2: PINA and or /PINB protein interactions except one study (Ramalingam, 2012) that reported weak protein-protein interactions of PINB-2V1 with PINA/PINB. *Pinb-2* variant 3 has been found to be a predominant in Australian cultivars (Chapter 3) and landraces (Chapter 4) and has been linked with increased grain hardness (Chen et al. 2013b). Thus, this study investigated the PINB-2V3: PINA/PINB interactions that might help to understand the role of *Pinb-2* genes in grain texture moderation. The interactions of PINB-2v3-1 with PINA and PINB were very low since no colony growth was observed and very little β-galactosidase activity (1.33 MU) was detected. These interactions can be considered negligible as compared to PINA: PINA, PINA: PINB and PINB: PINB interactions. These findings suggest that interaction between them is likely not required for lipid binding at starch granule surface that moderates the grain texture. An earlier

study by Ramalingam (2012) reported very low PINB-2v1-1: PIN interactions (zero growth on plates and less than 1.5 MU β-galactosidase activity)

A likely explanation for the very low interaction of PINB-2v1 and PINB-2v3 as compared to PIN: PIN interactions might be the difference in the TRD and HD of these proteins. If the TRD is critical for PPI the five tryptophan (Trp) and three basic residues in the TRD of PINA (FPVTWRWWKWWKG) may impart stronger hydrophobic and ionic interactions, respectively, compared to the TRD of PINB (FPVTWPTKWWKGG) with only three Trp and two basic residues, and that of PINB-2V1(FSIARLLKWWKGA), PINB-2V3 (FPISTLLKWWKGA) with only two Trp and two (V3)/three (V1) basic residues. It is important to mention here that although the PINB-2v1 and PINB-2v3, unlike the PINs have two highly hydrophobic residues, Leu and an Ile at the TRD, these residues do not appear to play role in PPI. Thus, it seems that Trp residues in the TRD help in association of the PIN proteins. Alfred et al. (2014) hypothesised that individual PINA and PINB proteins associate via their TRD with the endosperm polar lipids and starch granule surface. Later, the binding of TRD is enabled by formation of PIN homo- or hetero-dimers/oligomers via polar and/or hydrophobic interactions between residues on exposed loops (such as the HD) or helix surfaces. The HD (residue 105-115 of mature protein) of PINB-2v1 and PINB-2v3 (VQGNLGGIF) has I to V, R to N, F to I and L to F substitutions as compared to HD of PINB (IQGRLGGFL). The replacement of phenylalanine(F) with isoleucine(I) and leucine(L) with phenylalanine (F) does not appear to affect the hydrophobicity of the domain. However, the replacement of a basic residue Arginine (R) with a polar residue Asparagine (N) might lower the overall net positive charge of this domain. Decreased positive charge has been associated with reduced lipid binding and might affect the interaction of these proteins. Similarly, the replacement of a more hydrophobic residue Isoleucine (I) with lesser hydrophobic residue (Valine) leads to decreased hydrophobicity that might affect the PPI. Thus, based on above results, it is proposed that the basic residue Arginine and hydrophobic residue Valine are critical for PPI. Thus, it appears that the co-operative binding of PINA and PINB involved in moderation of grain texture does not likely involve the PINB-2v3 proteins though linked with minor role in grain hardness variation. PINB-2v3 proteins are most likely functional without the need to interact with PINA or PINB. The PINB-2 variant proteins seem to have greater significance as antimicrobial proteins and make minor contribution in affecting grain hardness. The low interactions may be due to the

different biological function of *Pin* and *Pinb-2* genes. *Pinb-2* genes are actively expressed in developing seeds (Wilkinson et al. 2008), while the *Pin* genes are most active during the middle stages of endosperm development and are found in the mature endosperm (Gautier et al. 1994, Capparelli et al. 2005). The *Pin/Pinb-2* gene expression at different stages may correspond to defence mechanisms of the wheat seed at different stages such as developing seeds (grain filling), mature endosperm and the young plant (seedlings).

It is well established that expression of *Pina and Pinb* genes is seed-specific (Zhang et al. 2009, Dhatwalia et al.2011). Transcriptome analysis of the wheat seeds has detected the expression of *Pina and Pinb* genes (Gilles et al.2012, Singh et al.2014, Nirmal et al.2016). Based on their expression pattern in wheat seeds the *Pins* seem to have an endosperm specific role. Wilkinson et al. (2008) noted *Pinb-2* transcripts in developing grains for the first time and later Giroux et al. (2013) reported that *Pinb-2* variants are expressed at low levels as compared to *Pin* genes in wheat seeds. Only one study so far by Chen et al. (2013b) reported the expression of *Pinb-2* genes in seeds (different stages during filling), leaves and root tissues. However, a major drawback of the above- mentioned study was the use of actin primers that were not found to span introns and exons and thus suggesting their non-specificity. This study investigated the gene expression of *Pinb-2* variants in leaf and root tissues of Australian wheat cultivars using intron flanking actin primers. *Pinb-2* variants were found to express in leaf and root tissue thus confirming the report of Chen et al. (2013b). However, the increased gene expression level of *Pinb-2v3-1a* was not found to be associated with grain hardness values. Chen et al. (2013b) reported such association in soft wheats (*Pinb-D1a*/*Pinb-D1a*) and the *Pinb-2v3-1a*, *Pinb-2v3-1*, *Pinb-2v2* cultivars investigated in this study had *Pina-D1a/Pinb-D1b* genotype. *Pinb-2* genes might express differentially between soft and hard wheats. However, per a recent study by Nirmal et al. (2016) there was no difference in expression of *Pinb-2* genes in mutant (*Pina-D1a/Pinb-D1b*, *Pina-D1b/Pinb-D1a*) and non-mutant *Pin* (*Pina-D1a/Pinb-D1a*) genotypes. Thus, the relationship between *Pinb-2* gene expression level and grain hardness is unclear and needs further investigation.

# CHAPTER 6

# Study of diversity of *Hin* genes in Australian barley germplasm and worldwide landraces

## 6.1 Abstract

Grain texture is an important trait affecting malting and processing quality of barley with relevance to its end-use. Hard textured barley is generally preferred for stock feed while soft is preferred for malting and brewing. Hordoindolines (HIN) are seed proteins of barley, similar to the wheat puroindoline (PIN) proteins. Wheat *Puroindoline* genes *Pina* and *Pinb* are the major genetic determinants of its grain texture, which have a major effect on its end-use properties. The *Hordoindoline* genes, *Hina* and the duplicates *Hinb-1* and *Hinb-2*, are orthologs of these, and the *Hinc* may be an orthologue of the wheat *Pinb-2* genes which make minor contributions to wheat texture. Significant information is available on the effects of diverse mutations in *Pin* genes on wheat grain texture, but information on contributions of genetic variants of *Hin* genes to barely grain texture is limited. The present study investigated the genetic diversity of *Hin* genes and encoded proteins in 12 Australian barley cultivars and 14 landraces chosen from important centres of diversity around the world. The seed samples used for determining SKCS grain hardness were provided by AWCC and Dr. Joe Panozzo. The cultivars represented a range of grain hardness (24.2-58.1 SKCS units) and the landraces ranged from 48.7- 94.4 SKCS units. In addition, grain texture of seeds visualised by Scanning Electron Microscopy (SEM) exhibited a tight binding of starch granules to the protein matrix. Molecular analyses undertaken by amplification of *Hin* genes and DNA sequencing indicated notable genetic diversity, with altogether seven haplotypes detected for *Hina*, six for *Hinb-1* and eight for *Hinb-2*. Interestingly, some of the haplotypes for all three genes were shared by the cultivars and/or landraces. Some single nucleotide polymorphisms led to substitutions at the functionally important tryptophan rich domain (TRD) and certain basic and hydrophobic residues in the putative proteins. This study suggested that *hordoindoline* sequence variation may play some role along with other factors that affect barley grain hardness.

## 6.2 Introduction

As detailed in Chapter 1, barley is an important cereal grain that has major contribution to the feed, food and malting industry. Soft textured barley is preferred for malting and brewing whereas hard textured for feed uses. Grain hardness is a key trait related to barley grain quality (Section 1.4.3). *Hordoindoline* genes are the genetic determinants of grain texture in barley and are orthologous to wheat *Puroindolines*. *Hordoindoline* genes comprise of *Hina*, *Hinb-1* and *Hinb-2* genes. The *Hin* genes show high homology to wheat *Pin* genes (*Hina-Pina*: 88%, *Hinb-1-Pinb*: 92%, *Hinb-2-Pinb*: 91%) (Section 1.4.7). Grain hardness in barley is also affected by traits such as grain weight, grain size, grain protein content, β-glucans and several QTLs related to grain hardness have been reported (reviewed in Walker et al. 2013). Wheat and barley belong to the same family (Poaceae), subfamily Pooideae and tribe Triticeae (Gaut, 2002); however, barley is typically harder than hard wheat. Unlike wheat, where a clear relationship has been demonstrated between single nucleotide polymorphisms (SNPs) or other mutations in the *Pin* genes and wheat grain texture (reviewed in Bhave and Morris, 2008a, Feiz et al. 2009a), literature on the impact of SNPs or other mutations in *Hina* and *Hinb* gene on barley grain hardness seems to be largely inconclusive. However, the association of *Hin* genes with barley grain hardness became evident when Takahashi et al. (2010) reported a *Hinb-2* null mutation with increased grain hardness (Genbank accession: AY644051 This study analysed barley cultivars and landraces representing a range of SKCS values to investigate possible associations between genetic variability and grain hardness. A selected set of lines were analysed by SEM to determine grain texture. Thus, this study aims to further clarify the relationship between genetic variations in *Hin* genes and grain hardness, so as to identify some SNPs in *Hin* genes that can impact grain hardness and identify mutations that might result in greater antimicrobial properties and thus be helpful to synthesize anti-microbial peptides (AMPs). Similar to Puroindolines, Hordoindolines have also been reported to be a source of bactericidal and fungicidal peptides (Phillips et al. 2011). The study investigated *Hin* genes of Australian barley cultivars and worldwide landraces. Twelve Australian barley cultivars (eight malt and four feed types), and 14 landraces (from Afghanistan, China, Ethiopia, Iran, Japan, Kyrgystan, Morocco, Nepal, North Korea and Russia) covering important centres of barley origin and diversity (Newman and Newman, 2008), were analysed for *Hin* gene sequence diversity.

## 6.3 Results

### 6.3.1 Genomic DNA extraction, SKCS grain hardness and amplification of *Hin* genes

Genomic DNA was extracted from 12 Australian barley cultivars and 14 landraces (Table 6.1). The method used for genomic DNA extraction from single seeds is described in Section 2.3.1. The seed samples provided by AWCC (landraces) and Dr. Joe Panozzo (cultivars) were used for genomic DNA extraction and determination of SKCS grain hardness (Section 2.2.4). The seed samples were not from replicated field trials and thus the environmental variation in SKCS grain hardness values could not be quantified. As these samples were not grown in the same field trial, a direct comparison of grain hardness values is not possible. The SKCS grain hardness values obtained in this study are not used as absolute values for comparison among cultivars/landraces. These values have been used for comparison in a broader context. Fox et al. (2007a) has however reported that the majority of the variance in grain hardness is attributed to genotype with minimal effects of environment. It was also suggested that the SKCS is a reliable method for determining barley grain hardness and is highly heritable. The SKCS values for the cultivars were found to range from 24.2 to 58.1 units, while those for the landraces ranged from 48.7 to 94.4 units. The cultivars SloopSA (24.2 units), Capstan (33.1 units) and Barque (36.7 units) had the three lowest values among all lines tested. In comparison, several landraces had values in the 85-95 units range, indicative of harder grains. The *Hina*, *Hinb-1* and *Hinb-2* genes were amplified by PCR using the primers listed in (Table 2.13). These primers were designed as reported by Takahashi et al. (2010) and amplify full length genes i.e. the coding sequence. The presence or absence of *Hina*, *Hinb-1* and *Hinb-2* genes was determined by analysing the PCR products with agarose gel electrophoresis. Unlike wheat, where the lack of amplification of a *Pina* PCR product is an indication of *Pina* null alleles, such phenomenon was not observed in barley and has not been reported so far. The *Hina*, *Hinb-1* and *Hinb-2* PCR products were found to be present for all the barley samples used in this study. Negative controls (no DNA template) were included in PCR (Figure 6.1 and Figure 6.2). All the cultivars and landraces gave positive amplification for *Hina*, *Hinb-1* and *Hinb-2* genes. Amplicons of size 634 bp, 535 bp and 576 bp were observed for *Hina*, *Hinb-1* and *Hinb-2* genes respectively. The PCR products were purified and used as template for direct sequencing using forward and reverse primer without any cloning since the genes occur as single copies on chromosome 5H of barley genome. All barley cultivars and landraces were subjected to two

independent rounds of PCR and sequencing to rule out any PCR introduced nucleotide changes.

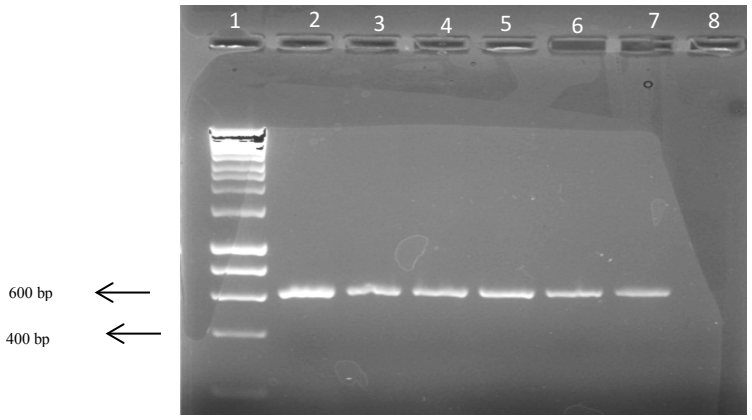



**Figure 6.1. Example of *Hina* PCR products of barley cultivars**
Lane 1: DNA molecular weight marker (Hyperladder 1), lane 2: Franklin, lane 3: Baudin, lane 4: Keel, lane 5: Hamelin, lane 6: Arivat, lane 7: Vlamingh, lane 8: negative control (no template).

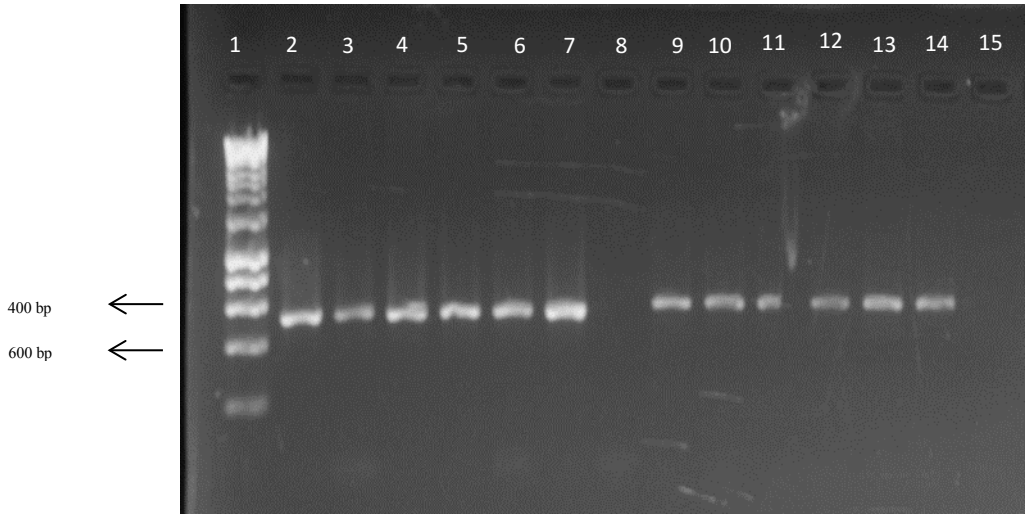



**Figure 6.2. Example of *Hinb-1* and *Hinb-2* PCR products of barley cultivars**
Lane 1: DNA molecular weight marker (Hyperladder 1); lane 2: *Hinb-1* PCR product in Franklin, lane 3: *Hinb-1* PCR product in Baudin, lane 4: *Hinb-1* PCR product in Keel, lane 5: *Hinb-1* PCR product in Hamelin, lane 6: *Hinb-1* PCR product in Arivat, lane 7: *Hinb-1* PCR product in Vlamingh, lane 8: negative control (no template), lane 9: *Hinb-2* PCR product in Franklin, lane 10: *Hinb-2* PCR product in Baudin, lane 11: *Hinb-2* PCR product in Keel, lane 12: *Hinb-2* PCR product in Hamelin, lane 13: *Hinb-2* PCR product in Arivat, lane 14: *Hinb-2* PCR product in Vlamingh, lane 15: negative control (no template)

### 6.3.2 *Hin* gene sequence diversity in barley cultivars and landraces

The *Hina*, *Hinb-1 and Hinb-2* sequences (AAV49987.1, AAV49986.1, AAV49985.1) reported by Caldwell et al. (2004) were used as reference for comparisons. Nucleotide variation was observed for *Hina*/*Hinb-1*/*Hinb-2* genes in the seven cultivars. Interestingly, nucleotide variation in all three genes was found in cultivars Stirling and Barque. No variation was found in any of the three *Hin* genes (*Hina, Hinb-1 and Hinb-2)* of five cultivars, i.e., Arapiles, Baudin, Buloke, Franklin, Vlamingh. Seven landraces, i.e., L400619 from North Korea, L402448 from Iran, L403089 from Russia, L403156 from China, L411152 from Morocco, L411813 from Kyrgystan and L491320 from Japan, were wild-type for all genes. Four landraces (L402433 from Afghanistan, L411144 from Morocco, L411145 from Morocco, and L411859 from Nepal) showed variations in all three genes, and the rest showed some SNPs. Interestingly, an insertion (INS210A) in *Hinb-1* resulting in a frame-shift mutation was detected in L400211 from Ethiopia. No other mutations such as insertions/deletions (INDELs) or frame-shifts were found in any other line. The results are summarised in Table 6.1 and detailed below.

**Table 6.1: *Hina*, *Hinb-1* and *Hinb-2* genes in barley cultivars and landraces**

| No. | Cultivar (Malt/Feed) | Origin | SKCS | *Hina* | *Hinb-1* | *Hinb-2* |
|---|---|---|---|---|---|---|
| 1 | Arapiles (Malt) | Australia | 46.9 | WT | WT | WT |
| 2 | Arivat (Feed) | Australia | - | M | WT | M |
| 3 | Barque (Feed) | Australia | 36.7 | M | M | M |
| 4 | Baudin (Malt) | Australia | 49.2 | WT | WT | WT |
| 5 | Buloke (Malt) | Australia | 58.1 | WT | WT | WT |
| 6 | Capstan (Feed) | Australia | 33.1 | M | M | WT |
| 7 | Franklin (Malt) | Australia | 54.6 | WT | WT | WT |
| 8 | Hamelin (Malt) | Australia | 55.4 | WT | WT | M |
| 9 | Keel (Feed) | Australia | 40.2 | M | WT | WT |
| 10 | SloopSA (Malt) | Australia | 24.2 | WT | WT | M |
| 11 | Stirling (Malt) | Australia | 56.4 | M | M | M |
| 12 | Vlamingh (Malt) | Australia | - | WT | WT | WT |
| **No.** | **Landrace Aus Number** | **Origin** | **SKCS** | ***Hina*** | ***Hinb-1*** | ***Hinb-2*** |
| 1 | L400097 | Ethiopia | 64.7 | M | M | WT |
| 2 | L400211 | Ethiopia | 87.4 | M | M | Not available |
| 3 | L400619 | North Korea | 94.4 | WT | WT | WT |
| 4 | L402433 | Afghanistan | 89.3 | M | M | M |
| 5 | L402448 | Iran | 87.9 | WT | WT | WT |
| 6 | L403089 | Russia | 85.2 | WT | WT | WT |
| 7 | L403156 | China | 84.7 | WT | WT | WT |
| 8 | L411144 | Morocco | 63.0 | M | M | M |
| 9 | L411145 | Morocco | 54.7 | M | M | M |

| | | | | | | |
|---|---|---|---|---|---|---|
| 10 | L411152 | Morocco | 48.7 | WT | WT | WT |
| 11 | L411859 | Nepal | 80.4 | M | M | M |
| 12 | L411813 | Kyrgystan | 58.4 | WT | WT | WT |
| 13 | L411890 | Nepal | 84.5 | M | WT | M |
| 14 | L491320 | Japan | 61.8 | WT | WT | WT |

WT- wild type, M- mutant

### 6.3.3 Nucleotide variations in *Hina* gene and amino acid substitutions in the putative protein

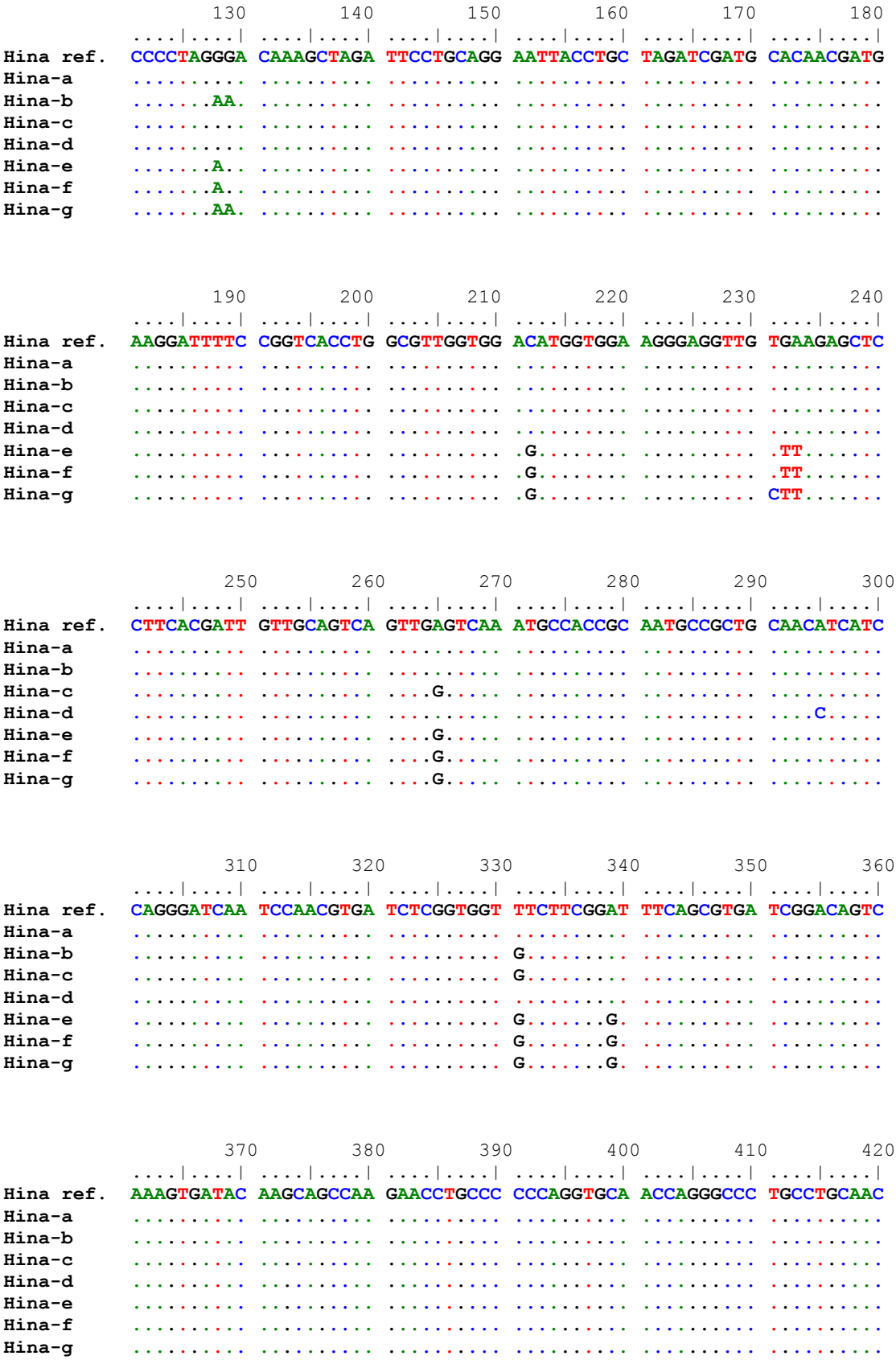
Variations in *Hina* genes were observed in five cultivars (Arivat, Barque, Capstan, Keel, Stirling), and seven landraces (L400097, L400211, L402433, L411144, L411145, L411859, L411890). A total of thirteen SNPs was found in the cultivars (Table 6.2), six of which (G128A, C212G, A265G, G232T, A295C, T331G) led to amino acid substitutions and seven (G21A, C54T, A114G, G129A, T231C, A233T, A339G) were silent mutations. Twelve SNPs were also found in the landraces, being shared with cultivars. One SNP A295C was unique to only cultivars. Five of the SNPs led to amino acid substitutions and six were silent. The SNP A295C (Keel) observed in this study has not been reported earlier.

A detailed analysis of the SNPs revealed seven *Hina* gene haplotypes or alleles (combinations of co-inherited nucleotides), designated as *Hina-a* to *Hina-g* (Fig. 6.3, Table 6.3). Three of these were shared between cultivars and landraces. The wild-type haplotype *Hina-a* (Caldwell et al. 2004) was shared by seven cultivars (Arapiles, Baudin, Buloke, Franklin, Hamelin, Sloop SA, Vlamingh) and seven landraces (L400619 L402448, L403089, L403156, L411152, L411813, L491320). The sequence of the haplotype designated here as *Hina-c,* was noted in two landraces (L411859, L411890), has been reported previously in domesticated barley *Hordeum vulgare subsp. vulgare* (Caldwell et al. 2006). Likewise, the sequence shared by cv. Arivat and Barque, both feed grade barley and designated as *Hina-g* has been reported in *Hordeum vulgare subsp. vulgare* cultivar Tolar (52.4 SKCS units at one location and 65.1 SKCS units at other location (Galassi et al. 2012). Four haplotypes, *Hina-b, Hina-d, Hina-e* and *Hina-f*, are novel. Identical haplotypes were observed in the landraces L411144 and L411145 (Morocco); L411859 and L411890 (Nepal) and L400097 and L400211 (Ethiopia).

Upon translation, the seven DNA haplotypes led to five putative protein types, designated as HINA-A to HINA-E (Fig. 6.4, Table 6.3). The amino acid residues were counted from the initiator methionine for convenience, rather than the first residue of the mature protein (Glu29) as undertaken for PINs in some publications. Nine cultivars and seven landraces exhibited HINA-A (identical to the reference sequence). All ten characteristic Cysteine residues were completely conserved, and the putative TRD (FPVTWRWWTWWKG) and HD (hydrophobic domain: IQRDLGGFFGF) were also conserved in all, except for a Thr71Arg substitution at the TRD in HINA-E. The TRD of HINs was predicted based on alignment with PINA-D1(Genbank accession: CAA49538) and PINB-D1(Genbank accession: CAA49537). The HD of HIN proteins was predicted based on alignment with HD of PIN proteins as predicted by Alfred et al. (2014). HINA-B contained a single substitution, Gly43Glu. Two substitutions, i.e., Ser89Gly and Phe111Val, were characteristics of HINA-C, while Ile99Leu was characteristic of HINA-D. Five substitutions (Gly43Glu, Thr71Arg, Glu78Leu, Ser89Gly, Phe111Val) were present in HINA-E, the most common type, shared by three cultivars and five landraces. The predicted pI of the putative mature proteins ranged from 8.39 to 8.72 (Table 6.3), confirming their basic nature. HINA-E had highest pI, probably due to its unique Arg71. HINA-B had the lowest pI (8.39) followed by HINA-C and HINA-D that depicted similar pI values (8.56).

```
                10         20         30         40         50         60
          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
Hina ref. ATGAAGGCCT TCTTCCTCGT GGGTCTGCTT GCTTTGGTAG CGAGCGCCGC CTTCGCGCAG
Hina-a    .......... .......... .......... .......... .......... ..........
Hina-b    .......... .......... .......... .......... .......... ..........
Hina-c    .......... .......... .......... .......... .......... ..........
Hina-d    .......... .......... .......... .......... .......... ..........
Hina-e    .......... .......... A......... .......... .......... ...T......
Hina-f    .......... .......... A......... .......... .......... ...T......
Hina-g    .......... .......... .......... .......... .......... ..........

                70         80         90        100        110        120
          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
Hina ref. TACGGAGAAG TTGTTGGCAG TTACGAGGGT GGTGCTGGTG GGGGTGGTGC TCAACAATGC
Hina-a    .......... .......... .......... .......... .......... ..........
Hina-b    .......... .......... .......... .......... .......... ..........
Hina-c    .......... .......... .......... .......... .......... ..........
Hina-d    .......... .......... .......... .......... .......... ..........
Hina-e    .......... .......... .......... .......... .......... ...G......
Hina-f    .......... .......... .......... .......... .......... ..........
Hina-g    .......... .......... .......... .......... .......... ..........
```


130 140 150 160 170 180
Hina ref. CCCCTAGGGA CAAAGCTAGA TTCCTGCAGG AATTACCTGC TAGATCGATG CACAACGATG
Hina-a ..........
Hina-b .......AA.
Hina-c ..........
Hina-d ..........
Hina-e .......A..
Hina-f .......A..
Hina-g .......AA.
190 200 210 220 230 240
Hina ref. AAGGATTTTC CGGTCACCTG GCGTTGGTGG ACATGGTGGA AGGGAGGTTG TGAAGAGCTC
Hina-a ..........
Hina-b ..........
Hina-c ..........
Hina-d ..........
Hina-e ..........
Hina-f ..........
Hina-g ..........
250 260 270 280 290 300
Hina ref. CTTCACGATT GTTGCAGTCA GTTGAGTCAA ATGCCACCGC AATGCCGCTG CAACATCATC
Hina-a ..........
Hina-b ..........
Hina-c ..........
Hina-d ..........
Hina-e ..........
Hina-f ..........
Hina-g ..........
310 320 330 340 350 360
Hina ref. CAGGGATCAA TCCAACGTGA TCTCGGTGGT TTCTTCGGAT TTCAGCGTGA TCGGACAGTC
Hina-a ..........
Hina-b ..........
Hina-c ..........
Hina-d ..........
Hina-e ..........
Hina-f ..........
Hina-g ..........
370 380 390 400 410 420
Hina ref. AAAGTGATAC AAGCAGCCAA GAACCTGCCC CCCAGGTGCA ACCAGGGCCC TGCCTGCAAC
Hina-a ..........
Hina-b ..........
Hina-c ..........
Hina-d ..........
Hina-e ..........
Hina-f ..........
Hina-g ..........

```
                430         440         450
           ....|....| ....|....| ....|....|
Hina ref.  ATCCCCAGCA CTACTGGCTA TTACTGGTGA
Hina-a     .......... .......... ..........
Hina-b     .......... .......... ..........
Hina-c     .......... .......... ..........
Hina-d     .......... .......... ..........
Hina-e     .......... .......... ..........
Hina-f     .......... .......... ..........
Hina-g     .......... .......... ..........
```

