Southern Cross University
ePublications@SCU

Theses

2011

Analysis of betaine aldehyde dehydrogenase encoding genes in wheat

Keshav Narayan Shrestha
Southern Cross University

Publication details
Shrestha, K 2011, 'Analysis of betaine aldehyde dehydrogenase encoding genes in wheat', MSc thesis, Southern Cross University, Lismore, NSW.
Copyright K Shrestha 2011

ePublications@SCU is an electronic repository administered by Southern Cross University Library. Its goal is to capture and preserve the intellectual output of Southern Cross University authors and researchers, and to increase visibility and impact through open access to researchers around the world. For further information please contact epubs@scu.edu.au.

Analysis of betaine aldehyde dehydrogenase encoding genes in wheat

Keshav Narayan Shrestha
M.Sc (Biotechnology)
Centre for Plant Conservation Genetics
Southern Cross University

Thesis submitted to fulfill the requirements of the Degree of Master of Science

May 2011

Declaration

I, Keshav Narayan Shrestha, declare that the work presented in this thesis is, to the best of my knowledge and belief, original, except as acknowledged in the text, and that the material has not been submitted, either in whole or in part, for a degree at this or any other university.

I acknowledge that I have read and understood the university's rule requirements, procedure and policy relating to my higher degree research award and to my thesis. I certify that I have complied with the rules, requirements, procedure and policy of the university.

Signed …….
Keshav Narayan Shrestha

Date …….

Abstract

Betaine aldehyde dehydrogenase (BADH) is an important enzyme which has dual roles in cereals influencing abiotic stress tolerance and rice fragrance. Most grass species have two BADH paralogs (*BADH1* and *BADH2)*. A mutation in some *BADH2* alleles introduces a termination codon which causes truncation of the protein and ultimately elevates the level of 2 acetyl-1-pyroline, leading to fragrance in rice and soybean. Bread wheat is hexaploid containing three genomes A, B and D. Identifying the potential number of *BADH* isozymes in wheat may indicate if it is possible to generate fragrant wheat by knocking out one, two or three homeologs. The hexaploid wheat variety Cadoux has three orthologs for both the *BADH* paralogs while its progenitors have a single ortholog. RT-PCR suggests only one homeolog of both *BADH1* and *BADH2* is expressed at two different time points (14 DPA and 30 DPA) in two different tissues (seeds and leaves) in hexaploid wheat variety Cadoux. Analysis of *BADH1* and *BADH2* gene expression in three other varieties in two different time points (14 DPA and 30 DPA) and two different tissues (seeds and leaves) also indicates that only the "A" genome homeolog is expressed. More research needs to undertaken to determine the exact function of this gene in wheat.

Acknowledgements

I would like to thank my principle supervisor Dr. Daniel Waters for giving me opportunity to undertake this project in Centre for Plant Conservation Genetics. I would like to thank him for his enormous help and direction in my project and in my personal life as well.

Next, I would like to thank my co-supervisor Dr. Nicole Rice for her valuable suggestion and direction provided during my research in Centre for Plant Conservation Genetics and critically analysis in my writing.

This project wouldn't have been completed without the support of my family, their love and their sacrifice. I would really like to dedicate this thesis in name of my mom whom I lost last year and want peace for her soul.

At Last but not the least, I would like to thanks my entire friends in post graduate room Shabana Kasem, Ardy, Sylvia Mye, Stuart, Helen and Tim for their support during my study and in personal help. I also would like to thank Mr. Dion Thompson from Centre for Photochemistry and Pharmacology for his valuable help in analysing the protein structure. It was always fun to share a lunch at the lunch table in T-block and the fun we share is unforgettable.

Publications

The following poster is based on this dissertation.

Shrestha K.N, Waters D.L.E, and Rice N
"BADH genes in wheat and their expression"
Plant and Animal Genome XIX Conference, January 15-19, 2011, San Diego, California, USA.

List of Abbreviations

µg	Micrograms
µl	Millilitres
µM	Micromolar
2AP	2- acetyl-1- pyrroline
AB-ald	4-aminobutyraldehyde
ALDH	Aldehyde dehydrogenase
APAL	Aminoaldhehyde-3- aminopropinaldehyde
AP-ald-	3-aminopropionaldheyde
BADH1	Betaine aldehyde dehydrogenase 1
BADH2	Betaine aldehyde dehydrogenase 2
badh2	Betaine aldehyde dehydrogenase 2
BBD1	Betaine aldehyde dehydrogenase 1in barley
BBD2	Betaine aldehyde dehydrogenase 2in barley
bp	Base pair
cDNA	Complimentary deoxyribonucleic acid
CMO	Choline monooxygenase
DMSP-ald	3-dimethyl sulfoniopropional aldehyde
DNA	Deoxyribonucleic acid
DPA	Days post anthesis
GABA	γ-aminobutyric acid

GABald	γ-aminobutyraldehyde
GB	Glycine betaine
GBAL	4- guanidiionobutyraldehyde
gDNA	Genomic DNA
GGBald	γ-guanidinobutyraldehyde
IPTG	Isopropyl β-D-thiogalactoside.
kbp	Kilo base pair
kDa	Kilo dalton
Km	Michaelis -Menten constant
LB media	Luria –Bertani media
MQ- water	Milli Q water
mRNA	Messenger ribonucleic acid
NaCl	Sodium chloride
NAD^+	Nicotinamide adenine dinuclotide
$NADP^+$	Nicotinamide adenine dinuclotide phosphate
^{O}C	Degree celsius
KCl	Potassium chloride.
PCR	Polymerase Chain Reaction
Ph	Pairing homologous
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RT-PCR	Reverse transcriptase PCR

SOC	Super optimal broth
Tm	Melting temperature.
UV	Ultra violet
X-gal	Bromo-chlor-indolyl-galactopyranoside

Table of contains

Declaration ii

Abstract iii

Acknowledgement iv

Publication v

Abbreviation used vi

Table of contents ix

List of figures xii

List of tables xiii

List of appendix xiv

Chapter One: Introduction

1.1 Betaine aldehyde dehydrogenase and rice fragrance 1

1.2 Betaine aldehyde dehydrogenase and plant abiotic stress tolerance 3

1.3 Structure of BADH 4

1.4 BADH substrate specificity 5

1.5 *BADH* gene expression 8

1.6 Effect of pH on *BADH* 8

1.7 Domestication of wheat 9

1.8 Wheat genome 10

Chapter Two: Identification of *BADH* homologs

2.1 Introduction 13

2.2 Materials and Methods 15

2.2.1 Sample collection and plants growth 15

2.2.2 DNA extraction 15

2.2.3 PCR primer design 15

2.2.4 PCR amplification 18

2.2.4.1 Reconditioning PCR 18

2.2.4.2 PCR cycling for screening positive bacterial clones 19

2.2.5 Purification of agarose gel slices 19

2.2.6 Cloning 19

2.2.6.1 LB media 19

2.2.6.2 SOC media 20

2.2.6.3 X-gal 20

2.2.6.4 IPTG 20

2.2.6.5 Cloning of PCR fragments 20

2.2.6.6 Gel slice cloning of PCR fragments 21

2.2.6.7 PCR screening of bacterial colonies 21

2.2.6.8 Templiphi amplification of colonies 21

2.2.7 Sequencing reaction protocol 22

2.2.8 Post- sequencing reaction purifications 22

2.2.9 Sequence analysis and alignment 22

2.3 Results 24

2.3.1 Genomic DNA extraction 24

2.3.2 *BADH1* and *BADH2* amplification 24

2.3.3 Gel slice purification and direct sequencing 25

2.3.4 Reconditioning PCR 26

2.3.5 Clone screening and sequencing 26

2.3.6 *BADH1* and *BADH2* amplification from progenitors 27

2.4 Discussion 41

Chapter Three: Expression of *BADH* homologs

3.1 Introduction 43

3.2 Materials and Methods 46

3.2.1 Sample collection and plants growth 46

3.2.2 RNA extraction 47

3.2.2.1 Pulverization of leaves and seeds 47

3.2.2.2 Modification of TRizol® and Qiagen RNA extraction 47

3.2.3 Quantification and quality of RNA 48

3.2.4 Primer design 48

3.2.5 Reverse transcriptase PCR (RT-PCR) 49

3.2.6 Agarose gel electrophorosis 50

3.2.7 Sequencing reaction protocol 50

3.2.8 Sequence analysis and alignment 51

3.2.9 Protein structure analysis 51

3.3 Results 52

3.3.1 RNA extraction 52

3.3.2 *BADH1* amplification 52

3.3.3 *BADH2* amplification 56

3.4 Discussion 65

Chapter Four: General discussion and conclusion 67

Bibliography 71

Appendix 85

List of Figures

Figure 1.1	Biochemical pathways to GB from choline via CMO and BADH.	3
Figure 1.2	Putative biochemical pathways in rice showing activity of BADH gene.	7
Figure 1.3	Domestication of hexaploid wheat.	10
Figure 2.1	ClustalW alignment of protein of different species of cereal with their amino acid sequences of BADH homologs.	17
Figure 2.2	PCR amplification of wheat *BADH1*.	24
Figure 2.3	PCR amplification of wheat *BADH2*.	25
Figure 2.4	Five cycles re-conditioning PCR of *BADH1* and *BADH2*.	26
Figure 2.5	PCR amplification of *BADH1* wheat progenitors.	28
Figure 2.6	PCR amplification of *BADH2* wheat progenitors.	29
Figure 2.7	Phylogenetic relationship of *BADH1* of hexaploid wheat varieties Cadoux and its progenitors.	30
Figure 2.8	Alignment of *BADH1* amplicon with hexaploid wheat variety cadoux and their progenitors.	31
Figure 2.9	ClustalW Alignment of BADH1 predicted amino acid sequences of hexaploid wheat varieties Cadoux and their progenitors.	34
Figure 2.10	Phylogenetic relationship of *BADH2* of hexaploid wheat variety Cadoux and its progenitors.	35
Figure 2.11	Alignment of *BADH2* amplicon with hexaploid wheat variety cadoux and their progenitors. .	36
Figure 2.12	ClustalW alignment of BADH2 predicted amino acid sequences of hexaploid wheat varieties cadoux and their progenitors.	37
Figure 2.13	ClustalW alignment of *BADH2* amplicons from hexaploid wheat variety Cadoux.	38
Figure 3.1	RT-PCR for *BADH1*.	53

Figure 3.2	ClustalW alignment of hexaploid wheat *BADH1* cDNA of Cadoux 14 DPA leaves with their progenitors.	54
Figure 3.3	Alignment of predicted amino acid sequences BADH1 cDNA of Cadoux of hexaploid wheat with their progenitors	55
Figure 3.4	RT-PCR for *BADH2*.	56
Figure 3.5	Alignment of hexaploid wheat *BADH2* cDNA Cadoux of 14 DPA leaves with their progenitors.	57
Figure 3.6	ClustalW alignment of predicted amino acid sequences BADH2 of hexaploid wheat with their progenitors.	58
Figure 3.7	ClustalW alignment of expressed nucleotide sequences of *BADH1* different species of different time (14 DPA and 30 DPA) and tissues (Leaves and seeds).	59
Figure 3.8	Alignment of predicted amino acid sequences expressed BADH1 different species of different time (14 DPA and 30 DPA) and tissues (Leaves and seeds).	62
Figure 3.9	Alignment of expressed nucleotide sequences of *BADH2* different species of different time (14 DPA and 30 DPA) and tissues (Leaves and seeds).	63
Figure 3.10	Alignment of predicted amino acid sequences expressed BADH2 different species of different time (14 DPA and 30 DPA and tissues (Leaves and seeds).	64

List of tables

Table 1.1	The classification of chromosome of wheat.	12
Table 2.1	Primers used for the amplification of BADH homologs.	17
Table 3.1	Reverse Transcriptase PCR primers.	49

List of appendices

Appendix i 85

Appendix ii 88

Appendix iii 91

Appendix iv 95

Chapter One: Introduction

1.1 Betaine aldehyde dehydrogenase and rice fragrance

Consumers have a preference for particular aromas and flavours and these preferences are important when they select foods (Delmundo and Juliano, 1981). Over the past 30 years, the compounds which give the flavour or aroma to cereals have been characterised, the most well known example being the aroma of basmati and jasmine rice which arises due to the presence of 2-acetyl-1-pyrroline (2AP) (Buttery et al., 1983b). This volatile compound is found in the aerial parts of the rice plant and is more concentrated in new growth, the tips of rice leaves contain more 2AP compare to the base of leaves while older leaves are less aromatic than young ones (Lorieux et al., 1996). The characteristic aroma of popcorn, green tea, pandanas leaves (Schieberle, 1991; Kumazawa and Masuda, 2002; Yoshihashi, 2002) and a number of other foods including beef and bread crusts (Yoshihashi et al., 2002) is due to the presence of 2AP. Pandanas leaves are added to non-fragrant rice during cooking to impart the 2AP aroma to the cooked rice.

Fragrant rice and non-fragrant rice are not defined by the presence or absence of 2AP, but by the concentration of 2AP (Buttery et al., 1983a). The concentration of 2AP in fragrant rice grain is 100 times that of non-fragrant rice (Widjaja et al., 1996). Although 2AP is the compound which determines whether or not a rice cultivar is fragrant, this is not the only compound which differs in concentration when comparing fragrant and non-fragrant rice. Fragrant rice has more indole than non-fragrant rice while non-fragrant rice has higher concentrations of other volatile compounds associated with the aroma of cooked rice such as 6-methyl-5-hepten-2-one, (E)-2-heptenal, 1-octen-3-ol, nonanal, (E)-2-octenal and (E) 2 4-

decadienal (Widjaja et al., 1996). The concentration of these compounds in combination with 2AP concentration determines rice fragrance quality.

The 2AP associated with rice fragrance results from a deletion in a putative betaine aldehyde dehydrogenase (*BADH*) encoding homolog, *BADH2* (Bradbury et al., 2005). Non-fragrant rice varieties have a functional *BADH2* allele and fragrant rice varieties have a mutant *BADH2* allele (*badh2*) which contains a premature termination codon. Non-fragrant rice varieties have a functional *BADH2* allele and fragrant rice varieties have a mutant *BADH2* allele (*badh2*) which contains a premature termination codon. When expressed in *Escherichia coli* (*E. coli*), the enzyme which carries the premature termination codon is non-functional (Bradbury *et al.*, 2008) while fragrant rice expressing transgenic *BADH2*, accumulates significantly reduced amounts of 2AP (Chen et al., 2008). Many plant species have two BADH paralogs, *BADH1* and *BADH2* (Fitzgerald et al., 2009). In contrast to BADH2, both fragrant and non-fragrant rice varieties contain a *BADH1* gene which produces a functional enzyme when expressed in *E. coli* (Bradbury *et al*., 2008).

Although the gene involved in fragrance is well characterised in rice, little is known about its orthologs in other cereals in relation to fragrance. BADH was initially defined as an enzyme that catalyses the conversion of betaine aldehyde to glycine betaine, a reaction which is dependent on oxidising the co–factor NAD^+ or $NADP^+$ (Weigel et al., 1986; Nakamura et al., 1997; Mori et al., 2002). Glycine betaine (GB) is accumulated by animals, some bacteria and in several higher plant families (Chenopodiaceace, Podiaceace, Adteraceae) (Ishitani et al., 1993). In plants, it is predominantly accumulated in the leaves and stems of a diverse range of dicotyledons and some monocotyledons (Ishitani et al., 1993). GB is produced in response to abiotic stresses including salinity, drought (Lerudulier et al., 1984) and cold temperatures (Kishitani et al., 1994). One hypothesis is that GB protects the cell from different stresses by

maintaining the osmotic balance between the intracellular and extracellular compartments of cells in addition to stabilising enzymes and biomolecules (Robinson and Jones, 1986).

BADH and choline monooxygenase (CMO) work together, CMO converts choline to betaine aldehyde and *BADH* converts betaine aldehyde to glycine betaine (Fig.1.1). These enzymes are compartmentalized within the chloroplast in plants (Weigel et al., 1986) with both reactions taking place independently of light but stimulated by the presence of and length of exposure to light (Hanson, 1985; Weigel, 1988).

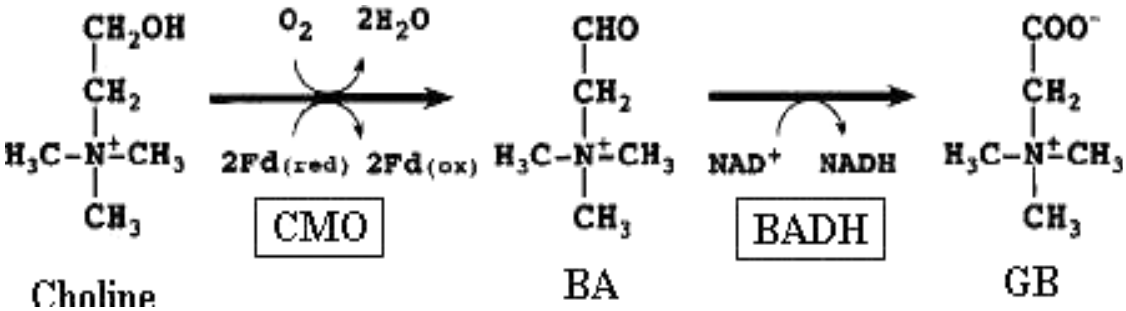

Figure 1.1 Biochemical pathways to GB from choline via CMO and BADH. Modified from (Sakamoto and Murata, 2002).

1.2 Betaine aldehyde dehydrogenase and plant abiotic stress tolerance

BADH is involved in abiotic stress tolerance in many crops by catalysing the formation of GB. Plant abiotic stress response is commonly mediated by increasing *BADH* gene activity (Nakamura et al., 2001) *BADH* transcript levels were more abundant compare to control plants in sorghum (Wood et al., 1996), spinach (Elizabeth A and Hanson, 1990) sugar beet (McCue and Hanson, 1992) and barley (Fujiwara et al., 2008) upon salt treatment.

Abiotic stress tolerance through GB accumulation can be enhanced in GB accumulating species. Wheat transformed with *BADH* from *Atriplex hortensis* was more tolerant to salt and freezing (Allard et al., 1998; Guo et al., 2000; Liang et al., 2009). Although rice does not

accumulate GB, expression of rice *BADH* in tobacco provided improved tolerance to salt and accumulation of glycine betaine to higher concentrations (Hasthanasombut et al., 2010). In rice there was no difference in *BADH1* expression in fragrant and non-fragrant rice in response to salt treatment. The concentration of *BADH2* transcripts were reduced in fragrant rice compared to non-fragrant rice without salt treatment (Fitzgerald et al., 2008) while yield is depressed with salt treatment to a greater extent in fragrant rice than non-fragrant rice (Fitzgerald et al., 2008). In addition, the concentration of 2AP rises when rice plants are subjected to salt and drought stress, suggesting the biosynthetic pathway which includes BADH2 is involved in abiotic stress tolerance (Yoshihashi et al., 2004).

1.3 Structure of BADH

Betaine aldehyde dehydrogenase (BADH, EC 1.2.1.8) belongs to the aldehyde dehydrogenase (ALDH) super family (Kovach et al., 2009). There is wide diversity in the structure of enzymes with BADH activity however, in the majority of plants, BADH is a dimeric protein (Weretilnyk and Hanson, 1989) while the *BADH* of animals and *E. coli* is tetrameric (Chern and Pietruszko, 1995; García Roberto Velasco, 1999; Mori et al., 2002). X-Ray diffraction studies of the pea ALDH isozymes at 2.4 A^0 and 2.5 A^0, found these to be dimeric proteins. Each monomer has three domains; the oligomerization domain, the enzyme binding domain and the catalytic domain (Brauner et al., 2003), similar to BADH isolated from mammalian livers (Johansson et al., 1998). Other examples include the BADH homodimer found in spinach and *Avena sativa* with molecular weights of 60-63 kDa and 61 kDa respectively. In contrast, amaranth has two subunits with molecular weights of 63kDa and 70kDa (Figueroa-Soto and Valenzuela-Soto, 2001). The amaranth protein is encoded by a single nuclear gene (Weretilnyk and Hanson, 1989) with two alleles giving a homodimer or hetrodimer product, depending upon whether or not the alleles are homozygous or heterozygous (Weretilnyk and Hanson, 1988).

The BADH isozymes localize in different sub-cellular compartments of chloroplast, peroxisomes of cytosol, depending on the species. Glycine betaine is synthesised in the cystol of chenopods (Hanson, 1985) which contrasts with the mammals which synthesise glycine betaine in the mitochondria (Johansson et al., 1998). The major BADH isozyme of spinach is localises to the chloroplast while the minor BADH is confined to the cytosol (Weigel et al., 1986). In contrast, both BADH1 and BADH2 from rice contain the tri-peptide C-terminal repeat, SKL, which targets both isozymes to the peroxisomes (Nakamura et al., 1997). Barley does not have this sequence and BBD2 is targeted to the cystol and BBD1 is targeted to the peroxisomes, suggesting they have different functions *in vivo*. In addition, BBD2 catalysis produces glycine betaine whereas BBD1 does not catalyse the formation of glycine betaine indicating they both have different *in vivo* functions (Fujiwara et al., 2008).

1.4 BADH substrate specificity

Enzymes coded by genes which are homologous to BADH have a range of *in vitro* activities. *BADH* has been found to have a relative broad substrate specificity, so the action of BADH is not necessarily only limited to the synthesis of glycine betaine (Trossat et al., 1997). For example, kinetic studies has found 3-dimethylsulfoniopropionaldehyde (DMSP-ald) is a better substrate for amaranth *BADH* than betaine aldehyde (Vojtechova et al., 1997). BADH from *Avena sativa and Zoysia tenuifolia* have shown activity towards a range of substrates including aminoaldehyde-3-aminopropionaldehyde (APAL), 4-aminobutyraldehyde (AB-ald) and 4-guanidiionobutyraldehyde (GBAL) (Livingstone et al., (Oishi and Ebina, 2005).

Genetic analysis has shown spinach BADH is encoded by a single nuclear gene with two alleles that give a homodimer or a hetrodimer with the same molecular mass but different

charges (Weretilnyk and Hanson, 1988). Barley BADH is composed of two isozymes, BBD1 and BBD2, both of which use NAD rather than NADP as a co-factor (Arakawa et al., 1990). BBD2 is the dominant enzyme in leaves while *BBD1* doesn't have affinity towards betaine aldehyde; rather it is specific to the amino aldehyde AB-ald and AP-ald (Fujiwara et al., 2008). Similarly, tobacco transformed with sugar beet BADH is more active on AP-aldehyde and AB-aldehyde than betaine aldehyde (Trossat et al., 1997). Enzymes from pea which show high homology with the N–terminal of sugar beet and spinach BADH, do not oxidize betaine but instead oxidise omega-aminoaldehydes to their corresponding omega-amino acids (Sebela et al., 2000; Brauner et al., 2003) .

Some plants including rice, tomato, arabidopsis and tobacco do not synthesise GB (Zhou et al., 2008) but have contain *BADH* homologs. Both BADH1 and BADH2 from rice have shown higher, yet different, affinity towards the amino aldehyde i.e. γ-aminobutyraldehyde (GABald) and γ -guanidinobutyraldehyde (GGBald) compared to betaine aldehyde. This suggests it is the accumulation of GGBald which results in the formation of 2AP in fragrant rice (Bradbury et al., 2008) (Figure 1.2). BADH2 has higher affinity for GABald with a K_m of 9 μM and high kcat/km of 68 m/s as compared to BADH1 which has K_m of 498 um and kcat/km 27 m/s. This then suggests that the two enzymes which were grouped together based on their nucleotide sequences, differ by their enzyme activity (Bradbury et al., 2008).