**Figure 6.3: DNA sequences of *Hina* haplotypes. Dots (.) indicate nucleotides identical to the top sequence in that particular group.**

```
                  10          20          30          40          50          60
           ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
HINA REF.  MKAFFLVGLL ALVASAAFAQ YGEVVGSYEG GAGGGGAQQC PLGTKLDSCR NYLLDRCTTM
HINA-A     .......... .......... .......... .......... .......... ..........
HINA-B     .......... .......... .......... .......... ..E....... ..........
HINA-C     .......... .......... .......... .......... .......... ..........
HINA-D     .......... .......... .......... .......... .......... ..........
HINA-E     .......... .......... .......... .......... ..E....... ..........

                  70          80          90         100         110         120
           ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
HINA REF.  KDFPVTWRWW TWWKGGCEEL LHDCCSQLSQ MPPQCRCNII QGSIQRDLGG FFGFQRDRTV
HINA-A     .......... .......... .......... .......... .......... ..........
HINA-B     .......... .......... .......... .......... .......... V.........
HINA-C     .......... .......... .........G .......... .......... V.........
HINA-D     .......... .......... .......... ........L. .......... ..........
HINA-E     .......... R.....L.. .........G .......... .......... V.........

                 130         140         150
           ....|....| ....|....| ....|....|
HINA REF.  KVIQAAKNLP PRCNQGPACN IPSTTGYYW*
HINA-A     .......... .......... .........*
HINA-B     .......... .......... .........*
HINA-C     .......... .......... .........*
HINA-D     .......... .......... .........*
HINA-E     .......... .......... .........*
```

**Figure 6.4: Putative protein sequences of reproducible *Hina* haplotypes. Dots (.) indicate residues identical to the top sequence in that particular group.**

**Table 6.2. SNPs found in *Hina* gene of barley cultivars and landraces**

| SNP (occuring in DNA haplotype/s) | Change in amino acid | Cultivar/ Landrace | References for other reports of the SNP |
|---|---|---|---|
| G21A (Hina-e, f) | silent | Capstan, L400097, L400211, L402433 | Turuspekov et al. (unpublished)* |
| C54T (Hina-e, f) | silent | Capstan, L400097, L400211, L402433 | Terasawa et al. (2012) |
| A114G (Hina-e) | Silent | Capstan. L400097, L400211 | Li et al. (2011) |
| G128A (Hina-b, e, f, g) | Gly43Glu#, polar-acidic | Arivat, Barque, Capstan Stirling, L400097, L400211, L402433, L411144, L411145, | Turuspekov et al. (2008), Galassi et al. (2012), Guzman and Alvarez (2014) |
| G129A (Hina-b ,g) | silent | Arivat, Barque, Stirling, L411144, L411145 | Galassi et al. (2012) |
| C212G (Hina-e, f, g) | Thr71Arg, polar-basic | Arivat, Barque, Capstan, L400097, L400211, L402433 L411144, L411145, | Beecher et al. (2001), Turuspekov et al. (2008), Li et al. (2011), Galassi et al. (2012) |
| T231C (Hina-g) | silent | Arivat, Barque, 411144, L411145 | Galassi et al. (2012) |
| G232T (Hina-e, f, g) | Glu78Leu acidic-nonpolar | Arivat, Barque, Capstan, L400097, L400211, L402433 L411144, L411145 | Beecher et al. (2001), Turuspekov et al. (2008), Li et al. (2011), Galassi et al. (2012) |
| A233T (Hina-e, f, g) | silent | Arivat, Barque, Capstan, L400097, L400211, L402433 L411144, L411145 | Galassi et al. (2012) |
| A265G (Hina-c, e ,f, g) | Ser89Gly polar-polar | Arivat, Barque, Capstan, L400097, L400211, L402433 L411144, L411145, L411859, L411890 | Turuspekov et al. (2008), Takahashi et al. (2010), Li et al. (2011), Galassi et al. (2012) |
| A295C (Hina-d) | Ile99Leu nonpolar-nonpolar | Keel | this work |
| T331G (Hina-b, c, e, f, g) | Phe111Val nonpolar-nonpolar | Arivat, Barque, Capstan, Keel L400097, L400211, L402433, L411144, L411145, L411859, L411890 | Turuspekov et al. (2008), Takahashi et al. (2010), Galassi et al. (2012) |
| A339G (Hina-e, f, g) | silent | Arivat, Capstan, Barque, L411144, L400097, L411145, L400211, L402433 | Caldwell et al. (2006), Turuspekov et al. (unpublished)**, Li et al. (2011), Galassi et al. (2012) |

*Genbank accession nos. DQ862190.1, DQ862163.1, DQ862148.1, DQ862213.1, GU591217.1, DQ862178.1

**Genbank accession nos. DQ862143.1, DQ862178.1, DQ862206.1)

# A naturally occurring hardness associated *Pinb* allele (*Pinb-D1r)* reported at same position by Ram et al. (2005)

**Table 6.3: *Hina* haplotypes and putative proteins in Australian cultivars and international landraces**

| DNA haplotype, Genbank accession, reference | Protein type | Cultivars | Landraces | SNPs occurring in the haplotype [a] | Characteristic amino acids | pI[b] | Hydrophobic residues[c] |
|---|---|---|---|---|---|---|---|
| *Hina-a* Wild type AY643843.1 (Caldwell et al. 2004) | HINA-A | Arapiles, Baudin, Buloke, Franklin, Hamelin, SloopSA Vlamingh | L400619 (N. Korea) L402448 (Iran) L403089 (Russia) L403156 (China) L411152 (Morocco) L411813 (Kyrgystan) L491320 (Japan) | NA | NA | 8.59 | 45 |
| *Hina-b* This work | HINA-B | Stirling | | **G128A**, G129A | Gly43Glu | 8.39 | 45 |
| *Hina-c* AY644142.1 (Caldwell et al. 2006) | HINA-C | | L411859 (Nepal), L411890 (Nepal) | **A265G, T331G** | Ser89Gly, Phe111Val | 8.56 | 45 |
| *Hina-d* This work | HINA-D | Keel | | **A295C** | Ile99Leu | 8.59 | 45 |
| *Hina-e* This work | HINA-E | Capstan | L400097 (Ethiopia), L400211 (Ethiopia) | G21A, C54T, A114G, **G128A**, **C212G**, **G232T**, A233T, **A265G**, **T331G**, A339G | Gly43Glu, Thr71Arg, Glu78Leu, Ser89Gly, Phe111Val | 8.72 | 46 |
| *Hina-f* This work | HINA-E | | L402433 (Afghanistan) | G21A, C54T, **G128A**, **C212G**, **G232T**, A233T**,** **A265G**, **T331G**, A339G | Gly43Glu, Thr71Arg, Glu78Leu, Ser89Gly, Phe111Val | 8.72 | 46 |
| *Hina-g* JN636836.1 (Galassi et al. 2012) | HINA-E | Arivat, Barque | L411144 (Morocco), L411145 (Morocco) | **G128A**, G129A, **C212G**, T231C**, G232T**, A233T, **A265G**, **T331G**, A339G | Gly43Glu, Thr71Arg, Glu78Leu, Ser89Gly, Phe111Val | 8.72 | 46 |

[a]SNPs shown in bold: those that result in amino acid substitutions; the rest are silent mutations. DNA haplotypes e, f and g lead to the same putative protein.
[b]Isoelectric point of mature protein, calculated starting at E29.
[c] Hydrophobic residues of mature protein were calculated and include F, I, W, L, V, M, Y, C, A (based on hydrophobicity index at pH 2 and 7; Nelson and Cox, 2005; http://www.sigmaaldrich.com/life-science/metabolomics/learning-center/amino-acid-reference-chart.html)

### 6.3.4 Nucleotide variations in *Hinb-1* gene and amino acid substitutions in the putative protein

Variations in *Hinb-1* were observed in three cultivars (Barque, Capstan and Stirling) and six landraces (L400097, L400211, L402433, L411144, L411145, L411859). The *Hinb-1* gene haplotypes and putative protein types are shown in Fig.6.5 and Fig. 6.6, respectively. A total of twelve SNPs was found in the cultivars (Table 6.4), seven of these (C72T, A138G, A231G, G264T, T291C, C408G, C429T) being silent. Of the three cultivars that showed *Hinb-1* variation, one contained a single SNP (Stirling: C72T) and second cultivar (Barque: A25C, C414G) depicted two SNPs while the third cultivar Capstan showed 10 SNPs (A25C, C47T, A138G, C176G, A231G, G264T, T291C, A314G, C408G, C429T). This is an interesting observation compared to *Hina* where four cultivars showed multiple SNPS, indicating greater diversity in *Hina* gene. Only one SNP (A25C) was shared between Barque and Capstan, unlike their *Hina* genes where three cultivars (Arivat, Barque and Capstan) shared seven SNPs (G128A, C212G, G232T, A233T, A265G, T331G, A339G). Twelve SNPs were noted altogether in the landraces (Table 6.4). All SNPs present in the cultivars, except for C72T (Stirling) and G180A (VB0104), also occurred in landraces, while one SNP G329C (L402433) was unique to landraces. Interestingly, an insertion (INS210A) in *Hinb-1* resulting in a frame-shift mutation (stop codon at position 77 of full length protein was detected in L400211 from Ethiopia.

A total of six *Hinb-1* haplotypes (*Hinb-1a* to *Hinb-1f*, Fig.6.5) were observed. The wild-type haplotype *Hinb-1a* (Caldwell et al. 2004) was shared by nine cultivars (Arapiles, Arivat, Baudin, Buloke, Franklin, Hamelin, Keel, SloopSA, Vlamingh) and eight landraces (L400619, L402448, L403089, L403156, L411152, L411813 L411890, and L491320). The haplotype *Hinb-1c* observed in landrace L402433 has been reported earlier in in domesticated barley *Hordeum vulgare subsp. vulgare* (Caldwell et al. 2006). Similarly, another haplotype *Hinb-1d* present in one cultivar (Barque) and three landraces (L411859, L411144, L411145) has been reported in *Hordeum vulgare subsp. vulgare* cultivar Naturel (57.2 SKCS units at one location and 51.2 SKCS units at other location) (Galassi et al. 2012). Three haplotypes, *Hinb-1b*, *Hinb-1e* and *Hinb-1f*, are novel. Identical haplotypes (*Hinb-1d*) were observed in the landraces L411859 (Nepal), L411144 and L411145 (Morocco). The haplotype *Hinb-1e* found in L400211 (Ethiopia) showed an 'A' insertion at nucleotide position 210. These six DNA haplotypes led to five

putative protein types, designated as HINB-1A to HINB-1E (Fig. 6.6, Table 6.5). Ten cultivars and eight landraces exhibited HINB-1A (identical to the reference sequence). All ten characteristic Cysteine residues were completely conserved, and the putative TRD (FPLTWPTKWWKG) was also conserved. Two substitutions Gln105Arg and Gly110Ala were noted in the putative HD (IQGKLGGIFGI). HINB-1B contained a single substitution i.e. M60I. Two substitutions, i.e., Ile9Leu and Gly110Ala, were characteristics of HINB-1B, while two others, Ile9Leu and Asp138Glu, were characteristic of HINB-1C (the most prevalent type shared by one cultivar and three landraces). Four substitutions (Ile9Leu, Thr16Ile, Thr59Arg, Gln105Arg,) were present in HINB-1D, shared by one cultivar and two landraces. A frame-shift mutation with premature stop codon at position 77 of full length protein was noted in HINB-1E. The predicted pI of the putative mature proteins ranged from 8.72 to 8.99 (Table 6.5) HINB-1Dhad highest pI, probably due to its unique Arg59 and Arg105. All other putative protein types (HINB-1A to HINB-1C) depicted similar pI values (8.72).

```
                    10         20         30         40         50         60
               ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
Hinb-1 ref.    ATGAAGACCT TATTCCTCCT AGCTATCCTT GCTCTTGTAG CAAGCACAAC CTTCGCGCAA
Hinb-1a        .......... .......... .......... .......... .......... ..........
Hinb-1b        .......... .......... .......... .......... .......... ..........
Hinb-1c        .......... .......... ....C..... .......... .......... ..........
Hinb-1d        .......... .......... ....C..... .......... .......... ..........
Hinb-1e        .......... .......... ....C..... .......... ......T... ..........
Hinb-1f        .......... .......... ....C..... .......... ......T... ..........

                    70         80         90        100        110        120
               ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
Hinb-1 ref.    TACTCAGTTG GCGGTGGTTA CAATGACGTT GGCGGAGGAG GCGGTTCTCA ACAATGCCCA
Hinb-1a        .......... .......... .......... .......... .......... ..........
Hinb-1b        .......... .T........ .......... .......... .......... ..........
Hinb-1c        .......... .......... .......... .......... .......... ..........
Hinb-1d        .......... .......... .......... .......... .......... ..........
Hinb-1e        .......... .......... .......... .......... .......... ..........
Hinb-1f        .......... .......... .......... .......... .......... ..........

                   130        140        150        160        170        180
               ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
Hinb-1 ref.    CAGGAGCGGC CGAACCTAGG CTCTTGCAAG GATTACGTGA TGGAGCGGTG TTTCACGATG
Hinb-1a        .......... .......... .......... .......... .......... ..........
Hinb-1b        .......... .......... .......... .......... .......... ..........
Hinb-1c        .......... .......... .......... .......... .......... ..........
Hinb-1d        .......... .......... .......... .......... .......... ..........
Hinb-1e        .......... .......G.. .......... .......... .......... .....G....
Hinb-1f        .......... .......G.. .......... .......... .......... .....G....
```

```
                  190         200         210         220         230         240
           ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
Hinb-1 ref.  AAGGATTTTC CACTTACCTG GCCCACAAAA TGGTGGAAGG GAGGCTGTGA ACAAGAGGTT
Hinb-1a      .......... .......... .......... .......... .......... ..........
Hinb-1b      .......... .......... .......... .......... .......... ..........
Hinb-1c      .......... .......... .......... .......... .......... ..........
Hinb-1d      .......... .......... .......... .......... .......... ..........
Hinb-1e      .......... .......... .......... .......... .......... G.........
Hinb-1f      .......... .......... .......... AT.GT.G.A. .GA.GCTGTG .GC.AGA.G.

                  250         260         270         280         290         300
           ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
Hinb-1 ref.  CGGGAGAAGT GTTGCCAGCA ACTGAGCCAG ATAGCACCAC AATGTCGCTG TGATGCTATC
Hinb-1a      .......... .......... .......... .......... .......... ..........
Hinb-1b      .......... .......... .......... .......... .......... ..........
Hinb-1c      .......... .......... .......... .......... .......... ..........
Hinb-1d      .......... .......... .......... .......... .......... ..........
Hinb-1e      .......... .......... ...T...... .......... .......... C.........
Hinb-1f      TC..GAG.AG TG.TG.CAGC .ACTTAG.CA GATAGCA.CA C.ATGTCGCT GCGATGCTAT

                  310         320         330         340         350         360
           ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
Hinb-1 ref.  CGGGGAGTGA TCCAAGGCAA GCTCGGTGGT ATCTTTGGCA TTGGGGGAGG TGATGTATTC
Hinb-1a      .......... .......... .......... .......... .......... ..........
Hinb-1b      .......... .......... .......... .......... .......... ..........
Hinb-1c      .......... .......... ........C. .......... .......... ..........
Hinb-1d      .......... .......... .......... .......... .......... ..........
Hinb-1e      .......... ...G...... .......... .......... .......... ..........
Hinb-1f      .C...GAGTG AT.CGA.GC. AGCTC.GT.G TATC..T.GC A.T....GA. GTGATGTA.T

                  370         380         390         400         410         420
           ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
Hinb-1 ref.  AAACAAATTC AGAGGGCCCA AATCCTCCCC TCAAAGTGCA ACATGGGCGC CGACTGTAAG
Hinb-1a      .......... .......... .......... .......... .......... ..........
Hinb-1b      .......... .......... .......... .......... .......... ..........
Hinb-1c      .......... .......... .......... .......... .......... ..........
Hinb-1d      .......... .......... .......... .......... .......... ...G......
Hinb-1e      .......... .......... .......... .......... .......G.. ..........
Hinb-1f      C..AC..A.T CAGA..G..C ..AT.CT... CTC..AGTGC .ACAT..G.G .CGACTGT.A

                   430         440
           ....|....| ....|....| ....|
Hinb-1 ref.  TTCCCTAGCG GCTATTACTG GTGA
Hinb-1a      .......... .......... ....
Hinb-1b      .......... .......... ....
Hinb-1c      .......... .......... ....
Hinb-1d      .......... .......... ....
Hinb-1e      .........T. .......... ....
Hinb-1f      G.T..CTAGT .GCTA.TACT .GTGA
```

**Figure 6.5: DNA sequences of *Hinb-1* haplotypes. Dots (.) indicate nucleotides identical to the top sequence in that particular group.**

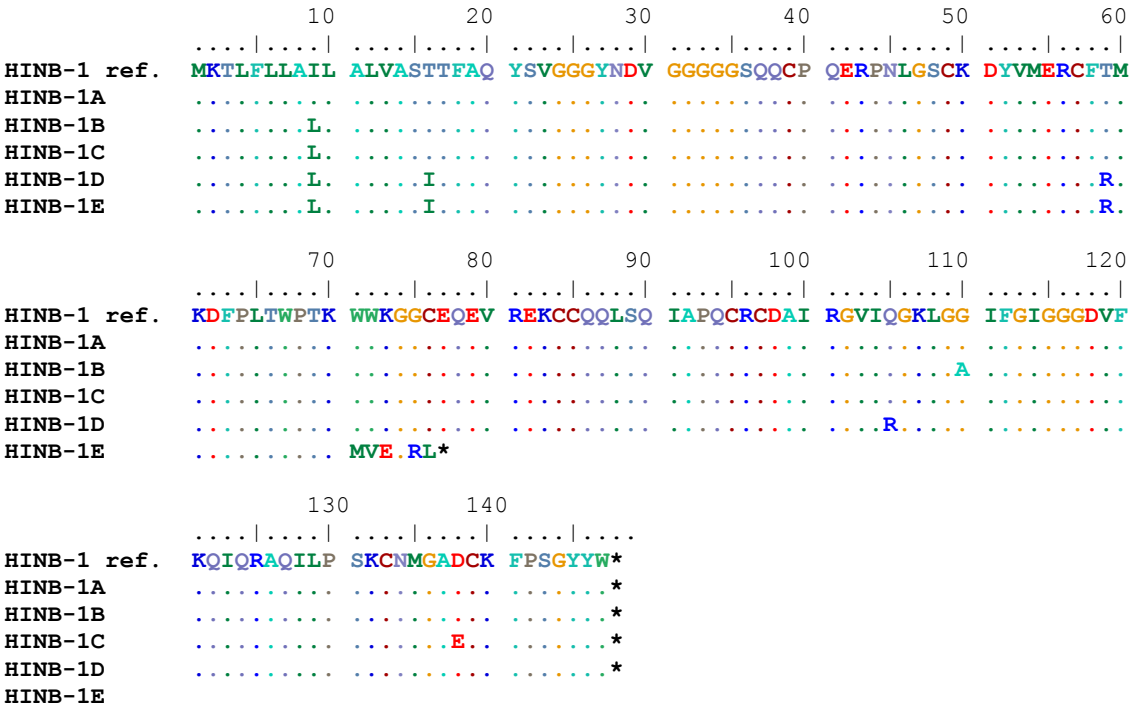



**Figure 6.6: Putative protein sequences of reproducible *Hinb-1* haplotypes. Dots (.) indicate residues identical to the top sequence in that particular group.**

**Table 6.4. SNPs found in *Hinb-1* gene of barley cultivars and landraces**

| SNP (occuring in DNA haplotype/s) | Change in amino acid | Cultivar/ Landrace | References for other reports of the SNP |
|---|---|---|---|
| A25C (Hinb-1 c, d, e, f) | Ile9Leu nonpolar-nonpolar | Barque, Capstan, L400097, L402433, L400211, L411144, L411145, L411859 | Turuspekov et al. (2008), Galassi et al. (2012) |
| C47T (Hinb-1 e, f) | Thr16Ile polar-nonpolar | Capstan, L400097, L400211 | Turuspekov et al. (2008), Galassi et al. (2012) |
| C72T (Hinb-1b) | silent | Stirling | this work |
| A138G (Hinb-1e, f) | silent | Capstan, L400097, L400211 | Galassi et al.(2012) |
| C176G (Hinb-1e, f) | Thr59Arg polar-basic | Capstan, L400097, L400211 | Turuspekov et al. (2008), Galassi et al. (2012) |
| INS*210A (Hinb-1f) | | L400211 | this work |
| A231G (Hinb-1e) | silent | Capstan, L400097, L400211 | Galassi et al. (2012) |
| G264T (Hinb-1e) | silent | Capstan, L400097, L400211 | Galassi et al. (2012) |
| T291C (Hinb-1e) | silent | Capstan, L400097, L400211 | Galassi et al. (2012) |
| A314G (Hinb-1e) | Gln105Arg Polar-basic | Capstan, L400097, L400211 | Galassi et al. (2012) |
| G329C (Hinb-1c) | Gly110Ala polar-nonpolar | L402433 | Caldwell et al. (2006) |
| C408G (Hinb-1e) | silent | Capstan, L400097, L400211 | Galassi et al. (2012) |
| C414G (Hinb-1d) | Asp138Glu acidic-acidic | Barque, L411859, L411144, L411145 | Takahashi et al. (2010), Galassi et al. (2012) |
| C429T (Hinb-1e) | silent | Capstan, L400097, L400211 | Turuspekov et al. (2008), Galassi et al. (2012) |

*INS-insertion

**Table 6.5: *Hinb-2* haplotypes found in Australian cultivars and international landraces**

| DNA haplotype, Genbank accession, reference | Protein type | Cultivars | Landraces | SNPs occurring in the haplotype [a] | Characteristic amino acids | pI[b] | Hydrophobic residues[c] |
|---|---|---|---|---|---|---|---|
| *Hinb-1a* Wild type AY643843.1 (Cadlwell et al.2004) | HINB-1A | Arapiles, Arivat, Baudin, Buloke, Franklin, Hamelin, Keel, SloopSA, Vlamingh | L400619 (N. Korea) L402448 (Iran) L403089 (Russia) L403156 (China) L411152 (Morocco) L411813 (Kyrgystan) L411890 (Nepal) L491320 (Japan) | NA | NA | 8.72 | 43 |
| *Hinb-1b* (this work) | HINB-1A | Stirling | - | C72T | | 8.72 | 43 |
| *Hinb-1c* AY643991.1 (Caldwell et al. 2006) | HINB-1B | - | L402433 (Afghanistan) | **A25C**, **G329C** | Ile9Leu, Gly110Ala | 8.72 | 44 |
| *Hinb-1d* JN636843.1 (Galassi et al. 2012) | HINB-1C | Barque | L411859 (Nepal) L411144(Morocco) L411145 (Morocco) | **A25C**, **C414G** | Ile9Leu, Asp138Glu | 8.72 | 43 |
| *Hinb-1e* (this work) | HINB-1D | Capstan | L400097 (Ethiopia) | **A25C, C47T**, A138G, **C176G**, A231G, G264T, T291C, **A314G**, C408G, C429T | Ile9Leu, Thr16Ile, Thr59Arg, Gln105Arg, | 8.99 | 43 |
| *Hinb-1f* (this work) | HINB-1E | | L400211 (Ethiopia) | **A25C, C47T**, A138G, **C176G, I**NS210A | Ile9Leu, Thr16Ile, Thr59Arg, stop77 | | |

[a]SNPs shown in bold: those that result in amino acid substitutions; the rest are silent mutations. Haplotypes a and b lead to the same putative protein.
[b]Isoelectric point of mature protein, calculated starting at D29. Haplotypes a and b lead to the same putative protein
[c] Hydrophobic residues of mature protein were calculated and include F, I, W, L, V, M, Y, C, A (based on hydrophobicity index at pH 2 and 7; Nelson and Cox, 2005; http://www.sigmaaldrich.com/life-science/metabolomics/learning-center/amino-acid-reference-chart.html)

### 6.3.5 Nucleotide variations in *Hinb-2* gene and amino acid substitutions in the putative protein

Variations in *Hinb-2* gene were observed in five cultivars (Arivat, Barque, Hamelin, SloopSA, Stirling), and five landraces (L402433, L411144, L411145, L411859, L411890). A total of eight *Hinb-2* haplotypes leading to five putative protein types were observed (Fig. 6.7 and Fig. 6.8). Seven SNPs were found in cultivars (Table 6.6), single SNPs were present in three out of five cultivars and only one SNP i.e. G419A was a shared SNP between two cultivars (Stirling and Barque). Five SNPs were found in landraces (Table 6.6) and all except one (C262T in L411859) were shared. Three SNPs in cultivars (A123G, G126A and G264A) and three SNPs in landraces (A123G, C225T, and C262T) were silent mutations. A total of eight haplotypes were observed and were designated as *Hinb-2a* to *Hinb-2h*. The haplotype *Hinb-2c* observed in cultivar Stirling has been reported earlier in a spring cultivar of barley *Hordeum vulgare subsp. Vulgare* (Genbank: DQ862366.1, Turuspekov et al. unpublished). Another haplotype, *Hinb-2f* present in one cultivar (Barque) and two landraces (L402433 and L411890) has been reported in in *Hordeum vulgare subsp. vulgare* cultivar Naturel (57.2 SKCS units at one location and 51.2 SKCS units at other location) (Galassi et al. 2012). Five haplotypes, *Hinb-2b*, *Hinb-2d*, *Hinb-2e*, *Hinb-2g* and *Hinb-2h* were identified in this study have not been reported earlier. Identical haplotypes were observed in the landraces L402433 (Afghanistan), L411890 (Nepal) and two landraces from Morocco (L411144, L411145) also depicted identical haplotypes.

These eight DNA haplotypes led to five putative protein types, designated as HINB-2A to HINB-2E (Fig. 6.8, Table 6.7). Ten cultivars and eight landraces exhibited HINB-2A (identical to the reference sequence). The ten signature Cysteine residues were conserved. No substitutions were noted in the TRD and HD. Single substitution was observed in three putative protein types i.e. Arg140Lys in HINB-2B, Ser143Thr in HINB-2C and Ser143Gly in HINB-2D whereas HINB-2E contained two substitutions i.e. Val137Ala and Arg140Lys. The predicted pI of the putative mature proteins ranged from 8.86 to 8.87 (Table 6.7). HINB-2A, HINB-2C and HINB-2D depicted similar pI value i.e. 8.87. HINB-2B and HINB-2E depicted a pI value of 8.86 probably due to R140K substitution. A frame-shift mutation resulting in an in-frame stop in *Hinb-2* and increased grain hardness (Takahashi et al. 2010) was not found in any cultivar or landrace.

```
                  10         20         30         40         50         60
                  ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
Hinb-2 ref.       ATGAAGACCT TATTCCTCCT AGCTCTCCTT GCTCTTGTAG CAAGCACAAC CTCCGCGCAA
Hinb-2a           .......... .......... .......... .......... .......... ..........
Hinb-2b           .......... .......... .......... .......... .......... ..........
Hinb-2c           .......... .......... .......... .......... .......... ..........
Hinb-2d           .......... .......... .......... .......... .......... ..........
Hinb-2e           .......... .......... .......... .......... .......... ..........
Hinb-2f           .......... .......... .......... .......... .......... ..........
Hinb-2g           .......... .......... .......... .......... .......... ..........
Hinb-2h           .......... .......... .......... .......... .......... ..........

                  70         80         90         100        110        120
                  ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
Hinb-2 ref.       TACTCAGTTG GTGGTGGTTA CAATGACGTT GGTGGAGGAG GTGGTTCTCA ACAATGCCCA
Hinb-2a           .......... .......... .......... .......... .......... ..........
Hinb-2b           .......... .......... .......... .......... .......... ..........
Hinb-2c           .......... .......... .......... .......... .......... ..........
Hinb-2d           .......... .......... .......... .......... .......... ..........
Hinb-2e           .......... .......... .......... .......... .......... ..........
Hinb-2f           .......... .......... .......... .......... .......... ..........
Hinb-2g           .......... .......... .......... .......... .......... ..........
Hinb-2h           .......... .......... .......... .......... .......... ..........

                  130        140        150        160        170        180
                  ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
Hinb-2 ref.       CAAGAGCGGC CAAACCTAGG TTCTTGCAAG GATTACGTGA TGGAGCGGTG TTTCACGATG
Hinb-2a           .......... .......... .......... .......... .......... ..........
Hinb-2b           .....A.... .......... .......... .......... .......... ..........
Hinb-2c           .......... .......... .......... .......... .......... ..........
Hinb-2d           .......... .......... .......... .......... .......... ..........
Hinb-2e           .......... .......... .......... .......... ..........
Hinb-2f           ..G....... .......... .......... .......... .......... ..........
Hinb-2g           ..G....... .......... .......... .......... .......... ..........
Hinb-2h           ..G....... .......... .......... .......... .......... ..........

                  190        200        210        220        230        240
                  ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
Hinb-2 ref.       AAGGATTTTC CAGTTACCTG GCCCACGAAG TGGTGGAAGG GAGGCTGTGA GCATGAGGTT
Hinb-2a           .......... .......... .......... .......... .......... ..........
Hinb-2b           .......... .......... .......... .......... .......... ..........
Hinb-2c           .......... .......... .......... .......... .......... ..........
Hinb-2d           .......... .......... .......... .......... .......... ..........
Hinb-2e           .......... .......... .......... .......... .......... ..........
Hinb-2f           .......... .......... .......... .......... .......... ..........
Hinb-2g           .......... .......... .......... .......... ....T..... ..........
Hinb-2h           .......... .......... .......... .......... .......... ..........

                  250        260        270        280        290        300
                  ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
Hinb-2 ref.       CGGGAGAAGT GTTGCCAGCA GCTGAGCCAG ATAGCACCAC ATTGTCGCTG CGATGCTATC
Hinb-2a           .......... .......... .......... .......... .......... ..........
Hinb-2b           .......... .......... .......... .......... .......... ..........
Hinb-2c           .......... .......... .......... .......... .......... ..........
Hinb-2d           .......... .......... .......... .......... .......... ..........
Hinb-2e           .......... .......... ...A...... .......... .......... ..........
Hinb-2f           .......... .......... .......... .......... .......... ..........
Hinb-2g           .......... .......... .......... .......... .......... ..........
Hinb-2h           .......... .......... .T........ .......... .......... ..........
```

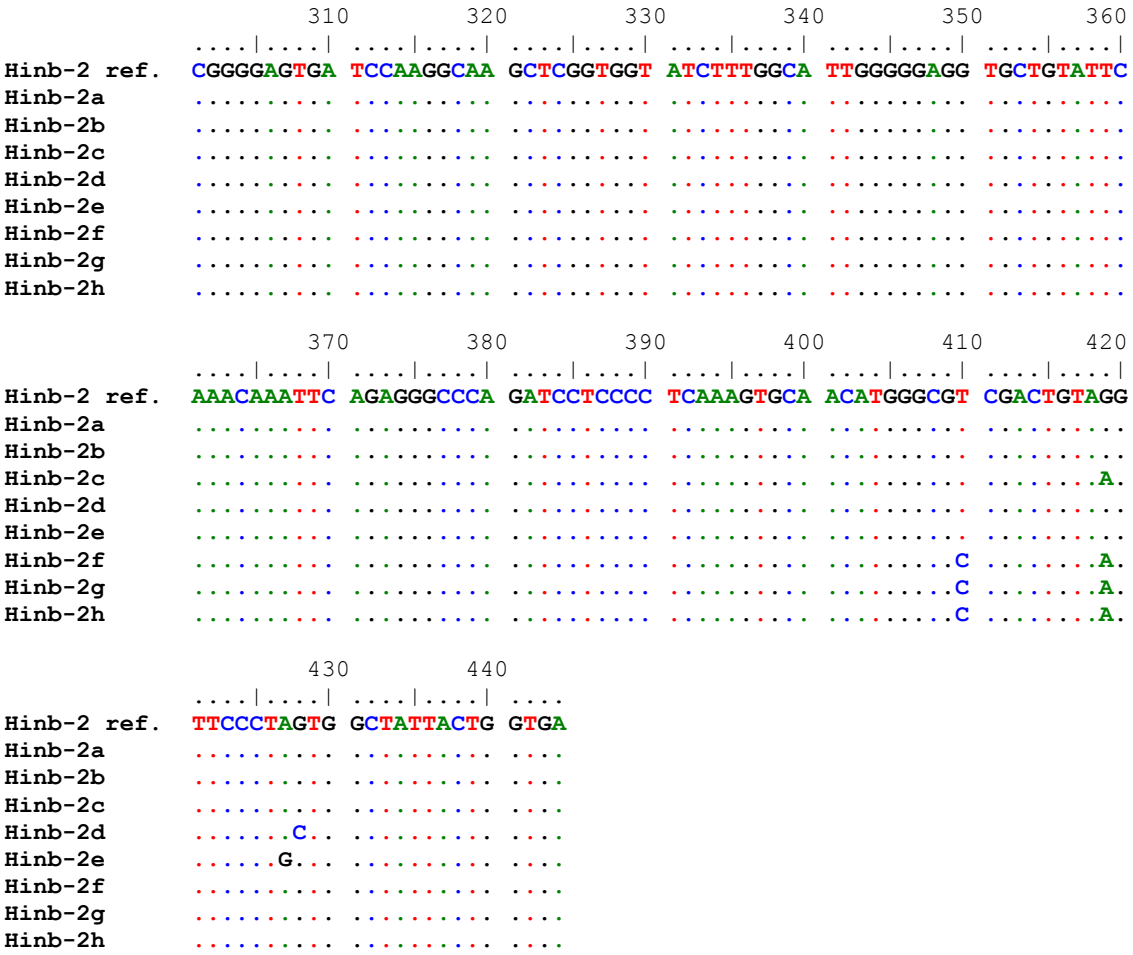



**Figure 6.7: DNA sequences of *Hinb-2* haplotypes. Dots (.) indicate nucleotides identical to the top sequence in that particular group.**

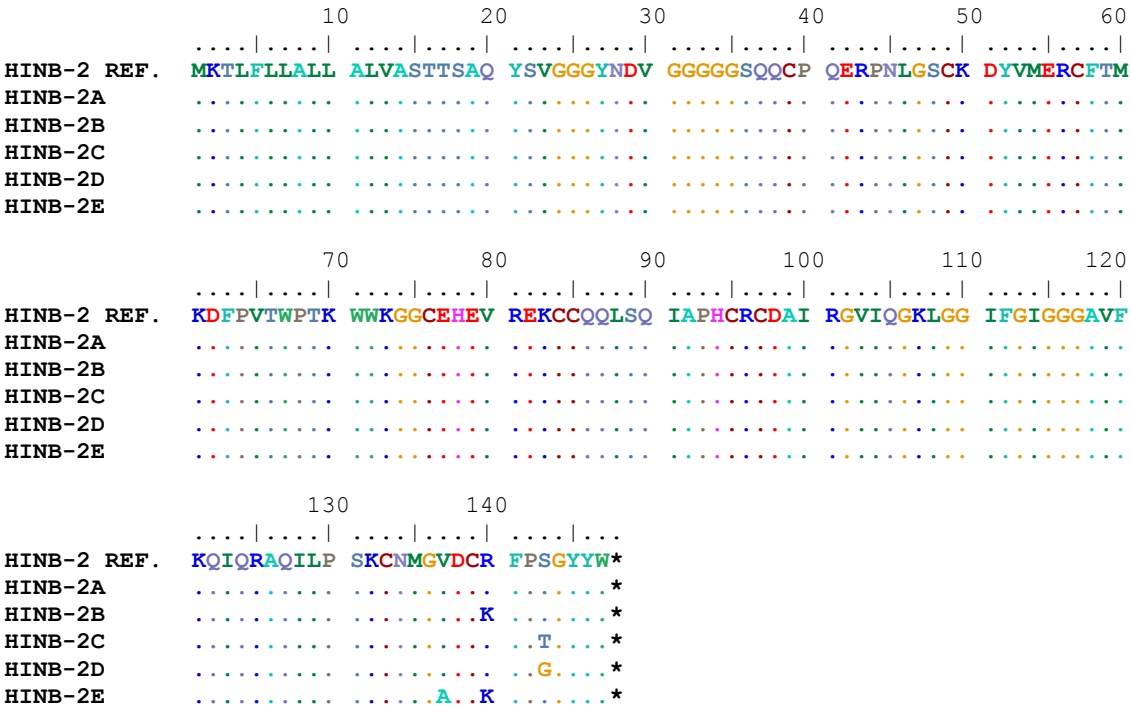



**Figure 6.8: Putative protein sequences of reproducible *Hinb-2* haplotypes. Dots (.) indicate residues identical to the top sequence in that particular group**

**Table 6.6. SNPs found in *Hinb-2* gene of barley cultivars and landraces**

| SNP (occuring in DNA haplotype/s) | Change in amino acid | Cultivar/ Landrace | References for other reports of the SNP |
|---|---|---|---|
| A123G (Hinb-2f, g, h) | silent | Barque, L402433, L411144, L411145, L411859, L411890, | Turuspekov et al. (2008), |
| G126A (Hinb-2b) | silent | Hamelin | this work |
| C225T (Hinb-2g) | silent | L411144, L411145 | Turuspekov et al. unpublished* Caldwell et al. (2006) |
| C262T (Hinb-2h) | silent | L411859 | Turuspekov et al. (2008), Galassi et al. (2012), Takahashi et al. (2010) |
| G264A (Hinb-2e) | silent | Arivat | Guzman and Alvarez (2014) |
| T410C (Hinb-2 f, g, h) | Val137Ala nonpolar-nonpolar | Barque, L402433, L411144, L411145, L411859, L411890, | Turuspekov et al. (2008), Galassi et al. (2012) |
| G419A (Hinb-2 c, f, g, h) | Arg140Lys basic-basic | Barque, Stirling L402433, L411144, L411145, L411859, L411890 | Turuspekov et al. (2008), Galassi et al. (2012) |
| A427G (Hinb-2e) | Ser143Gly polar-polar | Arivat | this work |
| G428C (Hinb-2d) | Ser143Thr# | Sloop SA | this work |

*Genbank accession no. DQ862360.1 # A naturally occurring hardness associated *Pinb* allele (*Pinb-D1w*) reported at same position by Chang et al. (2006)

**Table 6.7: *Hinb-2* haplotypes found in Australian cultivars and international landraces**

| DNA haplotype, Genbank accession, reference | Protein type | Cultivars | Landraces | SNPs occurring in the haplotype [a] | Characteristic amino acids | pI[b] | Hydrophobic residues[c] |
|---|---|---|---|---|---|---|---|
| *Hinb-2a* Wild type AY643843.1 (Caldwell et al.2004) | HINB-2A | Arapiles, Baudin, Buloke, Capstan Franklin, Keel, TLN/YADA, VB0104, Vlamingh | L400097 (Ethiopia) L400619 (N. Korea) L402448 (Iran) L403089 (Russia) L403156 (China) L411152 (Morocco) L411813 (Kyrgystan) L491320 (Japan) | | | 8.87 | 44 |
| *Hinb-2b* (this work) | HINB-2A | Hamelin | | G126A | | 8.87 | 44 |
| *Hinb-2c* DQ862366.1 (Turuspekov et al. unpublished) | HINB-2B | Stirling | | **G419A** | Arg140Lys | 8.86 | 44 |
| *Hinb-2d* (this work) | HINB-2C | Sloop SA | | **G428C** | Ser143Thr | 8.87 | |
| *Hinb-2e* (this work) | HINB-2D | Arivat | | G264A, **A427G** | Ser143Gly | 8.87 | 44 |
| *Hinb-2f* JN636845.1 (Galassi et al. 2012) | HINB-2E | Barque | L402433 (Afghanistan) L411890 (Nepal), | A123G, **T410C**, **G419A** | Val137Ala, Arg140Lys | 8.86 | 44 |
| *Hinb-2g* (this work) | HINB-2E | | L411144 (Morocco), L411145 (Morocco) | A123G, C225T, **T410C**, **G419A** | Val137Ala, Arg140Lys | 8.86 | 44 |
| *Hinb-2h* (this work) | HINB-2E | | L411859 (Nepal) | A123G, C262T, **T410C**, **G419A** | Val137Ala, Arg140Lys | 8.86 | 44 |

[a]SNPs shown in bold: those that result in amino acid substitutions; the rest are silent mutations. Haplotypes a, b lead to HINB-2A putative protein type. Haplotypes f, g, h lead to the HINB-2E putative protein

[b]Isoelectric point of mature protein, calculated starting at D29.

[c] Hydrophobic residues of mature protein were calculated and include F, I, W, L, V, M, Y, C, A (based on hydrophobicity index at pH 2 and 7; Nelson and Cox, 2005; http://www.sigmaaldrich.com/life-science/metabolomics/learning-center/amino-acid-reference-chart.html)

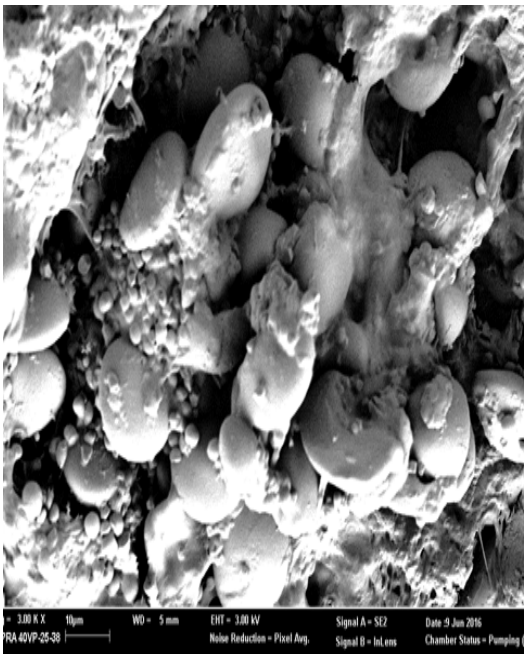
### 6.3.6 Investigation of barley grain hardness using scanning electron microscopy

Grain hardness was measured using SKCS, and SEM was used as an additional tool for grain texture studies. The structural differences in endosperms of selected barley cultivars and landraces were analysed by SEM by determining the amount of starch-protein association. The sample preparation for SEM (Section 2.11) was performed with consistency but the variations in seed size, type of seed fracture cannot be ignored. The cryo-fractured seed was dissected longitudinally and fixed onto glass slides using double-sided adhesive tape, exposing the starchy endosperm while mounting the sample.

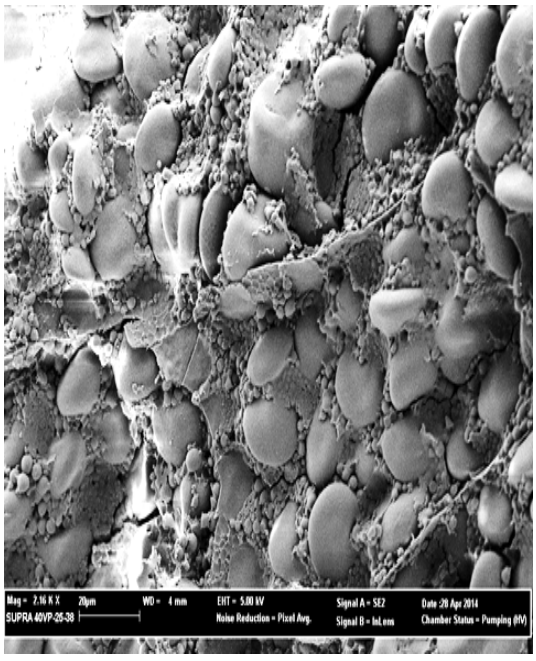
Both malt (Baudin, SloopSA, Vlamingh) and feed cultivars (Arivat, Barque) were chosen for analysis. The images are presented in Fig. 6.9. Baudin (49.2 SKCS units) was found to be wild type for all three *Hin* genes but is softer/has much lower SKCS value than the landrace L400619 (94.4 SKCS units) also with all three WT genes, while cv. Barque encoded the HINA-E protein with six substitutions (Gly43Glu, Thr71Arg, Glu78Leu, Ser89Gly, Ile99Leu, Phe111Val), as presented above. Another malt cultivar Vlamingh was also the wild type for all three *Hin* genes. The cultivar SloopSA had lowest SKCS value (24.2 SKCS units) among cultivars.

Of the landraces, two landraces L400619 (highest SKCS value: 94.4 SKCS units, wild type for all three *Hin* genes) and L411145 (54.7 SKCS units, containing substitutions: Gly43Glu, Thr71Arg, Glu78Leu, Ser89Gly, Ile99Leu, Phe111Val in HINA, Asp138Glu in HINB-1and Val137Ala, Arg140Lys in HINB-2). The greatest amount of starch-protein association was found in L400619 showing highest SKCS value of 94.4 SKCS units (Fig. 6.9 F). The large A-type starch granules were found to be embedded in dense layer of protein matrix such that smaller B-type starch granules were masked and barely visible. The starch- protein association was found to be relatively low in L411145 (54.7 SKCS units, Fig. 6.9 G) making the B-type starch granules visible indicating that the protein substitutions present in putative protein type might have some softening effect on grain texture. Comparatively weak starch associated protein was observed in malt-type cultivars Sloop SA (Fig. 6.9 A), Baudin (Fig. 6.9 B), Vlamingh (Fig. 6.9 C) strongly suggesting they may be softer than the feed cultivar Barque (Fig. 6.9 D). The malt-type cultivars appeared softer as compared to feed type cultivars Barque (Fig. 6.9 D) and Arivat (Fig. 6.9 E). The malt cultivar Sloop SA (least SKCS value: 24.2 SKCS units) showed least amount of starch-protein association. Scanning electron microscopy has been used earlier to examine good and poor malting quality barley cultivars showing that

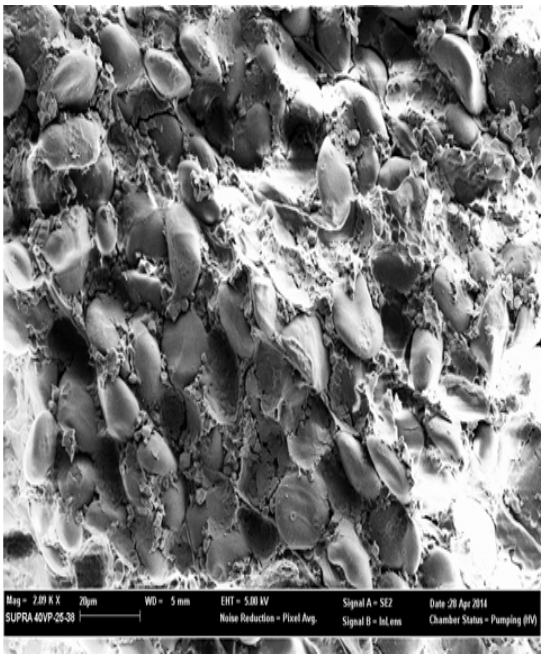
poor malting cultivars had harder grain (Brennan et al. 1996). Thus, SEM was found to be a useful tool for differentiating hard and soft barley types in addition to SKCS. However, as mentioned earlier in Chapter 3, Section 3.4.4 SEM cannot assess the environmental variation in grain texture.



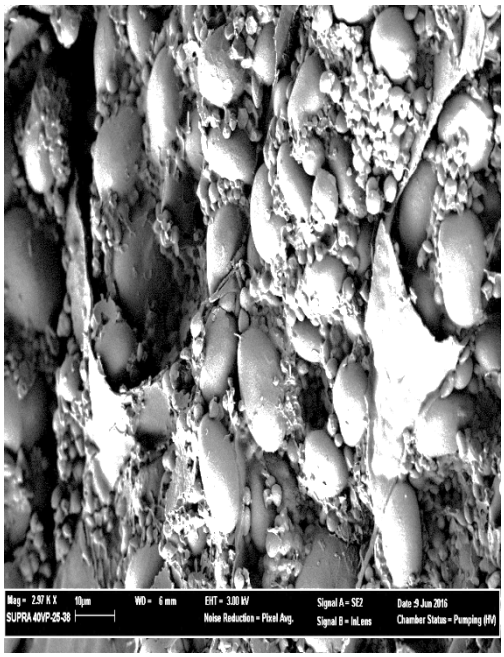

A Sloop SA (Malt, 24.2 SKCS units)



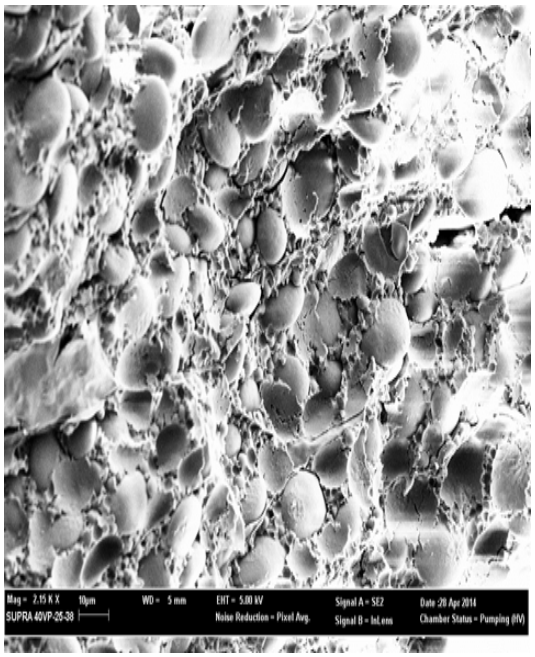

B Baudin (Malt, 49.2 SKCS units)



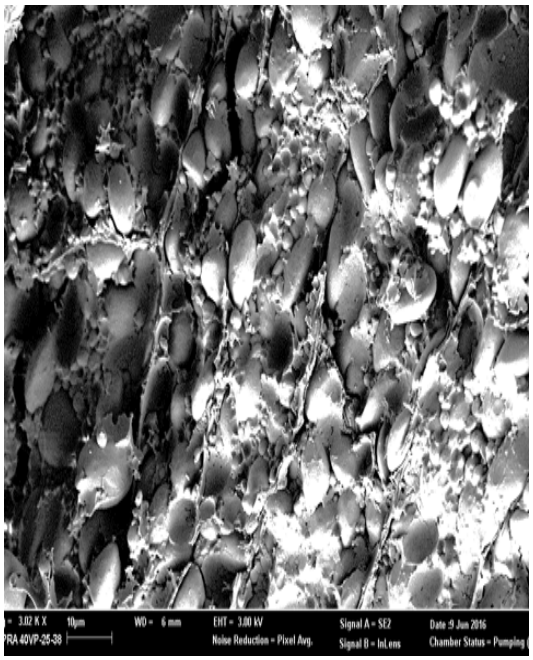

C Vlamingh (Malt, SKCS value not available)



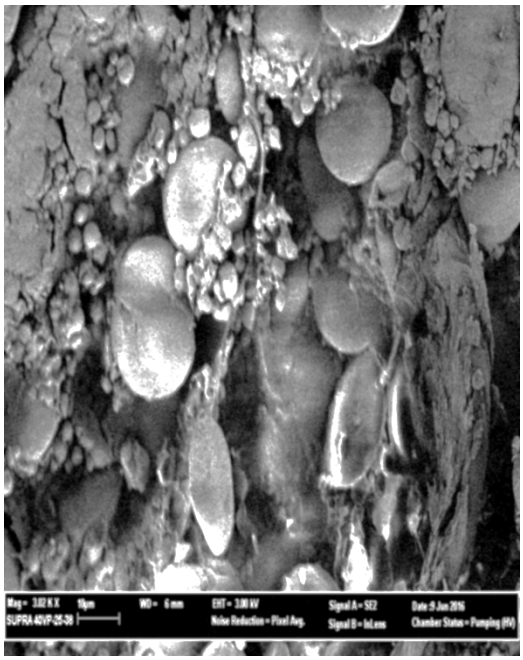
D Barque (Feed, 36.7 SKCS units)



E Arivat (Feed)



F L400619 (94.4 SKCS units)



G L411145 (54.7 SKCS units)

**Figure 6.9 SEM images of selected barley cultivars/ landraces.** Magnification: X 2000-3000, A: Sloop SA, B: Baudin, C: Vlamingh, D: Barque, E: Arivat, F: L400619, G: L411145

### 6.3.7 Barley grain protein content and its possible relationship with grain hardness

The grain protein percentage of barley cultivars and landraces was determined with the objective to determine a possible relationship with grain hardness. It was observed that increasing protein typically resulted in harder grain in cultivars and landraces (Table 6.8, Table 6.9). The cultivars with high SKCS value (Buloke, Stirling, Franklin) depicted a high protein content > 10 %. A range of 24.2-58.1 was observed for SKCS grain hardness in cultivars. Among cultivars majority of grains depicted a hardness range of 54.2-60.2 (Fig. 6.10). The grain protein depicted a range of 8.5-11.8 and mean value of 9.91 ± 1.14 respectively. For these samples correlation coefficient between grain hardness and protein was 0.63 ($p < 0.01$). However, the correlation was not statistically significant.

**Table 6.8: Grain protein content and hardness in barley cultivars**

| No. | Cultivar (Malt/Feed) | Origin | SKCS index | GPC (%) |
|---|---|---|---|---|
| 1 | Arapiles (Malt) | Australia | 46.9 | 9.7 |
| 2 | Arivat (Feed) | Australia | - | 10.7 |
| 3 | Barque (Feed) | Australia | 36.7 | 8.9 |
| 4 | Baudin (Malt) | Australia | 49.2 | 9.1 |
| 5 | Buloke (Malt) | Australia | 58.1 | 11.3 |
| 6 | Capstan (Feed) | Australia | 33.1 | 9.2 |
| 7 | Franklin (Malt) | Australia | 54.6 | 10.0 |
| 8 | Hamelin (Malt) | Australia | 55.4 | 9.1 |
| 9 | Keel (Feed) | Australia | 40.2 | 8.5 |
| 10 | SloopSA (Malt) | Australia | 24.2 | 9.1 |
| 11 | Stirling (Malt) | Australia | 56.4 | 11.6 |
| 12 | Vlamingh (Malt) | Australia | - | 11.8 |
| Range | | | 24.2-58.1 | 8.5-11.8 |
| Mean | | | 45.48 | 9.91 |
| Standard deviation | | | 10.89 | 1.14 |

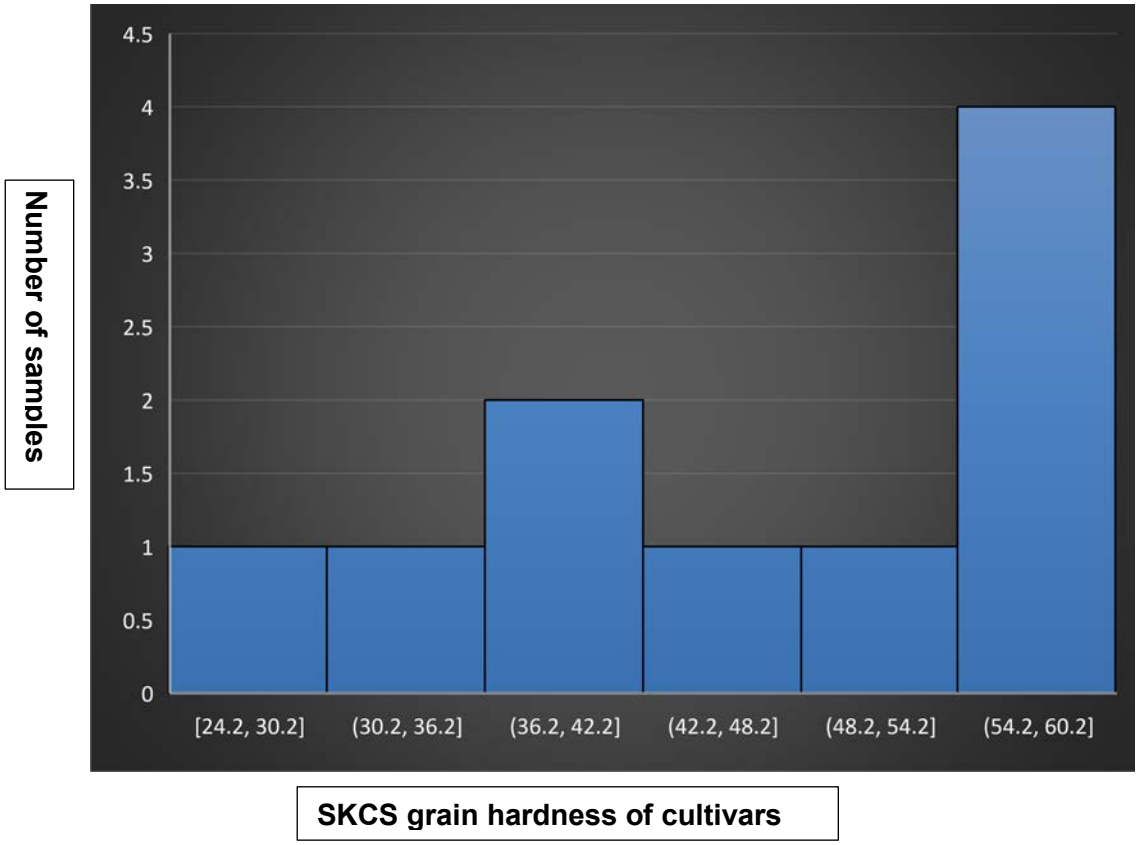



**Figure 6.10 Histograms representing grain hardness of barley cultivars.**
The bars depict grain hardness range intervals (e.g. 24.2-30.2)

In the case of landraces, the landrace (L 402448) with highest grain protein content (21%) depicted a high SKCS grain hardness value of 87.9. A range of 48.7-94.4 was observed for SKCS grain hardness with the majority of the grains having hardness value in the range of 83.7-90.7 (Fig. 6.11). The grain protein depicted a range of 11.6-21.1 and mean value of 14.51 ± 2.47 respectively. For these samples correlation coefficient between grain hardness and protein was 0.13 ($p < 0.01$). However, the correlation was not statistically significant. These findings were in line with previous studies that reported barley endosperm hardness and protein content were not correlated (Fox et al 2007b, Galmath et al. 2008). However, other studies have reported a significant positive correlation between these two traits (Fox et al. 2007a, Walker and Panozzo 2011). Thus, it is difficult to reach at a conclusion about the impact of grain protein content on barley grain hardness.

**Table 6.9: Grain protein content and hardness in barley landraces**

| No. | Landrace Aus Number | Origin | SKCS index | GPC (%) |
|---|---|---|---|---|
| 1 | L400097 | Ethiopia | 64.7 | 11.9 |
| 2 | L400211 | Ethiopia | 87.4 | 12.7 |
| 3 | L400619 | North Korea | 94.4 | 14.7 |
| 4 | L402433 | Afghanistan | 89.3 | 16.4 |
| 5 | L402448 | Iran | 87.9 | 21.1 |
| 6 | L403089 | Russia | 85.2 | 14.3 |
| 7 | L403156 | China | 84.7 | 14.2 |
| 8 | L411144 | Morocco | 63.0 | 16.9 |
| 9 | L411145 | Morocco | 54.7 | 13.7 |
| 10 | L411152 | Morocco | 48.7 | 15.8 |
| 11 | L411859 | Nepal | 80.4 | 11.6 |
| 12 | L411813 | Kyrgystan | 58.4 | 14.2 |
| 13 | L411890 | Nepal | 84.5 | 12.8 |
| 14 | L491320 | Japan | 61.8 | 12.9 |
| Range | | | 48.7-94.4 | 11.6-21.1 |
| Mean | | | 74.65 | 14.51 |
| Standard deviation | | | 15.24 | 2.47 |

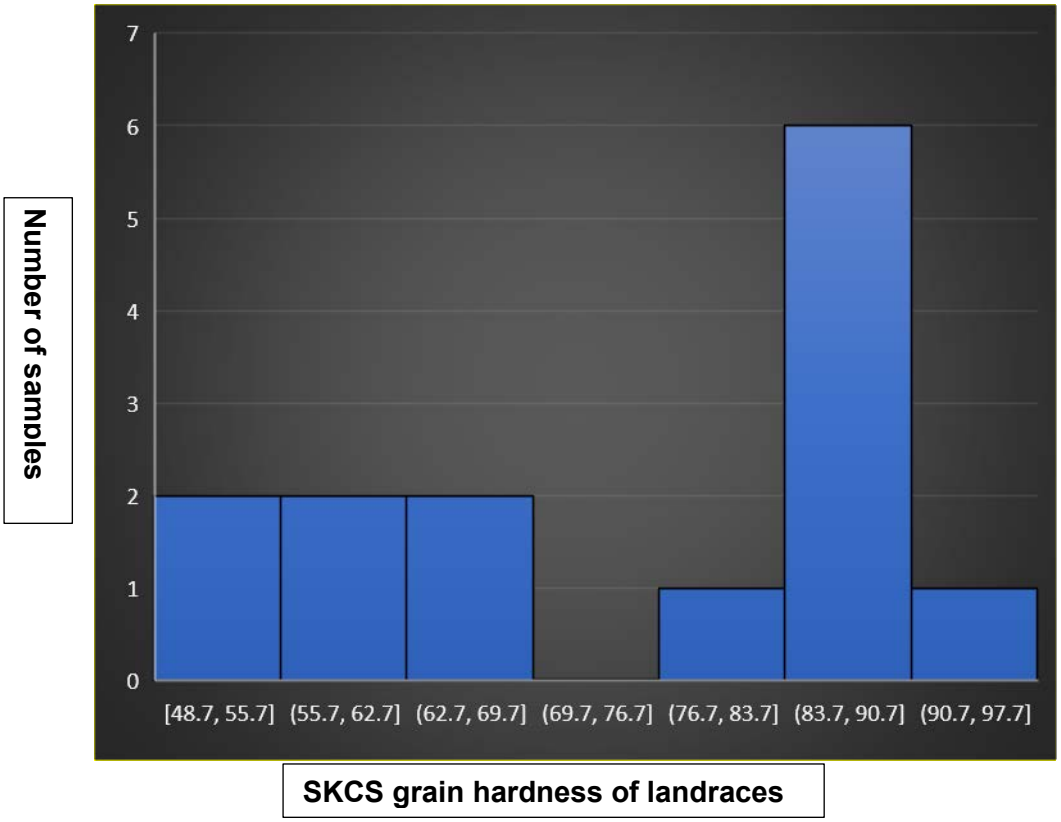



**Figure 6.11 Histograms representing grain hardness of barley landraces**
The bars depict grain hardness range intervals (e.g. 48.7- 55.7)

## 6.4 Discussion

### 6.4.1 *Hin* gene diversity in Australian barley cultivars and worldwide landraces

The investigation of *Hin* genes was undertaken to gain an understanding of the contribution of these genes to any textural properties of barley. Vast amount of information is available on the wheat *Pin* alleles and their association with texture, but such information in barley is scant. The role of *Hin* genes in influencing barley grain texture was ambiguous until Takahashi et al. (2010) reported a mutation in *Hinb*-2 gene identified from a Japanese barley line, Shikoku hadaka 84 and $F_2$ lines of the cross between Shikoku hadaka 84 (*Hinb-2b*)/ Shikoku hadaka 115 (*Hinb-2a*) with direct increase in grain hardness. The present study involved the molecular analysis of a number of Australian barley cultivars and landraces from different geographic origins from Asia and China, known to be the centres of origin and diversity of wheat and barley. Several sequence polymorphisms were detected that might contribute to the variation in structure and/or function of these proteins. The molecular analysis of *Hin* genes in cultivars/ landraces revealed nucleotide variation in one or more genes in 50% of the cultivars/landraces. The results of this study suggest that there is significant genetic variation in grain hardness genes in Australian cultivars and international landraces. Some of the earlier studies have reported low level of polymorphism in *Hin* genes in cultivars and conserved breeding material (Fox et al. 2007 b and Chen et al. 2008).