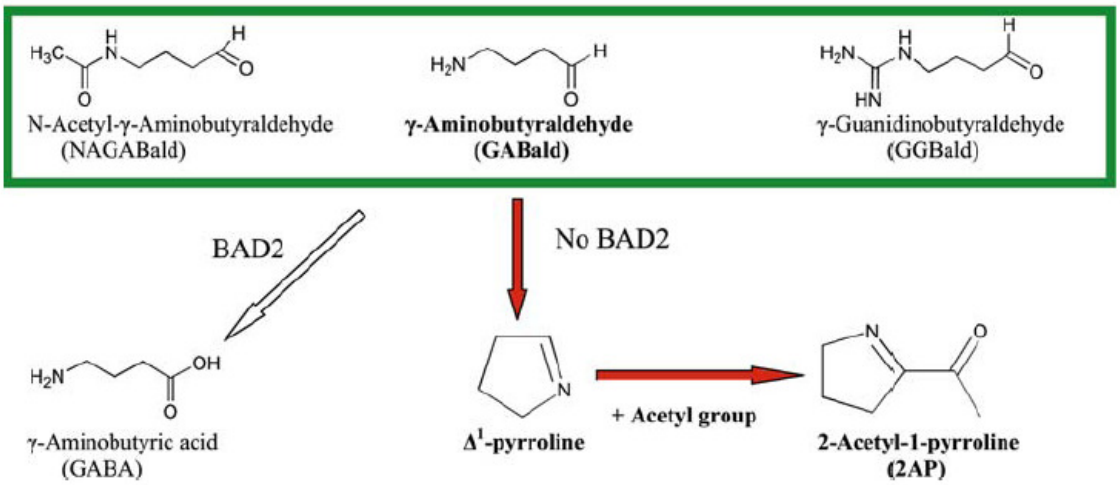

Figure 1.2 Putative biochemical pathways in rice showing the activities of *BAD (BADH)* gene. Non-functional *BAD2 (BADH2)* leads to the formation of 2-acetyl-pyroline and normal *BAD2* gives the product GABA. Modified from (Bradbury et al., 2008).

Although these plants do not naturally accumulate glycine betaine, transgenic tomato, tobacco and rice expressed *BADH* genes introduced from other sources and accumulated glycine betaine. For example, when tomatoes (Jia et al., 2002) and rice (Guo et al., 1997) were transformed with the *BADH* gene from *Atriplex hortensi,* a non-glycine betaine accumulating species, the transgenic plants were more tolerant to salt stress compared to the wild type, indicating the production of the gene product and the associated osmolytes. Transformation of tobacco with the *BADH* gene from *E. coli* resulted in bacterial *BADH* being correctly processed and biologically active allowing the plant to convert supplied betaine aldehyde to glycine betaine and exhibit an increased tolerance to salt stress (Holmstrom et al., 1994). Accumulation of physiologically active glycine betaine in these tobacco plants suggests the native BADH and the introduced BADH, although appearing to be similar enzymes on the basis of amino acid sequence, actually have very different *in vivo* enzyme activities.

1.5 BADH gene expression

BADH gene expression tends to be very sensitive to salt and abiotic stress in most plants and expression patterns vary between different plant species. In sorghum, gene expression in roots was comparatively lower than leaves and there was little difference in GB accumulation in control and salt treated plants (Yang et al., 2003). In wheat, the opposite was observed and there was an increase in accumulation of GB in roots in response to salt (Krishnamurthy and Bhagwat, 1990). Both the roots and leaves of sugar beet accumulate GB and the level of expression of *BADH* increases with increases in salt concentration (McCue and Hanson, 1992). Rice has two *BADH* homologs, *BADH1* and *BADH2*, but their expression patterns differ, *BADH2* is constitutively expressed whereas the expression of *BADH1* is not constitutive. The level of *BADH1* transcripts in fragrant and non-fragrant rice increases in response to salt treatment compared to *BADH2* transcripts, indicating *BADH1* is primarily responsible for salt tolerance in rice (Fitzgerald et al., 2008). Similarly, the expression level of the *BADH* homologs differs in barley and sorghum. The expression of barley *BADH2*, which is mainly expressed in leaf cytosol, is much higher than *BADH1*, which is most highly expressed in leaf peroxisomes, and the catalytic activity of BADH2 is 2000 fold higher than BADH1 when converting betaine aldehyde to GB (Fujiwara et al., 2008). Sorghum *BADH1* transcripts are more abundant than *BADH2* transcripts in leaves (Wood et al., 1996).

1.6 Effect of pH on BADH

A high optimum pH for BADH is reported in many species including rice. The optimum pH for barley BDD1 activity (Fujiwara et al., 2008) was 9.5 which was similar to rice BADH1 (Bradbury et al., 2008) and *Zoysia tenuifolia* ZBD1 (Oishi and Ebina, 2005) while spinach, showed maximum activity at pH 8.5. The activity of BADH2 of barley showed broad specificity from pH 6.5 to 10.5 (Fujiwara et al., 2008). The structure of GABald, the

substrate of BADH, is affected by pH exists almost entirely in the ring form of Δ^1-pyrolline at pH 9 to pH 10. The formation of 2AP is maximal at pH 8.0 (Blank et al., 2003), suggesting there is a relationship between pH, substrate and localisation.

1.7 Domestication of wheat

Wheat and barley are closely related species which belong to the Triticeae in the sub-family. Pooideae of the Poaceae family. The modern species commonly referred to as "wheat" is *Triticum aestivum* L.ssp. *aestivum*; henceforth abbreviated as *T. aestivum*. The two most common *Triticum* species utilised are *T. aestivum* and *Triticum turgidum* L.ssp. *durum* desf (durum wheat). The *T. aestivum* genome is allohexaploid composed of the genomes A, B and D (AABBDD; 2n=42) whereas *T. turgidum* ssp. *durum* is a tetraploid made up of A and B genomes (AABB; 2n =28). Common hexaploid wheat is derived from three different diploid species. The A genome is derived from einkorn wheat which represents both the wild and cultivated varieties and is generally known as *Tritium boeoticum* Bosis. Emend. Schiem. The B genome originated from *Aegilops speltoides* commonly known as wild or weedy goat grass. The D genome was derived from the wild or weedy grass *Aegilops tauschii* L (Sourdille et al., 2001; Feuillet et al., 2008).

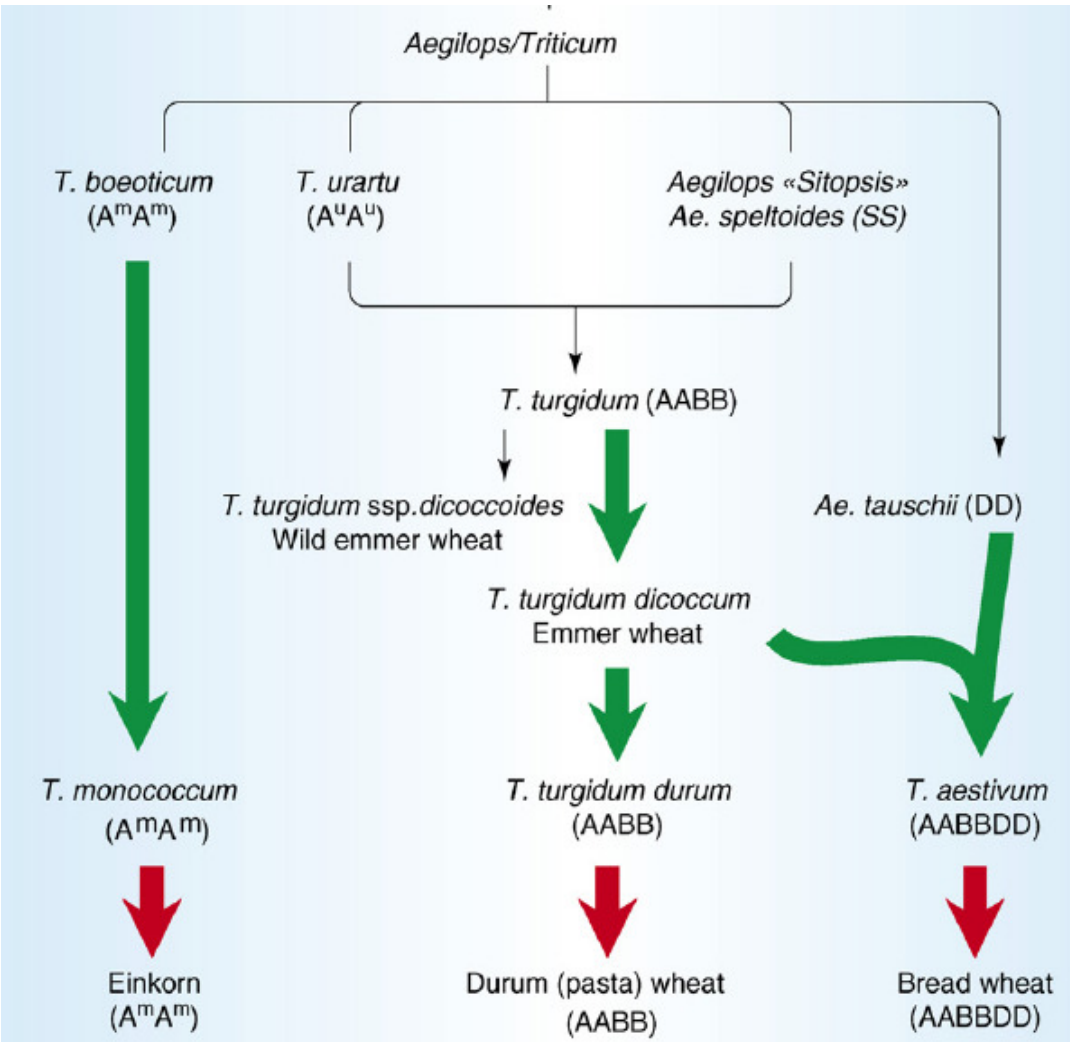

Figure 1.3 Domestication of hexaploid wheat. Modified from ((Feuillet et al., 2008)

1.8 Wheat genome

Polyploidy describes the situation where an organism carries more than one haploid genome. Autopolyploidy and allopolyploidy are the two major pathways which lead to the formation of polyploidy. Autopolyploidy arises through chromosome doubling within a single species while allopolyploidy arises from hybridization and the maintenance of the progenitor genomes within a single species (Chen and Ni, 2006b). Hexaploid wheat is an allopolyploid with three haploid genomes (A, B and D) which are derived from three different diploid species (Feuillet et al., 2008). Although wheat is hexaploid, recombination between homeologs does not take place due to the presence of the Ph1 (pairing chromosome) locus which is situated in long arm of chromosome 5(B) (Riley and Chapman, 1958; Riley and Kempanna, 1963; Sears, 1976).

The 21 haploid chromosomes of hexaploid wheat can be categorised into three groups of seven chromosomes (Table 1.1) representing the genome of each progenitor species (Okamoto, 1962). These homologous chromosomes have undergone little structural change since their integration in to the polyploid genome since they do not recombine or exchange DNA. They have similar genetic content and behave in large part as three parallel independent genomes derived from their progenitors. Chromosome V (5B) in *T. aestivum* plays an important role in maintaining the independence of the three different genomes and absence of this chromosome increases the formation of non-homologous pairing (Riley and Chapman, 1958; Riley and Kempanna, 1963).

Table 1.1 The classification of the chromosome of wheat into the genome and homologous groups (Okamoto, 1962).

Homoeologous Groups	Genomes		
	A	B	D
1	XIV(1A)	I(1B)	XVII (1D)
2	XIII(2A)	II(2B)	XX (2D)
3	XII(3A)	III(3B)	XVI(3D)
4	IV(4A)	VIII(4B)	XV(4D)
5	IX(5A)	V(5B)	XVIII(5D)
6	VI(6A)	X(6B)	XIX(5D)
7	XI (7A)	VII(7B)	XXI(7D)

The number and function of wheat *BADH* homologs is poorly characterised. If there is a possibility of developing fragrant wheat, research is required to identify wheat *BADH* homologs and understand their expression patterns. The overall objectives of this study were to determine the number of *BADH1* and *BADH2* orthologs in hexaploid wheat and their expression patterns in different tissues and time points.

Chapter 2: Identification of *BADH* homologs in hexaploid wheat and its progenitors

2.1 Introduction

Many diploid plant species including rice (Bradbury et al., 2005), barley (Fujiwara et al., 2008), soybean (Arikit et al., 2010) and spinach (Weigel et al., 1986) have two *BADH* paralogs, *BADH1* and *BADH2*. *Zoysia tenuifolia* is unusual because it has only one paralog, ZBD1 (Oishi and Ebina, 2005). In rice, *BADH1* is present in chromosome 4 and *BADH2* is present in chromosome 8. Fragrance in rice is a recessive character, which arises from an eight bp deletion in *BADH2* (Bradbury et al., 2005). Non-fragrant rice varieties have at least one functional *BADH2* allele and fragrant rice varieties are homozygous for the mutant *BADH2* (*badh2*) allele (Bradbury et al., 2005). Fragrance in soybean is due to a two bp deletion in exon 10 of *BADH2* (Arikit et al., 2010).

To date there has been no report of fragrance in wheat. It is therefore important to determine the *BADH* homolog number and their expression patterns in wheat because this will determine the feasibility of replicating fragrance in wheat and how this will impact wheat plant performance.

Rice, barley and wheat belong to the family Poaceae. Although the genome sizes of rice and wheat are different, they share the extensive synteny in a number of their genomic regions (Ahn et al., 1993). Unlike diploid rice, wheat is hexaploid which increases the complexity of the genetics and the potential interactions between the unknown numbers of *BADH* homologs.

The objective of this experiment was to determine the number of *BADH* homologs present in hexaploid wheat and their progenitors and assign *BADH* homologs to the A, B or D genomes in hexaploid wheat.

2.2 Materials and Methods

2.2.1 Sample collection and plant growth

Wheat *T.aestivum* variety Cadoux (ACO3 1002378) and species *T. monococcum* (ACO3 1002479), *T. speltoides*, (AC03 1002994) and *T. tauschii* (AC031003023) were obtained from the Australian Plant DNA Bank, Lismore, Australia (www.dnabank.com.au). Seeds were washed in 5% bleach (4% w/v chlorine present as sodium hypochlorite) for 5 min followed by washing with Milli Q water to remove traces of sodium hypochlorite.

Plants were grown in a controlled growth chamber at a temperature of 21-22°C with 16 h of daylight and 8 h of dark. Cold 6400k fluorescent lamps (Hygrow technologies, Australia) were used for vegetative growth on a 16/8 h light cycle, six lights per m^2 of plant space until the first flag leaf appeared when 2700k cold fluorescent lamps (Hygrow technologies. Australia) were substituted for 50% of the lamps.

2.2.2 DNA extraction

DNA was isolated from fresh and frozen leaf tissue. Fresh leaves were harvested from 1-2 week old plants. The leaves were wiped with 70% analytical reagent grade ethanol to remove any surface contamination and weighed. DNA was extracted using the Qiagen DNeasy Plant kit (Qiagen GMbH, Germany) according to the manufacturer's instructions. The extracted DNA concentration was determined by comparison with 1ng/µl, 2 ng/µl, 5 ng/µl and 10 ng/µl of standard lambda DNA (Roche Diagnostic Australia Pty. Ltd) by running in 1% agarose gel electrophorosis prior to being stored individually at -20°C.

2.2.3 PCR Primer design

BADH2 gene sequence was obtained from GenBank on the NCBI database (http://www.ncbi.nlm.nih.gov/) using the following accession numbers *Oryza sativa*,

(AP005537); *T. aestivum*, (AY050316); *Hordeum vulgare*, (AB063178) using BlastX. *BADH1* sequences were obtained from the NCBI databases with the following accession numbers *Hordeum vulgare* BAD1, (AB161712); *Oryza sativa* BAD1, (AB001348) using BlastX.

The sequences were then aligned with ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html) (Figure 2.1). Primers were designed using Primer3 input 0.4.0 (http://frodo.wi.mit.edu/primer3/) based on the consensus nucleotide sequences for *BADH1* and *BADH2* (Table 2.1).

```
                                                                    A
WheatBADH2        GCTAVLKPSELASVTCLELGDVCKEIGLPSGVLNIVTGLGHEAGAPLSSHPDVDKVAFTG 239
BarleyBADH2       GCTAVLKPSELASVTCLELGDVCKEVGLPSGVLNIVTGLGNEAGAPLSSHPDVDKVAFTG 239
OryzaBADH2        GCTAVLKPSELASVTCLELADVCKEVGLPSGVLNIVTGLGSEAGAPLSSHPGVDKVAFTG 238
HordeumBADH1      GCTAVLKPSELASLTCLELGAICEEIGLPSGVLNIITGLGPDAGAPIASHPHVDKIAFTG 240
OryzaBADH1        GCTAVLKPSELASLTCLELGGICAEIGLPPGVLNIITGLGTEAGAPLASHPHVDKIAFTG 240
                  **************:*****. :* *:***.*****:**** :****::*** ***:****
                                                                          B
WheatBADH2        SYATGQKIMVAAAPTVKPVTLELGGKSPIVVFDDV-DIDKAVEWTLFGCFWTNGQICSAT 298
BarleyBADH2       SYATGQKIMVAAAPTVKPVTLELGGKSPIVVFDDV-DIDKAVEWTLFGCFWTNGQICSAT 298
OryzaBADH2        SYETGKKIMASAAPMVKPVSLELGGKSPIVVFDDV-DVEKAVEWTLFGCFWTNGQICSAT 297
HordeumBADH1      STATGKTIMTAAAQMVKPVSLELGGKSPLVIFDDVADIDKAVEWAMFGCFFNGGQVCSAT 300
OryzaBADH1        STETGKRIMITASQMVKPVSLELGGKSPLIVFDDV-DIDKAVEWAMFGCFANAGQVCSAT 299
                  *  **: ** :*:  ****:********:::**** *::*****::**** . **:****
                                                                        C
WheatBADH2        SRLLIHKNIAKEFVDRMVAWSKNIKVSDPLEEGCRLGPVVSEGQYEKIKKFVANAKSEGA 358
BarleyBADH2       SRLLIHKNIAKEFVDRMVAWSKNIKVSDPLEEGCRLGPVVSEGQYEKIKKFVANAKTEGA 358
OryzaBADH2        SRLILHKKIAKEFQERMVAWAKNIKVSDPLEEGCRLGPVVSEGQYEKIKQFVSTAKSQGA 357
HordeumBADH1      SRLLLHEKIAERFLDRLVEWAKNIKISDPLEEGCRLGSVISKGQYEKIKKFISTARSEGA 360
OryzaBADH1        SRLLLHEKIAKRFLDRLVAWAKSIKISDPLEEGCRLGSVVSEGQYQKIMKFISTARCEGA 359
                  ***::*::**:.* :*:* *:*.**:***********.*:*:***:** :*::.*: :**

WheatBADH2        TILTGGVRPKHLEKGFFIEPTIITDINTSMEIWREEVFGPVLCVKEFSTEEEAIELANDT 418
BarleyBADH2       TILTGGVRPKHLEKGFFIEPTIITDINTSMEIWREEVFGPVLCVKEFSTEEEAIELANDT 418
OryzaBADH2        TILTGGVRPKHLEKGFYIEPTIITDVDTSMQIWREEVFGPVLCVKEFSTEEEAIELANDT 417
HordeumBADH1      TILHGGDRPKNLGKGFFIEPTIITGVSTSMQIWREEVFGPVICVKVFKTESEAVELANDT 420
OryzaBADH1        TILYGGARPQHLKRGFFIEPTIITNVSTSMQIWREEVFGPVICVKEFRTEREAVELANDT 419
                  *** ** **::* :**:*******.:.***:**********:*** * ** **:******
```

```
WheatBADH2        HYGLAGAVISGDRERCQRLAEEIDAGCIWVNCSQPCFCQAPWGGNKRSGFGRELGEGGID 478
BarleyBADH2       HYGLAGAVISGDRERCQR------------------------------------------ 436
OryzaBADH2        HYGLAGAVLSGDRERCQRLTEEIDAGIIWVNCSQPCFCQAPWGGNKRSGFGRELGEGGID 477
HordeumBADH1      HYGLAGGVISDDLERCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWGLE 480
OryzaBADH1        HYGLAGAVISNDLERCERISKAIQSGIVWINCSQPCFVQAPWGGNKRSGFGRELGQWGLD 479
                  ******.*:*. * ***:*
                            D
WheatBADH2        NYLSIKQVTEYTSDAPWGWYKAPAN- 503
BarleyBADH2       -----------------------------
OryzaBADH2        NYLSVKQVTEYASDEPWGWYKSPSKL 503
HordeumBADH1      NYLSVKQVTRYCKDELYGWYQRPSKL 506
OryzaBADH1        NYLSVKQVTKYCSDEPYGWYRPPSKL 505
```

Figure 2.1 ClustalW alignments of their amino acid sequences of *BADH1* and *BADH2* from three different cereal species. Boxed sequence with A and B superscript represents the forward and reverse primer respectively for BADH2 (BADH2P1). Boxed sequence with C and D superscript represents the forward and reverse primer respectively for BADH1 (BADH1P1).

Table 2.1 Primers used for the amplification of BADH homologs.

Homologs, Primer pairs	**Forward primer (5'-3')**	**Reverse primer (5'-3')**	Expected size of amplicon
BADH1, BADHP1	CAGCAAGAAGTGAAGGTGCTAC	GTACCTGGTGACTTGTTTCACG	~750 bp
BADH2, BADH2P1	CATCAGGTGTCTTAAACATTGTGAC	GGATAAGAAGACGAGATGTCGCACT	~1300 bp
BADH2, BADH2P2	CAGGTGTCTTAAACATTGTG	AGGACTTTTTCCACCAAGTTC	~410 bp
BADH2, BADH2P3	CTGCAGCTCCTACAGTCAAG	TGCATAGCTCCCGGTAAATGCAAC	~70 bp

2.2.4 PCR amplification

All amplifications were performed using a Palm-Cycler (Corbertt Research Pty. Ltd, Sydney, Australia). All PCR amplifications were carried out in triplicate along with the negative control. The negative control, contained water in place of genomic DNA.

Optimization of PCR amplification was undertaken with an annealing temperature gradient range of 50 ^{o}C to 70 ^{o}C and magnesium chloride ($MgCl_2$) concentration from 1 mM to 5 mM. Following optimization, all the PCRs were carried out in a 25 µl volume containing 0.4 µM of both forward and reverse primer (Sigma Aldrich, Sydney), 5 mM $MgCl_2$, 1x PCR buffer (Gibco BRL®, Invitrogen), 0.2 unit Platinum® Taq polymerase (Gibco BRL®, Invitrogen), 5 mM of equimolar dNTPs (Promega), 5-8% glycerol and 200 ng of genomic DNA. The PCR program was 95 ^{o}C denaturation for 3 min followed by 30 cycles of 95 ^{o}C for 30 s (denaturation), 60 s at 58 ^{o}C (annealing) and 72 ^{o}C for 1.5 min (extension). PCR products were separated on 1% agarose gel and the gel was stained with ethidium bromide (0.5µg/ml) for 10-15 min and destained in water for 15 min. After destaining, the DNA was visualized on a UV transilluminator.

2.2.4.1 Reconditioning PCR

Reconditioning PCR was undertaken prior to cloning in order to enrich the proportion of homoduplices. Heteroduplices form in the latter stages of PCR when PCR products which differ by sequence are present in a single PCR. If the PCR products are then cloned, the *E. coli* mismatch repair system randomly repairs the mismatches in the heteroduplices and this can generate novel artificial haplotypes (Learn and Grafstrom, 1989; Carraway and Marinus, 1993). By diluting the PCR product and then running a PCR of reduced cycles in the presence of excess primer, heteroduplex formation is minimised. This process is called reconditioning PCR and gives 95% homoduplex formation (Kanagawa, 2003).