### 6.4.2 Greater diversity observed in *Hina* compared to *Hinb-1* and *Hinb-2* in cultivars

Greater diversity was observed in *Hina* gene in barley cultivars as compared to *Hinb-1* and *Hinb-2* genes. The greatest number of shared SNPs (within cultivars) was found in *Hina* gene followed by *Hinb-1* and none in *Hinb-2* genes suggesting that *Hina* gene was more diverse than *Hinb-1* and *Hinb-2* genes. However, for the landraces used in this study, shared SNPs were present in all three *Hin* genes, indicating that landraces were more genetically diverse than common (in-bred) cultivars.

### 6.4.3 Point mutations in putative HINA protein and their possible implications on biological (plant defence) and technological (grain hardness) function

The putative HINA proteins showed six substitutions. A potentially important mutation in the TRD region i.e. Thr71Arg in putative protein type HINA-E, was found in three cultivars (Arivat, Barque: 36.7 SKCS units, Capstan: 33.1 SKCS units) and five landraces (L400097: 64.7 SKCS units, L400211: 87.4 SKCS units, L402433: 89.3 SKCS units, L411144: 63.0 SKCS units and L411145: 54.7 SKCS units). As mentioned earlier in Section 6.3.1, the SKCS grain hardness values are not used as absolute values for comparison. It has been proposed that the interaction between PINs and acidic headgroups of polar lipids leads to disruption of bacterial and fungal membranes, thus imparting PINs their antibacterial and antifungal properties (Jing et al. 2003; Philips et al. 2011; Miao et al. 2012). The level of tryptophan and arginine residues in the TRD is relatively higher and these are common residues in potent antimicrobial peptides (Chan et al. 2006) thus highlighting the importance of these residues in terms of antibacterial and antifungal activity of PINs. Purified HINA extracted from the cultivars/ landraces can be used to synthesize antimicrobial peptide (AMP) based on the TRD region i.e. FPVTWRWWTWWKG. High antimicrobial activity of HINA peptide (FPVTWRWWTWWKG-$NH_2$) has been reported earlier by Phillips et al. (2011) and Shagaghi et al. (2016). The Thr71Arg mutation might lead to increased antimicrobial activity of the peptide since the replacement of polar Thr with basic Arg leads to an increase in the predicted net positive charge of HINA and increased positive charge may result in stronger electrostatic forces between HINA and charged head groups of polar lipids of bacterial membranes resulting in disruption of biological membranes and thus increased antimicrobial activity (based on similar interactions in PINs, reviewed in Pauly et al. 2013). Further, Pasupuleti et al. 2008 reported that increasing the positive charge can significantly increase the antimicrobial activity of an AMP. Further, this substitution leads to increase in number of basic residues from two to three in HINA. The increased lipid binding capacity of HINs is not only important biologically (plant defence) but might have implications on the technological function i.e. grain hardness as in case of PINs where the polar lipids-PINs association at starch granule surface helps to stabilize the amyloplast membrane during grain development or putting it as 'the decreased lipid binding capacity of mutated PINs cannot stabilise the amyloplast membrane during grain desiccation and resulting in grain hardness (reviewed in Pauly et al. 2013). Thus, it seems

that higher affinity of HINs to polar lipids might play a role in moderating barley grain hardness.

Substitution of an acidic residue with hydrophobic residue i.e. Glu78Leu, near the TRD, was also observed in putative protein type, HINA-E. Substitution of an acidic residue Glu with nonpolar Leu would lead to increase in pI values and thus the increased positive charge might lead to stronger interaction with polar lipids of biological membranes. Thus, the substitution Glu78Leu might prove beneficial in terms of its increased lipid binding capacity that has implications for both plant defence and grain hardness.

Another substitution, Gly43Glu was observed in putative protein type HINA-E. Glutamic acid (E) is found in wild type PINB at same position and a frame-shift mutation at this position leading to grain hardness has been reported by Ram et al. (2005). A conservative mutation Ser89Gly was also observed in HINA-E. A substitution in the predicted hydrophobic domain (HD) of HINA, i.e. Phe111Val was also present in HINA-E. Mutagenesis and protein-protein interaction studies using the yeast 2-hybrid system (Alfred et al. 2014) suggest a role for the HD in protein-protein interactions (PPI) of wheat PINs. Phenylalanine is more hydrophobic than valine; hence this substitution may lead to a decrease in hydrophobicity, important for membrane-binding functionality. A unique mutation, Ile99Leu (not reported earlier), was found in HINA-D (cv. Keel) but both amino acids have similar hydrophobicity. It is worth mentioning that five (Gly43Glu, Thr71Arg, Glu78Leu, Ser89Gly, Phe111Val) out of six substitutions observed in HINA have been reported earlier (Table 6.2) with unknown effects on grain texture. Interestingly, these four potentially important substitutions (Gly43Glu, Thr71Arg, Glu78Leu, Phe111Val) were observed in three feed cultivars (Arivat, Barque, Capstan) and five landraces. The cultivars with HINA-E putative protein type showed lower SKCS values (Barque: 36.7 SKCS units, Capstan: 33.1 SKCS units) suggesting the substitutions might have softening effect on barley grain-hardness. However, considering the landraces (54.7, 63.0, 64.7, 87.4, 89.3 SKCS units) with HINA-E putative protein type it was not possible to draw conclusion regarding the impact of substitutions on grain texture. But the SEM analysis of two landraces L400619 (wild type) and L411145 (HINA-E) suggested that L411145 appeared softer indicating probable softening effect of the amino acid substitutions. Further, it is important to note that no single substitution was present in any cultivar/ landrace that would have given the opportunity to associate the genetic variation with grain hardness values. Only one single substitution, Gly43Glu, was noted

in cultivar Stirling (SKCS value 56.4) suggesting no association between genetic variation and grain hardness value since the wild type HINA-A showed a range of grain hardness values (24.2-58.1). Thus, the SKCS hardness values could not be directly associated with SNPs occurring in their *Hin* genes. It is thus concluded that two HINA substitutions (Thr71Arg, Glu78Leu) can be used to synthesize AMPs and the purified HINA-E containing all the important substitutions can also be used to make synthetic peptides.

### 6.4.4 Point mutations in putative HINB-1 protein and their possible implications

The putative HINB-1 proteins showed two substitutions, Ile9Leu and Thr16Ile, in the leader peptide. Substitutions with basic residue, i.e. Thr59Arg and Gln105Arg (in the HD), were also observed, which may lead to increase in net positive charge and greater antimicrobial activity as mentioned earlier. Two substitutions Gln105Arg and Gly110Ala noted in the HD might have implications on the HIN protein-protein interactions. A frame-shift mutation detected in L400211 from Ethiopia showed a high SKCS grain hardness value of 87.4 SKCS units indicative of hard grain. A *Hinb-2* null mutation reported earlier by Takahashi et al. (2010) has also been linked with increased grain hardness.

### 6.4.5 Point mutations in putative HINB-2 protein and their possible implications

Two substitutions (Val137Ala, Arg140Lys, Ser143Gly) noted in HINB-2 were conservative. However, a substitution Ser143Thr observed in cv. Sloop SA is an important finding, as Ser to Ile substitution at the same position in PINB of wheat has been associated with grain hardness (Chang et al. 2006). But interestingly, the SKCS hardness value of Sloop SA (wild type for HIN-A and HINB-1) was found to be low (24.2). This substitution has not been reported earlier.

In conclusion, the results indicate notable *Hin* gene diversity, with some of the substitutions being at the functionally important TRD and HD regions in the putative proteins. A clear relationship between SNPs in *Hin* genes and grain hardness could not be established but potentially important HIN substitutions were identified that could be used as AMPs. Thus, this study demonstrates that although barley grain hardness genes impact the grain quality, these genes alone cannot explain all the variation in grain hardness.

### 6.4.6 Relationship between *Hin* gene sequence variation and grain hardness based on SKCS data

Four cultivars (Arapiles, Baudin, Buloke, Franklin) that showed all three wild type *Hin* genes showed differences in the SKCS hardness values (46.9, 49.2, 58.1, 54.6 SKCS units). Similarly, landraces wild type for all three *Hin* genes (L400619, L402448, L403089, L403156, L411152, L411813, L491320) also showed differences for SKCS values (94.4, 87.9, 85.2, 84.7, 48.7, 58.4, 61.8 SKCS units) implying that a range of SKCS hardness values was observed both in wild type cultivars and landraces and a single value could not be used as a reference for comparison with the mutant types. Although there is high homology between *Hin* and *Pin* genes (Section 6.2), the *Hin* genes cannot be directly associated with grain hardness as in the case of some *Pin* mutations (reviewed in Bhave and Morris, 2008a) associated with hard texture. This is likely due to the complex nature of the barley grain hardness that is also controlled by other factors such as grain weight (Galassi et al.2012), diameter, seed size (Turuspekov et al. 2008), β-glucans (Gamlath et al. 2008; Pannozo et al. 2007; Psota et al. 2007).

A comprehensive analysis of genetic diversity in barley cultivars and landraces revealed several different *Hina*, *Hinb-1* and *Hinb-2* haplotypes. As discussed earlier some of the haplotypes like *Hina-e*, *Hina-f*, *Hina-g* (all leading to deduced protein type HINA-E), *Hinb-1f* (deduced protein type HINB-1E) *Hinb-2g*, *Hinb-2h* (deduced protein type HINB-2E) were found to be important because of implications on grain texture and/or antimicrobial properties. Previous studies (Caldwell et al. 2006; Galassi et al. 2012; Li et al. 2011; Turuspekov et al. 2008) reported *Hin* haplotypes with unknown effects on grain texture and one study by Takahashi et al. (2010) reported a mutation in *Hinb-2* with direct increase in grain hardness. This study compares different haplotypes observed in this work and postulates their likely implications on endosperm hardness and plant defence from biotic agents. Thus selection of certain haplotypes may serve as a useful tool for barley breeders aiming to improve barley quality.

# CHAPTER 7

# General discussion and conclusions

## 7.1 General discussion and conclusions

Wheat and barley are the third and fourth most-produced cereals, respectively in the world and are the most important grain crops in Australia based on quantity produced, area cultivated and revenue generated. These crops are highly nutritive for humans and animals. Grain hardness or endosperm texture is a commercially important trait that governs the end use of cereals. The variability in wheat grain texture enables production of different food products and eating qualities. Similarly, grain texture affects malting and processing quality of barley with relevance to its end-use. Hard textured barley is preferred for stockfeed while soft is preferred for malting and brewing. Together with protein percentage grain hardness is one of the most significant traits assessed for wheat and barley grain quality assessment and international trade. *Puroindoline* genes control grain hardness in wheat and *Hordoindoline* genes are the barley counterparts. Thus, the first chapter of this thesis reviewed the current available literature on this trait in wheat and barley. It described that the *puroindoline* genes (*Pina-D1* and *Pinb-D1)* located on chromosome 5D of common wheat (*Triticum aestivum* L) are the main genetic determinants of grain hardness. The Puroindolines (PIN) are small, highly basic proteins with a highly conserved cysteine backbone and a tryptophan rich domain (TRD). Both these genes are required in their wild-type form for conferring soft texture and any mutation or deletion in either or both *Pin* genes results in the 'hard' grain phenotype. The lipid binding properties of PINs have been associated with their grain softening effects and antimicrobial properties. Several mutations in *Pina*/*Pinb* genes with effect on grain hardness have been reported worldwide. Members of the gene copy called *Puroindoline b-2* (*Pinb-2*), present on all group 7 chromosomes, also have minor contributions to texture. It is known that genotypes of same *Pin* class (wild-type or mutant) vary in grain hardness. The *Pinb-2* genes might be the additional factors that contribute to grain hardness in wheat and explain the variation in grain hardness not contributed by the *Puroindoline* genes. The *Pinb-2* genes were discovered less than ten years ago, and exist as multigene family. Similar to Puroindolines, the PINB-2 proteins contain the ten characteristic cysteine residues and TRD. The literature reporting the identification, chromosomal location of *Pinb-2* variants, their allelic nature and association with grain yield and texture was reviewed. Other aspects like the *Pinb-2* gene expression and antimicrobial nature of PINB-2 peptides, like PIN peptides were also reviewed.

Chapter 1 also described the barley orthologues of *Puroindolines* i.e., *Hordoindolines* located on chromosome 5H. The *Hordoindoline* genes comprise of *Hina*, *Hinb-1*, *Hinb-2* genes. *Hinc* genes that may be an orthologue of the wheat *Pinb-2* genes have also been reported. Hordoindolines (HIN) are seed proteins of barley, similar to the wheat PIN proteins. Similar to PINs, HINs are predicted to be cysteine-rich basic proteins and contain a TRD. Significant information is available on the effects of diverse mutations in *Pin* genes on wheat grain texture, but information on contributions of genetic variants of *Hin* genes to barely grain texture is limited. Takahashi et al. (2010) first provided some evidence of association of *Hin* genes to texture, illustrated by a *Hinb-2* null mutant with increased grain hardness. However, the role of *Hin* genes in influencing grain texture is still not clear and needs investigation.

Thus, the main aims of this research project were:

1. To investigate the effect of any *Pinb-2* variations on grain hardness within same *Pin* genotype class in Australian wheat cultivars and worldwide landraces
2. To investigate the association of *Pinb-2v3-1a* (Val104Ala) with grain hardness
3. To investigate sequence diversity of the *Pinb-2* genes in selected wheat cultivars
4. To study protein-protein interactions occurring between PINA, PINB, and PINB-2v3 that may have effect on grain hardness
5. To study the gene expression of *Pinb-2* and its association with grain hardness
6. To investigate *Hin* gene diversity in Australian barley cultivars and worldwide landraces that may influence barley grain hardness.
7. To find novel *Hin* alleles that might have implications for lipid binding, food and feed and possibly also antimicrobial properties

The above aims were addressed using a number of experimental techniques as described in Chapter 2. Briefly, these included DNA extractions, RNA extractions, standard PCRs, nested PCR, sqRT-PCR, gene cloning, DNA sequencing and methods specific to yeast two-hybrid work. Bioinformatic methods such as DNA and protein sequence analysis and alignments, restriction mapping, Genbank and other databases were also used. Scanning electron microscopy was used for grain texture studies in wheat and barley. The first three aims were addressed in Chapter 3 and Chapter 4. Briefly, 22 Australian cultivars and 18 worldwide landraces were genotyped for *Pinb-2* variants. The cultivars

belonged to three different *Pin* genotype classes (*Pina-D1a/Pinb-D1a*, *Pina-D1a*/*Pinb-D1b, Pina-D1b*/*Pinb-D1a*) and varied for grain hardness. The full *Pinb-2* genotype diversity (study of all six variant alleles) was investigated in four cultivars. This chapter investigated *Pinb-2* gene diversity with regards to variation in grain hardness values between wheat cultivars with the same *Pin-D1* genotype. The results showed (i) the *Pinb-2v3* allele was predominant compared to *Pinb-2v2*, and these two alleles did not co-exist, supporting these being true alleles; (ii) the subtypes *Pinb-2v3-1* and *Pinb-2v3-1a* (Val104Ala) did not co-exist, confirming their allelic nature; (iii) *Pinb-2* variants: v1, v2 or v3, v4, v5, v6 occurred together in various combinations and are likely non-allelic; (iv) Val104Ala substitution in *Pinb-2v3* was observed in both soft (*Pina-D1a*/*Pinb-D1a*) and hard genotypes (*Pina-D1a*/*Pinb-D1f*, *Pina-D1a/Pinb-D1b*, *Pina-D1b*/*Pinb-D1a*) in landraces whereas in case of cultivars this substitution was found only in hard wheat genotypes (*Pina-D1a/Pinb-D1b*, *Pina-D1b/Pinb-D1a*); (v) occurrence/non-occurrence of *Pinb-2* variants (v2 or v3) could not explain the difference in SKCS hardness values of same *Pin* class (*PinaD1-a/Pinb-D1a*, *PinaD1a/Pinb-D1b*, *Pina-D1b/Pinb-D1a*) genotypes. No pattern was observed in relation to the SKCS values and the *Pinb-2v2*/*Pinb-2v3* alleles, in case of all three *Pin* classes; (vi) the *Pinb-2v3-1a* (Val104Ala) could not be associated with increased SKCS grain hardness values as reported in an earlier study by Chen et al. (2013a); (vii) the presence of six variant types and provided further evidence for the multigenic nature of the *Pinb-2* gene family (viii) the novel subtype *Pinb2v4-2b* found in cultivar Egret could be potentially useful as it contains Trp57Arg substitution that led to highest pI (9.27) of the putative protein; (ix) the identification of three landraces that had combination of grain hardness associated mutations in *Pin* genes and *Pinb-2v3* subtype with Val104Ala substitution that might be potentially useful for enhancing the Australian wheat germplasm

The next two aims (4, 5) were addressed in Chapter 5. *Pinb-2* variant 3 has been found to be predominant in Australian cultivars (Chapter 3) and landraces (Chapter 4) and has been linked with increased grain hardness (Chen et al. 2013a). Briefly, the association between PINA, PINB and PINB-2v3 was tested for physical interaction using the yeast two-hybrid system. An attempt was also made to investigate the gene expression of *Pinb-2* variants and its association with grain hardness. The results showed that (i) the interaction of PINB-2v3 with PIN proteins is weak suggesting that interaction between them is likely

not required for lipid binding at starch granule surface that moderates the grain texture; (ii) the lower number of tryptophan residues in the TRD of PINB-2v3 as compared to PINA and PINB may be the likely reason for weak interactions; (iii) based on comparison of hydrophobic domain of PIN and PINB-2 proteins it is proposed that the basic residue Arginine and hydrophobic residue Valine of the HD might be important for protein-protein interactions; iv) the PINB-2 variant proteins seem to have greater significance as antimicrobial proteins and make minor contribution in affecting grain hardness; (v) *Pinb-2* variants express in leaf and root tissues; (vi) the gene expression level of *Pinb-2v3-1a* (Val104Ala) was found higher as compared to *Pinb-2v2* and *Pinb-2v3-1*; (vii) the increased gene expression level of *Pinb-2v3-1a* was not found to be associated with grain hardness values.

The last two aims were addressed in Chapter 6. Briefly, this chapter investigated *Hina*, *Hinb-1* and *Hinb-2* gene diversity in 14 worldwide landraces and 12 Australian barley cultivars. The selection of cultivars/landraces was based on SKCS grain hardness value, utility and region. The results showed that: (i) nucleotide variation is present in one or more genes in 50% of the cultivars/landraces; (ii) a number of different *Hina*, *Hinb-1* and *Hinb-2* haplotypes (novel and published) were found in cultivars and landraces; (iii) a novel *Hinb-1* allele with insertion at nucleotide position 210 and leading to frame-shift mutation was discovered in landrace L400211 from Ethiopia (high SKCS value: 87.4 units ), suggesting that this mutation may result in increased barley grain hardness; (iv) some of the haplotypes were shared by cultivars and/ or landraces for all three *Hin* genes; (v) shared SNPs were present in all three *Hin* genes in landraces indicating greater diversity than cultivars as expected; (vi) greater amount of diversity was observed in *Hina* gene as compared to *Hinb-1* and *Hinb-2* genes; (vii) potentially important substitutions (Thr71Arg, Glu78Leu) with implications on lipid binding/antimicrobial properties were noted in cultivars/landraces; (viii) the SKCS hardness values could not be directly associated with SNPs occurring in their *Hin* genes except for L400211 that showed frame shift mutation in *Hinb-1* gene; (ix) SEM was found to be a useful tool for differentiating hard and soft barley types in addition to SKCS.

It is thus concluded that; (i) there was no pattern in relation to the grain hardness values and *Pinb*-2v2/*Pinb-2v3*/*Pinb-2v3-1a* alleles; (ii) unlike the *Pina-D1* and *Pinb-D1* genes, the expression of *Pinb-2* genes is not seed specific, as the expression was found in leaves and root tissues; (iii) the higher gene expression of *Pinb-2v3-1a* (Val104Ala) was not associated with grain hardness; (iv) the weak physical interactions of PINB-2v3 with PINs suggests that PINB-2 could have only minor influence on grain hardness; (vi) there is significant genetic variation in grain hardness genes in Australian barley cultivars and international landraces; (vi) although barley grain hardness genes impact the grain quality, these genes alone cannot explain all the variation in grain hardness; (vi) although there is high homology of 88-92% between the *Ha* locus component genes of barley and wheat there seems to be limited direct association between factors affecting barley and wheat grain hardness; (vii) direct correlation between barley and wheat grain hardness based only on mutations in *Hin/Pin* genes being unclear, due to higher SKCS values typically noted in wheat lines with mutant *Pina/Pinb* than those with wild-type genes, but similar SKCS values seen in barley lines with mutant or non-mutant *Hin* genes.

In wheat, it is well-established that approximately 60% of grain hardness variation is contributed by the *Ha* locus (Mikulikova, 2007) where the *Pin* genes are located. A number of *Pin* alleles are also well-established to lead to a hard texture in common wheat (Giroux and Morris 1997; Giroux and Morris 1998; Kumar et al. 2015; reviewed in Bhave and Morris 2008a). However, wheat grain hardness is also shown to be influenced by additional factors such as such as differences in *Pin* expression (Nirmal et al. 2016), environment and biochemical factors (Dessalegn et al. 2006; Gazza et al. 2008; Tranquilli et al. 2002; Turnbull and Rahman 2002). Minor loci (1A, 1B, 6D) other than *Ha* locus and several QTLs (1AS, 1BL, 5AS, 5DL, 6B) may be involved in influencing grain hardness (reviewed in Bhave and Morris, 2008b). Considering the *Hin* gene sequence variations together with the broad ranges seen in the SKCS and/or SEM results, a direct connection between the barley *hordoindoline* alleles and grain hardness seems unlikely, compared to that in wheat. However, the fact that barley is typically harder than wheat may also be related to other factors in barley grains such as glucans (Section 6.4.6 p 183). Apart from the *Hin* genes other gene effects have been associated with barley grain texture (Fox et al. 2007a; Walker et al. 2011; Walker et al. 2013; Mohammadi et al. 2014). Several QTLs associated with barley grain texture traits have been detected on

chromosomes 1HS, 2H, 3H, 4HL, 5HS, 5HL, 6H and 7H (reviewed in Walker and Pannozo, 2016). The combined effects of husk thickness, husk adherence to endosperm, arabinoxylan content, protein content and composition, grain diameter have also been suggested to influence grain hardness (Walker et al. 2011).

## 7.2 Future directions

The allelic diversity of *Pina* and *Pinb* genes is highly sought after because of its impact on grain texture and association with milling and processing qualities in bread wheat. The wheat textural classes with different end-uses are determined by different combinations of *Pina-D1* and *Pinb-D1* alleles. In the recent years, several new *Pina* and *Pinb* alleles have been discovered in different wheat germplasm (Ali et al. 2015, Ayala et al. 2016, Chen et al. 2013b, Kumar et al. 2015, Qamar et al. 2014). However, the reason for difference in the grain hardness of genotypes of same *Pin* class is not known. The *Pinb-2* genes could be one of the minor genetic factors involved in influencing grain hardness. This study did not find any association between *Pinb-2* variant type and sequence diversity with grain texture. The Val104Ala mutation in *Pinb-2v3* had been associated earlier with harder texture in soft wheats (Chen et al. 2013 a). However, no such association was not found in the hard wheats analysed in this study. Thus, the involvement of *Pinb-2* genes in grain texture remains unclear. Further studies are required to analyse the role of *Pinb-2* genes in influencing grain texture. The analysis of these genes in durum wheats could prove useful as there would be no effects of *Pina* and *Pinb* genes. Durum wheat recombinant lines involving *Pinb-2* variants can be investigated for effects on grain texture, as undertaken in case of *Pin* genes by Heinze et al. (2016).

Transgenic studies involving introduction of *Pinb-2* variants into different crops can be conducted like transgenic studies undertaken with *Pin* genes (Xia et al. 2008, Wada et al. 2010). These studies can be useful to understand the role of *Pinb-2* genes as grain texture moderators and antimicrobial defence agents. Further, the antimicrobial activity of PINB-2 peptides has been shown earlier (Phillips et al. 2011). Hence the potentially useful novel subtype *Pinb2v4-2b* (Trp57Arg substitution, pI 9.27) discovered in this study can be used to synthesize synthetic peptides and investigate its antimicrobial properties.

The investigation of promoter regions of *Pinb-2* genes might help understand their expression patterns and relation to effect on grain hardness and antimicrobial activities. This study found higher gene expression level of *Pinb-2v3-1a* (Val104Ala) as compared to *Pinb-2v2* and *Pinb-2v3-1*. However, the gene expression variation was not found to correlated to grain hardness values. Thus, the analysis of promoter regions of *Pinb-2* variants using inverse PCR might help to identify regions essential for tissue specific expression, as performed for the *Pinb* promoter (Digeon et al.1999). Work also needs to be undertaken to test any variation in *Pinb-2* gene expression under infections of plants by bacterial or fungal pathogens, as undertaken for *Pin* genes (Krishnamurthy et al. 2001). Increased resistance against *Magnaportha grisea* and *Rhizoctonia solani* which are causal agents of rice blast and sheath blight respectively was exhibited by the *Pin* genes expressed in transgenic rice (Krishnamurthy et al. 2001)

Use of the yeast two-hybrid system in this work has shown that the PINB-2v3-1 protein does not interact with PINA and PINB. Previous study in our lab had identified the hydrophobic domain as an important (but not essential) region for protein-protein interactions of PINs (Alfred et al.2014). Based on comparison of hydrophobic domain of PIN and PINB-2 proteins, it is hypothesized that the basic residue Arginine and hydrophobic residue Valine of the HD might be important for protein-protein interactions. However, this hypothesis needs to be tested. Site-directed mutagenesis PCR can be used to introduce substitutions in the protein and investigate the effect of this substitution on protein-protein interactions.

In the case of barley, the screening for new *Hin* alleles in search of diversity of grain texture would prove useful since hard textured barley is favourable for animal feed and soft is preferred for malting. In this study, a frameshift mutation in *Hinb-1* gene associated with hard texture was found in a landrace from Ethiopia. The landraces from Afghanistan, China, Ethiopia, Iran, Japan, Kyrgystan, Morocco, Nepal, North Korea and Russia which are the important centres of diversity of barley, thus offer a great resource for genetic variations with implications on grain texture. New alleles with influence on grain texture, or residues involved in lipid binding, can be identified. Synthetic peptides based on the putative protein type HINA-E with potentially important substitutions, i.e., Gly43Glu, Thr71Arg, Glu78Leu, Ser89Gly, and/or Phe111Val can be designed and tested for antimicrobial activity. High antimicrobial activity of peptide based on TRD of HINA (FPVTWRWWTWWKG) has been reported earlier in our lab by Phillips et al. (2011)

and Shagaghi et al. (2016). The *Hinc* genes (orthologues of *Pinb-2* genes) can be investigated for genetic variation as it might have implications on textural variations that have commercial significance. Genetic modification studies involving *Hin* alleles can be used to investigate their influence on grain texture and antimicrobial properties.

## Bibliography

Alfred RL, Palombo EA, Panozzo JF, Bhave, M. 2014. The co-operative interaction of Puroindolines in wheat grain texture may involve the hydrophobic domain. *Journal of Cereal Science*, 60(2): 323-330.

Ali I, Sardar Z, Rasheed A, Mahmood T. 2015. Molecular characterization of the *Puroindoline-a* and *b* alleles in synthetic hexaploid wheats and in silico functional and structural insights into *Pina-D1*. *Journal of Theoretical Biology* 376:1-7.

Allison MJ, Cowe IA, McHale R. 1976. A rapid test for the prediction of malting quality of barley. *Journal of the Institute of Brewing* 82: 166-167.

Amoroso MG, Longobardo L, Capparelli R. 2004. Real time RT-PCR and flow cytometry to investigate wheat kernel hardness: role of *Puroindoline* genes and proteins. *Biotechnology letters* 26(22): 1731-1737.

Australian Export Grain Innovation Centre 2015. www.aegic.org.au: Last accessed June, 2016.

Ayala M, Guzmán C, Peña RJ Alvarez JB. 2016. Genetic diversity and molecular characterization of *Puroindoline* genes (*Pina-D1* and *Pinb-D1*) in bread wheat landraces from Andalusia (Southern Spain). *Journal of Cereal Science* 71: 61-65.

Baik B-K, Newman CW, Newman RK. 2011. Food uses of barley. In: Ullrich SE (ed) Barley: Production, improvement and uses, John Wiley and sons, West Sussex, UK, pp 532-562.

Beecher B, Bowman J, Martin JM, Bettge AD, Morris CF, Blake TK, Giroux MJ. 2002. *Hordoindolines* are associated with a major endosperm-texture QTL in barley (*Hordeum vulgare*). *Genome* 45: 584–591.

Beecher B, Smidansky ED, See D, Blake TK, Giroux MJ. 2001. Mapping and sequence analysis of barley *Hordoindolines*. *Theoretical and Applied Genetics* 102: 833-840.

Bhave M, Morris CF. 2008a. Molecular genetics of *Puroindolines* and related genes: allelic diversity in wheat and other grasses. *Plant Molecular Biology* 66: 205-219.

Bhave M, Morris CF. 2008b. Molecular genetics of *Puroindolines* and related genes: regulation of expression, membrane binding properties and applications. *Plant Molecular Biology* 66: 221-231.

Bleidere M, Zinta G. 2012. Grain quality traits important in feed barley. *Proceedings of the Latvian Academy of Sciences* 66: 1-9.

Blochet J-E, Chevalier C, Forest E, Pebay-Peyroula E, Gautier M-F, Jourdrie P, Pézolet M, Marion D. 1993. Complete amino acid sequence of puroindoline, a new basic and cystiene-rich protein with a unique tryptophan-rich domain, isolated from wheat endosperm by Triton X-114 phase partitioning *FEBS Journal* 329:336-340.

Bowman JGP, Blake TK, Surber, LMM, Habernicht, TK, Daniels, JT. 1996. Genetic factors controlling digestibility of barley for ruminants. *Proceedings Animal Society of Animal Science* 47: 257-260.

Breseghello F, Finney PL, Gaines C, Andrews L, Tanaka J, Penner G, Sorrells ME. 2005. Genetic loci related to kernel quality differences between a soft and a hard wheat cultivar. *Crop science* 45(5): 1685-1695.

Brennan CS, Harris N, Smith D, Shewry PR. 1996. Structural differences in the mature endosperms of good and poor malting barley cultivars. *Journal of Cereal Science* 24:171-177.

Byung-Kee B, Ullrich SE. 2008. Barley for food: Characteristics, improvement and renewed interest. *Journal of Cereal Science* 48: 233-242.

Caldwell KS, Langridge P, Powell W. 2004. Comparative sequence analysis of the region harboring the *Hardness* locus in barley and its colinear region in rice. *Plant Physiology* 136: 3177–3190.

Caldwell KS, Russell J, Langridge P, Powell W. 2006. Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species, *Hordeum vulgare*. *Genetics* 172: 557–567.

Camm JP, Ellis RP, Morrison WR. 1990. Milling energy: an investigation into the biochemical basis of hardness in cereals. *Aspects of Applied Biology* 25:121-130.

Cane K, Spackman M, Eagles HA. 2004. *Puroindoline* genes and their effects on grain quality traits in southern Australian wheat cultivars. *Crop and Pasture Science*, 55(1): 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 Molecular Biology* 58:857-867.

Capparelli R, Borriello G, Giroux MJ, Amoroso MG. 2003. Puroindoline A-gene expression is involved in association of puroindolines to starch. *Theoretical and Applied Genetics* 107:1463-1468.

Chan DI, Prenner EJ, Vogel HJ. 2006. Trytophan- and arginine-rich antimicrobial peptides: structures and mechanisms of action. *Biochimica et Biophysica Acta* 1758:1184-1202.

Chandra G S, Proudlove M O, Baxter E D. 1999. The structure of barley endosperm – an important determinant of malt modification. *Journal of the Science of Food and Agriculture* 79:37–46.

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.

Chantret N, Salse J, Sabot F, Rahman S, Bellec A, Laubin B, Dubois I, Dossat C, Sourdille P, Joudrier P, Gautier M-F, Cattolico L, Beckert M, Aubourg S, Weissenbach J, Caboche M, Bernard M, Leroy P, Chalhoub B. 2005. Molecular

basis of evolutionary events that shaped the *Hardness* locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). *The Plant Cell* 17:1033-1045.

Chen F, Beecher BS, Morris CF. 2010a. Physical mapping and a new variant of *Puroindoline b-2* genes in wheat. *Theoretical and Applied Genetics* 120: 745-751.

Chen F, He Z, Xia X, Lillemo M, Morris CF. 2005. A new *Puroindoline b* mutation present in Chinese winter wheat cultivar Jingdong 11. *Journal of Cereal Science* 42:267-269.

Chen F, He ZH, Xia XC, Xia LQ, Zhang XY, Lillemo M, Morris CF. 2006. Molecular and biochemical characterisation of *Puroindoline a* and *b* alleles in Chinese landraces and historical cultivars. *Theoretical and Applied Genetics* 112:400-409.

Chen F, Li H, Cui D. 2013 a. Discovery, distribution and diversity of *Puroindoline-D1* genes in bread wheat from five countries (*Triticum aestivum* L.). *BMC Plant Biology*, 13(1): 125.

Chen F, Xu HX, Zhang FY, Xia XC, He ZH, Wang DW, Dong ZD, Zhan KH, Cheng XY, Cui DQ. 2011. Physical mapping of *Puroindoline b-2* genes and molecular characterization of a novel variant in durum wheat (*Triticum turgidum* L.). *Molecular Breeding* 28:153-161.

Chen F, Yu Y, Xia X, Z. 2007. Prevalence of a novel *Puroindoline b* allele in Yunnan endemic wheats (*Triticum aestivum* ssp. Yunnanense King). *Euphytica* 156: 39–46.

Chen F, Zhang F, Cheng X, Morris C, Xu H, Dong Z, Zhan K, Cui D. 2010b. Association of *Puroindoline b-B2* variants with grain traits, yield components and flag leaf size in bread wheat (*Triticum aestivum* L.) varieties of the Yellow and Huai Valleys of China. *Journal of Cereal Science* 52: 247-253.

Chen F, Zhang F, Li H, Morris CF, Cao Y, Shang X, Cui D. 2013b. Allelic variation and distribution independence of *Puroindoline b-B2* variants and their association with grain texture in wheat. *Molecular breeding*, 32(2): 399-409.

Chen M, Wilkinson M, Tosi P, He G, Shewry P. 2005. Novel puroindoline and grain softness protein alleles in *Aegilops* species with the C, D, S, M and U genomes. *Theoretical and Applied Genetics*, 111(6):1159-1166.

Chien CT, Bartel PL, Sternglanz R, Fields S. 1991. The two-hybrid system: A method to identify and clone genes for proteins that interact with a protein of interest. *Proceedings of the National Academy of Sciences of the United States of America* 88:9578-9582.

Clifton LA, Green RJ, Frazier RA. 2007a. *Puroindoline-b* mutations control the lipid binding interactions in mixed Puroindoline-a:Puroindoline-b systems.

*Biochemistry* 46:13929-13937.

Clifton LA, Lad MD, Green RJ, Frazier RA. 2007b. Single amino acid substitutions in puroindoline-b mutants influence lipid binding properties. *Biochemistry* 46:2260-2266.

Corona V, Gazza L, Zanier R, Pogna NE. 2001. Variation in friabilin composition as determined by A-PAGE fractionation and PCR amplification, and its relationship to grain hardness in bread wheat. *Journal of Cereal Science* 34: 243-250.

Cu ST, March TJ, Stewart S, Degner S, Coventry S, Box A, Stewart D, Skadhauge B, Burton RA, Fincher GB, Eglinton J. 2016. Genetic analysis of grain and malt quality in an elite barley population. *Molecular Breeding* 36(9):129.

Darlington HF, Tecsi L, Harris N, Griggs DL, Cantrell IC, Shewry PR. 2000. Starch granule associated proteins in barley and wheat. *Journal of Cereal Science* 31: 21–29.

Darlington HF, Rouster J, Hoffmann L, Halford NG, Shewry PR, Simpson DJ. 2001 Identification and molecular characterisation of *Hordoindolines* from barley grain. *Plant Molecular Biology* 47: 785–794.

Dessalegn T, Labuschagne MT, Van Deventer CS. 2006. Quality of Ethiopian durum wheat lines in two diverse environments. *Journal of Agronomy and Crop Science* 192(2): 147-150.

Digeon JF, Guiderdoni, E, Alary R, Michaux-Ferriere N, Joudrier P, Gautier MF. 1999. Cloning of a wheat puroindoline gene promoter by IPCR and analysis of promoter regions required for tissue-specific expression in transgenic rice seeds. *Plant Molecular Biology* 39: 1101-1112.

Dhatwalia VK, Sati OP, Tripathi MK, Kumar A. 2011. Puroindoline: antimicrobial wheat endosperm specific protein. *Journal of Agricultural Technology* 7:903-906.

Douliez JP, Michon T, Elmorjani K, Marion D. 2000. Mini Review: structure, biological and technological functions of Lipid Transfer Proteins and Indolines, the major lipid binding proteins from cereal kernels. *Journal of Cereal Science* 32: 1-20.

Dubreil L, Compoint J-P, Marion D. 1997. Interaction of Puroindolines with wheat flour polar lipids determines their foaming properties. *Journal of Agricultural and Food Chemistry* 45:108-116.

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-LTP1e1) of *Triticum aestivum* seeds. Relationships with their in vitro antifungal properties. *Plant Science* 138:121-135.

Edney MJ, Rossnagel BG, Endo Y, Ozawa, S, Brophy M. 2002. Pearling quality of Canadian barley varieties and their potential uses as rice extenders. *Journal of Cereal Science* 36, 295–305.

Feiz L, Martin JM, Giroux MJ. 2009a. Creation and functional analysis of new *Puroindoline* alleles in *Triticum aestivum*. *Theoretical and Applied Genetics* 118:247-257.

Feiz L, Wanjugi HW, Melnyk CW, Altosaar I, Martin JM, Giroux MJ. 2009b. Puroindolines co-localize to the starch granule surface and increase seed bound polar lipid content. *Journal of Cereal Science* 50: 91-98.

Food and agricultural organization of United nation (FAO): Barley post-harvest operations 2014. http://www.fao.org/fileadmin/user_upload/inpho/docs/Post_Harvest_Compendium_-_BARLEY.pdf; Accessed September 2015.

Fox G, Kelly A, Bowman J, Inkerman A, Poulsen D, Henry R 2009. Is malting barley better feed for cattle than feed barley. *Journal of the institute of brewing* 115 (2): 95-104.

Fox GP, Nguyen L, Bowman J, Poulsen D, Inkerman A, Henry RJ. 2007. Relationship between hardness genes and quality in barley (*Hordeum vulgare*). *Journal of the Institute of Brewing* 113(1): 87-95.

Fox GP, Osborne B, Bowman J, Kelly A, Cakir M, Poulsen D, Inkerman A, Henry R 2007. Measurement of genetic and environmental variation in barley (*Hordeum vulgare*) grain hardness. *Journal of Cereal Science* 46(1): 82–92.

Fox GP, Panozzo JF, Li CD, Lance RCM, Inkerman PA, Henry RJ. 2003. Molecular basis of barley quality. *Australian Journal of Agricultural Research* 54:1081–1101.

Galassi E, Gazzelloni G, Taddei F, Muccilli V, Gazza L, Pogna N 2012. Kernel texture and hordoindoline patterns in barley (*Hordeum vulgare*). *Molecular Breeding* 30:1–12.

Gautier M- F, Aleman M- E, Guirao A, Marion D, Joudrier P. 1994. *Triticum aestivum* Puroindolines, two basic cysteine rich seed proteins: cDNA sequence analysis and developmental gene expression. *Plant Molecular Biology* 25: 43-57.

Gautier M-F, Cosson P, Guirao A, Alary R, Joudrier P. 2000. *Puroindoline* genes are highly conserved in diploid ancestor wheats and related species but absent in tetraploid *Triticum* species. *Plant Science* 153: 81–91.

Gazza L, Nocente F, Ng PKW, Pogna NE. 2005. Genetic and biochemical analysis of common wheat cultivars lacking puroindoline a. *Theoretical and Applied Genetics* 110:470-478.

Gazza L, Taddei F, Conti S, Gazzelloni G, Muccilli V, Janni M, D'ovidio R, Alfieri M, Redaelli R, Pogna NE. 2015. Biochemical and molecular characterization of *Avena* indolines and their role in kernel texture. *Molecular Genetics and Genomics* 290(1): 39.

Gazza L, Taddei F, Corbellini M, Cacciatori P, Pogna NE. 2008. Genetic and environmental factors affecting grain texture in common wheat. Journal of Cereal Science 47(1): 52-58.

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. *Theoretical and Applied Genetics*109:1597-1603.

Geneix N, Dalgalarrondo M, Bakan B, Rolland-Sabaté A, Elmorjani K, Marion D. 2015. A single amino acid substitution in Puroindoline b impacts its self-assembly and the formation of heteromeric assemblies with puroindoline a. *Journal of Cereal Science*, 64:116-125.

Geng H, Beecher B S, He Z, Kiszonas AM, Morris CF. 2012. Prevalence of *Puroindoline D1* and *Puroindoline b-2* variants in U.S. Pacific Northwest wheat breeding germplasm pools, and their association with kernel texture. *Theoretical Applied Genetics* 124:1259-1269.

Geng H, Beecher BS, Pumphrey M, He Z, Morris CF. 2013. Segregation analysis indicates that *Puroindoline b-2* variants 2 and 3 are allelic in *Triticum aestivum* and that a revision to Puroindoline b-2 gene symbolization is indicated. *Journal of Cereal Science* 57:61-67.

Gillies S A, Futardo A, Henry R J. 2012. Gene expression in the developing aleurone and starchy endosperm of wheat. *Plant Biotechnology Journal* 10(6): 668-679.

Giroux MJ, Kim KH, Hogg AC, Martin JM, Beecher B. 2013 The *Puroindoline b-2* variants are expressed at low levels relative to the *Puroindoline D1* genes in wheat seeds. *Crop Science* 53 (3): 833-841.

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

Giroux MJ, Morris CF. 1998. Wheat grain hardness results from highly conserved mutations in the friabilin components Puroindoline a and b. *Proceedings in the National Academy of Sciences* 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 Science*, 40(2): 370-374.

Greenwell P, Schofield JD. 1986. A starch Granule protein associated with endosperm softness in wheat. *Cereal Chemistry* 63:379-380.

Guzmán C, Alvarez, JB. 2014. Molecular characterization of two novel alleles of *Hordoindoline* genes in *Hordeum chilense* Roem. et Schult. *Genetic resources and crop evolution* 61(2): 307-312.

Han F, Ullrich SE, Chirat S, Menteur S, Jestin L, Sarrafi O, Hayes PM, Jones BL, Blake TK, Wessenberg D, Kleinhofs A, Kilan A. 1995. Mapping of β-glucan and β-glucanase activity loci in barley and malt. *Theoretical and Applied Genetics* 91: 921-927.

Heinze K, Kiszonas A M, Murray J C, Morris C F, Lullien-Pellerin V. 2016. *Puroindoline* genes introduced into durum wheat reduce milling energy and change milling behavior similar to soft common wheats. *Journal of Cereal Science* 71: 183-189.

Hogg AC, Sripo T, Beecher B, Martin JM, Giroux MJ. 2004. Wheat Puroindolines interact to form friabilin and control wheat grain hardness. *Theoretical and Applied Genetics* 108:1089-1097.

Holopainen URM, Wilhelmson A, Salmenkallio-Marttila M, Peltonen-Sainio P, Rajala A, Reinikainen P, Kotaviita E, Simolin H, Home S. 2005. Endosperm structure affects the malting quality of barley (*Hordeum vulgare* L.). *Journal of Agricultural and Food Chemistry* 53(18): 7279–7287.

Ikeda TM, Ohnishi N, Nagamine T, Oda S, Hisatomi T, Yano H. 2005. Identification of new puroindoline genotypes and their relationship to flour texture among wheat cultivars. *Journal of Cereal Science* 41:1-6.

Ikeda TM, Cong H, Suzuki T, Takata K. 2010. Identification of new *Pina* null mutations among Asian common wheat cultivars. *Journal of Cereal Science* 51:235-237.

Iwami A, Osborne BG, Huynh HN, Anderssen RS, Wesley IJ, Kajiwara Y, Takashita H, Omori T. 2005. The Measurement of Structural Characteristics of Barley for Shochu Using Single-Kernel Characterization System 4100 Crush-Response Profiles. *Journal of the Institute of Brewing* 111(2):181-189.

Jagtap SS, Beardsley JMS, Forrest, JMS, Ellis, RP. 1993. Protein composition and grain quality in barley. *Aspects of Applied Biology (*United Kingdom).

Jang JY, Kim DG, Kim YO, Kim JS, Kang H. 2004 An expression analysis of a gene family encoding plasma membrane aquaporins in response to abiotic stresses in Arabidopsis thaliana. *Plant Molecular Biology* 54: 713-725.

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. *Journal of Bacteriology* 185: 4938-4947.

Jolly CJ, Rahman S, Kortt AA, Higgins TJV. 1993. Characterisation of the wheat Mr 15,000 'grain-softness protein' and analysis of the relationship between its accumulation and the whole seed and grain softness. *Theoretical Applied Genetics* 86: 589–597.

Inoue H, Nojima H, Okayama H. 1990. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* 96:23-28.

Kerckhoffs DA, Brouns F, Hornstra G, Mensink RP. 2002. Effect on the human lipoprotein profile of β-glucan, soy protein and isoflavones, plant sterols, garlic and tocotrienols. *The Journal of Nutrition* 32: 2494-2505.

Kim KH, Feiz L, Martin JM, Giroux MJ. 2012. Puroindolines are associated with decreased polar lipid breakdown during wheat seed development. *Journal of Cereal Science* 56:142-146.

Kooijman M, Orsel R, Hessing M, Hamer RJ, Bekkers ACAPA. 1997. Spectroscopic characterisation of the lipid-binding properties of wheat puroindolines. *Journal of Cereal Science* 26:145-159.

Kuchel H, Williams JK, Langridge P, Eagles HA, Jefferies, SP. 2007. Genetic dissection of grain yield in bread wheat. I. QTL analysis. *Theoretical and Applied Genetics* 115, 1029-1041.

Kumar R, Arora S, Singh K, Garg M. 2015. *Puroindoline* allelic diversity in Indian wheat germplasm and identification of new allelic variants. *Breeding Science* 65(4): 319-326.

Krishnamurthy K, Giroux MJ. 2001. Expression of wheat puroindoline genes in transgenic rice enhances grain softness. *Nature Biotechnology* 19:162-166.

Law CN, Young CF, Brown JWS, Snape JW, Worland AJ. 1978. The study of grain protein control in wheat using whole chromosome substitution lines. In: Seed Protein Improvment by Nuclear Techniques. Interanational Atomic Energy Agency. Vienna, Austria. p 483-502.

Lee MS, Jang CS, Lee SS, Kim JY, Lee B, Seong R C, Seo YW. 2006. Hordoindolines are predominantly expressed in the aleurone layer in late kernel development in barley. *Breeding Science* 56: 63-68.

Li GY, He ZH, Lillemo M, Sun QX, Xia XC. 2008. Molecular characterization of allelic variations at *Pina* and *Pinb* loci in Shandong wheat landraces, historical and current cultivars. *Journal of Cereal Science* 47: 510-517.

Li W, Jiang Q, Chen G, Pu Z, Liu Y, Wang J, Zheng Y, Wei Y. 2011. Comparative analysis of *Hina* gene sequences in wild *(Hordeum spontaneum*) and cultivated (*H. vulgare*) Barleys. *Agricultural Sciences in China* 10: 1313-1322.

Li Y, Mao X, Wang Q, Zhang J, Li X, Ma F, Sun F, Chang J, Chen M, Wang Y, Li K. 2014. Overexpression of *Puroindoline a* gene in transgenic durum wheat (*Triticum turgidum* ssp. *durum*) leads to a medium–hard kernel texture. *Molecular breeding*, 33(3): 545-554.

Lillemo M, Morris CF. 2000. A leucine to proline mutation in puroindoline b is frequently present in hard wheats from Northern Europe. *Theoretical and Applied Genetics* 100:1100-1107.

Lillemo M, Ringlund K. 2002. Impact of puroindoline b alleles on the genetic variation for hardness in soft x hard wheat crosses. *Plant Breeding* 121: 210-217.

Mather D E, Tinker N A, LaBerge D E, Edney M, Jones B L, Rossnagel B G, Legge W G, Briggs, K. G., Irvine R. B, Falk D. E (1997) Regions of the genome that affect grain and malt quality in a North American two-row barley cross. *Crop Science* 37:544-554.

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

Mattern PJ, Morris R, Schmidt JW, Johnson VA. 1973. Locations of genes for kernel properties in the wheat variety "Cheyenne" using chromosome substitution lines. In: Sears, E.R., Sears, L.M.S., eds. Proceedings of the 4th international wheat genetics symposium, Missouri Agricultural Experiment Station, Columbia, p 703–707.

Mayolle JE, Lullien-Pellerin V, Corbineau F, Boivin P, Guillard V 2012. Water diffusion and enzyme activities during malting of barley grains: a relationship assessment. *Journal of Food Engineering* 109(3): 358–365.

McIntyre CL, Mathews KL, Rattey A, Chapman S C, Drenth J, Ghaderi M, Reynolds M, Shorter R. 2010. Molecular detection of genomic regions associated with grain yield and yield-related components in an elite bread wheat cross evaluated under irrigated and rainfed conditions. *Theoretical and Applied Genetics* 120: 527-541.

Miao Y, Chen L, Wang C, Wang Y, Zheng Q, Gao C, Yang G, He G. 2012. Expression, purification and antimicrobial activity of puroindoline A protein and its mutants. *Amino Acids*:1-8.

Mikulìková D. 2007. The effect of friabilin on wheat grain hardness. *Czech Journal of Genetics and Plant Breeding* 43:35-43.

Mohammadi M, Endelman J, Nair S, Chao S, Jones S, Muehlbauer G, Ullrich S, Baik BK, Wise M, Smith K. 2014. Association mapping of grain hardness, polyphenol oxidase, total phenolics, amylose content, and β-glucan in US barley breeding germplasm. *Molecular breeding* 34(3).

Mohammadi M, Torkamaneh D, Hashemi M, Mehrabi R, Ebrahimi A. 2012. Fast and inexpensive DNA isolation from wheat (Triticum aestivum) and other small grains. Wheat Information Service. Electronic Newsletter for Wheat Researchers. 114: p17-20. (http://www.shigen.nig.ac.jp/ewis/article/html/118/article.html)

Mohler V, Schmolke M, Paladey E, Seling S, and Hartl L. 2012. Association analysis of *Puroindoline –D1* and *Puroindoline b-2* loci with 13 quality traits in European winter wheat (*Triticum aestivum* L.). *Journal of Cereal Science* 56: 623-628.

Moreira IS, Fernandes PA, Ramos MJ. 2007. Hot spots - A review of the protein-protein interface determinant amino-acid residues. *Proteins: Structure, Function and Genetics* 68:803-812.

Morris CF. 2002. Puroindolines: the molecular genetic basis of wheat grain hardness. *Plant Molecular Biology* 48:633-647.

Morris CF, Greenblatt GA, Bettge AD, Malkawi HI. 1994. Isolation and characterisation of multiple forms of friabilin. *Journal of Cereal Science* 21:167-174.

Morris CF, Lillemo M, Simeone MC. 2001. Prevalence of puroindoline grain hardness genotypes among historically significant North American spring and winter wheats. *Crop Science* 41:218-228.

Morrison WR, Greenwell P, Law CN, Sulaiman BD. 1992. Occurrence of friabilin, a low-molecular-weight protein associated with grain softness, on starch granules isolated from some wheats and related species. *Journal of Cereal Science* 15: 143-149.

Nagamine T, Sekiwa T, Yamaguchi E, Oozeki MK, Kato M. 2009. Relationship between quality parameters and SKCS hardness index in malting barley. *Journal of the Institute of Brewing* 115: 292–295.

Nair S, Knoblauch M, Ullrich S, Baik B-K. 2011. Microstructure of hard and soft kernels of barley. *Journal of Cereal Science* 54(3): 354–362.

Newman RK, Newman CW. 2008. Barley for Food and Health: Science, Technology, and Products, John Wiley and sons, New Jersey, USA, pp 56-80.

Nirmal R C, Furtado A, Wrigley C, Henry R J 2016. Influence of gene expression on hardness in wheat. *PloS one* 11(10): e0164746.

Oda S, Schofield JD. 1997. Characterisation of friabilin polypeptides. *Journal of Cereal Science* 26:29-36.

Osborne BG, Turnbull KM, Anderssen RS, Rahman S, Sharp PJ, Appels R. 2001. The *Hardness* locus in Australian wheat lines. *Crop and Pasture Science*, 52(12):1275- 1286.

Palmer GH. 1989. Cereals in malting and brewing. Cereal Science and Technology, G. H. Palmer, Ed., University Press: Aberdeen. 180-194.

Pan Z, Song W, Meng F, Xu L, Liu B, Zhu J. 2004. Characterisation of genes encoding wheat grain hardness form chinese cultivar GoaCheng 8901. *Cereal Chemistry* 81:287-289.

Panozzo JF, Eckermann PJ, Mather DE, Moody DB, Black CK, Collins HM, Barr AR, Lim P Cullis BR. 2007. QTL analysis of malting quality traits in two barley populations. *Australian Journal of Agricultural Research* 58(9): 858-866.

Park R, Wellings C, Bariana H, Bansal U. 2009. Australia cereal cultivar pedigree and seedling rust genotype information. *Cereal Rust Report Season* 7(2).

Pauly A, Pareyt B, Fierens E Delcour JA. 2013. Wheat (*Triticum aestivum* L. and *T. turgidum* L. *ssp. durum*) kernel hardness: I. Current view on the role of puroindolines and polar lipids. *Comprehensive Reviews in Food Science and Food Safety* 12(4): 413-426.

Phillips RL, Palombo EA, Panozzo JF Bhave M. 2011. Puroindolines, *Pin* alleles, Hordoindolines and Grain Softness Proteins are sources of bactericidal and fungicidal peptides. *Journal of cereal science* 53(1): 112-117.

Pickering PA, Bhave M. 2007. Comprehensive analysis of Australian hard wheat cultivars shows limited *puroindoline* allele diversity. *Plant Science* 172:371-379.

Pilen K, Zacharis, A, Leon J. 2003. Advanced backcross QTL analysis in barley (*Hordeum vulgare* L.). *Theoretical Applied Genetics* 107: 340-352.

Psota V, Vejrazka K, Famera O, Hrcka M. 2007. Relationship between grain hardness and malting quality of barley (*Hordeum vulgare* L.). *Journal of the Institute of Brewing* 113(1): 80–86.

Qamar ZU, Bansal UK, Dong CM, Alfred RL, Bhave M, Bariana HS. 2014. Detection of puroindoline (*Pina-D1* and *Pinb-D1*) allelic variation in wheat landraces. *Journal of Cereal Science* 60(3): 610-616.

Quarrie SA, Steed A, Calestani C, Semikbodskii A, Lebreton C, Chinoy C, Steele N, Pljevlajakusic D, Habash D Z, Farmer P, Saker L, Clarkson DT, Abugalieva A, Yessimbekova M, Turuspekov Y, Abugalieva S, Tuberosa R, Sanguineti MC, Hollington P A, Aragues, R, Royo A, Dodig D. 2005. A high density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring_ SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theoretical Applied Genetics* 110, 865-880.

Qureshi AA, Qureshi N, Wright JJK, Shen Z, Kramer, G, Gapor A, Chong, Y.H., Dewitt, G., Ong, A.S.H., Peterson, D.M., Bradlow, B.A. 1991. Lowering serum cholesterol in hypercholesterolemic humans by tocotrienols (palmvitee). *American Journal of Clinical Nutrition* 53, 1021–1026.

Rahman S, Jolly CJ, Skerritt JH, Wallosheck A. 1994. Cloning of a wheat 15-kDa grain softness protein (GSP) - GSP is a mixture of Puroindoline-like polypeptides. *European Journal of Biochemistry* 223:917-925.

Ram S, Jain N, Shoran J, Singh R. 2005. New frame shift mutation in puroindoline B in Indian wheat cultivars Hyb65 and NI5439. *Journal of Plant Biochemistry and Biotechnology* 14:45-48.

Ramalingam A. 2012. Genetic variability and protein-protein interactions of puroindolines in relation to wheat grain texture and antimicrobial properties. In: Faculty of Life and Social Science. Melbourne, Australia: Swinburne University of Technology.

Ramalingam A, Palombo EA, and Bhave M. 2012. The *Pinb-2* genes in wheat comprise a multigene family with great sequence diversity and important variants. *Journal of Cereal Science* 56: 171-180.

Rosicka-Kaczmarek J, Stasiuk M, Nebesny E, Komisarczyk A. 2015. Fluorimetric studies of the interactions of wheat puroindolines with polar lipids on the surface starch granules. *Journal of Cereal Science*, 66: 53-58.

Rouve´s, S., Boef, C., Zwickert-Menteur, S., Gautier, M.F., Bernard, M., Joudrier, P., Jestin, L. 1996. Locating supplementary RFLP markers on barley chromosome 7 and synteny with homeologous wheat group 5. *Plant Breeding* 115, 511–513.

Salamini F, Özkan H, Brandolini A, Schäfer-Pregl R, Martin W 2002. Genetics and geography of wild cereal domestication in the near east. *Nature Reviews Genetics* 3: 429-441.

Sambrook J, Russell DW. 2001. Molecular cloning: a laboratory manual, 3rd ed: Cold Spring Harbor, N.Y.

Schwartz R, Clark AG, Istrail S. 2004. Inferring piecewise ancestral history from haploid sequences. *Lecture Notes in Bioinformatics* 2983: 62-73.

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. *Journal of Cereal Science* 44: 21-33.

Singh A, Mantri S, Sharma M, Chaudhury A, Tuli R, Roy J. 2014. Genome-wide transcriptome study in wheat identified candidate genes related to processing quality, majority of them showing interaction (quality x development) and having temporal and spatial distributions. *BMC Genomics* 15(1): 29.

Shagaghi N, Alfred RL, Clayton AH, Palombo EA Bhave M. 2016. Anti-biofilm and sporicidal activity of peptides based on wheat Puroindoline and barley Hordoindoline proteins. *Journal of Peptide Science* 22(7): 492-500.

Suprunova T, Krugman T, Fahima T, Chen G, Shams I, Korol A, Nevo E. 2004 Differential expression of dehydrin genes in wild barley, *Hordeum spontaneum*, associated with resistance to water deficit. *Plant Cell Environment* 27: 1297-1308.

Swanston JS. 1995. Effects on barley grain size, texture and modification during malting associated with three genes on chromosome 1. *Journal of Cereal Science* 22: 157–161.

Symes KJ. 1965. The inheritance of grain hardness in wheat as measured by the particle size index. *Australian Journal of Agricultural Research* 16:113-123.

Tanaka H, Morris CF, Haruna M, Tsujimoto H. 2008. Prevalence of puroindoline alleles in wheat varieties from eastern Asia including the discovery of a new SNP in *Puroindoline b*. *Plant Genetic Resources* 6:142-152.

Tanchak MA, Schernthaner JP, Giband M, Altosaar I. 1998. Tryptophanins: isolation and molecular characterization of oat cDNA clones encoding proteins structurally related to Puroindoline and wheat grain softness proteins. *Plant Science* 137: 173-184.

Terasawa Y, Takata K, Anai T, Ikeda TM. 2012. Identification and distribution of *Puroindoline b-2* variant gene homologs in *Hordeum*. *Genetica* 141: 359-368.

Thomas W T B, Powel W, Swanston J S, Ellis R P, Chamlers K J, Barua UM, Jack P, Lea V, Forster B P, Waugh R. 1996. Quantitative trait loci for germination and malting quality characters in a spring barley cross. *Crop Science* 36: 265-273.