Trials of five cycles, three cycles and one cycle of reconditioning PCR were performed. Although three cycles are recommended (Kanagawa, 2003), five cycles of PCR yielded bands of appropriate concentration (Figure 2.4) when visualized using UV light.

The PCR program was 95 ^{o}C denaturation for 3 min followed by five cycles of 95 ^{o}C for 30 s (denaturation), 60 s at 58 ^{o}C (annealing) and 72 ^{o}C for 1.5 min (extension). PCR products were separated on 1% agarose gel and gel was stained with ethidium bromide (0.5µg/ml) for 10-15 min and destained in the solution containing water for 15 min. After destaining the DNA was visualized on a UV transilluminator.

2.2.4.2 PCR cycling for screening positive bacterial clones

The PCR program was 94 ^{o}C denaturation for 4 min followed by 35 cycles of 94 ^{o}C for 30 s (denaturation), 40 s at 55 ^{o}C (annealing) and 72 ^{o}C for 3 min (extension). The 3rd step was carried out at 72 ^{o}C for five min followed by hold of four min at 4 ^{o}C.

2.2.5 Purification of agarose gel slices

PCR products were visualized with a hand held UV light and cut from the gel. The gel slice obtained was purified with a Qiagen (Qiagen GMbH, Germany) gel extraction kit.

2.2.6 Cloning

2.2.6.1 LB Media

LB media were prepared according to technical manual instructions provided with the pGEM®-T Easy Vector Systems (Promega Corporation, USA). The ingredients were then autoclaved at 121 ^{o}C for half an hour and allowed to cool in room temperature. The media were stored at 4 ^{o}C for further use.

2.2.6.2 SOC Media

SOC media were prepared according to technical manual instructions provided by pGEM®-T Easy Vector Systems (Promega Corporation, USA) then autoclaved at 121 ^{o}C for half an hour. Filter sterilized 2M $MgCl_2$ and 2M glucose were then added to 100 ml of autoclaved LB media, which gives final concentration of 20 mM each and stored at 4 ^{o}C.

2.2.6.3 X-gal

5-bromo-4-chloro-3-indolyl–beta-d-galactoside (X-gal) (100 mg) was dissolved in 2 ml of N, N'–dimethyl-formamide, and aliquoted in two 1 ml tube covered with aluminum foil and stored at -20 ^{o}C. The working concentration was 80µg/ml.

2.2.6.4 IPTG

IPTG (1.2 g) was dissolved in 50 ml of milli Q water, filter sterilized and aliquots of 10 ml stored at 4 ^{o}C and 0.5 mM was used as working concentration.

2.2.6.5 Cloning of PCR fragments

Cloning of PCR products was carried out using the pGEM®-T Easy Vector System (Promega Corporation, USA). Each ligation reaction of 10 µl contained 5 µl of 2X Rapid ligation buffer, T4 ligase (3 Weiss units/µl) (0.01 Weiss unit of T4 DNA ligase is defined as the amount of enzyme required to catalyse the ligation of greater than 95% of the *Hind* III fragments of 1µg of lambda at 16 degree in 20 min), 1 µl of pGEM-T Easy Vector 50 ng, 3 µl of the PCR product. The positive control contained 1 µl of the control insert DNA and 1 µl of DNA ligase. The negative control contains all the reagents except T4 ligase and water was added to make up to 10 µl. Transformation was carried out according to manufacturer's manual (Promega Corporation, USA).

2.2.6.6 Gel slice cloning of PCR fragments

Cloning was carried out using the reagents obtained from the pGEM-T Easy Vector System® (Promega Corporation, USA). Each 11 µl reaction contained 5 µl of 2X Rapid ligation buffer, T4 ligase (3 Weiss units/µl), 1µl of pGEM-T Easy Vector (50 ng), and 4 µl of low melting gel sliced re-conditioning PCR product. The positive control contained 2 µl of control insert DNA and 1 µl of DNA ligase (3 Weiss units/µl) (Promega Corporation, USA). The negative control contained all the reagents except T4 ligase and water was added to make the volume 11 µl. Transformation was carried out according to the manufactures protocol (Promega Corporation, USA).

2.2.6.7 PCR screening of bacterial colonies

White colonies were picked and dissolved in 100 ml TE buffer pH 8.0 and PCR screening was undertaken using M13 Forward and Reverse primers as stated in page 19 (PCR cycling for screening positive bacterial clones).

2.2.6.8 Templiphi amplification of colonies

TempliPhi™ (GE Healthcare, Australia) uses rolling circle amplification by bacteriophage lambda Phi 29 DNA polymerase for the amplification of bacterial DNA. Initially, TempliPhi™ denaturation buffer and premix were thawed on ice. Colonies were picked with steel needles were placed in 5 µl of denaturation buffer. The plate was agitated so that all colonies dissolved in the buffer. The sample was denatured at 95 ^{o}C for 5 min cooled on ice and then 5 µl premix was added and incubated overnight at 30 ^{o}C. The following day the samples were denatured at 95 ^{o}C for 10 min then diluted with 40 µl of Milli Q water.

2.2.7 Sequencing reaction protocol

Sequencing was carried out using Big Dye Terminator V3 (Applied Biosystems) according to the manufacturer's instructions. Briefly, 2 µl of Big Dye terminator, 2 µl of AB Buffer sequencing buffer, primer (0.41µM), 5 µl of template and 2 µl of Milli Q water were added in PCR tubes.

2.2.8 Post -sequencing reaction purifications

Sequencing reactions were transferred to 1.5 ml tubes and 3.2mM EDTA, 0.07M sodium acetate (pH 5.0) and 5 µl of MQ water and 60 µl of 100% ethanol was added to the tube and incubated at room temperature in the dark for 15-20 min and then centrifuged for 30 min at 14,000 rpm. The supernatant was discarded and the DNA pellet was washed with 100 µl of 70% ethanol. Following centrifugation for 5 min at 14,000 rpm, the supernatant was discarded and dried for 5 min at room temperature. The samples were submitted to Southern Cross Plant Genomics (Southern Cross University, Lismore) for the separation of the sequence reactions.

2.2.9 Sequence analysis and alignment

Sequences were analysed using Sequencher 4.10.1 (Gene Code Corporation, USA) with default settings. SNPs were identified by visual inspection of chromatograms.

Cadoux BADH homologs were assigned to each genome (A, B and D) on the basis of the phylogenetic relationship between the hexaploid wheat progenitor BADH DNA sequence and Cadoux *BADH* homologs DNA sequence. Alignment was carried out with ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) using the Neighbor-Joining clustering with default settings. A bootstrap value of 1000 was used to generate the node value. TreeView™ Win 32 versions 1.6.6 was used with default settings in order to generate the phylogenetic

relation of hexaploid wheat and their progenitors. For the prediction of amino acid sequences, ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and MEGA 4.0™ (Tamura et al., 2007) were used with default settings.

2.3 Results

2.3.1 Genomic DNA extraction

Genomic DNA extracted from the young leaves was colourless and readily re-dissolved in water and exhibited no shearing and minimal contamination as analysed by agarose gel electrophoresis.

2.3.2 *BADH1* and *BADH2* amplification

BADH1 and *BADH2* of the hexaploid wheat variety Cadoux were amplified with the primer pairs BADH1P1 and BADHP2 respectively and visualized on a UV transilluminator. A *BADH1* fragment of the expected size of ~750 bp (Figure 2.2) and a BADH2 fragment of ~1300 bp (Figure 2.3) were amplified.

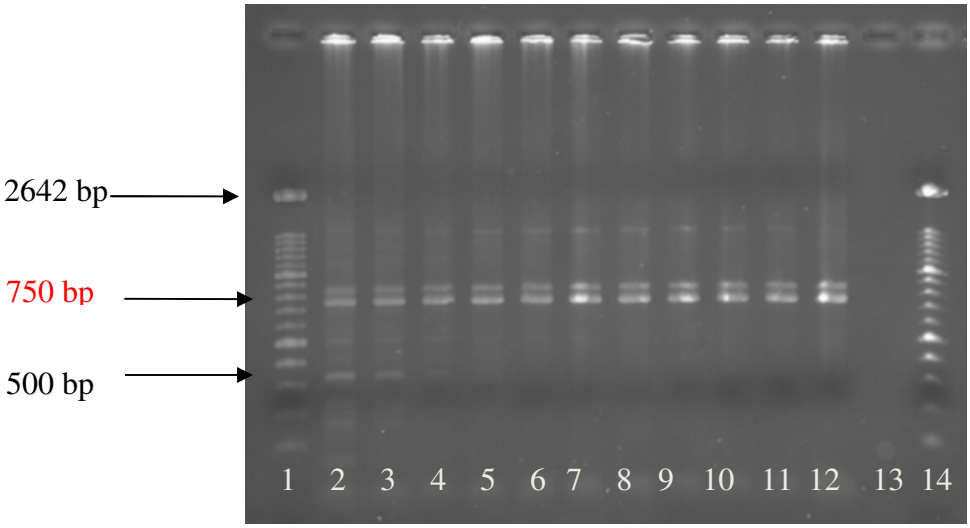

Figure 2.2 *BADH1* fragments from wheat which corresponds to AA353 to AA491 of barley. Lane 1 and 14 is Molecular marker XIV (Roche Diagnostics, Australia). Lanes 2-12 are *BADH1* amplicons of hexaploid wheat variety Cadoux. Lane 13 is negative control.

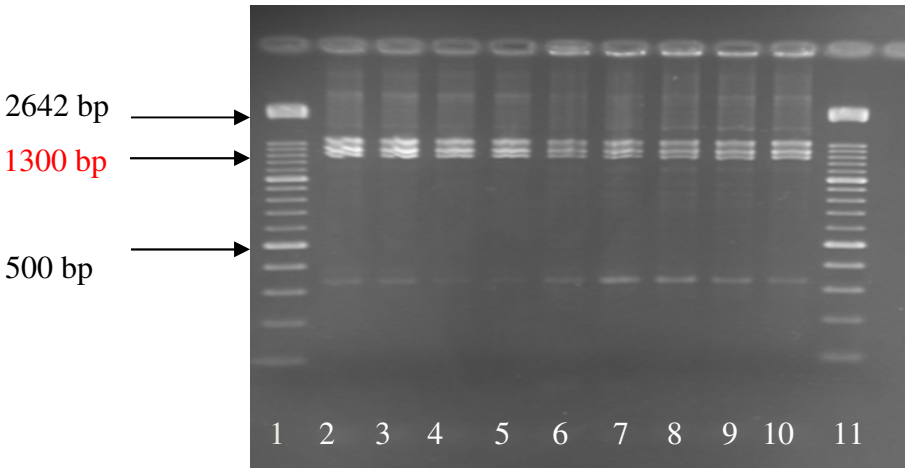

Figure 2.3 PCR amplified wheat *BADH2* gene fragments which code for wheat BADH2 protein (AY050316) between AA226 and AA302. Lane 1 and 11 is Molecular marker XIV (Roche Diagnostics Australia). Lanes 2 to 10 are *BAHD2* of hexaploid wheat variety Cadoux.

2.3.3 Gel slice purification and direct sequencing of hexaploid wheat variety Cadoux

PCR of hexaploid wheat gDNA using primers BADH2P1 generated three bands as visualized by ethidium bromide staining of an agarose gels (Figure 2.3). The bands were cut from the gel, purified and sequenced directly. The smaller band yielded sequence data which showed 40% identity with wheat *BADH2* (AAL05264.1) and barley *BADH2* (ABC86863.1) and 82% identity with rice (ADW27189.1) with E-values 4e-06 for barley and wheat and 0.005 for rice as determined by BlastX in NCBI. The remaining two bands generated sequence data which suggesting there was cross contamination during the cutting of the bands. These two bands were much closer to each other and were difficult to separate. Further gel splice purification and direct sequencing was not undertaken for *BADH1* because attempts using this approach

to sequence the *BADH2* fragments did not give unambiguous results. Cloning was necessary to obtain sequences from these bands.

2.3.4 Reconditioning PCR

BADH1 and *BADH2* bands were visualized on a UV transilluminator after reconditioning PCR trials (Figure 2.4).

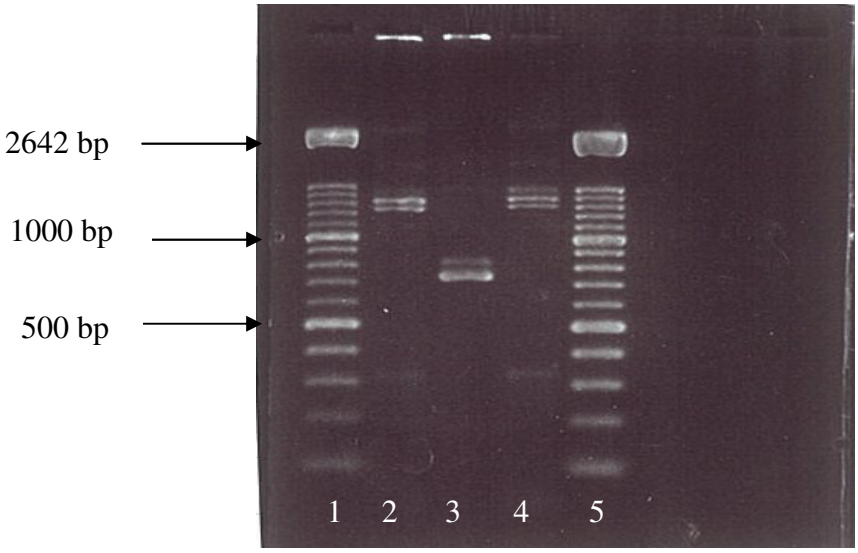

Figure: 2.4 Five cycle reconditioning PCR for *BADH1* and *BADH2* of hexaploid wheat variety Cadoux. Lane 1 and 4 is 100 bp Molecular marker XIV, lane 2 and 4 are *BADH2*, lane 3 is *BADH1*.

2.3.5 Clone screening and sequencing

Purified *BADH1* and *BADH2* PCR products were cloned and the colonies screened by PCR with M13 forward and reverse primers. Forty five colonies were screened for *BADH1* and 45 colonies were screened for *BADH2*. Of the screened colonies which had an insert of the right size following electrophoresis, 30 were sequenced and the sequences aligned. Inspection of aligned sequences identified three different consensus sequences for both *BADH1* and *BADH2*. These consensus sequences were used to interrogate GenBank by BlastX. Both the *BADH1* and *BADH2* genes were found to be very similar to rice and barley *BADH1* and

BADH2 genes. Alignment of *BADH2* with intron sequence removed revealed 100% identity with wheat *BADH2* (AAL05264.1) with an E-value of 4e-07, 100%, identity with barley *BADH2* (BAB62846.1) with an E-value of 4e-07 and 92% identity with rice *BADH2* (AB184118.1) with an E-value of 1e-06 at the nucleotide level. Whereas at the amino acid level, BADH2 showed 100% identity with barley (AB063178.1), 98% identity (100% positive) with wheat (AY050316.1) and 84% identity (92% positive) with rice (AK060461.1) with E–values of 5e-20, 2e-19 and 3e-16 respectively (0% gaps). Similarly, *BADH1* showed 97% identity, with barley (AB063179.1) with an E-value of 1e-72 and 83% identical with rice (AK103582.1) with an E-value of 9e-64. At the amino acid level, wheat BADH1 was 98% identity with barley BADH1 (BAB62847.1) and 84% identity with rice BADH1 (ABB83473.1) with E-values of 6e-73 and 2e-62 respectively (0% gaps).

2.3.6 *BADH1* and *BADH2* amplification from progenitors

BADH1 fragments of the expected size of ~750 bp were amplified from the wheat progenitor species *T. monococcum*, *T. speltoides* and *T. tauschii* (Figure 2.5) with the primer pair BADH1P1 (Table 2.1). A faint non-specific band in the region of 300 bp in lane 3 was not sequenced since its size was significantly less than the expected band size.

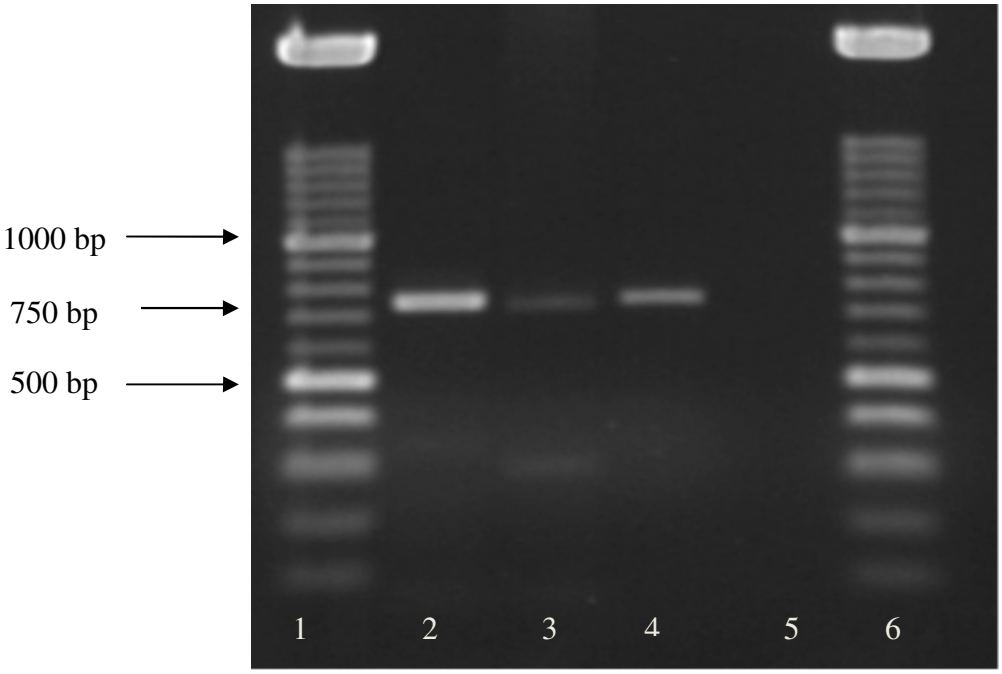

Figure 2.5 *BADH1* fragments from wheat progenitors which correspond to AA353 to 491AA of barley. Lane 1 and 6 are 100 bp Molecular Marker XIV (Roche Diagnostics, Australia). Lane 2 is *T. monococcum*, lane 3 is *T. speltoides*, and lane 4 is *T. tauschii*. Lane 5 is negative control.

The wheat progenitors *BADH2* were amplified with the primer pair BADH2P2 (Table 2.1) and gave a band of ~750 bp as expected for all the three progenitors species (*T. monococcum*, *T. speltoides* and *T. tauschii*) when visualized on a UV transilluminator (Figure 2.6). There was a faint band a ~300bp. There was band ~300 bp in lane 3. This was not near the expected size and was therefore most likely a non-specific PCR product.

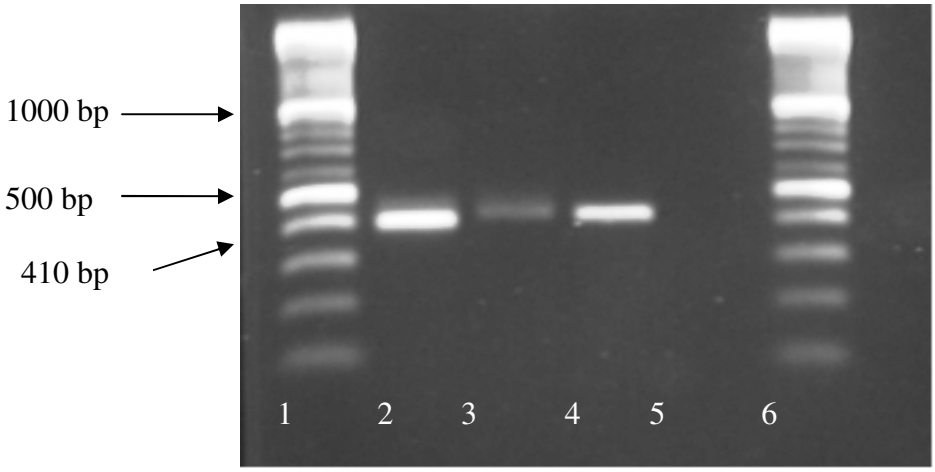

Figure 2.6: *BADH2* gene fragments from wheat progenitors. Lane 1 and 6 is 100 bp ladder (Molecular marker XIV (Roche Diagnostics, Australia). Lane 2 is *T. monococcum*, lane 3 is *T. speltoides*, and lane 4 is *T. tauschii*. Lane 5 is negative control.

The A genome is believed to be derived from *T. monococcum*, the B genome from *T. speltoides* and the D genome from *T. tauschii* (Feuillet et al., 2008). *BADH1* gene fragments from hexaploid wheat were assigned to each of the genomes A, B and D by phylogenetic analysis of *BADH1* gene fragments amplified from hexaploid wheat variety Cadoux and their progenitors (Figure 2.7). The progenitors are not the immediate progenitors of hexaploid wheat (variety Cadoux) and because of this the *BADH1* amplicon derived from genome B might have diverged significantly from the *T. speltoides* ancestral copy of the gene.

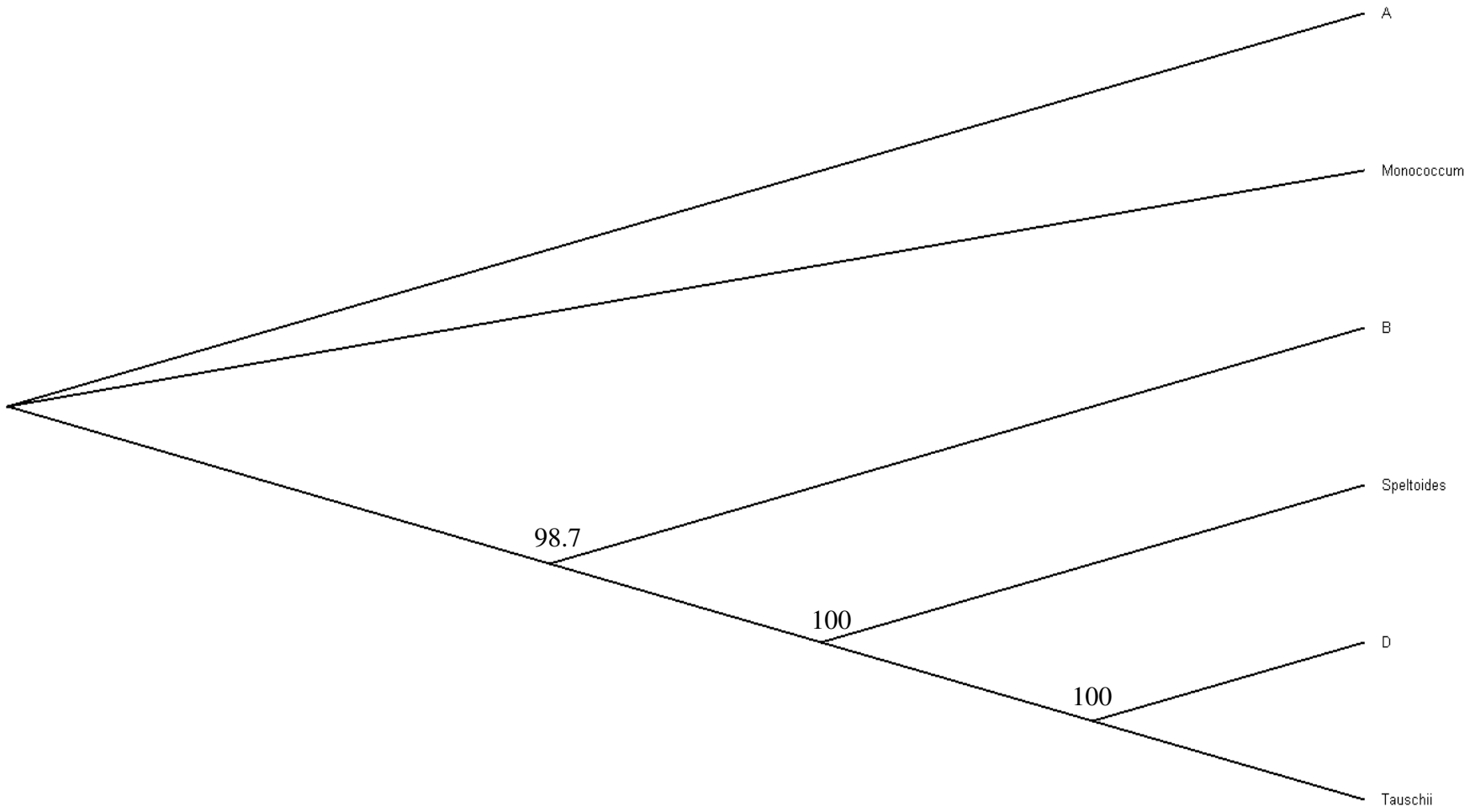

Figure 2.7 Phylogenetic relationships of *BADH*1 derived hexaploid wheat variety Cadoux's genome and its progenitors based on nucleotide sequences with TreeView ™. Analysis was carried out in ClustalW using Neighbor-Joining method.