Tranquilli G, Heaton J, Chicaiza O, Dubcovsky J. 2002. Substitutions and deletions of genes related to grain hardness in wheat and their effect on grain texture. *Crop Science* 42:1812-1817.

Tranquilli G, Lijavetzky D, Muzzi G, Dubcovsky J. 1999. Genetic and physical characterisation of grain texture-related loci in diploid wheat. *Molecular and General Genetics* 262:846-850.

Turnbull KM, Rahman S. 2002. Endosperm texture in wheat. *Journal of Cereal Science* 36 (3): 327-337.

Turuspekov Y, Beecher B, Darlington Y, Bowman J, Blake TK, Giroux MJ. 2008. *Hardness* locus sequence variation and endosperm texture in spring barley. *Crop Science* 48:1007–1019.

Vejrazka K, Psota V, Ehrenbergerova J, Hrstkova P. 2008. Relationship between grain milling energy and malting quality of barley. *Cereal Research Communication* 36: 97-105.

Walker CK, Panozzo JF, Ford R, Eckermann P, Moody D, Lehmensiek A, Appels R. 2011. Chromosomal loci associated with endosperm hardness in a malting barley cross. *Theoretical and Applied Genetics* 122(1): 151-162.

Walker, C.K., Ford, R., Munoz-Amatrian M., Panozzo, J.F. 2013. The detection of QTLs in barley associated with endosperm hardness, grain size and malting quality using rapid phenotyping tools. *Theoretical Applied Genetics* 126: 2533-2551.

Wang L, Li G, Xia X, He Z, Mu P. 2008. Molecular characterisation of *Pina* and *Pinb* allelic variations in Xinjiang landraces and commercial wheat cultivars. *Euphytica* 164(3): 745-752.

Wanjugi HW, Hogg AC, Martin JM, Giroux MJ. 2007. The role of Puroindoline a and b individually and in combination on grain hardness and starch association. *Crop Science* 47:67-76.

Wilkinson M, Wan Y, Tosi P, Leverington M, Snape J, Mitchell RAC, Shewry PR. 2008. Identification and genetic mapping of variant forms of *Puroindoline b* expressed in developing wheat grain. *Journal of Cereal Science* 48:722-728.

Xia L, Chen F, Chen X, Morris CF. 2005. Occurrence of Puroindoline alleles in Chinese winter wheats. *Cereal Chemistry* 38:38-43.

Yamazaki WT. 1972. A modified particle-size index test for kernel texture in soft wheat. *Crop Science*. 12: 116.

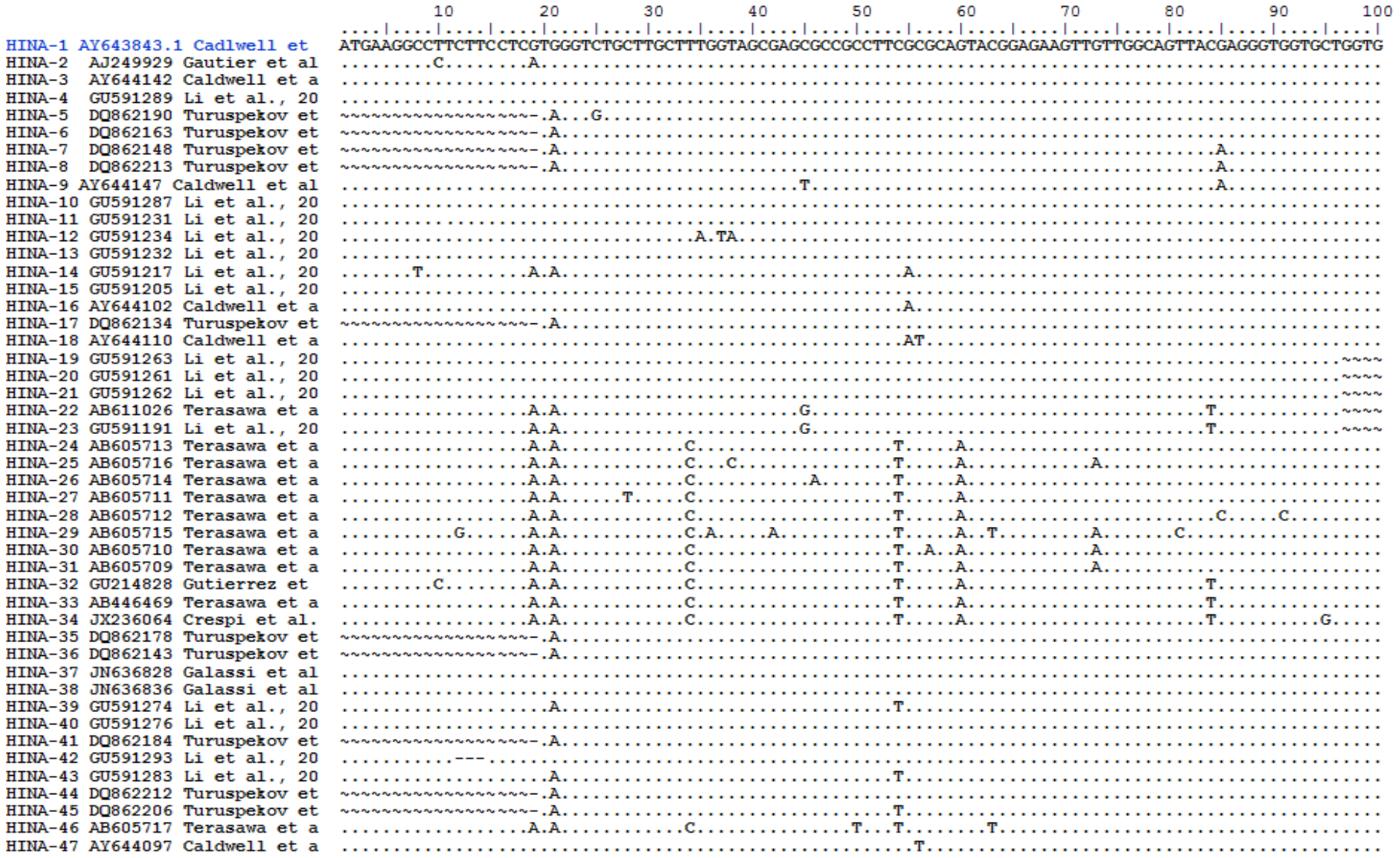
Yanaka, M, Takata K, Terasawa Y, Ikeda TM. 2011. Chromosome 5H of *Hordeum* species involved in reduction in grain hardness in wheat genetic background. *Theoretical and Applied Genetics* 123(6): 1013-1018.

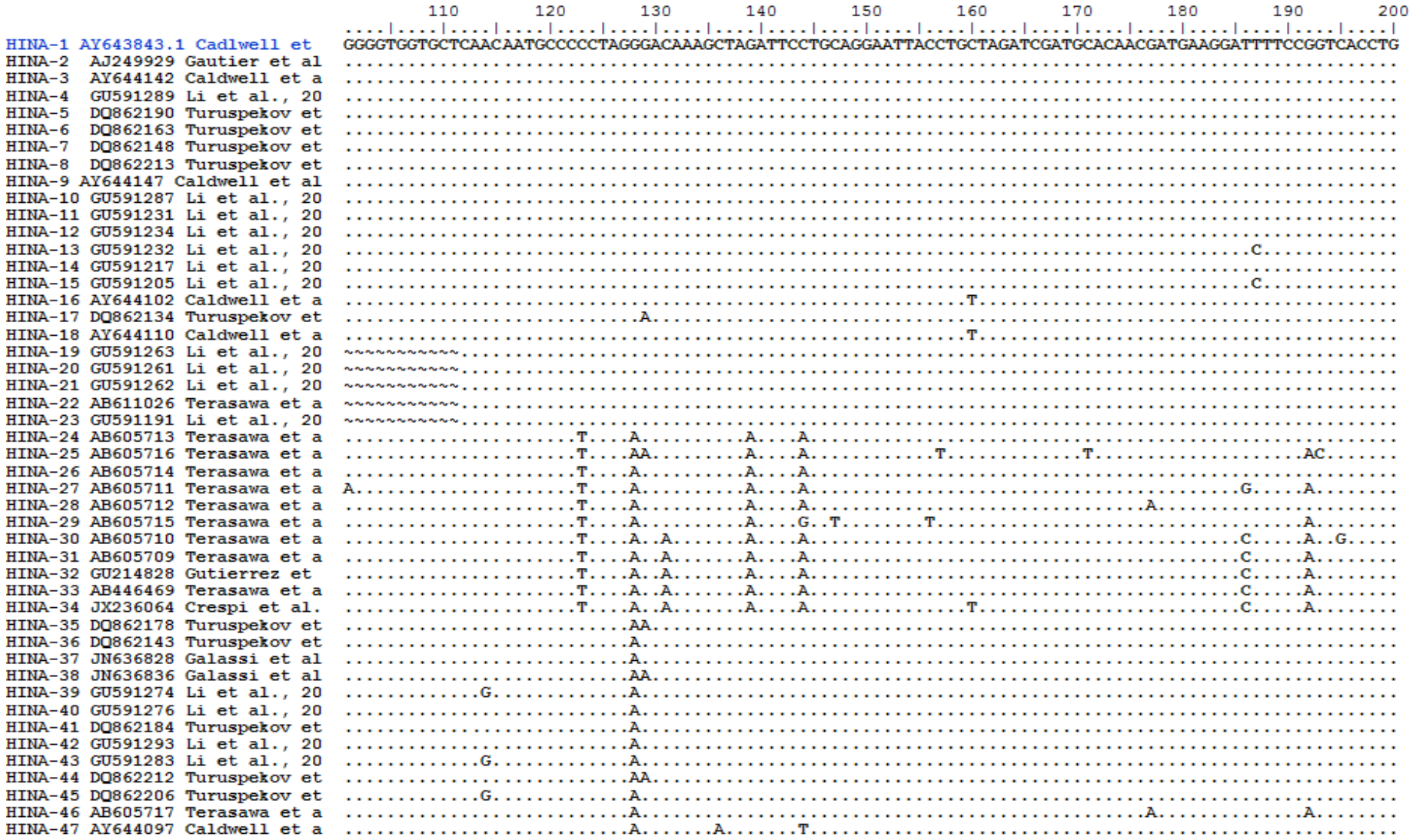
Zhang J, Martin JM, Beecher B, Morris CF, Curtis Hannah L, Giroux MJ. 2009. Seed-specific expression of the wheat puroindoline genes improves maize wet milling yields. *Plant Biotechnology Journal* 7(8): 733-743.

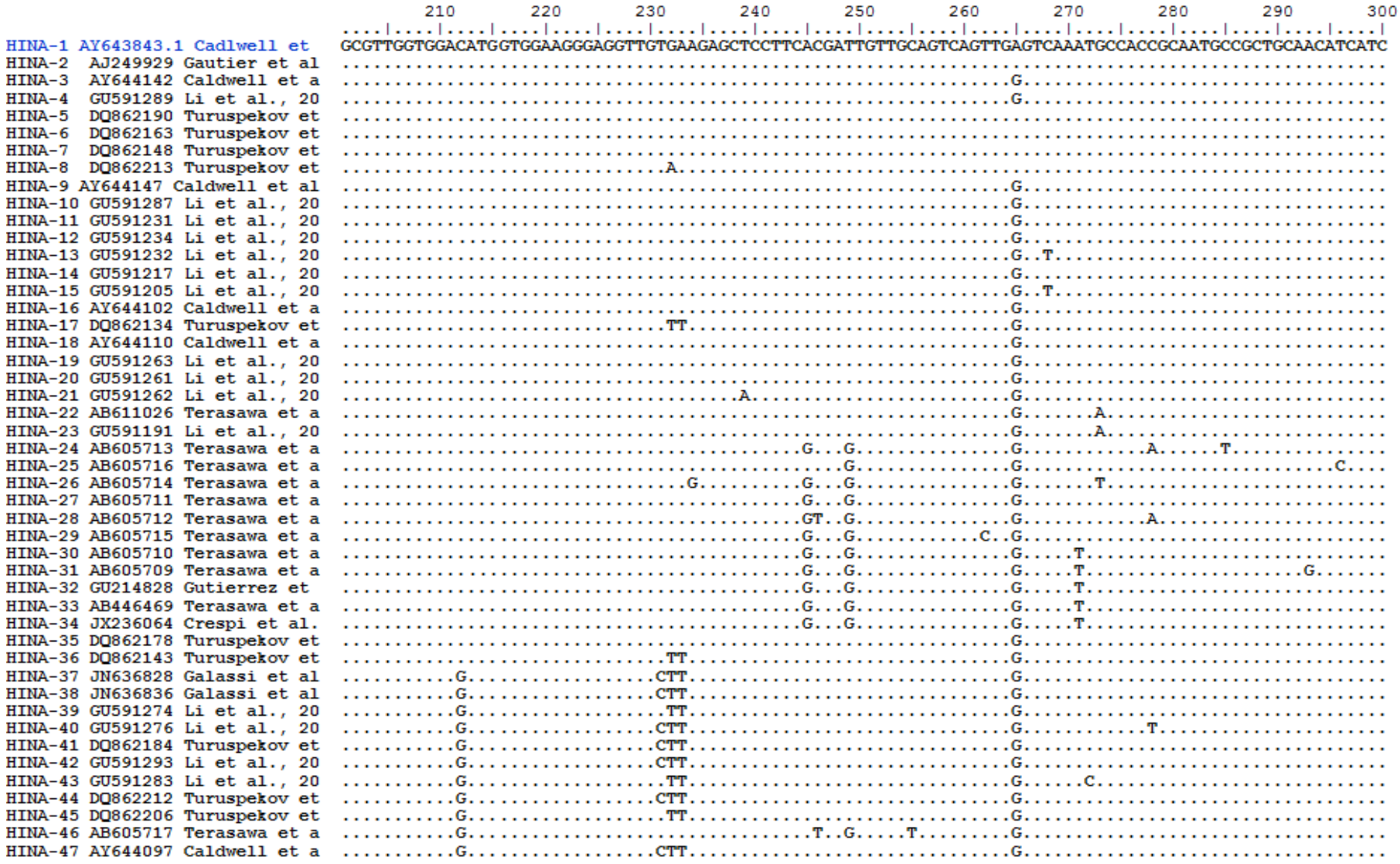
Zohary D, Hopf M. 2000. Domestication of plants in the old world: the origin and spread of cultivated plants in West Asia, Europe, and the Nile Valley, 3rd edn. New York, Oxford University Press.

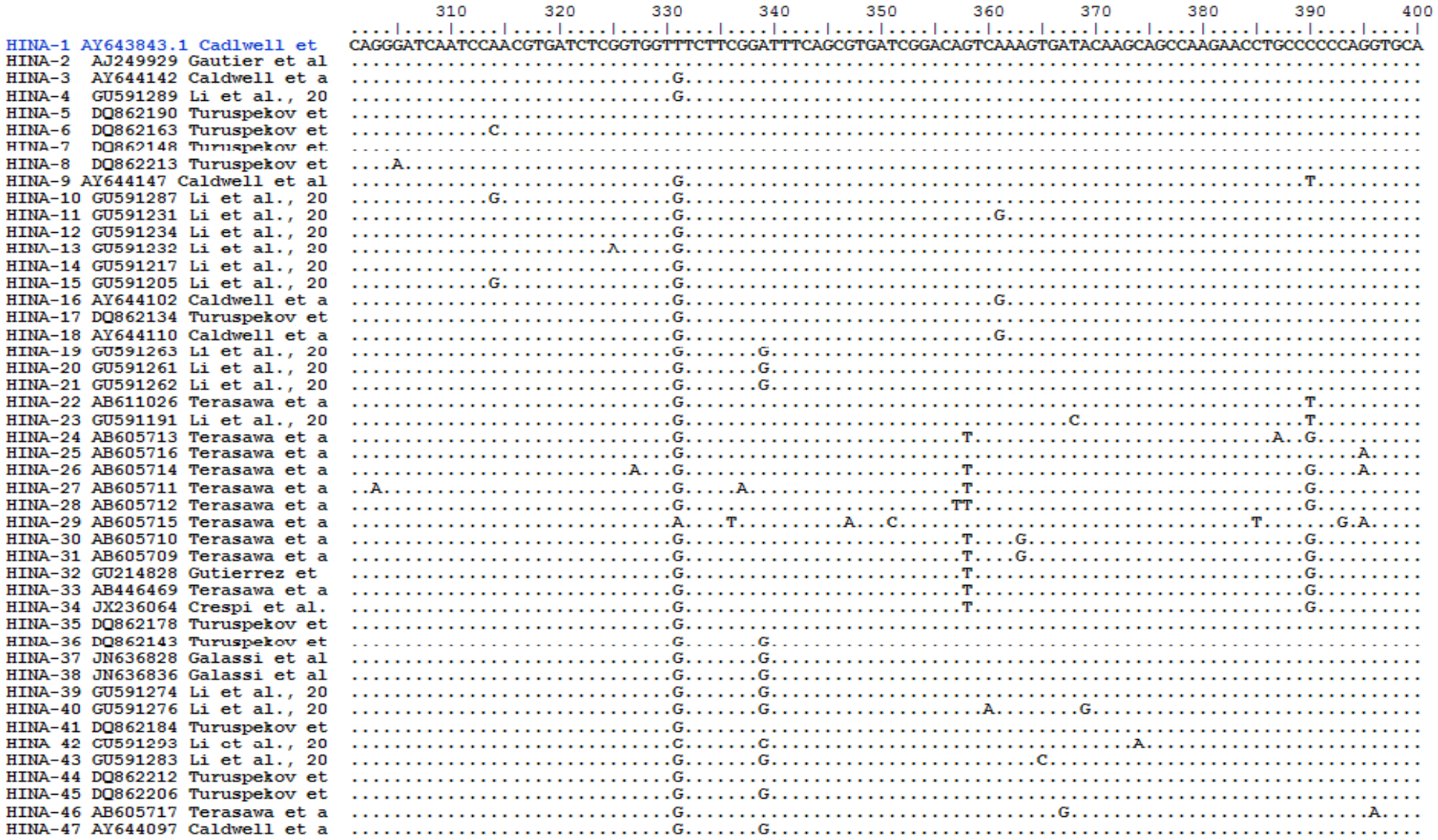
# Appendices

Appendix I: DNA and putative protein alignments for HIN sequences (Chapter 1)




110 120 130 140 150 160 170 180 190 200
HINA-1 AY643843.1 Cadlwell et GGGGTGGTGCTCAACAATGCCCCCTAGGGACAAAGCTAGATTCCTGCAGGAATTACCTGCTAGATCGATGCACAACGATGAAGGATTTTCCGGTCACCTG
HINA-2 AJ249929 Gautier et al
HINA-3 AY644142 Caldwell et a
HINA-4 GU591289 Li et al., 20
HINA-5 DQ862190 Turuspekov et
HINA-6 DQ862163 Turuspekov et
HINA-7 DQ862148 Turuspekov et
HINA-8 DQ862213 Turuspekov et
HINA-9 AY644147 Caldwell et al
HINA-10 GU591287 Li et al., 20
HINA-11 GU591231 Li et al., 20
HINA-12 GU591234 Li et al., 20
HINA-13 GU591232 Li et al., 20
HINA-14 GU591217 Li et al., 20
HINA-15 GU591205 Li et al., 20
HINA-16 AY644102 Caldwell et a
HINA-17 DQ862134 Turuspekov et
HINA-18 AY644110 Caldwell et a
HINA-19 GU591263 Li et al., 20
HINA-20 GU591261 Li et al., 20
HINA-21 GU591262 Li et al., 20
HINA-22 AB611026 Terasawa et a
HINA-23 GU591191 Li et al., 20
HINA-24 AB605713 Terasawa et a
HINA-25 AB605716 Terasawa et a
HINA-26 AB605714 Terasawa et a
HINA-27 AB605711 Terasawa et a
HINA-28 AB605712 Terasawa et a
HINA-29 AB605715 Terasawa et a
HINA-30 AB605710 Terasawa et a
HINA-31 AB605709 Terasawa et a
HINA-32 GU214828 Gutierrez et
HINA-33 AB446469 Terasawa et a
HINA-34 JX236064 Crespi et al.
HINA-35 DQ862178 Turuspekov et
HINA-36 DQ862143 Turuspekov et
HINA-37 JN636828 Galassi et al
HINA-38 JN636836 Galassi et al
HINA-39 GU591274 Li et al., 20
HINA-40 GU591276 Li et al., 20
HINA-41 DQ862184 Turuspekov et
HINA-42 GU591293 Li et al., 20
HINA-43 GU591283 Li et al., 20
HINA-44 DQ862212 Turuspekov et
HINA-45 DQ862206 Turuspekov et
HINA-46 AB605717 Terasawa et a
HINA-47 AY644097 Caldwell et a


210 220 230 240 250 260 270 280 290 300
HINA-1 AY643843.1 Cadlwell et GCGTTGGTGGACATGGTGGAAGGGAGGTTGTGAAGAGCTCCTTCACGATTGTTGCAGTCAGTTGAGTCAAATGCCACCGCAATGCCGCTGCAACATCATC
HINA-2 AJ249929 Gautier et al
HINA-3 AY644142 Caldwell et a
HINA-4 GU591289 Li et al., 20
HINA-5 DQ862190 Turuspekov et
HINA-6 DQ862163 Turuspekov et
HINA-7 DQ862148 Turuspekov et
HINA-8 DQ862213 Turuspekov et
HINA-9 AY644147 Caldwell et al
HINA-10 GU591287 Li et al., 20
HINA-11 GU591231 Li et al., 20
HINA-12 GU591234 Li et al., 20
HINA-13 GU591232 Li et al., 20
HINA-14 GU591217 Li et al., 20
HINA-15 GU591205 Li et al., 20
HINA-16 AY644102 Caldwell et a
HINA-17 DQ862134 Turuspekov et
HINA-18 AY644110 Caldwell et a
HINA-19 GU591263 Li et al., 20
HINA-20 GU591261 Li et al., 20
HINA-21 GU591262 Li et al., 20
HINA-22 AB611026 Terasawa et a
HINA-23 GU591191 Li et al., 20
HINA-24 AB605713 Terasawa et a
HINA-25 AB605716 Terasawa et a
HINA-26 AB605714 Terasawa et a
HINA-27 AB605711 Terasawa et a
HINA-28 AB605712 Terasawa et a
HINA-29 AB605715 Terasawa et a
HINA-30 AB605710 Terasawa et a
HINA-31 AB605709 Terasawa et a
HINA-32 GU214828 Gutierrez et
HINA-33 AB446469 Terasawa et a
HINA-34 JX236064 Crespi et al.
HINA-35 DQ862178 Turuspekov et
HINA-36 DQ862143 Turuspekov et
HINA-37 JN636828 Galassi et al
HINA-38 JN636836 Galassi et al
HINA-39 GU591274 Li et al., 20
HINA-40 GU591276 Li et al., 20
HINA-41 DQ862184 Turuspekov et
HINA-42 GU591293 Li et al., 20
HINA-43 GU591283 Li et al., 20
HINA-44 DQ862212 Turuspekov et
HINA-45 DQ862206 Turuspekov et
HINA-46 AB605717 Terasawa et a
HINA-47 AY644097 Caldwell et a


310 320 330 340 350 360 370 380 390 400
HINA-1 AY643843.1 Cadlwell et CAGGGATCAATCCAACGTGATCTCGGTGGTTTCTTCGGATTTCAGCGTGATCGGACAGTCAAAGTGATACAAGCAGCCAAGAACCTGCCCCCCAGGTGCA
HINA-2 AJ249929 Gautier et al
HINA-3 AY644142 Caldwell et a
HINA-4 GU591289 Li et al., 20
HINA-5 DQ862190 Turuspekov et
HINA-6 DQ862163 Turuspekov et
HINA-7 DQ862148 Turuspekov et
HINA-8 DQ862213 Turuspekov et
HINA-9 AY644147 Caldwell et al
HINA-10 GU591287 Li et al., 20
HINA-11 GU591231 Li et al., 20
HINA-12 GU591234 Li et al., 20
HINA-13 GU591232 Li et al., 20
HINA-14 GU591217 Li et al., 20
HINA-15 GU591205 Li et al., 20
HINA-16 AY644102 Caldwell et a
HINA-17 DQ862134 Turuspekov et
HINA-18 AY644110 Caldwell et a
HINA-19 GU591263 Li et al., 20
HINA-20 GU591261 Li et al., 20
HINA-21 GU591262 Li et al., 20
HINA-22 AB611026 Terasawa et a
HINA-23 GU591191 Li et al., 20
HINA-24 AB605713 Terasawa et a
HINA-25 AB605716 Terasawa et a
HINA-26 AB605714 Terasawa et a
HINA-27 AB605711 Terasawa et a
HINA-28 AB605712 Terasawa et a
HINA-29 AB605715 Terasawa et a
HINA-30 AB605710 Terasawa et a
HINA-31 AB605709 Terasawa et a
HINA-32 GU214828 Gutierrez et
HINA-33 AB446469 Terasawa et a
HINA-34 JX236064 Crespi et al.
HINA-35 DQ862178 Turuspekov et
HINA-36 DQ862143 Turuspekov et
HINA-37 JN636828 Galassi et al
HINA-38 JN636836 Galassi et al
HINA-39 GU591274 Li et al., 20
HINA-40 GU591276 Li et al., 20
HINA-41 DQ862184 Turuspekov et
HINA-42 GU591293 Li et al., 20
HINA-43 GU591283 Li et al., 20
HINA-44 DQ862212 Turuspekov et
HINA-45 DQ862206 Turuspekov et
HINA-46 AB605717 Terasawa et a
HINA-47 AY644097 Caldwell et a

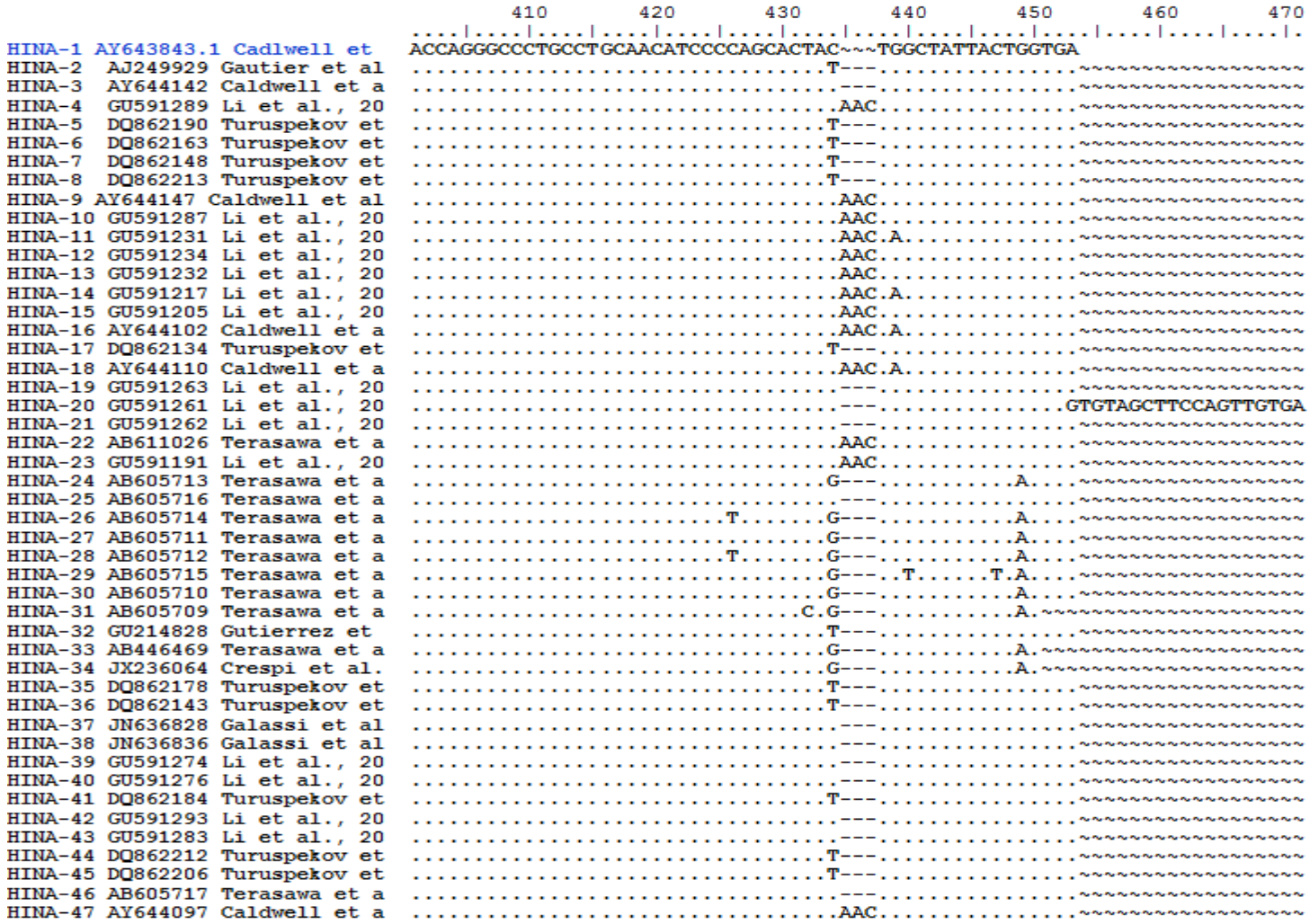



**Fig. I a: Alignment of *Hina* DNA sequences**

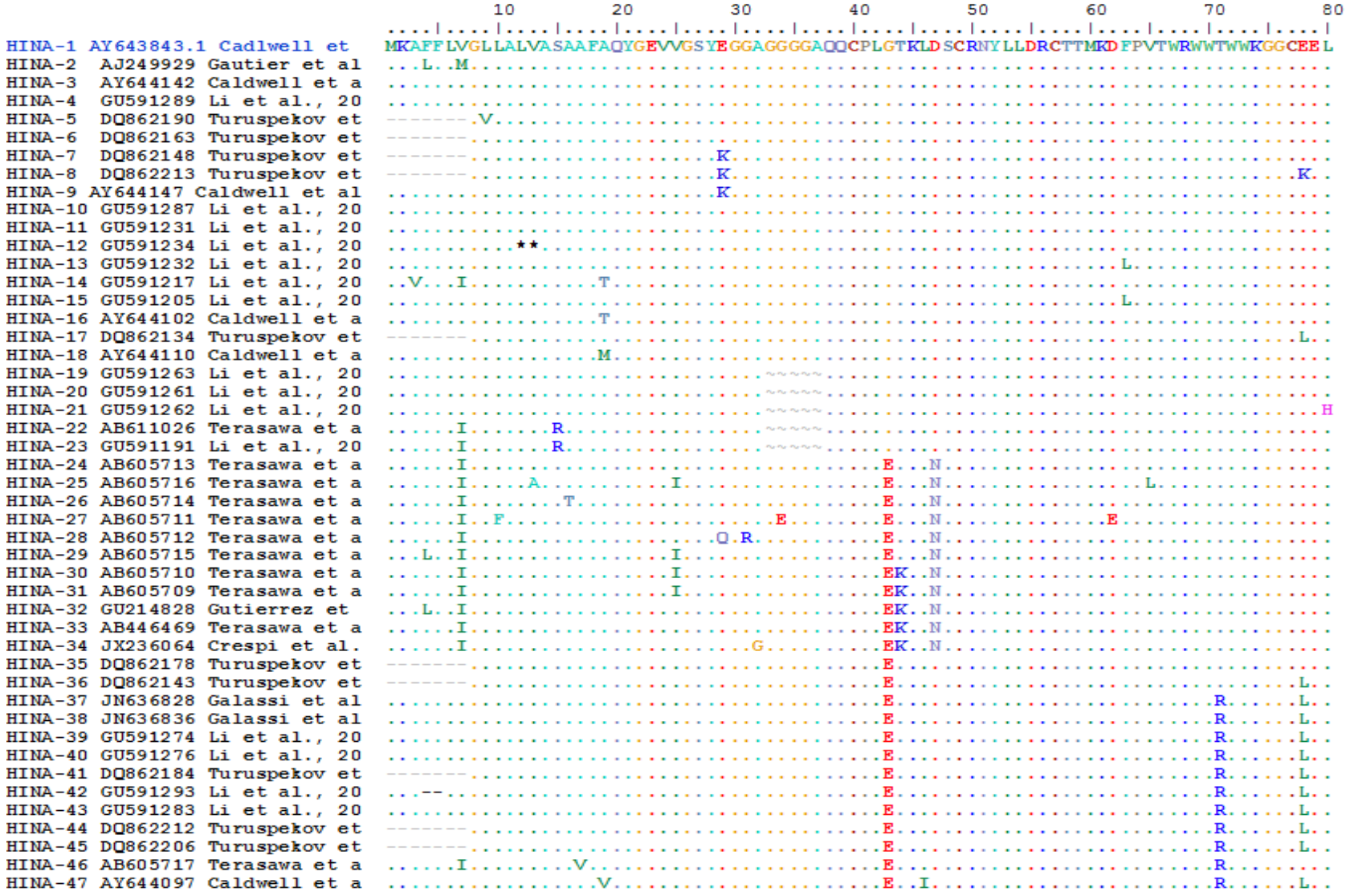
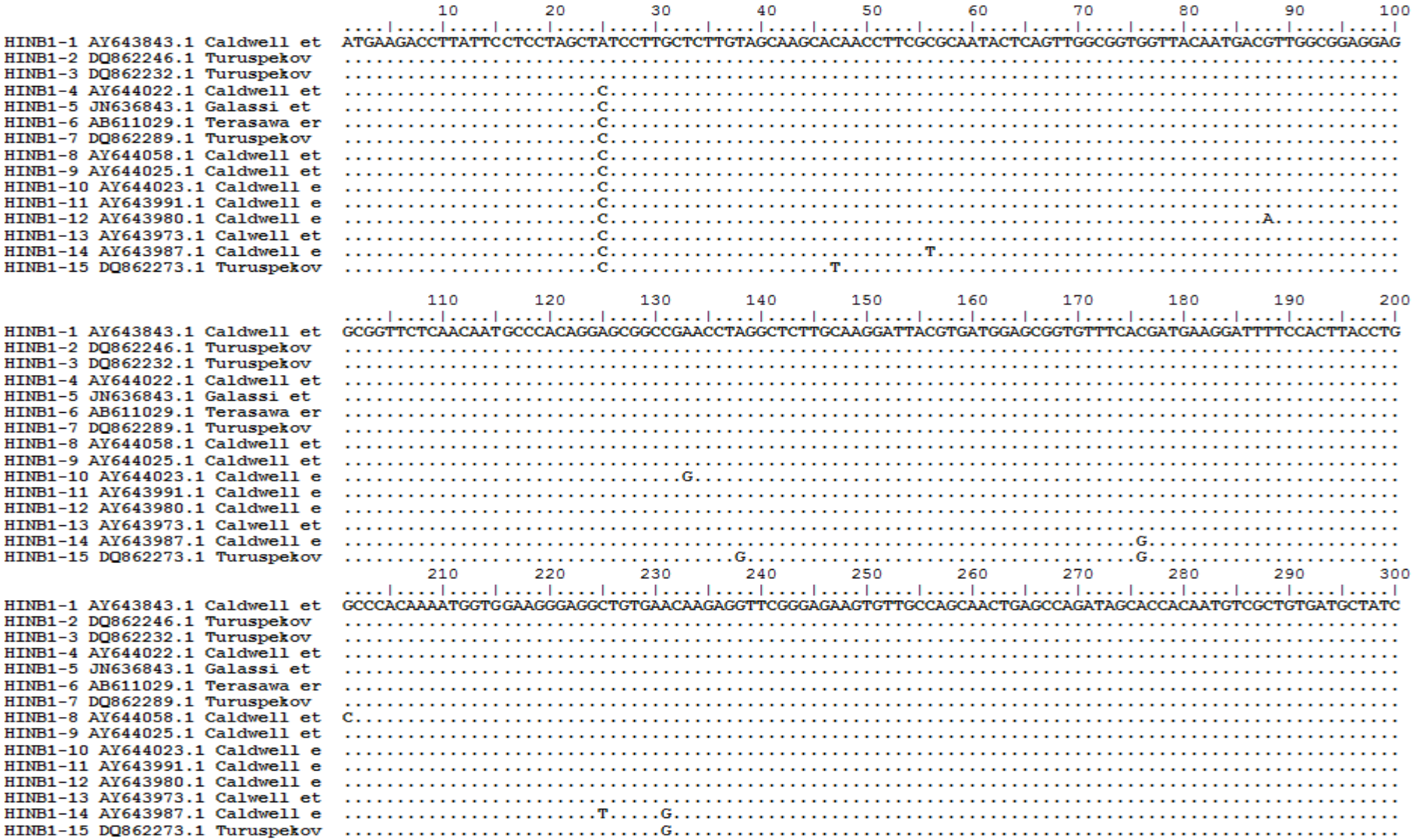

10 20 30 40 50 60 70 80
HINA-1 AY643843.1 Cadlwell et MKAFFLVGLLALVASAAFAQYGEVVGSYEGGAGGGGAQQCPLGTKLDSCRNYLLDRCTTMKDFPVTWRWWTWWKGGCEEL
HINA-2 AJ249929 Gautier et al
HINA-3 AY644142 Caldwell et a
HINA-4 GU591289 Li et al., 20
HINA-5 DQ862190 Turuspekov et
HINA-6 DQ862163 Turuspekov et
HINA-7 DQ862148 Turuspekov et
HINA-8 DQ862213 Turuspekov et
HINA-9 AY644147 Caldwell et al
HINA-10 GU591287 Li et al., 20
HINA-11 GU591231 Li et al., 20
HINA-12 GU591234 Li et al., 20
HINA-13 GU591232 Li et al., 20
HINA-14 GU591217 Li et al., 20
HINA-15 GU591205 Li et al., 20
HINA-16 AY644102 Caldwell et a
HINA-17 DQ862134 Turuspekov et
HINA-18 AY644110 Caldwell et a
HINA-19 GU591263 Li et al., 20
HINA-20 GU591261 Li et al., 20
HINA-21 GU591262 Li et al., 20
HINA-22 AB611026 Terasawa et a
HINA-23 GU591191 Li et al., 20
HINA-24 AB605713 Terasawa et a
HINA-25 AB605716 Terasawa et a
HINA-26 AB605714 Terasawa et a
HINA-27 AB605711 Terasawa et a
HINA-28 AB605712 Terasawa et a
HINA-29 AB605715 Terasawa et a
HINA-30 AB605710 Terasawa et a
HINA-31 AB605709 Terasawa et a
HINA-32 GU214828 Gutierrez et
HINA-33 AB446469 Terasawa et a
HINA-34 JX236064 Crespi et al.
HINA-35 DQ862178 Turuspekov et
HINA-36 DQ862143 Turuspekov et
HINA-37 JN636828 Galassi et al
HINA-38 JN636836 Galassi et al
HINA-39 GU591274 Li et al., 20
HINA-40 GU591276 Li et al., 20
HINA-41 DQ862184 Turuspekov et
HINA-42 GU591293 Li et al., 20
HINA-43 GU591283 Li et al., 20
HINA-44 DQ862212 Turuspekov et
HINA-45 DQ862206 Turuspekov et
HINA-46 AB605717 Terasawa et a
HINA-47 AY644097 Caldwell et a


**Fig. I b: Alignment of HINA putative protein sequences**


10 20 30 40 50 60 70 80 90 100
HINB1-1 AY643843.1 Caldwell et ATGAAGACCTTATTCCTCCTAGCTATCCTTGCTCTTGTAGCAAGCACAACCTTCGCGCAATACTCAGTTGGCGGTGGTTACAATGACGTTGGCGGAGGAG
HINB1-2 DQ862246.1 Turuspekov
HINB1-3 DQ862232.1 Turuspekov
HINB1-4 AY644022.1 Caldwell et
HINB1-5 JN636843.1 Galassi et
HINB1-6 AB611029.1 Terasawa er
HINB1-7 DQ862289.1 Turuspekov
HINB1-8 AY644058.1 Caldwell et
HINB1-9 AY644025.1 Caldwell et
HINB1-10 AY644023.1 Caldwell e
HINB1-11 AY643991.1 Caldwell e
HINB1-12 AY643980.1 Caldwell e
HINB1-13 AY643973.1 Calwell et
HINB1-14 AY643987.1 Caldwell e
HINB1-15 DQ862273.1 Turuspekov
110 120 130 140 150 160 170 180 190 200
HINB1-1 AY643843.1 Caldwell et GCGGTTCTCAACAATGCCCACAGGAGCGGCCGAACCTAGGCTCTTGCAAGGATTACGTGATGGAGCGGTGTTTCACGATGAAGGATTTTCCACTTACCTG
HINB1-2 DQ862246.1 Turuspekov
HINB1-3 DQ862232.1 Turuspekov
HINB1-4 AY644022.1 Caldwell et
HINB1-5 JN636843.1 Galassi et
HINB1-6 AB611029.1 Terasawa er
HINB1-7 DQ862289.1 Turuspekov
HINB1-8 AY644058.1 Caldwell et
HINB1-9 AY644025.1 Caldwell et
HINB1-10 AY644023.1 Caldwell e
HINB1-11 AY643991.1 Caldwell e
HINB1-12 AY643980.1 Caldwell e
HINB1-13 AY643973.1 Calwell et
HINB1-14 AY643987.1 Caldwell e
HINB1-15 DQ862273.1 Turuspekov
210 220 230 240 250 260 270 280 290 300
HINB1-1 AY643843.1 Caldwell et GCCCACAAAATGGTGGAAGGGAGGCTGTGAACAAGAGGTTCGGGAGAAGTGTTGCCAGCAACTGAGCCAGATAGCACCACAATGTCGCTGTGATGCTATC
HINB1-2 DQ862246.1 Turuspekov
HINB1-3 DQ862232.1 Turuspekov
HINB1-4 AY644022.1 Caldwell et
HINB1-5 JN636843.1 Galassi et
HINB1-6 AB611029.1 Terasawa er
HINB1-7 DQ862289.1 Turuspekov
HINB1-8 AY644058.1 Caldwell et
HINB1-9 AY644025.1 Caldwell et
HINB1-10 AY644023.1 Caldwell e
HINB1-11 AY643991.1 Caldwell e
HINB1-12 AY643980.1 Caldwell e
HINB1-13 AY643973.1 Calwell et
HINB1-14 AY643987.1 Caldwell e
HINB1-15 DQ862273.1 Turuspekov

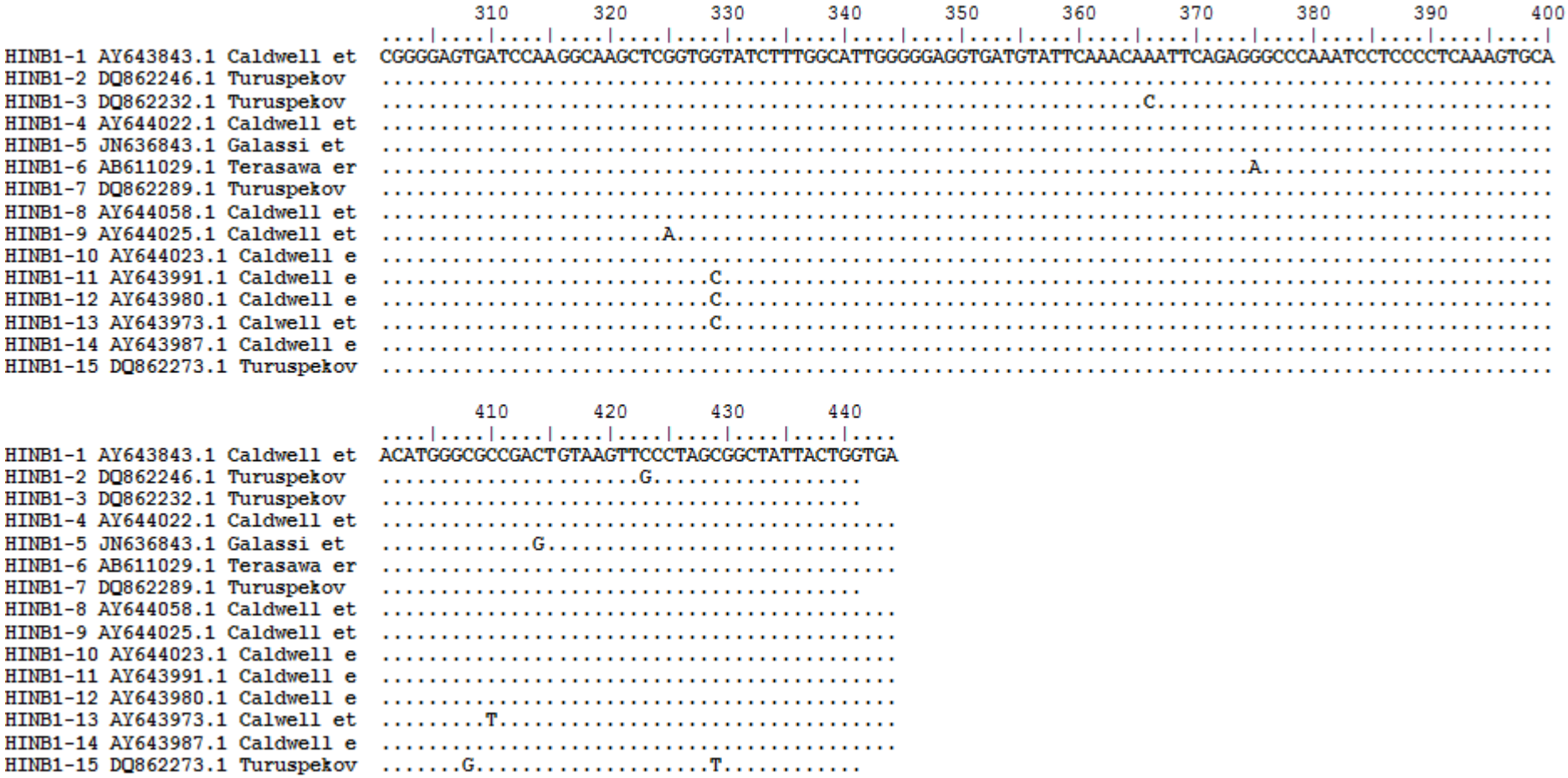



**Fig. I c: Alignment of *Hinb-1* DNA sequences**

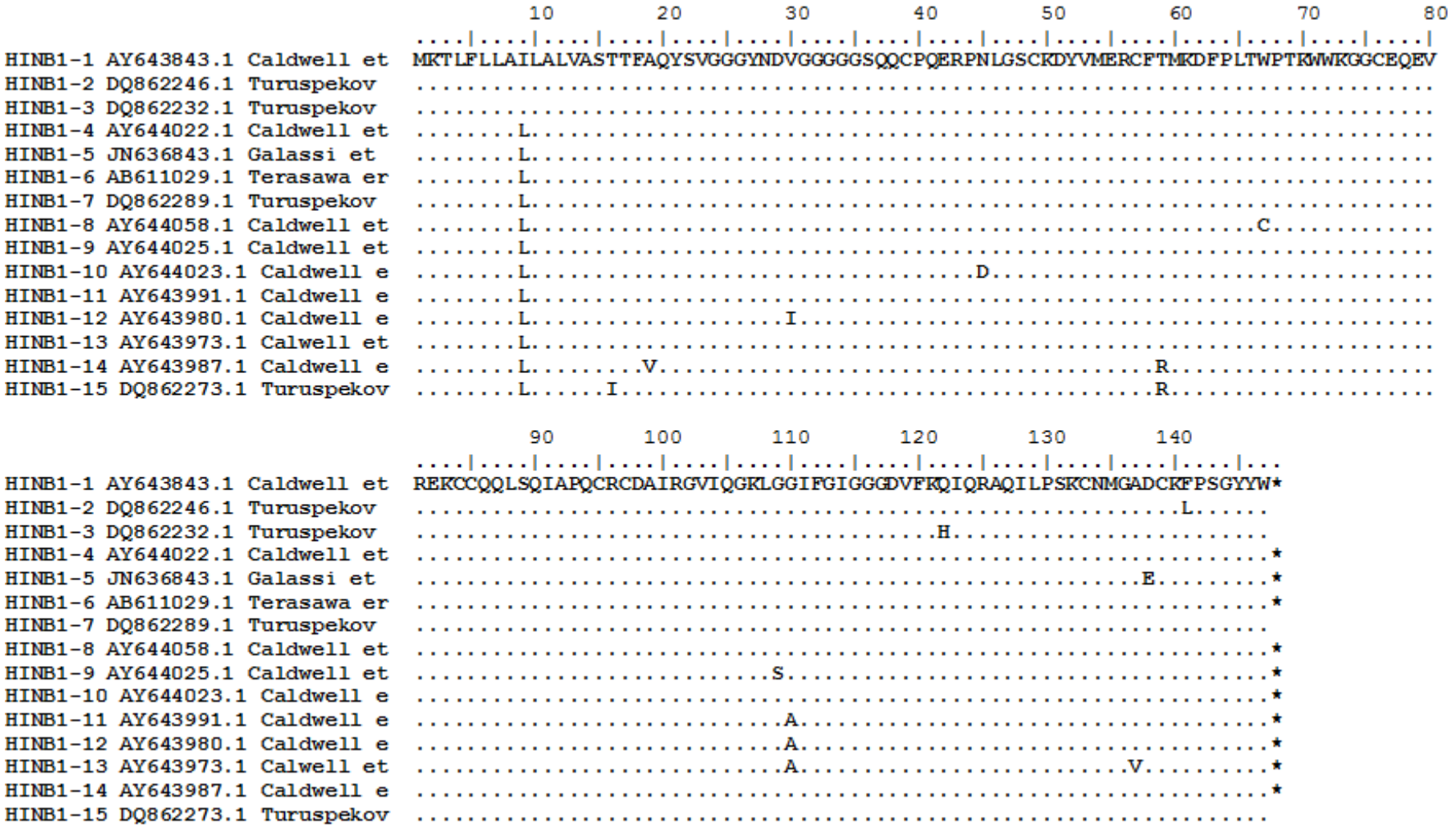
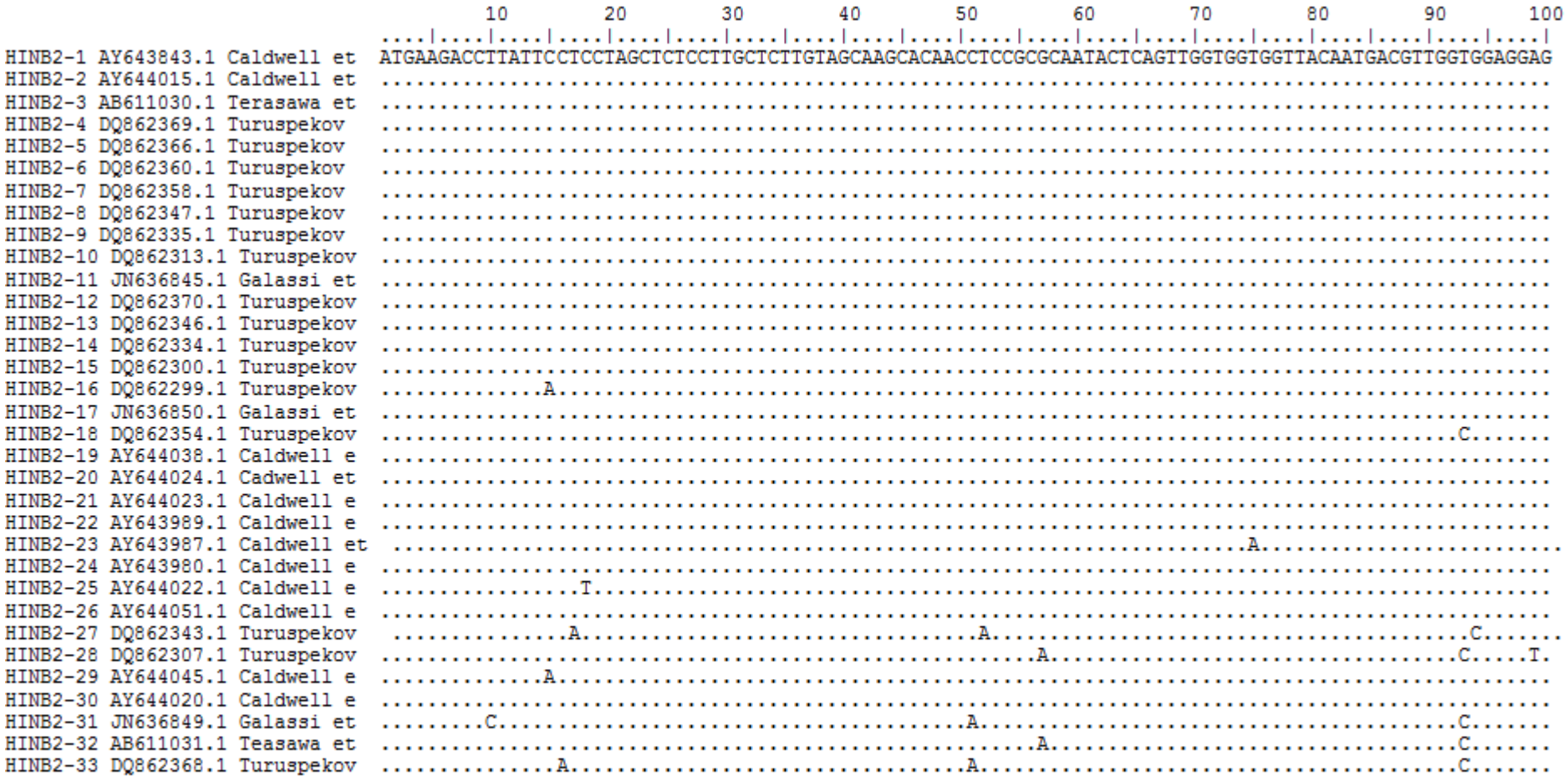
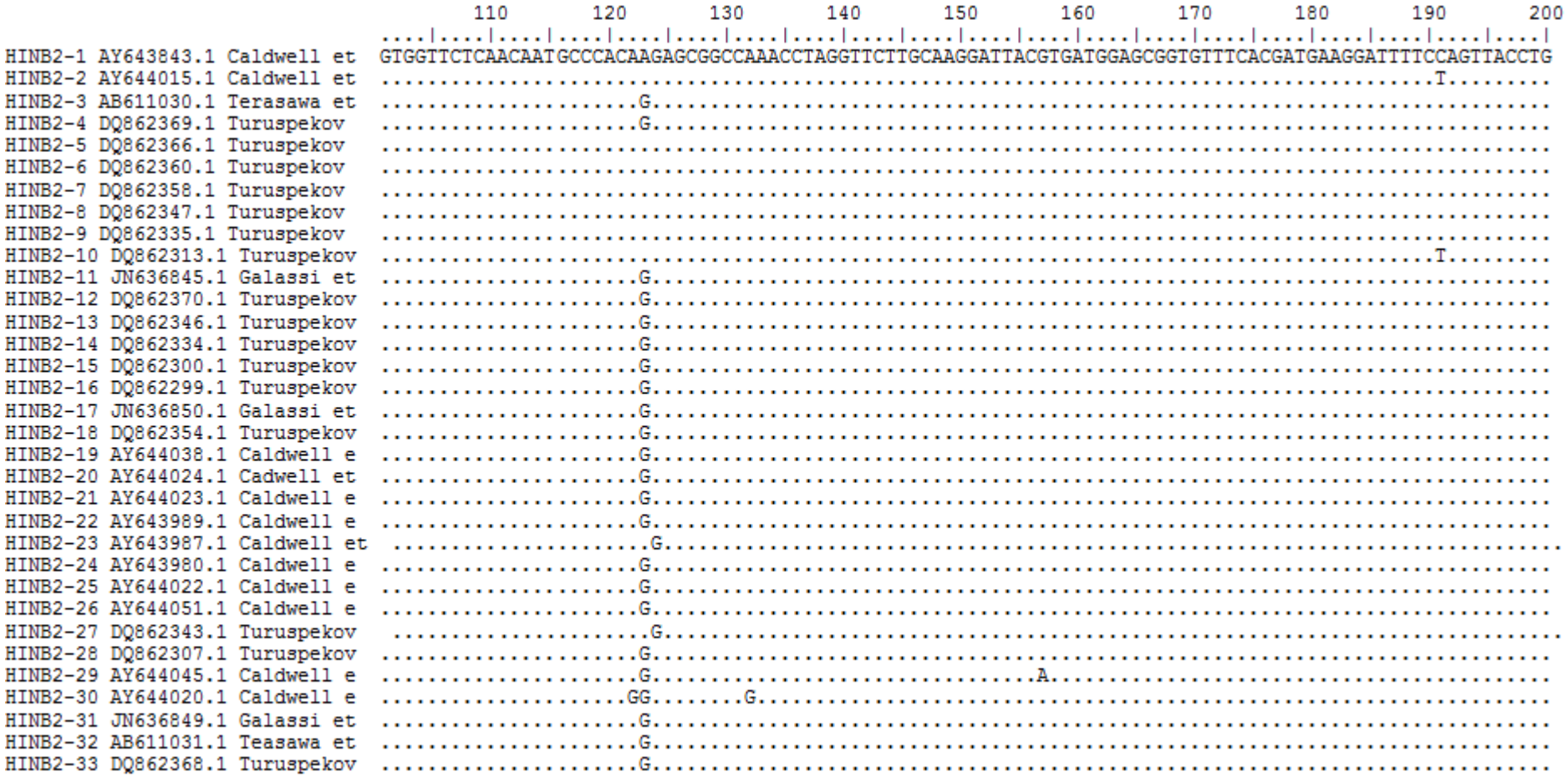
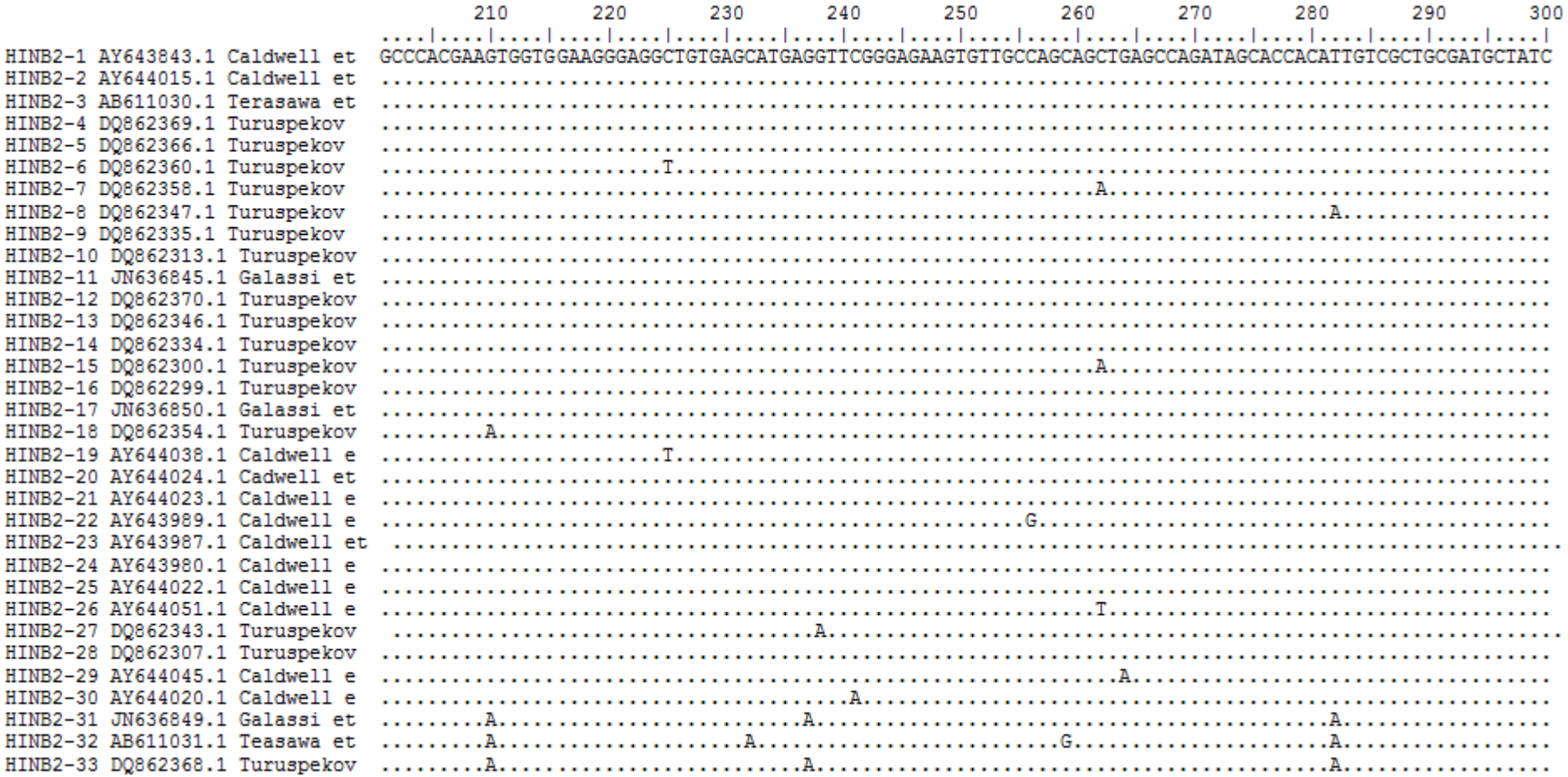