ClustalW alignment of *BADH1* shows that there are high levels of similarity among the three genomes and their progenitors (Figure 2.8). There was seven SNP in exon 12, two SNP in exon 13 and six SNP in exon 14 between the hexaploid wheat homeologs and their progenitors. The C/G SNP between homeologs at nucleotide position 397 of *BADH1* amplified with primer BADH1P1 (Table 2.1) changes the amino acid from glutamine to glutamic acid. BlastN comparison of these sequences with nucleotide sequences deposited in NCBI (http://www.ncbi.nlm.nih.gov) showed 97% identity with barley *BADH1* (BAB62847.1) with an E-value of 6e-73 and 83% identity with rice *BADH1* (ABB83473.1) with an E-value of 2e-62. TblastN comparisons of these amino acid sequences deposited in NCBI showed 97% identity with barley (AB063179.1) with an E-value of 1e-72 and 84% identical with rice (AK103582.1) with an E-value of 9e-64, confirming the sequences amplified with the primer BADH1P1 were the correct paralogs of *BADH1*.

```
A               AGGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGGTAAGAATTGGTCATTACATTA- 59
B               AGGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGGTAAGCATTGGTCATTACATTA- 59
D               AGGTGCTACAATTTTGCATGGGGGTGACCGACCAAAGGTAAGCATTGATCACTACATTAT 60
Monococcum      AGGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGGTAAGCATTGGTCATTACATTA- 59
Speltoides      AGGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGGTAAGCATTGGTCATTACATTA- 59
Tauschii        AGGTGCTACAATTTTGCATGGGGGTGACCGACCAAAGGTAAGCATTGATCACTACATTAT 60
                ********************* ******************* **** *** *******

A               ----CATGCTCTTGCCTCTCGGTTATTCATCCAAGCTACGTTTGTGAAATGGCAGCATCT 115
B               ----CATGCTCTTGTCTCTCGGTTATTCATCCAAGCTACGTTTGTGAAATGGCAGCATCT 115
D               ATTACATGCTCTTGTCTCTCGGTTATTCATCCAAGCTACGTTTGTGAAATGGCAGCATCT 120
Monococcum      ----CATGCTCTTGCCTCTCGGTTATTCATCCAAGCTACGTTTGTGAAATGGCAGCATCT 115
Speltoides      ----CATGCTCTTGTCTCTCGGTTATTCATCCAAGCTACGTTTGTGAAATGGCAGCATCT 115
Tauschii        ATTACATGCTCTTGTCTCTCGGTTATTCATCCAAGCTACGTTTGTGAAATGGCAGCATCT 120
                    ** ******** ********************************************

A               CGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGACGTTAGCACATCAATGCAAAT 175
B               CGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGATGTTAGCACATCAATGCAAAT 175
D               CGGAAAAGGGTTCTTTATTGAACCTACTATTATAACAGACGTTAGCACATCAATGCAAAT 180
Monococcum      CGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGACGTTAGCACATCAATGCAAAT 175
Speltoides      CGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGACGTTAGCACATCAATGCAAAT 175
Tauschii        CGGAAAAGGGTTCTTTATTGAACCTACTATTATAACAGACGTTAGCACATCAATGCAAAT 180
                *************************** *********** ********************
```

Figure 2.8 Alignment of primer pair BADHP1 *BADH1* amplicons derived from hexaploid wheat variety Cadoux and its progenitor species.

```
A               TTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGTGA 235
B               TTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGCGA 235
D               TTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGTGTCAAAGTATTTAAGACAGAGAGCGA 240
Monococcum      TTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGTGA 235
Speltoides      TTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGCGA 235
Tauschii        TTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGTGTCAAAGTATTTAAGACAGAGAGCGA 240
                ********************************************** *********** **

A               AGCTGTAGAGCTTGCAAATGATACCCAGTGAGTTACCTGATTTGCGCAAGCCTGAAAAAA 295
B               AGCTGTAGAGCTTGCAAATGATACCCAGTGAGTTACCTGATT-GCGCAAGCCTGAAAAAA 294
D               AGCCGTAGAGCTTGCAAATGATACCCAGTGAGTTAACTGATTTGCGCAAGCCTGAAGAAA 300
Monococcum      AGCTGTAGAGCTTGCAAATGATACTCAGTGAGTTACCTGATTTGCGCAAGCCTGAAAAAA 295
Speltoides      AGCTGTAGAGCTTGCAAATGATACCCAGTGAGTTAACTGATTTGCGCAAGCGTGAAAAAA 295
Tauschii        AGCCGTAGAGCTTGCAAATGATACCCAGTGAGTTAACTGATTTGCGCAAGCCTGAAGAAA 300
                *** ******************** ********** ****** ******** **** ***

A               AA--GCCCGTGCTTATRTCGTTCCCTTGACTTCYCTACCATTTTGTTTTCAGCTATGGCT 353
B               AAATGCCCGTGCTTATATCGTTCCCTTGACTTCTCTGCCATTTTGTTTTCAGCTATGGCT 354
D               AAACTTCCATGCTTATATCACTCCCTTGACTTCCCTACCATTTTATTTTCAGCTATGGCT 360
Monococcum      AA--GCCCGTGCTTATATCGTTCCCTTGACTTCTCTACCATTTTGTTTTCAGCTATGGCT 353
Speltoides      AA-CTCCCATGCTTACATCATTCCCTTGACTTCCCTACCATTTTGTTTTCAGCTATGGCT 354
Tauschii        AAACTTCCATGCTTATATCACTCCCTTGACTTCCCTACCATTTTATTTTCAGCTATGGCT 360
                **    ** ******  **  ************ ** ******* ***************

A               TGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGCATTGCAAAGGTAGATT 413
B               TGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGCATTGCAAAGGTAGATT 414
D               TGGCTGGTGGTGTGATCTCTGATGATCTAGAGAGGTGTGAGCGCATTGCAAAGGTAGATT 420
Monococcum      TGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGCATTGCAAAGGTAGATT 413
Speltoides      TGGCTGGTGGTGTGATCTCTGATGATCTAGAGAGGTGTGAGCGCATTGCAAAGGTAGATT 414
Tauschii        TGGCTGGTGGTGTGATCTCTGATGATCTAGAGAGGTGTGAGCGCATTGCAAAGGTAGATT 420
                ****************************** ******************************

A               CAAG------------CACGCCCGTAAATTTGTAGCGTGTGATGCAATTGAAGTATTCGA 461
B               CAAG------------TGCGCCTGTAAATTTGTAGCATGTGATGCAATCGAAGTATTCGA 462
D               CAAGTTGAACTCTGAACATGCACGTAAATTTTGAATGTGTGATGAGAT---------CGA 471
Monococcum      CAAG------------CACGCCCGTAAATTTGTAGCGTGTGATGCAATTGAAGTATTCGA 461
Speltoides      CAAG------------CACGCCCGTAAATTTGTAGCGTGTGATGCAAT---------CGA 453
Tauschii        CAAGTTGAACTCTGAACATGCACGTAAATTTTGAACGTGTGATGAGAT---------CGA 471
                ****               **  ********  *   *******  **         ***

A               AATATTCAGTGAGCTGATAGGTAAACGTACAATCCCACACCTGCAGGTTATTCACTCAGG 521
B               AATATTCGGTGAGCTGATAGGTAAACGTACAATCCCACACCTGCAGGTTATTCACTCAGG 522
D               AGTATTTGGTGAGCTGATAGGTAAACGTACAAACCCACACCTGCAGGTTATTCACTCGGG 531
Monococcum      AATATTCAGTGAGCTGATAGGTAAACGTACAATCCCACACCTGCAGGTTATTCACTCAGG 521
Speltoides      ATTATTCGGTGAGCTGATAGGTAAACGTACAATCCCACACCTGCAGGTCATTCACTCTGG 513
Tauschii        AGTATTTGGTGAGCTGATAGGTAAACGTACAAACCCACACCTGCAGGTTATTCACTCGGG 531
                * ****  ************************ *************** ******** **

A               CATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTTCAAGCTCCGTGGGGAGGGAACAA 581
B               CATTGTTTGGATAAACTGCTCGCAACCGACCCTTGTTCAAGCTCCGTGGGGAGGGAACAA 582
D               CATTGTTTGGATAAACTGCTCGCAACCGACTCTGGTTCAAGCTCCGTGGGGAGGGAACAA 591
Monococcum      CATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTTCAAGCTCCGTGGGGAGGGAACAA 581
Speltoides      CATTGTTTGGATAAACTGCTCGCAACCGACTCTGGTTCAAGCTCCGTGGGGAGGGAACAA 573
Tauschii        CATTGTTTGGATAAACTGCTCGCAACCGACTCTGGTTCAAGCTCCGTGGGGAGGGAACAA 591
                ****************************** ** *************************
```

Figure 2.8 continued Alignment of primer pair BADHP1 *BADH1* amplicons derived from hexaploid wheat variety Cadoux and its progenitor species.

```
A               GCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGTGAGTT-ACTTCACGAGCGCTATGGT 640
B               GCGGAGTGGTTTTGGCCGGGAGCTAGGAGAATGGTGAGTT-ACTTCATGAGCGCTATGGT 641
D               GCGTAGTGGTTTTGGCCGGGAGCTAGGAGAATGGTGAGTTGACTTCATGAGCGCTATGGT 651
Monococcum      GCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGTGAGTT-ACTTCACGAGCGCTATTGT 640
Speltoides      GCGTAGTGGTTTTGGACGGGAGCTAGGAGAATGGTGAGTTGACTAAATGAGCGCTATGGT 633
Tauschii        GCGTAGTGGTTTTGGCCGGGAGCTAGGAGAATGGTGAATTGACTTCATGAGCGCTATGGT 651
                *** ** ******** ********************* ** ***  * ********* **

A               TTGCAAACTGCACTGCCCCTGGAAGTATCCTCTGAACTGTCTTTGCTGTTTTGCTTTCAG 700
B               TCGCAAACTGCACTGCCCCTGAAAGTATCTTCTGAACTGTCT--GCTGTTTTGCTTTCAG 699
D               TCACAAACTGCA--------GCTAGTATCCTCTGAACCATCTTTGCTGTTTTGCTTTCAG 703
Monococcum      TCGCAAACTGCACTGCCCCTGGAAGTATCCTCTGAACTGTCTTTGCTGTTTTGCTTTCAG 700
Speltoides      TCACAAACTGCA--------GCTAGTATCCTCTGAACCGTCTTTGCTGTTTTGCTTTCAG 685
Tauschii        TCACAAACTGCA--------GCTAGTATCCTCTGAACCATCTTTGCTGTTTTGCTTTCAG 703
                *  *********        *  ****** *******  ***  ****************

A               GGGCCTCGAGAAC--------------- 713
B               GGGCCTCGAGAACT-------------- 713
D               GGGCCTCGAG------------------ 713
Monococcum      GGGCCTCGAGAAC--------------- 713
Speltoides      GGGCCTCGAGAACTACCTGAGCGTGAAA 713
Tauschii        GGGCCTCGAG------------------ 713
                *********
```

Figure 2.8 continued Alignment of primer pair BADHP1 *BADH1* amplicons derived from hexaploid wheat variety Cadoux and its progenitor species.

ClustalW alignment showed there was no difference in the amino acid sequences (Figure 2.9) except at position 76 where the amino acids are either glutamine of glutamic acid. Genome A, B and *T.monococcum* have glutamine whereas genome D, *T. speltoides* and *T. tauschii* contain glutamic acid. Comparison with pea ALDH 3D structure suggests this change in amino acid has little effect on protein conformation.

```
A               ARSEGATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAV 60
B               ARSEGATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAV 60
D               ARSEGATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAV 60
Monococcum      -------ILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAV 53
Speltoides      ARSEGATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAV 60
Tauschii        ----GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAV 56
                       ***************************************************

A               ELANDTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGREL 120
B               ELANDTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGREL 120
D               ELANDTHYGLAGGVISDDLERCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGREL 120
Monococcum      ELANDTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGREL 113
Speltoides      ELANDTHYGLAGGVISDDLERCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGREL 120
Tauschii        ELANDTHYGLAGGVISDDLERCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGREL 116
                *******************:****************************************

A               GEWGLENYLSVKQVTR 136
B               GEWGLENYLSVKQVTR 136
D               GEWGLENYLSVKQVTR 136
Monococcum      GEWGLENYLSVKQVTR 129
Speltoides      GEWGLENYLSVKQVTR 136
Tauschii        GEWGLENYLSVK---- 128
                ************
```

Figure 2.9 ClustalW alignments of *BADH1* amino acid sequences of hexaploid wheat and their progenitors. Change in amino acid is colour in red.

Phylogenetic analysis of *BADH2* gene fragments from hexaploid wheat variety Cadoux and progenitor's species was used to assign *BADH2* gene fragments to genomes A, B and D (Figure 2.10).

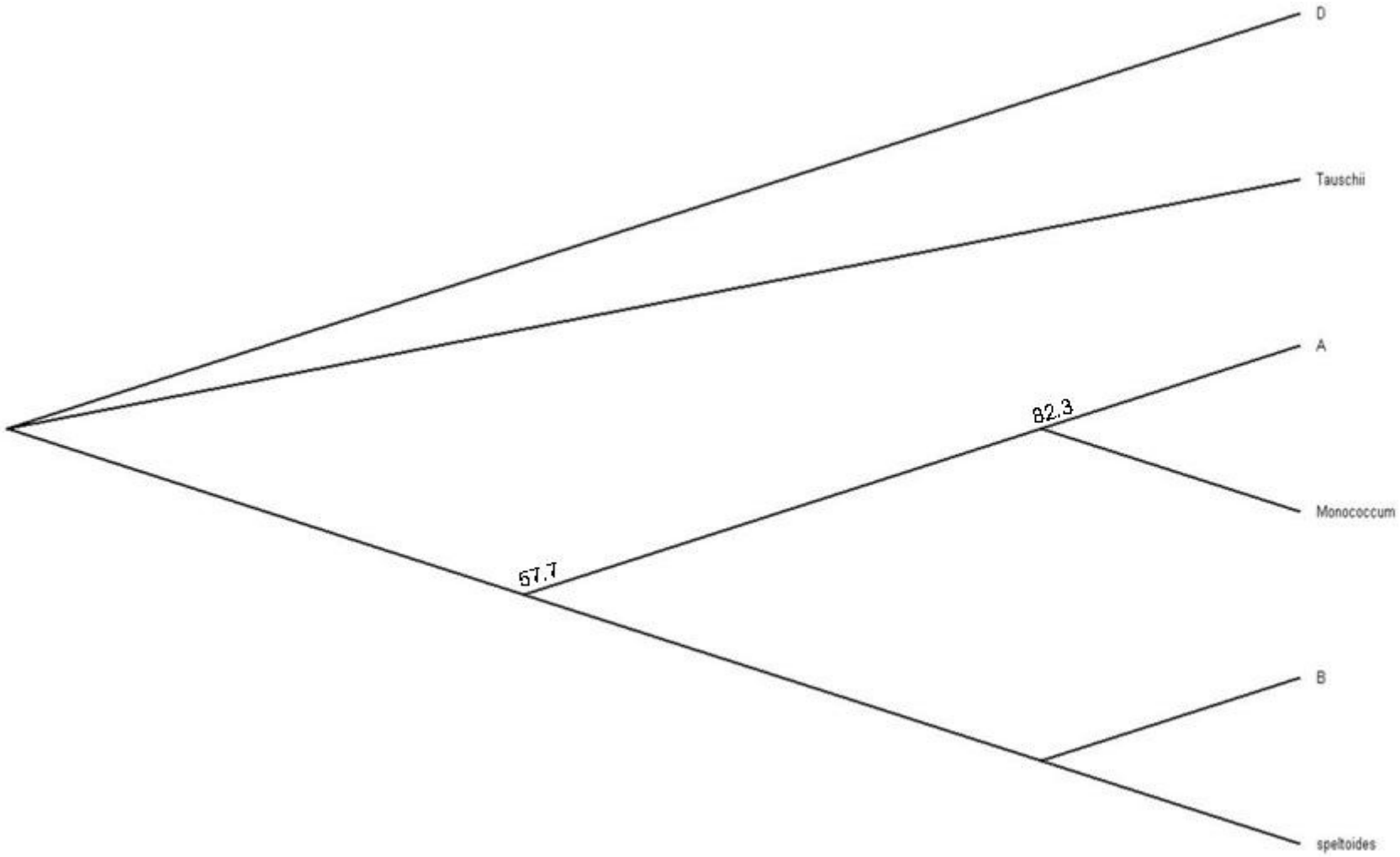

Figure 2.10 Phylogenetic relationships of BADH2 derived hexaploid wheat variety Cadoux's genomes and its progenitors based on nucleotide sequences with TreeView^{TM}. Bootstrap values are provided for clades with greater than 50% supports. Analysis was carried out in ClustalW.

The *BADH2* sequences derived from genomes A, B and D showed high homology with other betaine aldehyde dehydrogenase encoding genes from other species as measured by BlastN in NCBI databases with 100% identity with wheat (AAL05264.1), 100% with barley (BAB62846.1) and 92% with rice (AB184118.1). The total length of *BADH2* amplified was expected to be around 1300 bp. The A genome ortholog was found to be 1385 bp, the B genome 1374 bp and D genome 1297 bp. BADH2P1 primer pair did not generate a PCR product in the progenitor species so another pair of primers, BADH2P2 (Table 2.1), was designed and used to amplify these sequences in the progenitor species. This amplicon was targeted to exon 7 and was around 410 bp in the progenitor species. There were no differences found between the homeologs of hexaploid wheat variety Cadoux and its progenitors in this exon.

```
A               TCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCCTG 60
B               TCTTAAACATTGTGACTGGATTAGGTCATGAAGCTGGCGCTCCTTTGTCGTCACACCCTG 60
D               TCTTAAACATTGTGACTGGATTAGGTCATGAAGCTGGCGCTCCTTTGTCGTCACACCCTG 60
Monococcum      TCTTAAACATTGTGACTGGATTAGGTCATGAAGCTGGCGCTCCTTTGTCGTCACACCCTG 60
Speltoides      TCTTAAACATTGTGACTGGATTAGGTCA-GAAGCTGGCGCTCCTTTGTCGTCACACCCTG 59
Tauschii        TCTTAAACATTGTGACTGGATTAGGTCATGAAGCTGGCGCTCCTTTGTCGTCACACCCTG 60
                ************************** * ******************** **********

A               ACGTCGACAAGGTACATATATTTGTCCTATATCTTTTTGTGATCCATCTATGTCGATGGA 120
B               ATGTCGACAAGGTACATATATTTGTCCTATATCTTTTTGTGATCCATCTATGTCGATGGA 120
D               ATGTCGACAAGGTACATATATTTGTCCAATATCTTTTTGTGATCCATCTATGTCGATGGA 120
Monococcum      ACGTCGACAAGGTACATATATTTGTCCTATACCTTTTTGTGATCTATCTATG-------- 112
Speltoides      ATGTCGACAAGGTACATATATTTGTCCTATATCTTTTTGTGATCCATCTATGTCGATGGA 119
Tauschii        ATGTCGACAAGGTACATATATTTGTCCTATATCTTTTTGTGATCCATCTATGTCGATGGA 120
                * ************************* *** ************ *******

A               GTATATTCCTCACAATACATGGCTTATTTGAGCTTGTAGGTTGCATTTACCGGGAGCTAT 180
B               GTATATTCCTCACAATACATGGCTTATTTGTGCTTTTAGGTTGCATTTACCGGGAGCTAT 180
D               GTATATTCCTCACAATACATGGCTTATTTGTGCTTTTAGGTTGCATTTACCGGGAGCTAT 180
Monococcum      GGTTATTCCTCACAATACATGGCTTATTTGATCTTTTAGGTTGCATTTACCGGGAGCTAT 172
Speltoides      GTATATTCCTCACAATACATGGCTTATTTGTGCTTTTAGGTTGCATTTACCGGGAGCTAT 179
Tauschii        GTATATTCCTCACAATACATGGCTTATTTGTGCTTTTAGGTTGCATTTACCGGGAGCTAT 180
                *  **************************  *** ************************
```

Figure 2.11 Alignment of primer pair BADH2P2 *BADH*2 amplicons derived from hexaploid wheat variety Cadoux and its progenitor species.

```
A               GCAACCGGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGGTTTGTTCCACATGT 240
B               GCAACCGGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGGTTTGT-CCACATGT 239
D               GCAACCGGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGGTTTGTTCCACATGT 240
Monococcum      GCAACCGGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGGTTTGTTCCACATGT 232
Speltoides      GCAACCGGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGGTTTGTTCCACATGT 239
Tauschii        GCAACCGGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGGTTTGTTCCACATGT 240
                ************************************************ ********

A               CCCTATTCTTTTAATCTTTGTAGTAATTTATCACCAAGTATCCTTTTTTCCCCTCCTTTA 300
B               CTCTATTTTTTAAATCTTTGTAGTAATTTATCACCAAGTATCTGTTTTCCCCCTCCTTTA 299
D               CTCTATTTTTTCAATCTTTGTAGTAATGTATCACCAAGTATCTGTTTTTCCCCTCCTTTA 300
Monococcum      CTCTATTTTT------------GAAAATTTTCACCAAATATCCCGTTTTTTCC-CCTCTA 279
Speltoides      CTATATTTTTTTAATCTTTGTAGTAATGTATCACCAAGTATCTGTTTTCCCCCTCCTTTA 299
Tauschii        CTCTATTTTTTAAATCTTTGTAGTAATGTATCACCAAGTATCTGTTTTTCCCCTCCTTTA 300
                *  **** **            * **  * ******* ****   ***   ** *** **

A               GGTTTTTAAAATATTGACCAAGATACTAC-ACCGCTCTCCTGATAGCAGACCAA-CTCAC 358
B               GATTTTTAGAATATTGACCAAGATACTACGAGCGCTCTCCTGATAGCAGACCATTCTCAC 359
D               GATTTTTAGAATATTGACCAAGATACTACGAGCGCTCTCCTGATAGCAGACCATTCTCAC 360
Monococcum      GATTTTTAGAATATTTACTAAGATACTATGAGCGCTCTCCTGATAGCAGACCATTCTCAC 339
Speltoides      GATTTTTAGAATATTGACCAAGATACTACGAGTGCTCTCCTGATAGCAGACCATTCTCAC 359
Tauschii        GATTTTTAGAATATTGACCAAGATACTACGAGCGCTCTCCTGATAGCAGACCATTCTCAC 360
                * ****** ****** ** *********  *  ********************  *****

A               CTTTTGTGTTGCATCTTTTCTTGACAGCCTGTTACATTGGAACTTG 404
B               CTTTTGTGTTGCATCTTTTCTTGACAGCCTGTTACATTGGAACTTG 405
D               CTTTTGTGTTGCATCTTTTCTTGACAGCCTGTTACATTGGAACTTG 406
Monococcum      CTTTTGTGTTGCGTATTTTCTTGACAGCCTGTTACATTGGAACTTG 385
Speltoides      CTTTTGTGTTGCATCTTTTCTTGACAGCCTGTTACATTGGAACTTG 406
Tauschii        CTTTTGTGTTGCATCTTTTCTTGACAGCCTGTTACATTGGAACTTG 406
                ************ * ******************************
```

Figure 2.11 continued Alignment of primer pair BADH2P2 *BADH*2 amplicons derived from hexaploid wheat variety Cadoux and its progenitor species.

Although there were four SNP at the nucleotide level, translation of *BADH2* nucleotide sequence found one variation in amino acid sequence (Figure 2.11). The A genome has asparagine whereas the other two homeologs along with *T. monococcum, T. speltoides* and *T. tauschii* has histidine.

```
A               LNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLEL 51
B               LNIVTGLGHEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLEL 51
D               LNIVTGLGHEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLEL 51
Monococcum      LNIVTGLGHEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLEL 51
Speltoides      LNIVTGLGHEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLEL 51
Tauschii        LNIVTGLGHEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLEL 51
                ********:******************************************
```

Figure 2.12 ClustalW alignment of BADH2 amino acid sequence of hexaploid wheat homeologs and their progenitors.

The size of each *BADH2* homeolog was 1385 bp for A, 1374 bp for B, and 1297 bp for genome D. Alignment of the three homeologs (Figure 2.10) shows that they are very similar in exons six to nine (only part of exon 6 and 9 were amplified). There were three SNPs between the homeologs in the amplified region of exon 6 and a single SNP in exon 8 (Table 2.3). There was no difference in exon 7. The major difference was in intron 9 where a deletion occurred which was responsible for the size difference between homeologs as seen in agarose gel electrophoresis. The A/C SNP at position 36 changed the amino acid from asparagine to histidine which, unlike asparagine, has an aromatic ring. Although these two amino acids are not polar, histidine may have a positive charge.

```
A                CATCAGGTGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCT 60
B                CATCAGGTGTCTTAAACATTGTGACTGGATTAGGTCATGAAGCTGGCGCTCCTTTGTCGT 60
D                CATCAGGTGTCTTAAACATTGTGACTGGATTAGGTCATGAAGCTGGCGCTCCTTTGTCGT 60
                 ********************************** ********************** *

A                CACACCCTGACGTCGACAAGGTACATATATTTGTCCTATATCTTTTTGTGATCCATCTAT 120
B                CACACCCTGATGTCGACAAGGTACATATATTTGTCCTATATCTTTTTGTGATCCATCTAT 120
D                CACACCCTGATGTCGACAAGGTACATATATTTGTCCAATATCTTTTTGTGATCCATCTAT 120
                 ********** ************************* ************************

A                GTCGATGGAGTATATTCCTCACAATACATGGCTTATTTGAGCTTGTAGGTTGCATTTACC 180
B                GTCGATGGAGTATATTCCTCACAATACATGGCTTATTTGTGCTTTTAGGTTGCATTTACC 180
D                GTCGATGGAGTATATTCCTCACAATACATGGCTTATTTGTGCTTTTAGGTTGCATTTACC 180
                 ************************************** **** ***************

A                GGGAGCTATGCAACCGGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGGTTTGT 240
B                GGGAGCTATGCAACCGGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGGTTTGT 240
D                GGGAGCTATGCAACCGGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGGTTTGT 240
                 ************************************************************

A                TCCACATGTCCCTATTCTTTTAATCTTTGTAGTAATTTATCACCAAGTATCCTTTTTTCC 300
B                -CCACATGTCTCTATTTTTTAAATCTTTGTAGTAATTTATCACCAAGTATCTGTTTTCCC 299
D                TCCACATGTCTCTATTTTTTCAATCTTTGTAGTAATGTATCACCAAGTATCTGTTTTTCC 300
                  ********* ***** *** *************** **************  **** **

A                CCTCCTTTAGGTTTTTAAAATATTGACCAAGATACTAC-ACCGCTCTCCTGATAGCAGAC 359
B                CCTCCTTTAGATTTTTAGAATATTGACCAAGATACTACGAGCGCTCTCCTGATAGCAGAC 359
D                CCTCCTTTAGATTTTTAGAATATTGACCAAGATACTACGAGCGCTCTCCTGATAGCAGAC 360
                 ********** ****** ******************** * *******************

A                CAA-CTCACCTTTTGTGTTGCATCTTTTCTTGACAGCCTGTTACATTGGAACTTGGTGGA 418
B                CATTCTCACCTTTTGTGTTGCATCTTTTCTTGACAGCCTGTTACATTGGAACTTGGTGGC 419
D                CATTCTCACCTTTTGTGTTGCATCTTTTCTTGACAGCCTGTTACATTGGAACTTGGTGGC 420
                 **  ******************************************************
```

Figure 2.13 ClustalW alignments of *BADH2* amplicons from hexaploid wheat variety Cadoux.

```
A                AAAAGTCCTATTGTAGTATTTGATGATGTCGACATTGACAAAGGTACATACTGC---ATG 475
B                AAAAGTCCTATTGTAGTATTTGATGATGTCGACATTGACAAAGGTACATGCTGCTGCATG 479
D                AAAAGTCCTATTGTAGTATTTGATGATGTCGACATTGACAAAGGTACATGCTGCTGCATG 480
                 ************************************************* ****   ***

A                CTGAGATTTATCCAGT-CAAGGT---TTGTTCCAATTCG--AAATGC-----ATGGTGTT 524
B                CTGAGACTTGTCCAATTCAAAATGCCCTGTGTGGAATCACTAAATACTGACCATGGTGTT 539
D                CTGAGACTTGTCCAATTCGAAATGCCCTATGCGAAATCACTTCATACTG--CATGGTGTT 538
                 ****** ** **** * * *  *    * *    * **     ** *     ********

A                ACACACCCCAAAGCCTTCATTTGGTGTCAGGGTCTTGTCATGATAAATAATTGTTAGAAC 584
B                ACACACCCCAAAGCCTTTATTTGGTGTCTGGGGCTTTTCATGATAAATGATTGTTAGAAC 599
D                ACACACCTCAAAGCCTTCATTTGGTGTCTGGGGTTTGTCATGATAAATGATTGTTATTAC 598
                 ******* ********* ********** ***  ** *********** *******  **

A                TTAGGATTAGTAACCTCTAAAATGATTAAAAAGAATAATGTCTATAATTTCACCAGACGT 644
B                TTCGGATTAGCAACCTGTAAAATGATTAAAAAGAATGATGTCTATGATTTCACCAGACTT 659
D                TTAGGATTAGTAACCTCTAAAATGATTAAAAAGAATGATGTCTATAATTTCACCAGACTT 658
                 ** ******* ***** ******************* ******** ************ *

A                TCTGCTTGTTCAAACTGGCGTTATTTTTTATGGAGGGTACATACAAGCTGAATAATGAAA 704
B                TCTGCTTGTTCGAACTGGCGTAATTTTTTATTGAGGGTACATACAAGCTGAATAATGAAA 719
D                TCTGCTTGTTCAAACTGGTGTAATTTTTTGTGGAGGGTACATACAAGCTGAATAACGAAA 718
                 *********** ****** ** ******* * ********************** ****

A                CGCAGGATACATTTGGTTCCTAGATAAGCTGGGCTGTAGCCCAAGTAACTTATCCCGGCT 764
B                CGAAGGATACATTTGGTTCCTAGCTAAGCTGGGCTAGAGCCCAAGTAACTTATCCCGGCT 779
D                CGAAGAATA----------------------------------------------------- 727
                 ** ** ***

A                TTGTTGTGCTGGCCTTGGCTAGCTATTTGGTTTAATCTGACTAATCCTAGCTTCCCCTAA 824
B                TTGTTGTGCTAGCCTTGGCTAGCTGTTTGGTTTAATCTGACTAATCCTAGCTT------- 832
D                -----------GCCTTGGCTAGCTATTTGGTTTAATCTGACTAATCGTAGCTT------- 769
                            ************* ********************* ******

A                AAAAAGACTAATCCTAGCTTTTGGAGGATAGTTACGTTTTGTAGGGAAGCCCCCTCAGAG 884
B                --------------------TTGGAGGATAGTTAAGTTTTGTAGGGAAGCCCCCTCAGAG 872
D                ---------------------TTGGAGGATAGCTACTTTTTGTAGGGAAGCCCCCTCAAAG 809
                                      *********** **  ********************** **

A                ATCCATATGAGTTACCAAAGG-GCTGTTCTTGTGTGCGATCCTTGCCCAAAAGTGCAAGC 943
B                ATCCATATGAATTACCAAAAAAGCTGTTCTGGTGTGTGATTCTTAGCCAAAAGTGCAAGC 932
D                ATCCATATGAGTTACCAAAGA-GTTGTTCTTGTGTGTGATTCTTGCCCAAAAGTGCAAGC 868
                 ********** ********   * ****** ***** *** ***  **************

A                GGATTCGGCTCACGCAACTACTTTTGAAACCATCGTAGGCTACGTCGCTGCCAAATAGTA 1003
B                GGATTCGACTCACGCAGCTAATTATGAAACTATCCTAGGCTACATTGCCGCCGAAGAGTA 992
D                GGATTCGGCTCACGCAACTACTTTAGAAACCATCC-------------CGCCAAATAGTA 915
                 ******* ******** *** **  ***** ***               *** ** ****

A                AATTCTGGTTGTTCCAAAACGCACAACCTTGCTGGGCTAGTCTAGGGGTTTTCTTGCTTG 1063
B                AATTCTGGTTGTTCCAAAATGCACAACCTTGCTGGGCTAGTCTAGGGGTTTTCTTGCTTG 1052
D                AATTCTGGCTGTTCCAAAACGCACACCCTTGCTTGGCTAGTCTAGGGGTTTTCTTGCTTG 975
                 ******** ********** ***** ******* **************************

A                AGGCCCAAACATACAGGAAATTATTTCATCACAAATTTTTGGTGACAAGCAAGGCACTTA 1123
B                AGGCCCAAACCTACCGGAAATTATTTCATCACTGCTTTTTGGTGACAAGCAAGGCACTTA 1112
D                AGGCCCAAACATACCGGAAATTGTTTCATCACTAATTTTTGGTGACAAGCAAGGCAGTTA 1035
                 ********** *** ******* *********   ********************* ***
```

Figure 2.13 continued ClustalW alignment of *BADH2* amplicons from hexaploid wheat variety Cadoux.

```
A                TTTTGAACTGATTATTAAAACATGATTAGTATGCTTTAGCTCTGGAAACCTTTACGATCT 1183
B                TTTTGAACTGATTATAAAAACATGAGTAGTCTGCTTTAGCTCTGGAAACCTTTACGATCT 1172
D                TTTTGAACTGATTATTAAAACATGAGTAGTATGCTTTAGCTCTGGAAACCTTTACGATCT 1095
                 *************** ********* **** ****************************

A                AAGTGTTTCTAGACAGTGATTCACTCAAGTTAGTTGATACATACATTTTGATTGAATAAT 1243
B                AAATGTTTCTAGACAGTGATTCACTCAAGTTAGTTGATACATAAATTTTGATTGAATAAT 1232
D                AAATGTTTCTATACAGTGAATCACTCAAGTTAGTTGATACATACATTTTGATTGAATAAT 1155
                 ** ******** ******* ************************ ****************

A                ATGATTCTGTGAATTACCTTTTGTGACCTTTAGTTAGTTTTGATTGTTGATCCTTCTATT 1303
B                ATGATTCTGTTAATTATCTTTTGCGACCTTTAGTTAAATTTGATTATTGATCCTTCTATT 1292
D                ATGATTCTGTTAAGTACCTTTTGTGACCTTTAGTTTGATTTGATTGCTGATCCTTCTATT 1215
                 ********** ** ** ****** ***********   *******  *************

A                TAAACAGCTGTTGAGTGGACTCTATTTGGGTGCTTTTGGACCAACGGTCAGATTTGCAGT 1363
B                TAAACAGCTGTTGAGTGGACTCTATTTGGGTGCTTTTGGACCAACGGTCAGATTTGCAGT 1352
D                TAAACAGCTGTTGAGTGGACTCTATTTGGGTGCTTTTGGACCAACGGTCAGATTTGCAGT 1275
                 ************************************************************

A                GCGACATCTCGTCTTCTTATCC 1385
B                GCGACATCTCGTCTTCTTATCC 1374
D                GCGACATCTCGTCTTCTTATCC 1297
                 **********************
```

Figure 2.13 continued ClustalW alignment of *BADH2* amplicons from hexaploid wheat variety Cadoux.

2.4 Discussion

Unlike rice which is an aquatic plant, wheat is subject to drought stress. *BADH* homologs are associated with drought stress in many crops including barley and sorghum. This work has found both *BADH* paralogs, *BADH1* and *BADH2* are present in hexaploid wheat as well as its progenitors. Hexaploid wheat variety Cadoux contains three homeologs for each *BADH* paralog whereas their progenitors contain single orthologs.

Phylogenetic analysis suggests hexaploid wheat derived these genes from each of its progenitors. The SNPs between the progenitors and hexaploid wheat variety Cadoux suggest these homologs have diverged from the progenitor genes since domestication. However, in comparison with the *waxy* loci where the progenitor species and hexaploid wheat differed by 11 amino acids (Clark et al., 1991; Yan et al., 2000), the *BADH* homologs are relatively conserved.

Wheat accumulates glycine betaine up to a concentration of 20 µmol/g dry weight in leaves (Hitz and Hanson, 1980). However, it is not known which BADH paralogs are involved in glycine betaine synthesis. In barley, BBD1 showed very low affinity to betaine aldehyde compared to BBD2 and is involved in oxidation of omega aminoaldehyde (Fujiwara et al., 2008). Expression analysis of transcripts under different salt treatments may indicate which paralog is involved in providing salt tolerance.

There were no differences in the size of the exons of hexaploid wheat homeologs and their progenitors. Differences were seen in the introns of the *BADH2* homoelogs. Intron 6 and intron 7 were the same size. Intron 8 of *BADH2* was variable, whereas the A genome homeolog was 849 bp, B genome 837 bp and D genome 759 bp. Introns have higher levels of polymorphism than exons (Bryan et al., 1999). The deduced amino acid sequences of these

homeologs and progenitor genes had the same number of amino acids. There were some SNPs, however these SNPs most likely do not change protein conformation.

Wheat contains three *BADH1* and three *BADH2* homeologs. RNA expression analysis was required to determine if one, two of three *BADH1* or *BADH2* homeologs were expressed in hexaploid wheat.

CHAPTER 3: Expression of *BADH* homologs in hexaploid wheat

3.1 **Introduction**

Hexaploid wheat contains three genomes and homeolog expression patterns change in response to genetic and epigenetic conditions (Shitsukawa et al., 2007). In synthetic and natural hexaploid wheat, a change in gene expression is common (He et al., 2003). Gene silencing in polyploidy is mainly due to gene regulation which is not associated with chromosomal loss (He et al., 2003; Chague et al.) but is associated with the cytosine methylation and retro elements which mediate antisense silencing (Levy and Feldman, 2004). Retro elements can activate adjacent genes at the same time as making a gene undergo silencing (Kashkush et al., 2003; Levy and Feldman, 2004). The gene which undergoes silencing may become a pseudo gene or mutate to give a new character to the species (Adams et al., 2004). Inactivation or modification of gene expression is seen in newly synthesized amphiploid plants due to methylation leading to the dosage compensation and ultimately bringing the new variations (Liu et al., 1998).

The expression pattern of any one gene can also fluctuate in response to the environment as well as the genotype of plants (Tonsor et al., 2005)**.** The expression patterns of genes in natural polyploids such as cotton, wheat and *Arabidopsis suecica* have shown silencing of one copy, or a strong expression bias towards the other copy, or the equal contribution of the homeologs and the expression pattern varies greatly within an organ, tissue and between time points (Adams et al., 2003; Nomura et al., 2005; Bottley et al., 2006).

Expression of the same gene varies during the developmental stages of particular organs (Adams et al., 2003; Adams et al., 2004) and the complete partition of expression between homeologs can occur in two different organs. Reciprocal gene expression of homeologs is observed in cotton where one homeolog is expressed in one organ and another homeolog is expressed in others and this phenomenon can be seen soon after polyploid formation and is repeatable (Adams et al., 2003). Furthermore, the gene expression pattern of some genes can vary among different individuals of the same gene family (Chen and Ni, 2006a).

There is rapid gene silencing associated with cytosine methylation in synthetic allotetraploid wheat relative to its two diploid progenitors (Kashkush et al., 2002). The D genome undergoes silencing at higher frequency compared to the A and B genomes. Newly synthesized hexaploid wheat under goes more gene silencing than tetraploid wheat (He et al., 2003). The expression pattern in newly synthesized allohexaploid differs from the progenitors, indicating the pattern of gene expression is altered during polyploid formation. However, it is hypothesized the majority of genes are expressed from all three genomes in hexaploid wheat but the close similarity of homeologs makes it difficult to differentiate their expression (Wendel, 2000).

Hexaploid wheat variety Cadoux contains three homeologs for both *BADH* paralogs. It is important to know which homeolog is being expressed among these three genomes of hexaploid wheat in order to answer the question "Is fragrance in wheat possible?" It may be knocking out two homeologs will not make any difference if only one homeolog is expressed and is sufficient to break down substrate. Alternatively, it may be possible to calibrate fragrance and abiotic stress tolerance by knocking out one, two or three homeologs if all three homeologs are expressed.

The expression patterns of *BADH* homologs in two different tissues (leaves and seeds) at two time points (14DPA and 30DPA) of four different varieties were determined by using RT-PCR. RT-PCR provides a rapid and sensitive method for analysis of gene expression and to determine the presence of absence of transcripts.

3.2 Materials and methods

3.2.1 Sample collection and plant growth

Hexaploid wheat varieties Bob White, NW 51A, Cadoux and Banks were obtained from the Australian Plant DNA Bank, Lismore, Australia (www.dnabank.com.au). The husks were removed from the seeds and the seeds washed in 5% bleach (4% w/v chlorine) for five min followed by washing with Milli Q water to remove traces of sodium hypochlorite.

Plants were grown in a controlled growth chamber at a temperature of 21-22°C with 16 h of daylight and 8 h of dark. Cold 6400k fluorescent lamps (Hygrow technologies, Australia) were used for vegetative growth on a 16/8 h light cycle, six lights per m^2 of plant space until the first flag leaf appeared when 2700k cold fluorescent lamps (Hygrow technologies. Australia) were substituted for 50% of the lamps.

Five seeds were placed in an 8.0 L pot with equal amounts of peat moss, perlite (size P500) and vermiculite (size 2). Osmocote extract (macro nutrients plus some trace elements) 500g/100 L of mix and Dolomite (calcium, magnesium use to act as a buffer agent for pH) 100g/100 L mix were used as fertilizers. Further fertilizer was used once the wheat plant reached the flag leaf stage (10 ml Maxicrop to 10L water weekly). The 14 DPA and 30 DPA seeds were identified based on the development of the embryo morphology and seeds were collected along with the leaves when the samples reached 14 DPA and 30 DPA, snap frozen in liquid nitrogen and stored individually at -80 ^{o}C.

3.2.2 RNA Extraction

3.2.2.1 Pulverization of leaves and seeds

Leaves and seeds stored at -80 ^{o}C were used for the pulverization. Care was taken not to thaw before pulverization. Pulverization was carried out by tissue lyser (MEP Instruments Pty Ltd, Australia) in liquid nitrogen. The conditions for the pulverization were 25 hertz for 60 s to get a fine powder. This was stored in -80 ^{o}C or used directly for RNA extraction.

3.2.2.2 Modified TRizol® and Qiagen RNA extraction

A combination of TRizol® and Qiagen RNA extraction procedures were used to obtain total RNA for gene expression analysis. Pulverized seeds (500 mg) and leaves (250 mg) and 2 ml of the TRizol® (Invitrogen, Australia) (cat no 15596-026) were mixed gently in 5 ml centrifuge tube for 1-2 min tubes and allowed to stand at room temperature for 5-10 min followed by centrifugation at 5525 rpm for 20 min at 4 ^{o}C.

The supernatant was transferred to another 5 ml centrifuge tube along with 0.2% of chloroform per 1 ml of TRizol reagent. The tubes were allowed to stand for about 5 min at room temperature and then centrifuged at 5525 rpm for 20 min at 4 ^{o}C. The supernatant was collected without disturbing the interface or the organic pellet and placed in a fresh tube.

Absolute ethanol (0.5 ml) was added to the supernatant and mixed thoroughly and 650 µl added to a RNAse free column (RNeasy® Plant mini kit; Qiagen GMbH, Germany) and centrifuged at >8000rpm for 15 s. All subsequent steps were according to the manufacturer's instructions (Qiagen GMbH, Germany).

3.2.3 Quantification and Quality of RNA

Purity, integrity and quantity of RNA were determined by capillary electrophorosis on an Agilent Bioanalyser 2100 according to manufacturer's instructions (Agilent Technologies, Australia). Agilent 2100 expert software was used to calculate RNA integrity number (RIN). RIN is an algorithm for designating integrity values of RNA measurements. Though RNA is a stable molecule, it is digested by RNase. Low RIN indicates degraded RNA which yields spurious gene expression data (Auer et al., 2003). Absorbance at 260/280 nm ratio was used to check the quality of RNA (Beckman Coulter® DU®730 UV/Vis spectrophotometer). For the spectrophotometric measurements, 5µl of RNA sample and 495 µl of TE buffer were added to a 500 µl cuvette.

3.2.4 Primer design

The primers listed in Table 3.0 were used for RT-PCR analysis of *BADH1* and *BADH2* gene expression. *BADH* primers were designed based on exon regions conserved between the three homeologs of hexaploid wheat variety Cadoux.

Table: 3.1. Reverse Transcriptase PCR primers

Homolog, Primer pairs	Forward primer (5'-3')	Reverse primer(5'-3')	Expected size of amplicon
BADH1, BADH1P1	CAGCAAGAAGTGAAGGTGCTAC	GTACCTGGTGACTTGTTTCACG	~410 bp
BADH2, BADH2P2	CAGGTGTCTTAAACATTGTG	AGGACTTTTTCCACCAAGTTC	~190 bp
Actin primer	CTGGGATGACATGGGGAAAATA TGGCA	GCCTTGGAGATCCACATCTGCTG GA	~850 bp

3.2.5 Reverse transcriptase PCR (RT-PCR)

All RT-PCR amplifications were performed using a Palm-Cycler (Corbertt Research Pty. Ltd, Sydney, Australia). RT-PCR reactions were carried out in a volume of 25 µl containing 0.4 µM forward and reverse primer (Sigma Aldrich, Sydney), 2x reaction mix buffer, 0.5 µl Superscript^{TM} one step RT-PCR with Platinum Taq (Invitrogen, Australia cat. No. 10928-042), 5 mM of equimolar dNTPs (Promega) and 100 ng of total RNA. The RT-PCR program was 50 ^{o}C for 30 min for the synthesis of cDNA followed by a 94 ^{o}C for denaturation for two min, 35 cycles of 94 ^{o}C for 15 s (denaturation), 30 s at 57 ^{o}C (annealing) and 72 ^{o}C for 1.5 min (extension). Two negative controls were used for RT-PCR. In the first negative control Superscript^{TM} one step RT-PCR with Platinum Taq was replaced with 0.2 µl of Platinum Taq (Invitrogen, Australia). In the second negative control, RNA which was replaced with RNase free water.