**Fig. I d: Alignment of HINB-1 putative protein sequences**


10 20 30 40 50 60 70 80 90 100
HINB2-1 AY643843.1 Caldwell et ATGAAGACCTTATTCCTCCTAGCTCTCCTTGCTCTTGTAGCAAGCACAACCTCCGCGCAATACTCAGTTGGTGGTGGTTACAATGACGTTGGTGGAGGAG
HINB2-2 AY644015.1 Caldwell et
HINB2-3 AB611030.1 Terasawa et
HINB2-4 DQ862369.1 Turuspekov
HINB2-5 DQ862366.1 Turuspekov
HINB2-6 DQ862360.1 Turuspekov
HINB2-7 DQ862358.1 Turuspekov
HINB2-8 DQ862347.1 Turuspekov
HINB2-9 DQ862335.1 Turuspekov
HINB2-10 DQ862313.1 Turuspekov
HINB2-11 JN636845.1 Galassi et
HINB2-12 DQ862370.1 Turuspekov
HINB2-13 DQ862346.1 Turuspekov
HINB2-14 DQ862334.1 Turuspekov
HINB2-15 DQ862300.1 Turuspekov
HINB2-16 DQ862299.1 Turuspekov
HINB2-17 JN636850.1 Galassi et
HINB2-18 DQ862354.1 Turuspekov
HINB2-19 AY644038.1 Caldwell e
HINB2-20 AY644024.1 Cadwell et
HINB2-21 AY644023.1 Caldwell e
HINB2-22 AY643989.1 Caldwell e
HINB2-23 AY643987.1 Caldwell et
HINB2-24 AY643980.1 Caldwell e
HINB2-25 AY644022.1 Caldwell e
HINB2-26 AY644051.1 Caldwell e
HINB2-27 DQ862343.1 Turuspekov
HINB2-28 DQ862307.1 Turuspekov
HINB2-29 AY644045.1 Caldwell e
HINB2-30 AY644020.1 Caldwell e
HINB2-31 JN636849.1 Galassi et
HINB2-32 AB611031.1 Teasawa et
HINB2-33 DQ862368.1 Turuspekov

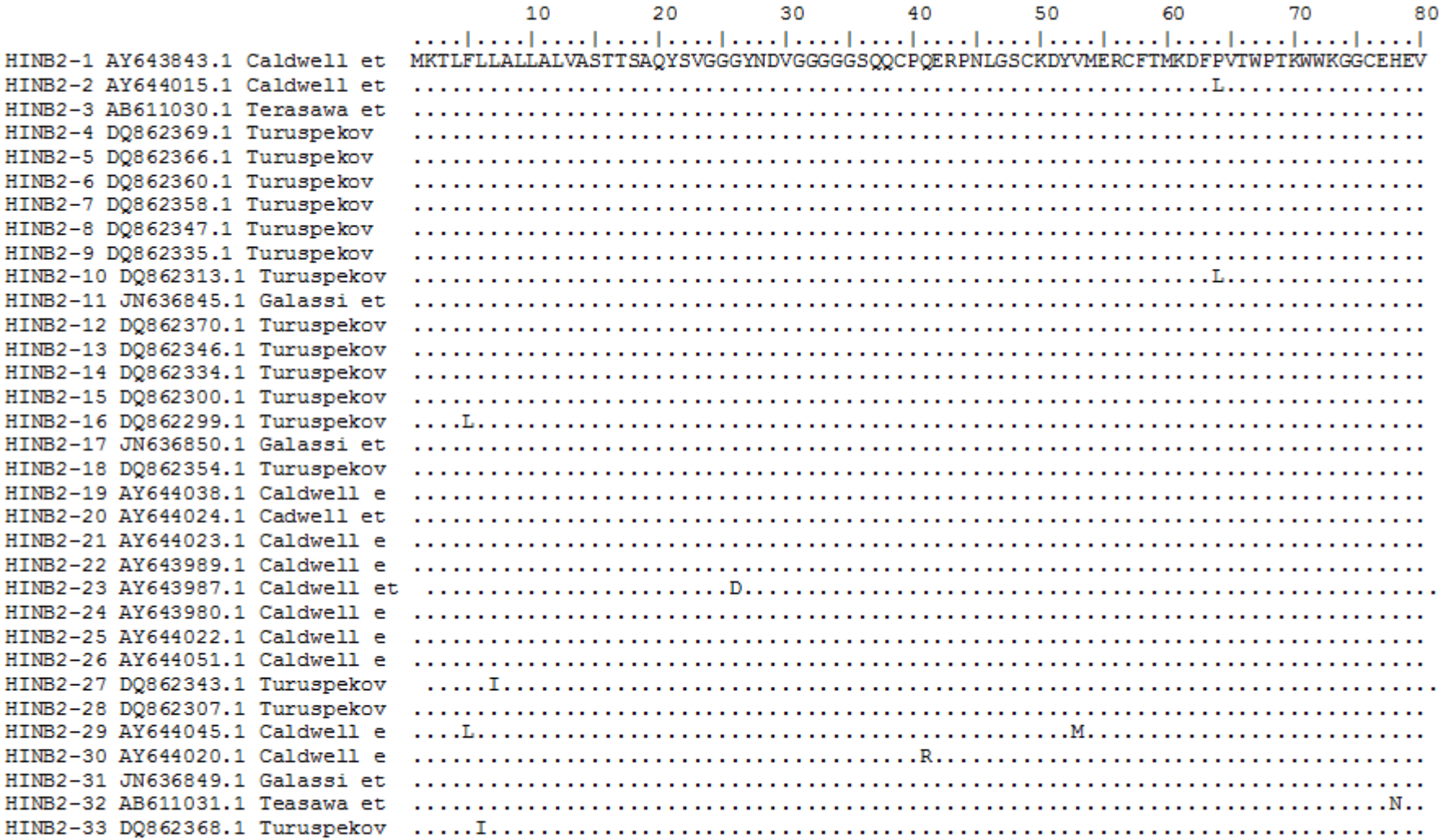

110 120 130 140 150 160 170 180 190 200
HINB2-1 AY643843.1 Caldwell et GTGGTTCTCAACAATGCCCACAAGAGCGGCCAAACCTAGGTTCTTGCAAGGATTACGTGATGGAGCGGTGTTTCACGATGAAGGATTTTCCAGTTACCTG
HINB2-2 AY644015.1 Caldwell et .....T.....
HINB2-3 AB611030.1 Terasawa et .....G.....
HINB2-4 DQ862369.1 Turuspekov .....G.....
HINB2-5 DQ862366.1 Turuspekov .....
HINB2-6 DQ862360.1 Turuspekov .....
HINB2-7 DQ862358.1 Turuspekov .....
HINB2-8 DQ862347.1 Turuspekov .....
HINB2-9 DQ862335.1 Turuspekov .....
HINB2-10 DQ862313.1 Turuspekov .....T.....
HINB2-11 JN636845.1 Galassi et .....G.....
HINB2-12 DQ862370.1 Turuspekov .....G.....
HINB2-13 DQ862346.1 Turuspekov .....G.....
HINB2-14 DQ862334.1 Turuspekov .....G.....
HINB2-15 DQ862300.1 Turuspekov .....G.....
HINB2-16 DQ862299.1 Turuspekov .....G.....
HINB2-17 JN636850.1 Galassi et .....G.....
HINB2-18 DQ862354.1 Turuspekov .....G.....
HINB2-19 AY644038.1 Caldwell e .....G.....
HINB2-20 AY644024.1 Cadwell et .....G.....
HINB2-21 AY644023.1 Caldwell e .....G.....
HINB2-22 AY643989.1 Caldwell e .....G.....
HINB2-23 AY643987.1 Caldwell et .....G.....
HINB2-24 AY643980.1 Caldwell e .....G.....
HINB2-25 AY644022.1 Caldwell e .....G.....
HINB2-26 AY644051.1 Caldwell e .....G.....
HINB2-27 DQ862343.1 Turuspekov .....G.....
HINB2-28 DQ862307.1 Turuspekov .....G.....
HINB2-29 AY644045.1 Caldwell e .....G.....A.....
HINB2-30 AY644020.1 Caldwell e .....GG.....G.....
HINB2-31 JN636849.1 Galassi et .....G.....
HINB2-32 AB611031.1 Teasawa et .....G.....
HINB2-33 DQ862368.1 Turuspekov .....G.....


210 220 230 240 250 260 270 280 290 300
HINB2-1 AY643843.1 Caldwell et GCCCACGAAGTGGTGGAAGGGAGGCTGTGAGCATGAGGTTCGGGAGAAGTGTTGCCAGCAGCTGAGCCAGATAGCACCACATTGTCGCTGCGATGCTATC
HINB2-2 AY644015.1 Caldwell et
HINB2-3 AB611030.1 Terasawa et
HINB2-4 DQ862369.1 Turuspekov
HINB2-5 DQ862366.1 Turuspekov
HINB2-6 DQ862360.1 Turuspekov
HINB2-7 DQ862358.1 Turuspekov
HINB2-8 DQ862347.1 Turuspekov
HINB2-9 DQ862335.1 Turuspekov
HINB2-10 DQ862313.1 Turuspekov
HINB2-11 JN636845.1 Galassi et
HINB2-12 DQ862370.1 Turuspekov
HINB2-13 DQ862346.1 Turuspekov
HINB2-14 DQ862334.1 Turuspekov
HINB2-15 DQ862300.1 Turuspekov
HINB2-16 DQ862299.1 Turuspekov
HINB2-17 JN636850.1 Galassi et
HINB2-18 DQ862354.1 Turuspekov
HINB2-19 AY644038.1 Caldwell e
HINB2-20 AY644024.1 Cadwell et
HINB2-21 AY644023.1 Caldwell e
HINB2-22 AY643989.1 Caldwell e
HINB2-23 AY643987.1 Caldwell et
HINB2-24 AY643980.1 Caldwell e
HINB2-25 AY644022.1 Caldwell e
HINB2-26 AY644051.1 Caldwell e
HINB2-27 DQ862343.1 Turuspekov
HINB2-28 DQ862307.1 Turuspekov
HINB2-29 AY644045.1 Caldwell e
HINB2-30 AY644020.1 Caldwell e
HINB2-31 JN636849.1 Galassi et
HINB2-32 AB611031.1 Teasawa et
HINB2-33 DQ862368.1 Turuspekov

```
                                  310       320       330       340       350       360       370       380       390       400
                                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
HINB2-1 AY643843.1 Caldwell et    CGGGGAGTGATCCAAGGCAAGCTCGGTGGTATCTTTGGCATTGGGGGAGGTGCTGTATTCAAACAAATTCAGAGGGCCCAGATCCTCCCCTCAAAGTGCA
HINB2-2 AY644015.1 Caldwell et    ....................................................................................................
HINB2-3 AB611030.1 Terasawa et   ..........................................................A.........................................
HINB2-4 DQ862369.1 Turuspekov    ....................................................................................................
HINB2-5 DQ862366.1 Turuspekov    ....................................................................................................
HINB2-6 DQ862360.1 Turuspekov    ....................................................................................................
HINB2-7 DQ862358.1 Turuspekov    ....................................................................................................
HINB2-8 DQ862347.1 Turuspekov    ....................................................................................................
HINB2-9 DQ862335.1 Turuspekov    ....................................................................................................
HINB2-10 DQ862313.1 Turuspekov   ....................................................................................................
HINB2-11 JN636845.1 Galassi et    ....................................................................................................
HINB2-12 DQ862370.1 Turuspekov   ....................................................................................................
HINB2-13 DQ862346.1 Turuspekov   ....................................................................................................
HINB2-14 DQ862334.1 Turuspekov   ..........................................................A.........................................
HINB2-15 DQ862300.1 Turuspekov   ....................................................................................................
HINB2-16 DQ862299.1 Turuspekov   ....................................................................................................
HINB2-17 JN636850.1 Galassi et    ..........................................................A.........................................
HINB2-18 DQ862354.1 Turuspekov   ....................................................................................................
HINB2-19 AY644038.1 Caldwell e    ....................................................................................................
HINB2-20 AY644024.1 Cadwell et    ..................................................................................T..
HINB2-21 AY644023.1 Caldwell e    T...................................................................................................
HINB2-22 AY643989.1 Caldwell e    ....................................................................................................
HINB2-23 AY643987.1 Caldwell et     ....................................................................................................
HINB2-24 AY643980.1 Caldwell e    ....................................................................................................
HINB2-25 AY644022.1 Caldwell e    ....................................................................................................
HINB2-26 AY644051.1 Caldwell e    ....AGTGA
HINB2-27 DQ862343.1 Turuspekov     ....................................................................................................
HINB2-28 DQ862307.1 Turuspekov   ....................................................................................................
HINB2-29 AY644045.1 Caldwell e    ....................................................................................................
HINB2-30 AY644020.1 Caldwell e    ..................................................................................G.................
HINB2-31 JN636849.1 Galassi et    ..........................................................A.........................................
HINB2-32 AB611031.1 Teasawa et   ..........................................................A.........................................
HINB2-33 DQ862368.1 Turuspekov   ..........................................................A.........................................
```

```
                                  410       420       430       440
                          ....|....|....|....|....|....|....|....|....
HINB2-1 AY643843.1 Caldwell et  ACATGGGCGTCGACTGTAGGTTCCCTAGTGGCTATTACTGGTGA
HINB2-2 AY644015.1 Caldwell et  ........................................
HINB2-3 AB611030.1 Terasawa et  ........................................
HINB2-4 DQ862369.1 Turuspekov   ........................................
HINB2-5 DQ862366.1 Turuspekov   ................A.......................
HINB2-6 DQ862360.1 Turuspekov   ........................................
HINB2-7 DQ862358.1 Turuspekov   ........................................
HINB2-8 DQ862347.1 Turuspekov   ........................................
HINB2-9 DQ862335.1 Turuspekov   .........C..............................
HINB2-10 DQ862313.1 Turuspekov  ........................................
HINB2-11 JN636845.1 Galassi et  .........C........A.......................
HINB2-12 DQ862370.1 Turuspekov  .........C..............................
HINB2-13 DQ862346.1 Turuspekov  ................A.......................
HINB2-14 DQ862334.1 Turuspekov  ........................................
HINB2-15 DQ862300.1 Turuspekov  ........................................
HINB2-16 DQ862299.1 Turuspekov  ........................................
HINB2-17 JN636850.1 Galassi et  .........C........A.......................
HINB2-18 DQ862354.1 Turuspekov  ........................................
HINB2-19 AY644038.1 Caldwell e  .........C........A.......................
HINB2-20 AY644024.1 Cadwell et
HINB2-21 AY644023.1 Caldwell e  .........C........A.......................
HINB2-22 AY643989.1 Caldwell e  .........C........A.......................
HINB2-23 AY643987.1 Caldwell et  .........C........A.......................
HINB2-24 AY643980.1 Caldwell e  .........C........A.....A.................
HINB2-25 AY644022.1 Caldwell e  .........C......C.A.......................
HINB2-26 AY644051.1 Caldwell e
HINB2-27 DQ862343.1 Turuspekov   ........................................
HINB2-28 DQ862307.1 Turuspekov  ............T...........................
HINB2-29 AY644045.1 Caldwell e  .........C........A.......................
HINB2-30 AY644020.1 Caldwell e  .........C........A.......................
HINB2-31 JN636849.1 Galassi et  .........C........A.......................
HINB2-32 AB611031.1 Teasawa et  .........C...T....A.......................
HINB2-33 DQ862368.1 Turuspekov  .........C........A.......................
```

**Fig. I e: Alignment of *Hinb-2* DNA sequences**


10 20 30 40 50 60 70 80
HINB2-1 AY643843.1 Caldwell et MKTLFLLALLALVASTTSAQYSVGGGYNDVGGGGGSQQCPQERPNLGSCKDYVMERCFTMKDFPVTWPTKWWKGGCEHEV
HINB2-2 AY644015.1 Caldwell et .....L.....
HINB2-3 AB611030.1 Terasawa et .....
HINB2-4 DQ862369.1 Turuspekov .....
HINB2-5 DQ862366.1 Turuspekov .....
HINB2-6 DQ862360.1 Turuspekov .....
HINB2-7 DQ862358.1 Turuspekov .....
HINB2-8 DQ862347.1 Turuspekov .....
HINB2-9 DQ862335.1 Turuspekov .....
HINB2-10 DQ862313.1 Turuspekov .....L.....
HINB2-11 JN636845.1 Galassi et .....
HINB2-12 DQ862370.1 Turuspekov .....
HINB2-13 DQ862346.1 Turuspekov .....
HINB2-14 DQ862334.1 Turuspekov .....
HINB2-15 DQ862300.1 Turuspekov .....
HINB2-16 DQ862299.1 Turuspekov ....L.....
HINB2-17 JN636850.1 Galassi et .....
HINB2-18 DQ862354.1 Turuspekov .....
HINB2-19 AY644038.1 Caldwell e .....
HINB2-20 AY644024.1 Cadwell et .....
HINB2-21 AY644023.1 Caldwell e .....
HINB2-22 AY643989.1 Caldwell e .....
HINB2-23 AY643987.1 Caldwell et .....D.....
HINB2-24 AY643980.1 Caldwell e .....
HINB2-25 AY644022.1 Caldwell e .....
HINB2-26 AY644051.1 Caldwell e .....
HINB2-27 DQ862343.1 Turuspekov .....I.....
HINB2-28 DQ862307.1 Turuspekov .....
HINB2-29 AY644045.1 Caldwell e ....L.....M.....
HINB2-30 AY644020.1 Caldwell e .....R.....
HINB2-31 JN636849.1 Galassi et .....
HINB2-32 AB611031.1 Teasawa et .....N..
HINB2-33 DQ862368.1 Turuspekov .....I.....

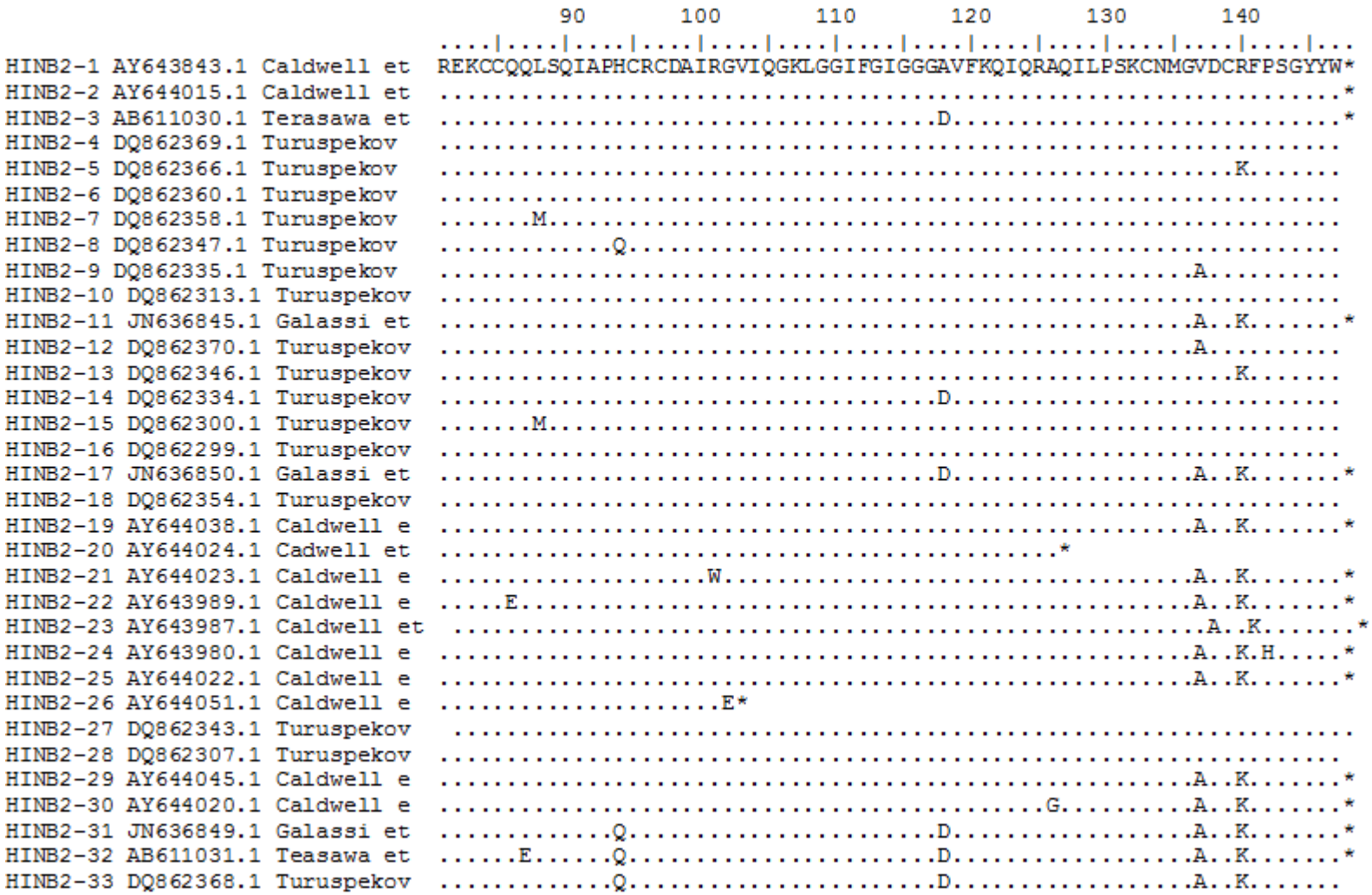



**Fig. I f: Alignment of HINB-2 putative protein sequences**

Appendix II pGEMT-Easy vector system

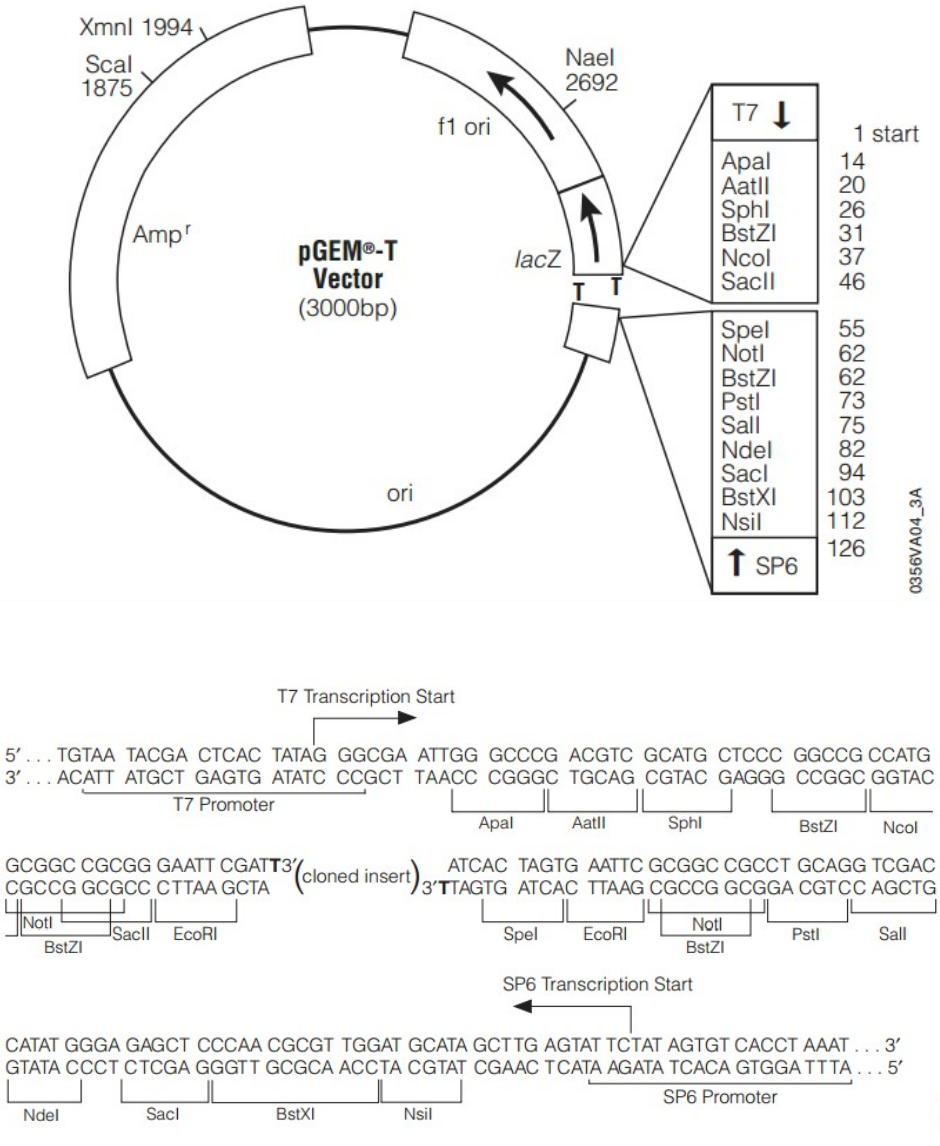
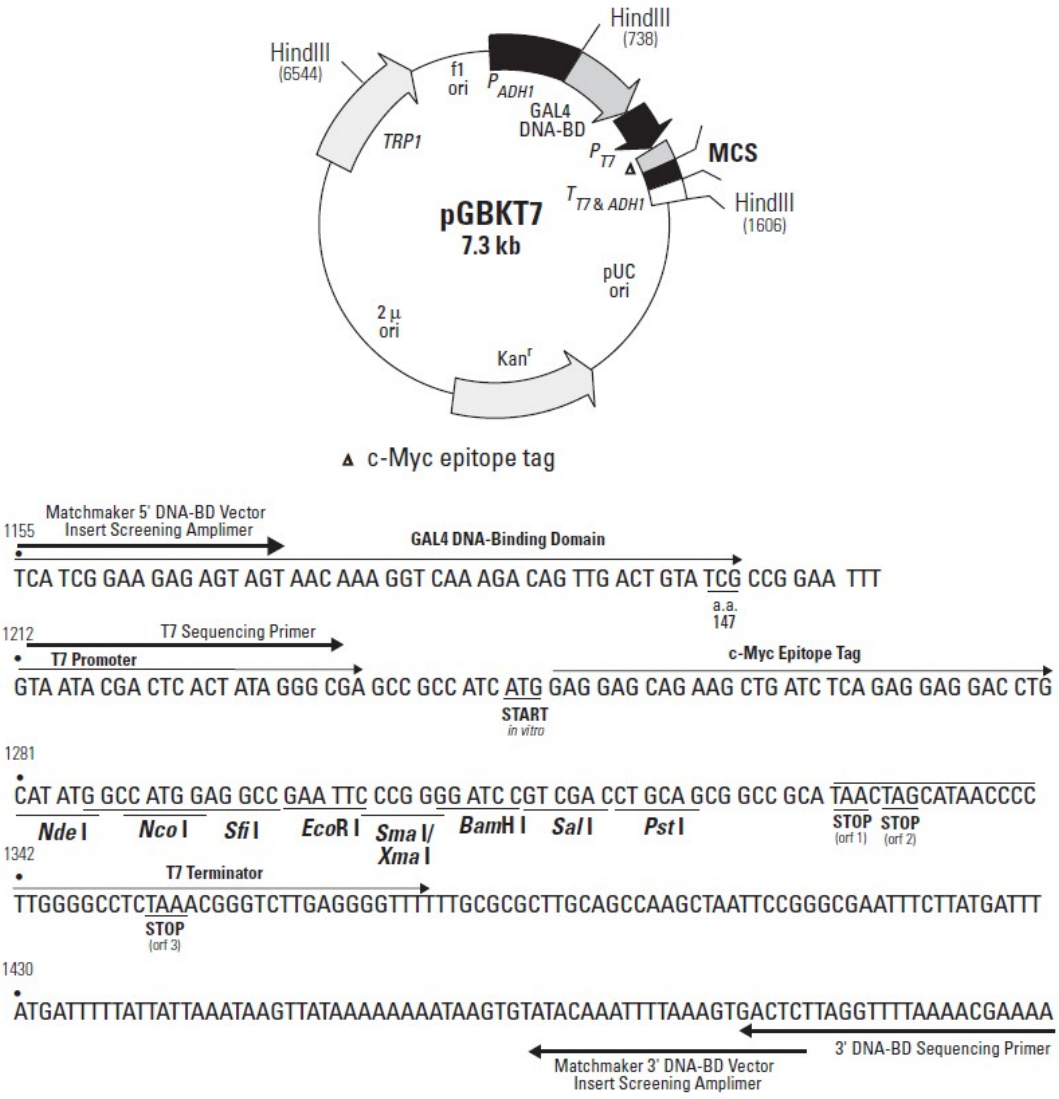



**Figure The promoter and multiple cloning sequence of the pGEM®-T Easy Vector.** The top strand shown corresponds to the RNA synthesized by T7 RNA polymerase. The bottom strand corresponds to the RNA synthesized by SP6 RNA polymerase.

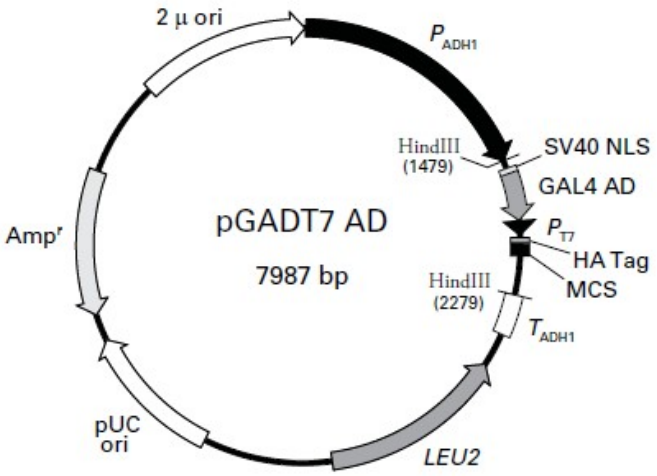
## Appendix III Y2H bait and prey vectors

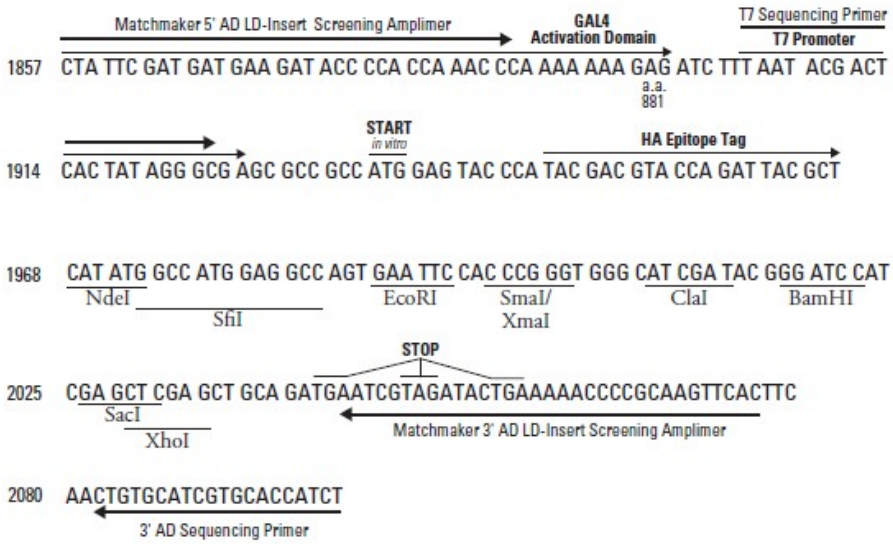
### pGBKT7 (bait) vector




**Restriction Map and Multiple Cloning Site (MCS) of pGBKT7.** Unique restriction sites are in bold.

pGADT7 (prey) vector




**pGADT7 AD Vector Map and Multiple Cloning Site (MCS).**