3.2.6 Agarose gel electrophorosis

cDNA concentration, quality and purity were assessed by gel electrophoresis. cDNA was separated using 2% analytical agarose gel (Omnigel-Low, Edwards Instruments Co. Narellan, NSW) in 0.5x TBE buffer (Tris Borate EDTA) with MWM XIV (Molecular weight maker IV, Roche Diagnostics, Australia) and 6X gel loading buffer II at 5-8 volt/cm. cDNA concentration was determined by comparison with 1ng/µl, 2ng/µl, 5 ng/µl and 10ng/µl of lambda DNA (Roche Diagnostic Australia Pty. Ltd).

Gels were stained with ethidium bromide solution for 10-15 min (0.50 µg/ml) followed by 10 min in water to de-stain. The DNA was visualized on a UV light.

3.2.7 Sequencing reaction protocol

Sequencing was carried out using Big Dye Terminator V3 (Applied Biosystems) in both forward and reverse directions according to the manufacturer's instructions. Briefly, 1 µl of Big Dye terminator, 2 µl of AB Buffer sequencing buffer, primer (0.41µM), 5 µl of template and 2 µl of Milli Q water was added in PCR tubes. Sequencing reactions were transferred to 1.5 ml tubes and 3.2 mM EDTA, 0.07M sodium acetate (pH 5.0) and 5 µl of MQ water and 60 µl of 100% ethanol was added to the tube and incubated at room temperature in the dark for 15-20 min and then centrifuged for 30 min at 14,000 rpm. The supernatant was discarded and the DNA pellet washed with 100 ul of 70% ethanol. Following centrifugation for 5 min at 14,000 rpm, the supernatant was discarded and dried for 5 min at room temperature. The samples were submitted to Southern Cross Plant Genomics (Southern Cross University, Lismore) for the separation of the sequence reactions.

3.2.8 Sequence analysis and alignment

Sequences were analysed using Sequencher 4.10.1 (Gene Code Corporation, USA) with default settings. SNPs were identified by visual inspection of chromatograms.

Cadoux BADH homologs were assigned to each genome (A, B and D) on the basis of the phylogenetic relationship between the hexaploid wheat progenitor BADH DNA sequence and Cadoux BADH homologs DNA sequence. Alignment was carried out with ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) using the Neighbor-Joining clustering with default settings. A bootstrap value of 1000 was used to generate the node value. TreeView™ Win 32 versions 1.6.6 was used with default settings in order to generate the phylogenetic relation of hexaploid wheat and their progenitors. For the prediction of amino acid sequences, ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and MEGA 4.0™ (Tamura et al., 2007) were used with default settings.

3.2.9 Protein structure analysis

Ligand Explorer 3.9 (http://www.rcsb.org/pdb/home/home.do) and Swiss-PDB Viewer 4.0.1 (http://spdbv.vital-it.ch/) was used to analyze the BADH1 and BADH2 sequences relative to the 3-D crystal structure of pea ALDH, 3IWJ (http://www.rcsb.org/pdb/explore/explore.do?structureId=3IWJ) and 31WK (http://www.rcsb.org/pdb/explore.do?structureId=3IWK).

3.3 Results

3.3.1 RNA extraction

Extracted RNA was colourless and readily re-dissolved in RNAse free water. RNA extracted from the leaves and seeds was of good quality with no shearing and minimal contamination as analysed by agarose gel electrophoresis and Agilent Bioanalyser 2100 (Agilent Technologies, Australia). RNA extracted with modified TRizol® Method gave the RNA Integrity number (RIN) of 5.0 to 8.5.

3.3.2 *BADH1* amplification

The *BADH1* homeologs were amplified with the primer pair *BADH1P1* (Table 3.1). RT-PCR resulted in the production of a single band as visualized on 2% agarose gel after staining with ethidium bromide. The band visualized was of the expected size of 410 bp (Figure 3.1) Agarose gel electrophorosis showed a single band indicating either only one genome was expressed or all three genomes expressed cDNA sequences of the same size or very similar size which agarose gel electrophoresis could not resolve. Sequence analysis showed a clear chromatogram trace when viewed using Sequencher® 4.10.1 (Gene Codes Corporation. USA). The three homeologs were very similar but could be identified by diagnostic SNP. A single peak at each SNP position in the chromatogram implies only one homoelog is present in the RT-PCR product. If more than one homoelog was expressed then two or three peaks would be evident at each SNP position in the chromatogram, making it difficult to read the chromatogram. Therefore, only one homoelogs of *BADH1* was expressed in the wheat varieties investigated at both the time points (14 DPA and 30 DPA) in both tissues (leaves and seeds). BlastX interrogation of the NCBI (http://www.ncbi.nlm.nih.gov/) database found high similarity with the betaine aldehyde dehydrogenase gene of related species at nucleotide level for *BADH1* with barley 97% *BADH1* (BAB62847.1) and 83% identity with rice *BADH1*

(ABB83473.1 with E-value of 6e-73 and 2e-62 respectively. TblastN also showed high identity at amino acid level with 98% identity with barley (AB063179.1) with an E-value 1e-72 and 84% identical with rice (AK103582.1) with an E-value 9e-64.

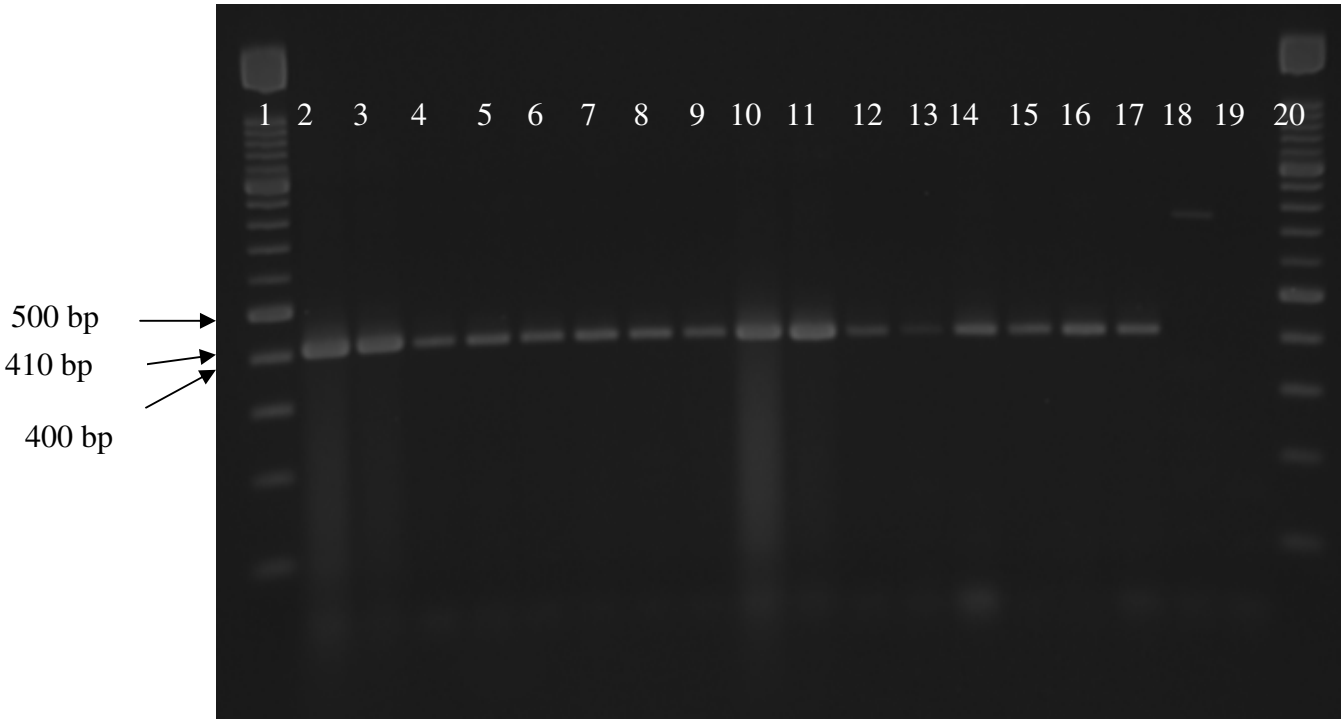

Figure: 3.1 *BADH1* RT-PCR analyses. Lanes 1 and 20 is 100 bp ladder (Molecular Marker XIV, Roche Diagnostics Australia). Lanes 2, 4, 6, 8, are the 14 DPA leaves and Lane 3, 5, 7, 9 are the 30 DPA leaves RT-PCR product of Caudox, NW 51A, Bob White and Banks respectively. Lanes 10, 12, 14, 16 are 14 DPA seeds and 11, 13, 15, 17 are 30 DPA seeds RT-PCR product of Caudox, NW 51A, Bob White and Banks respectively. Lanes 18 and 19 are the negative control.

ClustalW alignment of hexaploid wheat *BADH1* cDNA sequence with *BADH1* gDNA sequence from hexaploid wheat and its progenitors found it was 99% similar with homeolog A when aligned in ClustalW, implying only one homeolog is expressed at both time points and in both tissues (14 DPA and 30 DPA leaves and seeds) in hexaploid wheat.

```
A               AATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATTGAACCTAC 60
B               AATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATTGAACCTAC 60
D               AATTTTGCATGGGGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATTGAACCTAC 60
Monococcum      AATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATTGAACCTAC 60
Speltoides      AATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATTGAACCTAC 60
Tauschii        AATTTTGCATGGGGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATTGAACCTAC 60
cDNA            AATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATTGAACCTAC 60
                ************ ************************************************

A               AATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGT 120
B               AATTATAACAGATGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGT 120
D               TATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGT 120
Monococcum      AATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGT 120
Speltoides      AATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGT 120
Tauschii        TATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGT 120
cDNA            AATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGT 120
                ***********************************************************

A               CATCTGTGTCAAAGTATTCAAGACAGAGAGTGAAGCTGTAGAGCTTGCAAATGATACCCA 180
B               CATCTGTGTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAATGATACCCA 180
D               CATCTGTGTCAAAGTATTTAAGACAGAGAGCGAAGCCGTAGAGCTTGCAAATGATACCCA 180
Monococcum      CATCTGTGTCAAAGTATTCAAGACAGAGAGTGAAGCTGTAGAGCTTGCAAATGATACTCA 180
Speltoides      CATCTGTGTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAATGATACCCA 180
Tauschii        CATCTGTGTCAAAGTATTTAAGACAGAGAGCGAAGCCGTAGAGCTTGCAAATGATACCCA 180
cDNA            CATCTGTGTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAATGATACCCA 180
                ****************** *********** ***** ******************** **

A               CTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGCATTGCAAA 240
B               CTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGCATTGCAAA 240
D               CTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTAGAGAGGTGTGAGCGCATTGCAAA 240
Monococcum      CTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGCATTGCAAA 240
Speltoides      CTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTAGAGAGGTGTGAGCGCATTGCAAA 240
Tauschii        CTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTAGAGAGGTGTGAGCGCATTGCAAA 240
cDNA            CTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGCATTGCAAA 240
                ************************************* **********************

A               GGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTTCAAGCTCC 300
B               GGTAATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACCCTTGTTCAAGCTCC 300
D               GGTAATTCACTCGGGCATTGTTTGGATAAACTGCTCGCAACCGACTCTGGTTCAAGCTCC 300
Monococcum      GGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTTCAAGCTCC 300
Speltoides      GGTCATTCACTCTGGCATTGTTTGGATAAACTGCTCGCAACCGACTCTGGTTCAAGCTCC 300
Tauschii        GGTTATTCACTCGGGCATTGTTTGGATAAACTGCTCGCAACCGACTCTGGTTCAAGCTCC 300
cDNA            GGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTTCAAGCTCC 300
                *** ******** ********************************* ** ***********

A               GTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGCCTCGAGAA 360
B               GTGGGGAGGGAACAAGCGGAGTGGTTTTGGCCGGGAGCTAGGAGAATGGGGCCTCGAGAA 360
D               GTGGGGAGGGAACAAGCGTAGTGGTTTTGGCCGGGAGCTAGGAGAATGGGGCCTCGAGAA 360
Monococcum      GTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGCCTCGAGAA 360
Speltoides      GTGGGGAGGGAACAAGCGTAGTGGTTTTGGACGGGAGCTAGGAGAATGGGGCCTCGAGAA 360
Tauschii        GTGGGGAGGGAACAAGCGTAGTGGTTTTGGCCGGGAGCTAGGAGAATGGGGCCTCGAGAA 360
cDNA            GTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGCCTCGAGAA 360
                ****************** ** ******** ******************************
```

Figure 3.2 ClustalW alignment of Cadoux *BADH1* 14DPA leaf cDNA with Cadoux *BADH1* and *BADH1* progenitor gDNA. The symbol "*" denotes identical nucleotides and blank spaces non-identical nucleotides.

```
A               CTACTTGAGCGTGAAA 376
B               CTACTTGAGCGTGAAA 376
D               CTACCTGAGCGTGAAA 376
Monococcum      CTACTTGAGCGTGAAA 376
Speltoides      CTACCTGAGCGTGAAA 376
Tauschii        CTACCTGAGCGTGAAA 376
cDNA            CTACTTGAGCGTGAAA 376
                **** ***********
```

Figure 3.2 continued ClustalW alignment of Cadoux *BADH1* 14DPA leaf cDNA with Cadoux *BADH1* and *BADH1* progenitor gDNA. The symbol "*" denotes identical nucleotides and blank spaces non-identical nucleotides.

BADH1 amino acid sequence was predicted using MEGA 4.0™ and aligned by ClustalW with the predicted amino acid sequence of each progenitor species. This alignment shows there is one amino acid change. Genomes A and B of hexaploid wheat and *T. monococcum* have glutamine at position 76 of the alignment whereas genome D, *T. speltoides* and T. *tauschii* contain glutamic acid.

```
A               GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
B               GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
D               GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
Monococcum      ---ILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 57
Speltoides      GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
Tauschii        GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
cDNA            GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
                   ******************************************************

A               DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
B               DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
D               DTHYGLAGGVISDDLERCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
Monococcum      DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 117
Speltoides      DTHYGLAGGVISDDLERCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
Tauschii        DTHYGLAGGVISDDLERCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
cDNA            DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
                ***************:**********************************************

A               LENYLSVK 128
B               LENYLSVK 128
D               LENYLSVK 128
Monococcum      LENYLSVK 125
Speltoides      LENYLSVK 128
Tauschii        LENYLSVK 128
cDNA            LENYLSVK 128
                ********
```

Figure: 3.3 Alignment of predicted amino acid sequence of hexaploid wheat BADH1 with its progenitors. The symbol "*" denotes identical amino acids.

3.3.3 *BADH2* amplification

The *BADH2* homeologs were amplified with the primer pairs BADH2P2 (Table 3.1). RT-PCR resulted in the production of single band as visualized on a 2% agarose gel stained with ethidium bromide. The band visualized was of the expected size of ~190 bp (Figure 3.4).

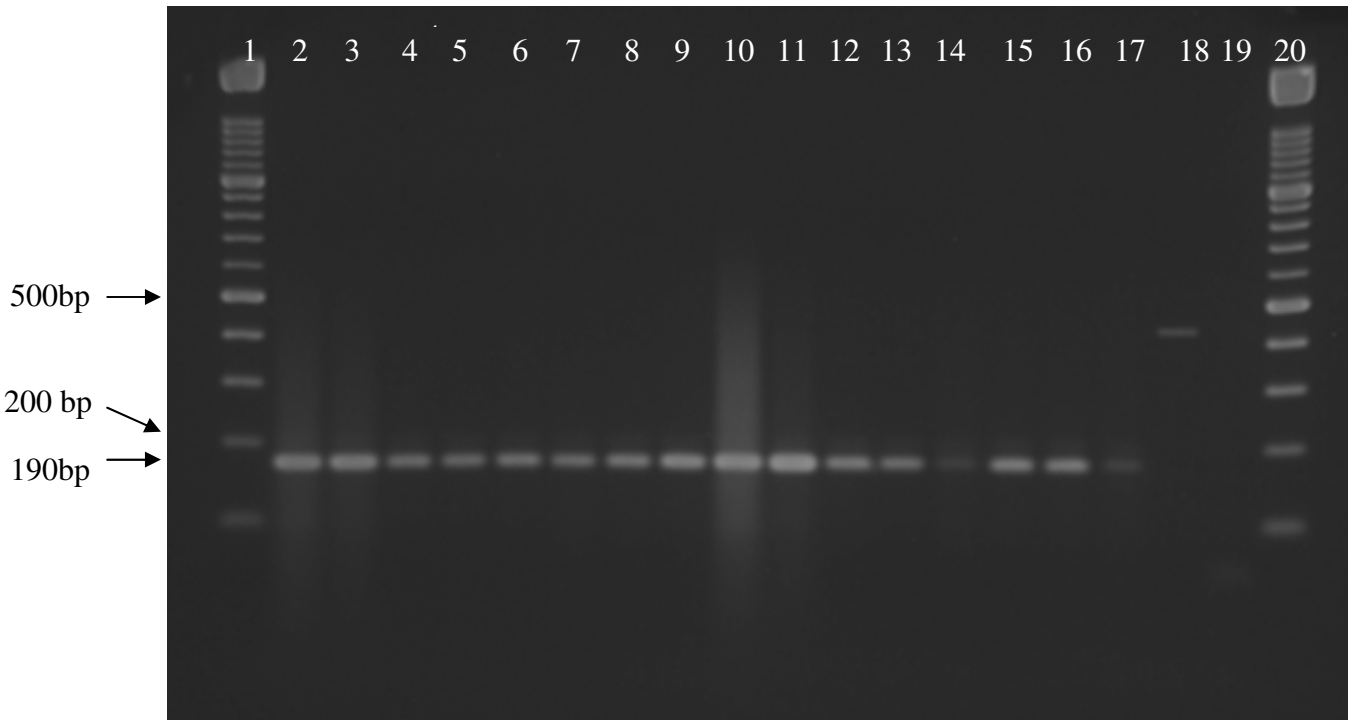

Figure 3.4 *BADH2* RT-PCR analyses. Lanes 1 and 20 is 100 bp ladder (Molecular Markers XIV, Roche Diagnostics Australia). Lanes 2, 4, 6 and 8 are 14 DPA leaves and Lane 3, 5, 7, 9 are the 30 DPA leaves RT-PCR product of Caudox, NW 51A, Bob White and Banks respectively. Lane 10, 12, 14, 16 are 14 DPA seeds and 11, 13, 15, 17 are 30DPA seeds RT-PCR product of Caudox, NW 51A, Bob White and Banks respectively. Lane 18 and 19 are the negative control.

These amplicons were sequenced and analysed by Sequencher® 4.10.1 (Gene Codes Corporation. USA). The three homeologs are very similar but can be identified by diagnostic SNP. A single peak at each SNP position in the chromatogram implies only one homoelog is

present in the RT-PCR product. If more than one homoelog was expressed then two or three peaks would be evident at each SNP position in the chromatogram, making it difficult to read the chromatogram. Therefore, only one homoelogs is expressed in the wheat varieties investigated.

A single peak was evident at all nucleotides indicating only one homeolog is expressed among the three homeologs present of *BADH2*. BlastX using these sequences of the NCBI database found high identity with the other cereal species. At the nucleotide level there was 100% identity with wheat *BADH2* (AAL05264.1), E-value of 4e-07; 100% identity with barley *BADH2* (BAB62846.1), E-value of 4e-07; and 92% identity with barley *BADH2* (AB184118.1), E-value of 1e-06. At the amino acid level, BADH2 showed 100% identity with barley (AB063178.1), 98% with wheat (AY050316.1) and with rice (AK060461.1) with E–values of 5e-20, 2e-19 and 3e-16 respectively. This indicates the gene amplified by the BADH2P2 primers is *BADH2* and only one homeolog is expressed at both time points (14 DPA and 30 DPA) and tissues (leaves and seeds).

ClustalW alignment of hexaploid wheat variety Cadoux 14 DPA leaf *BADH2* cDNA with *BADH2* from the progenitor species shows there is identity with the genome A *BADH2* homeolog indicating only genome A is expressed and genome B and D *BADH2* are silenced in hexaploid wheat variety Caudox 14 DPA and 30 DPA leaves.

```
A               TCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCCTG 60
B               TCTTAAACATTGTGACTGGATTAGGTCATGAAGCTGGCGCTCCTTTGTCGTCACACCCTG 60
D               TCTTAAACATTGTGACTGGATTAGGTCATGAAGCTGGCGCTCCTTTGTCGTCACACCCTG 60
Monococcum      TCTTAAACATTGTGACTGGATTAGGTCATGAAGCTGGCGCTCCTTTGTCGTCACACCCTG 60
Speltoides      TCTTAAACATTGTGACTGGATTAGGTCATGAAGCTGGCGCTCCTTTGTCGTCACACCCTG 60
Tauschii        TCTTAAACATTGTGACTGGATTAGGTCATGAAGCTGGCGCTCCTTTGTCGTCACACCCTG 60
cDNA            TCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCCTG 60
                ************************** ********************** **********
```

Figure 3.5 Alignment of hexaploid wheat variety Cadoux 14 DPA leaf *BADH2* cDNA and progenitor gDNA sequence.

```
A               ACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGCTG 120
B               ATGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGCTG 120
D               ATGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGCTG 120
Monococcum      ACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGCTG 120
Speltoides      ATGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGCTG 120
Tauschii        ACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGCTG 120
cDNA            ACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGCTG 120
                * ************************************************************

A               CAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTG 156
B               CAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTG 156
D               CAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTG 156
Monococcum      CAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTG 156
Speltoides      CAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTG 156
Tauschii        CAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTG 156
cDNA            CAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTG 156
                ************************************
```

Figure 3.5 continued Alignment of hexaploid wheat variety Cadoux 14 DPA leaf *BADH2* cDNA and progenitor gDNA sequence.

```
A               LNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLEL 51
B               LNIVTGLGHEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLEL 51
D               LNIVTGLGHEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLEL 51
Monococcum      LNIVTGLGHEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLEL 51
Speltoides      LNIVTGLGHEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLEL 51
Tauschii        LNIVTGLGHEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLEL 51
cDNA            LNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLEL 51
                ********:*****************************************
```

Figure 3.6 ClustalW alignment of BADH2 amino acid sequence of hexaploid wheat variety Cadoux predicted from 14 DPA leaf cDNA with progenitor BADH2 amino acid sequence.

BADH2 cDNA of hexaploid wheat variety Caudox leaf shows identity with the A homeologs suggesting only the A genome copy of BADH2 is expressed in the leaves at two time points (14 DPA and 30 DPA). However, some SNPs were evident at the nucleotide level of the three homeologs and progenitor sequences (Table 3.4). There were no differences at the amino acid level of *T. monococcum, T. speltoides* and *T. tauschii,* indicating this gene and its gene product are highly conserved across the different species. Only one amino acid change is seen as indicated in red colour in Figure 3.6. Asparagine is evident in the A genome homeolog and in expressed RNA while Histidine is present in the other two homeologs.

Taking the sequences of Cadoux which were constant across all time points and tissues as the reference, there were a number of SNPs between species in expressed *BADH1* mRNA (Figure 3.7) . However, when they were aligned with ClustalW, all expressed sequences were more similar to the genome A homelog of Cadoux indicating that in all genotypes sampled only the genome A *BADH1* homeolog is expressed. *BADH1* from NW51A 14 day's post-anthesis seeds was the most divergent with five SNPs while Bob White 14 days post-anthesis seeds had three SNPs, each of which were the same as NW51A relative to Cadoux.

```
Caudox14DPAleaves        GGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
Caudox14DPAseeds         GGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
Caudox30DPAleaves        GGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
Caudox30DPAseeds         GGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
BobWhite14DPAleaves      GGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
BobWhite14DPAseeds       GGTGCTACAATTTTGCATGGGGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
BobWhite30DPAleaves      GGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
BobWhite30DPAseeds       GGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
Banks14DPAleaves         GGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
Banks14DPAseeds          GGTGCTACAATTTTGCATGGGGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
Banks30DPAleaves         GGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
Banks30DPAseeds          GGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
NW5114DPAleaves          GGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
NW51A14DPAseeds          GGTGCTACAATTTTGCATGGGGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
NW51A30DPAleaves         GGTGCTCCAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
NW51A130DPAseeds         GGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGGGTTCTTTATT 60
                         ****** ************* ****************************************

Caudox14DPAleaves        GAACCTACAATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
Caudox14DPAseeds         GAACCTACAATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
Caudox30DPAseeds         GAACCTACAATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
Caudox30DPAleaves        GAACCTACAATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
BobWhite14DPAleaves      GAACCTACAATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
BobWhite14DPAseeds       GAACCTACAATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
BobWhite30DPAleaves      GAACCTACAATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
BobWhite30DPAseeds       GAACCTACAATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
Banks14DPAleaves         GAACCTACAATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
Banks14DPAseeds          GAACCTACAATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
Banks30DPAleaves         GAACCTACAATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
Banks30DPAseeds          GAACCTACAATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
NW5114DPAleaves          GAACCTACAATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
NW51A14DPAseeds          GAACCTACTATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
NW51A30DPAleaves         GAACCTACAATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
NW51A130DPAseeds         GAACCTACAATTATAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTT 120
                         ******** ***************************************************
```

Figure 3.7 ClustalW alignments of *BADH1* nucleotide sequences of two time points (14DPA and 30 DPA) and two tissues (seeds and leaves) of hexaploid wheat varieties.

```
Caudox14DPAleaves        GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGTGAAGCTGTAGAGCTTGCAAAT 180
Caudox14DPAseeds         GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGTGAAGCTGTAGAGCTTGCAAAT 180
Caudox30DPAseeds         GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGTGAAGCTGTAGAGCTTGCAAAT 180
Caudox30DPAleaves        GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGTGAAGCTGTAGAGCTTGCAAAT 180
BobWhite14DPAleaves      GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGTGAAGCTGTAGAGCTTGCAAAT 180
BobWhite14DPAseeds       GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAAT 180
BobWhite30DPAleaves      GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGTGAAGCTGTAGAGCTTGCAAAT 180
BobWhite30DPAseeds       GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGTGAAGCTGTAGAGCTTGCAAAT 180
Banks14DPAleaves         GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAAT 180
Banks14DPAseeds          GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAAT 180
Banks30DPAleaves         GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAAT 180
Banks30DPAseeds          GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAAT 180
NW5114DPAleaves          GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAAT 180
NW51A14DPAseeds          GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAAT 180
NW51A30DPAleaves         GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAAT 180
NW51A130DPAseeds         GGACCAGTCATCTGTGTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAAT 180
                         ************************************* ********************

Caudox14DPAleaves        GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
Caudox14DPAseeds         GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
Caudox30DPAleaves        GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
Caudox30DPAseeds         GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
BobWhite14DPAleaves      GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
BobWhite14DPAseeds       GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
BobWhite30DPAleaves      GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
BobWhite30DPAseeds       GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
Banks14DPAleaves         GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
Banks14DPAseeds          GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
Banks30DPAleaves         GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
Banks30DPAseeds          GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
NW5114DPAleaves          GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
NW51A14DPAseeds          GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
NW51A30DPAleaves         GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
NW51A130DPAseeds         GATACCCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGC 240
                         ************************************************************

Caudox14DPAleaves        ATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTT 300
Caudox14DPAseeds         ATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTT 300
Caudox30DPAleaves        ATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACTCTGGTT 300
Caudox30DPAseeds         ATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACTCTGGTT 300
BobWhite14DPAleaves      ATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTT 300
BobWhite14DPAseeds       ATTGCAAAGGTTATTCACTCGGGCATTGTTTGGATAAACTGCTCGCAACCGACTCTGGTT 300
BobWhite30DPAleaves      ATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTT 300
BobWhite30DPAseeds       ATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTT 300
Banks14DPAleaves         ATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTT 300
Banks14DPAseeds          ATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACTCTGGTT 300
Banks30DPAleaves         ATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTT 300
Banks30DPAseeds          ATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACTCTGGTT 300
NW5114DPAleaves          ATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTT 300
NW51A14DPAseeds          ATTGCAAAGGTTATTCACTCGGGCATTGTTTGGATAAACTGCTCGCAACCGACTCTGGTT 300
NW51A30DPAleaves         ATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACACTGGTT 300
NW51A130DPAseeds         ATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTT 300
                         ******************* ******************************** ******
```

Figure 3.7 continued ClustalW alignments of BADH1 nucleotide sequences of two time points (14DPA and 30 DPA) and two tissues (seeds and leaves) of hexaploid wheat varieties.

```
Caudox14DPAleaves        CAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
Caudox14DPAseeds         CAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
Caudox30DPAleaves        CAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
Caudox30DPAseeds         CAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
BobWhite14DPAleaves      CAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
BobWhite14DPAseeds       CAAGCTCCGTGGGGAGGGAACAAGCGTAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
BobWhite30DPAleaves      CAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
BobWhite30DPAseeds       CAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
Banks14DPAleaves         CAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
Banks14DPAseeds          CAAGCTCCGTGGGGAGGGAACAAGCGTAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
Banks30DPAleaves         CAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
Banks30DPAseeds          CAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
NW5114DPAleaves          CAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
NW51A14DPAseeds          CAAGCTCCGTGGGGAGGGAACAAGCGTAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
NW51A30DPAleaves         CAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
NW51A130DPAseeds         CAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGGAGAATGGGGC 360
                         ************************* ********************************

Caudox14DPAleaves        CTCGAGAACTACTTGAGCGTGAAACA 386
Caudox14DPAseeds         CTCGAGAACTACTTGAGCGTGAAACA 386
Caudox30DPAleaves        CTCGAGAACTACTTGAGCGTGAAACA 386
Caudox30DPAseeds         CTCGAGAACTACTTGAGCGTGAAACA 386
BobWhite14DPAleaves      CTCGAGAACTACTTGAGCGTGAAACA 386
BobWhite14DPAseeds       CTCGAGAACTACTTGAGCGTGAAACA 386
BobWhite30DPAleaves      CTCGAGAACTACTTGAGCGTGAAACA 386
BobWhite30DPAseeds       CTCGAGAACTACTTGAGCGTGAAACA 386
Banks14DPAleaves         CTCGAGAACTACTTGAGCGTGAAACA 386
Banks14DPAseeds          CTCGAGAACTACTTGAGCGTGAAACA 386
Banks30DPAleaves         CTCGAGAACTACTTGAGCGTGAAACA 386
Banks30DPAseeds          CTCGAGAACTACTTGAGCGTGAAACA 386
NW5114DPAleaves          CTCGAGAACTACTTGAGCGTGAAACA 386
NW51A14DPAseeds          CTCGAGAACTACCTGAGCGTGAAACA 386
NW51A30DPAleaves         CTCGAGAACTACTTGAGCGTGAAACA 386
NW51A130DPAseeds         CTCGAGAACTACTTGAGCGTGAAACA 386
                         ************ *************
```

Figure 3.7 continued ClustalW alignments of BADH1 nucleotide sequences of two time points (14DPA and 30 DPA) and two tissues (seeds and leaves) of hexaploid wheat varieties.

Although there were five SNPs in NW51A 14DPA seeds, there was little difference in BADH1 amino acid sequence (Figure 3.8). Only one amino acid difference among the experimental varieties was noted, NW51A 30 days post anthesis in leaves, where the amino acid threonine is replaced with proline (Figure 3.8.) as indicated in red.

```
Caudox14DPAleaves        GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
Caudox14DPAseeds         GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
Caudox30DPAleaves        GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
Caudox30DPAseeds         GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
BobWhite14DPAleaves      GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
BobWhite14DPAseeds       GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
BobWhite30DPAleaves      GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
BobWhite30DPAseeds       GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
Banks14DPAleaves         GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
Banks14DPAseeds          GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
Banks30DPAleaves         GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
Banks30DPAseeds          GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
NW51A14DPAleaves         GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
NW51A14DPAseeds          GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
NW51A30DPAleaves         GAPILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
NW51A30DPAseeds          GATILHGGDRPKHLGKGFFIEPTIITDVSTSMQIWREEVFGPVICVKVFKTESEAVELAN 60
                         **.************************************************************

Caudox14DPAleaves        DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
Caudox14DPAseeds         DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
Caudox30DPAleaves        DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
Caudox30DPAseeds         DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
BobWhite14DPAleaves      DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
BobWhite14DPAseeds       DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
BobWhite30DPAleaves      DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
BobWhite30DPAseeds       DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
Banks14DPAleaves         DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
Banks14DPAseeds          DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
Banks30DPAleaves         DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
Banks30DPAseeds          DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
NW51A14DPAleaves         DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
NW51A14DPAseeds          DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
NW51A30DPAleaves         DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
NW51A30DPAseeds          DTHYGLAGGVISDDLQRCERIAKVIHSGIVWINCSQPTLVQAPWGGNKRSGFGRELGEWG 120
                         ************************************************************

Caudox14DPAleaves        LENYLSVK 128
Caudox14DPAseeds         LENYLSVK 128
Caudox30DPAleaves        LENYLSVK 128
Caudox30DPAseeds         LENYLSVK 128
BobWhite14DPAleaves      LENYLSVK 128
BobWhite14DPAseeds       LENYLSVK 128
BobWhite30DPAleaves      LENYLSVK 128
BobWhite30DPAseeds       LENYLSVK 128
Banks14DPAleaves         LENYLSVK 128
Banks14DPAseeds          LENYLSVK 128
Banks30DPAleaves         LENYLSVK 128
Banks30DPAseeds          LENYLSVK 128
NW51A14DPAleaves         LENYLSVK 128
NW51A14DPAseeds          LENYLSVK 128
NW51A30DPAleaves         LENYLSVK 128
NW51A30DPAseeds          LENYLSVK 128
                         ********
```

Figure: 3.8 ClustalW alignments of BADH1 predicted amino acid sequence of two time points (14DPA and 30 DPA) and two tissues (seeds and leaves) of hexaploid wheat varieties.

There was no nucleotide differences between any of the *BADH2* sequences expressed in any of the hexaploid wheat varieties which were investigated in this experiment.

```
Caudox14DPAleaves        TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60

Caudox14DPAseeds         TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60
Caudox30DPAleaves        TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60
Caudox30DPAseeds         TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60
BobWhite14DPAleaves      TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60
Bobwhite14DPAseeds       TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60
Bobwhite30DPAleaves      TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60
Bobwhite30DPAseeds       TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60
Banks14DPAleaves         TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60
Banks14DPAseeds          TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60
Banks30DPAleaves         TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60
Banks30DPAseeds          TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60
NW51A14DPAleaves         TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60
NW51A14DPAseeds          TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60
NW51A30DPAleaves         TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60
NW51A30DPAseeds          TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGTCCTCACACCC 60
                         ************************************************************

Caudox14DPAleaves        TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
Caudox14DPAseeds         TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
Caudox30DPAleaves        TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
Caudox30DPAseeds         TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
BobWhite14DPAleaves      TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
Bobwhite14DPAseeds       TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
Bobwhite30DPAseeds       TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
Bobwhite30DPAleaves      TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
Banks14DPAleaves         TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
Banks14DPAseeds          TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
Banks30DPAleaves         TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
Banks30DPAseeds          TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
NW51A14DPAleaves         TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
NW51A14DPAseeds          TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
NW51A30DPAleaves         TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
NW51A30DPAseeds          TGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGC 120
                         ************************************************************

Caudox14DPAleaves        TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
Caudox14DPAseeds         TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
Caudox30DPAleaves        TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
Caudox30DPAseeds         TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
BobWhite14DPAleaves      TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
Bobwhite14DPAseeds       TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
Bobwhite30DPAleaves      TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
Bobwhite30DPAseeds       TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
Banks14DPAleaves         TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
Banks14DPAseeds          TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
Banks30DPAleaves         TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
Banks30DPAseeds          TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
NW51A14DPAleaves         TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
NW51A14DPAseeds          TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
NW51A30DPAleaves         TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
NW51A30DPAseeds          TGCAGCTCCTACAGTCAAGCCTGTTACATTGGAACTTGGTGGAAAAA 167
                         ***********************************************
```

Figure: 3.9 *BADH2* ClustalW alignment of nucleotide of expressed cDNA of the samples of two time points (14DPA and 30 DPA) and two tissues (seeds and leaves) of hexaploid wheat varieties.

```
Caudox14DPAleaves     VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGK- 56
Caudox14DPAseeds      VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGKS 56
Caudox30DPAleaves     VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGK- 55
Caudox30DPAseeds      VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGKS 56
BobWhite14DPAleaves   VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGKS 56
BobWhite1DPA4seeds    VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGKS 56
BobWhite30DPAleaves   VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGKS 56
BobWhite30DPAseeds    VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGKS 56
Banks14DPAleaves      VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGKS 56
Banks14DPAseeds       VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGKS 56
Banks30DPAleaves      VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGKS 56
Banks30DPAseeds       VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGKS 56
NW51A14DPAleaves      VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGKS 56
NW51A14DPAseeds       VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGKS 56
NW51A30DPAleaves      VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGKS 56
NW51A30DPAseeds       VLNIVTGLGNEAGAPLSSHPDVDKVAFTGSYATGQKIMVAAAPTVKPVTLELGGK- 55
                      *******************************************************
```

Figure 3.10 ClustalW alignments of BADH2 amino acid sequences as predicted from cDNA at two time points (14DPA and 30 DPA) and two tissues (seeds and leaves) of hexaploid wheat varieties.

3.4 Discussion

Gene expression patterns in polyploids differ between gene family members (Chen and Ni, 2006) and between different developmental stages and tissues (Adams et al., 2003; Adams et al., 2004). Many homeologs in polyploids display reciprocal expression or silencing of gene expression due to epigenetic factors including gene methylation. Gene silencing is more frequent than gene loss in polyploids (Koh et al.) and about 25% of genes in established polyploids have undergone gene silencing (Bottley et al., 2006). Although the mechanism of *BADH* homeolog regulation is not yet known, only the A genome homeolog of *BADH* (*BADH1 or BADH2*) gene is expressed in 14 DPA and 30 DPA leaves and seeds in the hexaploid wheat varieties sampled in this study.

Barley and wheat both accumulate glycine betaine in response to drought (Hitz and Hanson, 1980). Tblastn comparison of deduced amino acid sequence of wheat *BADH* showed 100% identity with barley *BADH* paralogs. This high level of similarity in amino acid sequences with barley indicates wheat BADH paralogs might also localise to two different sites in the wheat plant. Transformation of wheat with non-native *BADH* in has improved the stability of chloroplasts and increased photosynthesis through over accumulation of glycine betaine (Wang et al., 2010).

It will be interesting to know which paralog of wheat BADH provides drought tolerance as one of the paralogs is found to be responsible for drought resistance in barley (Fujiwara et al., 2008) and rice (Fitzgerald et al., 2008). Furthermore, these paralogs showed different affinity towards substrate. BBD1 paralogs of barley showed very little affinity towards betaine aldehyde compared to BBD2 (Fujiwara et al., 2008). While in the case of rice, both paralogs showed greater affinity to GABald than betaine aldehyde.

The expression of *BADH* homologs in four different hexaploid wheat varieties at two different time points and two different tissue types was investigated and it was found *BADH2* gene expression did not differ between samples. The varieties were from different geographical locations and according to (Kovach et al., 2009), rice *BADH2* exhibited polymorphism across a geographical range. In contrast to this, wheat *BADH2* doesn't show any polymorphism indicating wheat *BADH2* is under stronger purifying natural selection relative to rice.

Among four hexaploid wheat varieties, only BADH1 differed at the amino acid level in the expressed sequences (mRNA) while BADH2 was identical among the varieties. NW51A BADH1 amino acid sequence differed from the other varieties by one amino acid, it had proline at position 3 (Figure 3.8). While the other varieties had threonine. At the level of gDNA, *BADH1* differed between homeologs of hexaploid wheat variety Cadoux and its progenitor species sequences, homeolog A and B and *T. monococcum* had glutamine at position 76. While homeolog D and *T. speltoides* had glutamic acid (Figure 2.8). Both amino acid changes lay at the periphery of the protein and so these differences in deduced amino acid is highly unlikely to affect conformation of BADH1 when this was compared with the known 3D structure of pea ALDH. Overall the results indicate that two BADH homologs might have different biochemical function *in vivo* and more investigation is required to confirm their function in plants.

Chapter 4: General discussion and conclusion

BADH is responsible for fragrance in rice (Bradbury et al., 2008; Shi et al., 2008) and soybean (Arikit et al.; Juwattanasomran et al., 2010) and the aim of this project was to identify the number of BADH homologs in hexaploid wheat and their progenitors and determine their expression patterns in hexaploid wheat.

In hexaploid wheat, there are three homeologs for each *BADH* paralog in contrast to their diploid progenitors where only one ortholog is present. Gene expression in polyploids change according to the developmental stages of the tissue (Adams et al., 2003; Adams et al., 2004). Therefore gene expression in two different tissues (leaves and seeds) and two different time points (14 DPA and 30 DPA) were investigated. Although three *BADH1* and *BADH2* homeologs were present, only one of each is expressed in these two tissues (leaves and seeds) and time points (14 DPA and 30 DPA).

If all three *BADH2* homeologs were expressed in hexaploid wheat, it may be possible to knock out one of the homeologs without disturbing normal *BADH2* function. However, the data here suggests this is not possible and knocking out the one homeolog which is expressed may have a significant impact on plant performance, given that fragrance in rice is associated with salt sensitivity (Fitzgerald et al, 2010). However, it is difficult to predict the exact function of this gene and its effect on the plant performance.

Partial *BADH2* gene sequencing revealed this gene is under some sort of external selection pressure because this gene is highly conserved among these samples, in contrast to *BADH1* in which SNPs were detected. Expression of a single *BADH* homeolog means the remaining two homoelogs may be silenced or may not be up-regulated or might have become pseudo genes after polyploid formation.

Hexaploid wheat has a highly stable genome and exhibits quasi diploid behaviour genetically by maintaining triplicate parallel genomes which do not recombine (He et al., 2003). Genomic re-arrangement through gene loss substitutions, insertion or deletions is minimal in wheat (Bryan et al., 1999). There is minimal polymorphism between *BADH* homologs, so gene silencing due to gene deletion or insertion is highly unlikely. Therefore, epigenetic factors such as hypermethylation of DNA cytosine (Lee and Chen, 2001), deacetylation, methylation or other modification of histones or change in chromatin structure could be the reason for gene silencing (Richards and Elgin, 2002). Other possibilities include synthesis of antisense transcripts' generated by readout transcription of a retrotransposon which mediated the silencing of and adjacent gene (He et al., 2003) or small RNAs which interact and leads to gene silencing (Comai et al., 2003).

The data here suggests there is no chromosomal loss of *BADH* homologs, however, only one genome is expressed, supporting the hypothesis that silencing and not gene loss occurs during polyploid formation (He et al., 2003; Chague et al., 2010). Further research is required to indentify the silencing mechanism of these two homoelogs. Several epigenetic factors could be the reasons for gene silencing and play important role in expression of progenitor's gene in polyploid formation and speciation (Lee and Chen, 2001). DNA methylation which interferes with transcription is the most common cause of gene silencing in eukaryotes including *A. thaliana* (Lee and Chen, 2001)and would therefore be the first mechanism to explore in relation to *BADH* silencing. Following this, investigation of promoter structure of these genes could be important as well.

Rice is an aquatic plant and undergoes relatively little drought stress and does not accumulate glycine betaine (Ishitani et al., 1993). Unlike rice, wheat is exposed to more drought stress and accumulates glycine betaine up to 20 µmol/g dry weight in leaves and the concentration increases several fold with an increase in water and salt stress (Hanson and Nelsen, 1978;

Hitz and Hanson, 1980). Malfunction of *BADH2* is responsible for fragrance in rice (Bradbury et al., 2008; Shi et al., 2008) and soybean (Arikit et al.; Juwattanasomran et al., 2010). However, knocking out this ortholog in wheat doesn't necessarily mean it will produce fragrant wheat that would be suitable for cultivation since it may play an important role in abiotic or drought tolerance and knocking out this gene could have a detrimental effect on plant performance.

Many studies have shown that plant BADH catalyse a range of substrates more effectively than betaine aldehyde (Trossat et al., 1997; Livingstone et al., 2003). Enzymes with aminoaldehydes dehydrogenase activity which are very similar to betaine aldehyde dehydrogenase do not show any affinity to betaine aldehyde (Sebela et al., 2000). BADH paralogs of barley have a higher V_{max}/k_m with omega-amino aldehyde compared to betaine aldehyde (Nakamura et al., 2001) and similar results were seen in other plants and human and in some bacteria (Chern and Pietruszko, 1995; Trossat et al., 1997; Livingstone et al., 2003). However, mangrove orthologs did not show any affinity towards omega amino aldehyde (Hibino et al., 2001). Fragrance in rice appears to be due to the accumulation of 4-aminobutyraldehyde /Δ^1 –pyrroline, which is a precursor of 2AP. Both isozymes showed greater affinity and high catalytic efficiency towards 3-aminopropionaldehyde and GGBald. However, BADH2 had the greatest affinity for 4-aminobutyraldehyde which can spontaneous cyclise to form Δ^1 –pyrroline from which the fragrant compound 2AP is formed in fragrant rice (Bradbury et al., 2008). There is no report of fragrant wheat; however, accumulation of glycine betaine is reported in wheat suggesting BADH is correctly processed. An alternate substrate for BADH could be 3-aminopropionaldehyde, 4-aminoguanidinobutyraldehyde and 4-aminobutyraldehyde which exist in equilibrium with Δ^1 -pyrroline (Trossat et al., 1997; Livingstone et al., 2003) apart from betaine aldehyde which is the precursor of 2AP.

Unlike rice, wheat accumulates glycine betaine and given the range substrate specificities of BADH enzymes, wheat BADH could have different substrate specificity to rice and BADH enzymes of other species or it could be bi-functional and use wide range of substrates. In conclusion, more work needs to be carried out at biochemical for substrate specificity to know how this BADH homologs works and its physiochemical role in plants.

Bibliography

Adams, K.L., Percifield, R., and Wendel, J.F. (2004). Organ-specific silencing of duplicated genes in a newly synthesized cotton allotetraploid. Genetics **168,** 2217-2226.

Adams, K.L., Cronn, R., Percifield, R., and Wendel, J.F. (2003). Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. Proceedings of the National Academy of Sciences of the United States of America **100,** 4649-4654.

Ahn, S., Anderson, J.A., Sorrells, M.E., and Tanksley, S.D. (1993). Homoelogous relationships of Rice,Wheat and Maize chromosomes. Molecular & General Genetics **241,** 483-490.

Allard, F., Houde, M., Krol, M., Ivanov, A., Huner, N.P.A., and Sarhan, F. (1998). Betaine improves freezing tolerance in wheat. Plant Cell Physiology **39,** 1194-1202.

Arakawa, K., Katayama, M., and Takabe, T. (1990). Levels of Betaine and Betaine Aldehyde Dehydrogenase Activity in the Green Leaves, and Etiolated Leaves and Roots of Barley. Plant Cell Physiology **31,** 797-803.

Arikit, S., Yoshihashi, T., Wanchana, S., Uyen, T.T., Huong, N.T.T., Wongpornchai, S., and Vanavichit, A. (2010). Deficiency in the amino aldehyde dehydrogenase encoded by GmAMADH2, the homologue of rice Os2AP, enhances 2-acetyl-1-pyrroline biosynthesis in soybeans (*Glycine max* L.). Plant Biotechnology Journal**,** 1-13.

Auer, H., Lyianarachchi, S., Newsom, D., Klisovic, M.I., Marcucci, u., and Kornacker, K. (2003). Chipping away at the chip bias: RNA degradation in microarray analysis. Nature Genetics **35,** 292-293.

Bachem, C.W.B., Van Der Hoeven, R.S., De Bruijn, S.M., Vreugdenhil, D., Zabeau, M., and Visser, R.G.F. (1996). Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: Analysis of gene expression during potato tuber development. The Plant Journal **9,** 745-753.

Blank, I., Devaud, S.p., Matthey-Doret, W., and Robert, F. (2003). Formation of odorants in maillard model systems based on l-proline as affected by pH. Journal of Agricultural and Food Chemistry **51,** 3643-3650.

Bottley, A., Xia, G.M., and Koebner, R.M.D. (2006). Homoeologous gene silencing in hexaploid wheat. Plant Journal **47,** 897-906.

Bradbury, L.M.T., Fitzgerald, T.L., Henry, R.J., Jin, Q.S., and Waters, D.L.E. (2005). The gene for fragrance in rice. Plant Biotechnology **3,** 363-370

Bradbury, L.M.T., Gillies, S.A., Brushett, D.J., Waters, D.L.E., and Henry, R.J. (2008). Inactivation of an aminoaldehyde dehydrogenase is responsible for fragrance in rice. Plant Molcular Biology **68,** 439-449.

Brauner, F., Sebela, M., Snegaroff, J., Pec, P., and Meunier, J.C. (2003). Pea seedling aminoaldehyde dehydrogenase: primary structure and active site residues. Plant Physiology and Biochemistry **41,** 1-10.

Bryan, G.J., Stephenson, P., Collins, A., Kirby, J., Smith, J.B., and Gale, M.D. (1999). Low levels of DNA sequence variation among adapted genotypes of hexaploid wheat. Theoretical and Applied Genetics **99,** 192-198.

Buttery, R.G., Juliano, B.O., and Ling, L.C. (1983a). Identification of rice aroma compound 2-acetyl-1- pyrroline in pandan leaves. Chemistry and Industry**,** 478-478.

Buttery, R.G., Ling, L.C., Juliano, B.O., and Turnbaugh, J.G. (1983b). Cooked rice aroma and 2-acetyl-1-pyrroline. Jouranl of Agricultural and Food Chemistry **31,** 823-826.

Carraway, M., and Marinus, M.G. (1993). Repair of heteroduplex DNA-molecules with multibase loops in *Escherichia-coli*. Jouranal of Bacteriology **175,** 3972-3980.

Cha-um, S., Kirdmanee, C., and Supaibulwatana, K. (2004). Biochemical and physiological response of thai jasmine rice(Oryza L.ssp. *indica* cv. KDML105) to salt stress. ScienceAsia 247-253.

Chague, V., Just, J., Mestiri, I., Balzergue, S., Tanguy, A.M., Huneau, C., Huteau, V., Belcram, H., Coriton, O., Jahier, J., and Chalhoub, B. (2010). Genome-wide gene expression changes in genetically stable synthetic and natural wheat allohexaploids. New Phytologist **187,** 1181-1194.

Chen, S.H., Yang, Y., Shi, W.W., Ji, Q., He, F., Zhang, Z.D., Cheng, Z.K., Liu, X.N., and Xu, M.L. (2008). Badh2, encoding betaine aldehyde dehydrogenase, inhibits the biosynthesis of 2-acetyl-1-pyrroline, a major component in rice fragrance. Plant Cell **20,** 1850-1861.

Chen, Z.J., and Ni, Z. (2006). Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. Bioessays **28,** 240-252.

Chern, M.K., and Pietruszko, R. (1995). Human aldehyde dehydrogenase e3 isozyme is a betaine aldehyde dehydrogenase. Biochemical and Biophysical Research Communications **213,** 561-568.

Clark, J.R., Robertson, M., and Ainsworth, C.C. (1991). Nucleotide-sequencing of a wheat (*Triticum-aestivum*) cDNA clone encoding the waxy protein. Plant Molcular Biology **16,** 1099-1101.

Comai, L., Madlung, A., Josefsson, C., and Tyagi, A. (2003). Do the different parental 'heteromes' cause genomic shock in newly formed allopolyploids? Philosophical Transactions of Royal Society of London Sereries B-Biology Science **358,** 1149-1155.

Delmundo, A.M., and Juliano, B.O. (1981). Consumer preference and properties of raw and cooked milled rice. Journal of Texture Studies **12,** 107-120.

Elizabeth A, W., and Hanson, A.D. (1990). Molecular cloning of a plant betaine-aldehyde dehydrogenase an enzyme implicated in adaptation to salinity and drought. Proceedings of the National Academy of Sciences of the United States of America **87,** 2745-2749.

Feuillet, C., Langridge, P., and Waugh, R. (2008). Cereal breeding takes a walk on the wild side. Trends in Genetics **24,** 24-32.

Fitzgerald, T.L., Waters, D.L.E., and Henry, R.J. (2008). The effect of salt on betaine aldehyde dehydrogenase transcript levels and 2-acetyl-1-pyrroline concentration in fragrant and non-fragrant rice (*Oryza sativa*). Plant Science **175,** 539-546.

Fujiwara, T., Hori, K., Ozaki, K., Yokota, Y., Mitsuya, S., Ichiyanagi, T., Hattori, T., and Takabe, T. (2008). Enzymatic characterization of peroxisomal and cytosolic betaine aldehyde dehydrogenases in barley. Physiologia Plantarum **134,** 22-30.

Guo, B.H., Zhang, Y.M., Li, H.J., Du, L.Q., Li, Y.X., Zhang, J.S., Chen, S.Y., and Zhu, Z.Q. (2000). Transformation of wheat with a gene encoding for the betaine aldehyde dehydrogenase (BADH). Acta Botanica Sinicia **42,** 279-283.

Guo, Y., Zhang, L., Xiao, G., Cao, S.Y., Gu, D.M., Tian, W.Z., and Chen, S.Y. (1997). Expression of betaine aldehyde dehydrogenase gene and salinity tolerance in rice transgenic plants. Science in China Series C-Life Sciences **40,** 496-501.

Hanson, A.D., and Nelsen, C.E. (1978). Betaine-accumulation and formate-C-14 metabolism in water stressed barley leaves. Plant Physiology **62,** 305-312.

Hanson, A.D.M., A. M.Grumet, R. Bode, J. Jamieson,G.C.Rhodes, D. (1985). Betaine synthesis in chenopods: Localization in chloroplasts. Proceedings of the National Academy of Sciences of the United States of America **82,** 3678-3682.

Hasthanasombut, S., Ntui, V., Supaibulwatana, K., Mii, M., and Nakamura, I. (2010). Expression of Indica rice <i>OsBADH1</i> gene under salinity stress in transgenic tobacco. Plant Biotechnology Reports **4,** 75-83.

He, P., Friebe, B.R., Gill, B.S., and Zhou, J.M. (2003). Allopolyploidy alters gene expression in the highly stable hexaploid wheat. Plant Molcular Biology **52,** 401-414.

Hibino, T., Meng, Y.L., Kawamitsu, Y., Uehara, N., Matsuda, N., Tanaka, Y., Ishikawa, H., Baba, S., Takabe, T., Wada, K., and Ishii, T. (2001). Molecular cloning and functional

characterization of two kinds of betaine-aldehyde dehydrogenase in betaine-accumulating mangrove *Avicennia marina* (Forsk.) Vierh. Plant Molcular Biology **45,** 353-363.

Hitz, W.D., and Hanson, A.D. (1980). Determination of glycine betaine by pryolysis-gas chromatography in cereals and grasses. Phytochemistry **19,** 2371-2374.

Holmstrom, K.O., Welin, B., Mandal, A., Kristiansdottir, I., Teeri, T.H., Lamark, T., Strom, A.R., and Palva, E.T. (1994). Production of the *Escherichia-coli* betaine-aldehyde dehydrogenase, an enzyme required for the synthesis of the osmoprotectant glycine betaine, in transgenic plants. . Plant Journal **6,** 749-758.

Ishitani, M., Arakawa, K., Mizuno, K., Kishitani, S., and Takabe, T. (1993). Betaine aldhehyde dehydrogenase in the gramineae-levels in leaves of both betiaine-accumulating and nonaccumulating cereal plants. Plant Cell Physiology **34,** 493-495.

Jia, G.X., Zhu, Z.Q., Chang, F.Q., and Li, Y.X. (2002). Transformation of tomato with the BADH gene from *Atriplex* improves salt tolerance. Plant Cell Reports **21,** 141-146.

Johansson, K., El-Ahmad, M., Ramaswamy, S., Hjelmqvist, L., Jornvall, H., and Eklund, H. (1998). Structure of betaine aldehyde dehydrogenase at 2.1 angstrom resolution. Protein Science **7,** 2106-2117.

Juwattanasomran, R., Somta, P., Kaga, A., Chankaew, S., Shimizu, T., Sorajjapinun, W., and Srinives, P. (2010). Identification of a new fragrance allele in soybean and development of its functional marker. Molecular Breeding**,** 1-9.

Kanagawa, T. (2003). Bias and artifacts in multitemplate polymerase chain reactions (PCR). Journal of Bioscience and Bioengineering. **96,** 317-323.

Kashkush, K., Feldman, M., and Levy, A.A. (2002). Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. Genetics **160,** 1651-1659.

Kashkush, K., Feldman, M., and Levy, A.A. (2003). Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. Nature Genetics **33,** 102-106.

Kishitani, S., Watanabe, K., Yasuda, S., Arakawa, K., and Takabe, T. (1994). Acummulation of glycinebetaine during cold-acclimation and freezing tolerance in leaves in winter and spring barley plants. Plant Cell Environment **17,** 89-95.

Koh, J., Soltis, P.S., and Soltis, D.E. (2010). Homeolog loss and expression changes in natural populations of the recently and repeatedly formed allotetraploid *Tragopogon mirus* (Asteraceae). BMC Genomics **11,** 16.

Kovach, M.J., Calingacion, M.N., Fitzgerald, M.A., and McCouch, S.R. (2009). The origin and evolution of fragrance in rice (*Oryza sativa* L.). Proceedings of the National Academy of Sciences of the United States of America **106,** 14444-14449.

Krishnamurthy, R., and Bhagwat, K.A. (1990). Accumulation of choline and glycine betaine in salt stress wheat seedlings. Current science **59,** 111-112.

Kumazawa, K., and Masuda, H. (2002). Identification of potent odorants in different green tea varieties using flavor dilution technique. Jouranl of Agricultural and Food Chemistry **50,** 5660-5663.

Learn, B.A., and Grafstrom, R.H. (1989). Methyl-directed repair of framshift heteroduplexes in cell-extracts from Escherichia-coli. Journal of Bacteriology **171,** 6473-6481.

Lee, H.S., and Chen, Z.J. (2001). Protein-coding genes are epigenetically regulated in arabidopsis polyploids. Proceedings of the National Academy of Sciences of the United States of America **98,** 6753-6758.

Lerudulier, D., Strom, A.R., Dandekar, A.M., Smith, L.T., and Valentine, R.C. (1984). Molecular -biology of osmoregulation. Science **224,** 1064-1068.

Levy, A.A., and Feldman, M. (2004). Genetic and epigenetic reprogramming of the wheat genome upon allopolyploidization. Biological Journal of the Linnean Society **82,** 607-613.

Liang, C., Zhang, X.Y., Luo, Y., Wang, G.P., Zou, Q., and Wang, W. (2009). Overaccumulation of glycine betaine alleviates the negative effects of salt stress in wheat. Russian Journal of Plant Physiology **56,** 370-376.

Liu, B., Vega, J.M., and Feldman, M. (1998). Rapid genomic changes in newly synthesized amphiploids of Triticum and Aegilops. II. Changes in low-copy coding DNA sequences. Genome **41,** 535-542.

Livingstone, J.R., Maruo, T., Yoshida, I., Tarui, Y., Hirooka, K., Yamamoto, Y., Tsutui, N., and Hirasawa, E. (2003). Purification and properties of betaine aldehyde dehydrogenase from *Avena sativa*. Journal of Plant Research **116,** 133-140.

Lorieux, M., Petrov, M., Huang, N., Guiderdoni, E., and Ghesquière, A. (1996). Aroma in rice: genetic analysis of a quantitative trait. Theoretical and Applied Genetics **93,** 1145-1151.

McCue, K.F., and Hanson, A.D. (1992). Salt-inducible betaine aldehyde dehydrogenase from sugar beet: cDNA cloning and expression. Plant Molecular Biology **18,** 1-11.

Mori, N., Fuchigami, S., and Kitamoto, Y. (2002). Purification and properties of betaine aldehyde dehydrogenase with high affinity for NADP from *Arthrobacter globiformis*. Journal of Biosciencess and Bioengineering **93,** 130-135.

Nakamura, T., Normura M, Mori H, Jagendrof AT, Ueda A, and Takabe, T. (2001). An isozyme of betaine aldehyde dehydrogenase in barley. Plant Cell Physiol. **42,** 1088-1092.

Nakamura, T., Yokota, S., Muramoto, Y., Tsutsui, K., Oguri, Y., Fukui, K., and Takabe, T. (1997). Expression of a betaine aldehyde dehydrogenase gene in rice, a glycienbetaine nonaccumulator, and possible localization of its protein in peroxisomes. The Plant Journal **11,** 1150-1120.

Niu, X.L., Tang, W., Huang, W.Z., Ren, G.J., Wang, Q.L., Luo, D., Xiao, Y.Y., Yang, S.M., Wang, F., Lu, B.R., Gao, F.Y., Lu, T.G., and Liu, Y.S. (2008). RNAi-directed downregulation of OsBADH2 results in aroma (2-acetyl-1-pyrroline) production in rice (*Oryza sativa* L.). BMC Plant Biology **8,** 10.

Nomura, T., Ishihara, A., Yanagita, R.C., Endo, T.R., and Iwamura, H. (2005). Three genomes differentially contribute to the biosynthesis of benzoxazinones in hexaploid wheat.

Proceedings of the National Academy of Sciences of the United States of America **102,** 16490-16495.

Oishi, H., and Ebina, M. (2005). Isolation of cDNA and enzymatic properties of betaine aldehyde dehydrogenase from *Zoysia tenuifolia*. Journal of Plant Physiology **162,** 1077-1086.

Okamoto, M. (1962). Identification of the chromosomes of common wheat belonging to the A and B genomes. Canadian of Journal of Genetics and Cytology **4,** 31-37.

Richards, E.J., and Elgin, S.C.R. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. Cell **108,** 489-500.

Riley, R., and Chapman, V. (1958). Genetic control of the cytologically diploid behaviour of hexaploid wheat. Nature **182,** 713-715.

Riley, R., and Kempanna, C. (1963). The homoeologous nature of the non-homologous meiotic pairing in *Triticum aestivum* deficient for chromosome V (5B). Heredity **18,** 287-306.

Robinson, S.P., and Jones, G.P. (1986). Accumulation of glycinebetaine in chloroplast provides osmotic adjustment during salt stress. Australian Journal of Plant Physiology **13,** 659-668.

Schieberle, P. (1991). Primary odorants in popcorn. Journal of Agricultral and Food Chemistry **39,** 1141-1144.

Sears, E.R. (1976). Genetic-control of chromosome-pairing in wheat. Annual Review of Genetics **10,** 31-51.

Sebela, M., Brauner, F., Radova, A., Jacobsen, S., Havlis, J., Galuszka, P., and Pec, P. (2000). Characterisation of a homogeneous plant aminoaldehyde dehydrogenase. Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology **1480,** 329-341.

Shi, W., Yang, Y., Chen, S., and Xu, M. (2008). Discovery of a new fragrance allele and the development of functional markers for the breeding of fragrant rice varieties. Molecular Breeding **22,** 185-192.

Shitsukawa, N., Tahira, C., Kassai, K.-i., Hirabayashi, C., Shimizu, T., Takumi, S., Mochida, K., Kawaura, K., Ogihara, Y., and Murai, K. (2007). Genetic and epigenetic alteration among three homoeologous genes of a class E MADS box gene in hexaploid wheat. Plant Cell **19,** 1723-1737.

Sourdille, P., Tavaud, M., Charmet, G., and Bernard, M. (2001). Transferability of wheat microsatellites to diploid Triticeae species carrying the A, B and D genomes. Theoretical and Applied Genetics **103,** 346-352.

Stelmakh, A.F. (1992). Genetic effects of Vrn genes on heading date and agronomic traits in bread wheat. Euphytica **65,** 53-60.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution **24,** 1596-1599.

Tonsor, S.J., Alonso-Blanco, C., and Koornneef, M. (2005). Gene function beyond the single trait: natural variation, gene effects, and evolutionary ecology in *Arabidopsis thaliana*. Plant Cell Environment **28,** 2-20.

Trossat, C., Rathinasabapathi, B., and Hanson, A.D. (1997). Transgenically expressed betaine aldehyde dehydrogenase efficiently catalyzes oxidation of dimethylsulfoniopropionaldehyde and omega-aminoaldehydes. Plant Physiology **113,** 1457-1461.

Vojtechova, M., Hanson, A.D., and MunozClares, R.A. (1997). Betaine-aldehyde dehydrogenase from amaranth leaves efficiently catalyzes the NAD-dependent oxidation of dimethylsulfoniopropionaldehyde to dimethylsulfoniopropionate. Archives of Biochemistry and Biophysics **337,** 81-88.

Wang, G.P., Zhang, X.Y., Li, F., Luo, Y., and Wang, W. (2010). Overaccumulation of glycine betaine enhances tolerance to drought and heat stress in wheat leaves in the protection of photosynthesis. Photosynthetica **48,** 117-126.

Weigel, P., Weretilnyk, E.A., and Hanson, A.D. (1986). Betaine aldehyde oxidation by spinach chloroplasts. Plant Physiology **82,** 753-759.

Weigel, P.L., L. Hanson, A.D. (1988). Choline oxidation by intact spinach chloroplasts. Plant Physiology **86,** 54-60.

Wendel, J.F. (2000). Genome evolution in polyploids. Plant Molcular Biology **42,** 225-249.

Weretilnyk, E.A., and Hanson, A.D. (1988). Betaine aldehyde dehydrogenase polymorphism in spinach- genetic and biochemical characterization. Biochemistry Genetics. **26,** 143-151.

Weretilnyk, E.A., and Hanson, A.D. (1989). Betaine aldehyde dehydrogenase from spinach leaves: purification, in vitro translation of the mRNA, and regulation by salinity. Archives of Biochemistry and Biophysics **271,** 56-63.

Widjaja, R., Craske, J.D., and Wootton, M. (1996). Comparative studies on volatile components of non-fragrant and fragrant Rices. Journal of the Science of Food and Agriculture **70,** 151-161.

Wood, A.J., Saneoka, H., Rhodes, D., Joly, R.J., and Goldsbrough, P.B. (1996). Betaine aldehyde dehydrogenase in sorghum - molecular cloning and expression of two related genes. Plant Physiology **110,** 1301-1308.

Yan, L.L., Bhave, M., Fairclough, R., Konik, C., Rahman, S., and Appels, R. (2000). The genes encoding granule-bound starch synthases at the waxy loci of the A, B, and D progenitors of common wheat. Genome **43,** 264-272.

Yang, W.J., Rich, P.J., Axtell, J.D., Wood, K.V., Bonham, C.C., Ejeta, G., Mickelbart, M.V., and Rhodes, D. (2003). Genotypic variation for glycinebetaine in sorghum. Crop Science **43,** 162-169.

Yoshihashi, T. (2002). Quantitative analysis on 2-acetyl-1-pyrroline of an aromatic rice by stable isotope dilution method and model studies on its formation during cooking. In Journal of Food Science, pp. 619-622.

Yoshihashi, T., Huong, N.T.T., and Inatomi, H. (2002). Precursors of 2-acetyl-1-pyrroline, a potent flavor compound of an aromatic rice variety. Jouranl of Agricultural and Food Chemistry **50,** 2001-2004.

Yoshihashi, T., Nguyen, T.T.H., and Kabaki, N. (2004). Area dependency of 2-acetyl-1-pyrroline content in an aromatic rice variety, Khao Dawk Mali 105. Jarq-Japan Agricultural Research Quarterly **38,** 105-109.

Zhou, S.F., Chen, X.Y., Zhang, X.G., and Li, Y.X. (2008). Improved salt tolerance in tobacco plants by co-transformation of a betaine synthesis gene BADH and a vacuolar Na+/H+ antiporter gene SeNHX1. Biotechnology Letters **30,** 369-376.

Appendix i

*BADH*1 Genomic DNA sequences of hexaploid wheat and their progenitors

T. monococcum

AAGAAGTGAAGGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGGTAAGCAT
TGGTCATTACATTACATGCTCTTGCCTCTCGGTTATTCATCCAAGCTACGTTTGTG
AAATGGCAGCATCTCGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGAC
GTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGTG
TCAAAGTATTCAAGACAGAGAGTGAAGCTGTAGAGCTTGCAAATGATACTCAGT
GAGTTACCTGATTTGCGCAAGCCTGAAAAAAAAGCCCGTGCTTATATCGTTCCCT
TGACTTCTCTACCATTTTGTTTTCAGCTATGGCTTGGCTGGTGGTGTGATCTCTGA
TGATCTACAGAGGTGTGAGCGCATTGCAAAGGTAGATTCAAGCACGCCCGTAAA
TTTGTAGCGTGTGATGCAATTGAAGTATTCGAAATATTCAGTGAGCTGATAGGTA
AACGTACAATCCCACACCTGCAGGTTATTCACTCAGGCATTGTTTGGATAAACTG
CTCGCAACCGACCCTGGTTCAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTT
TGGCCGGGAGCTAGGAGAATGGTGAGTTACTTCACGAGCGCTATTGTTCGCAAA
CTGCACTGCCCCTGGAAGTATCCTCTGAACTGTCTTTGCTGTTTTGCTTTCAGGGG
CCTCGAGAACTACTTGAGCGTGAAACAAGTCACCAGGT

T .speltoides

CAGCAAGAAATGAAGGTGCTACAATTTTGCATGGGGGTGACCGACCAAAGGTAA
GCATTGATCACTACATTATATTACATGCTCTTGTCTCTCGGTTATTCATCCAAGCT
ACGTTTGTGAAATGGCAGCATCTCGGAAAAGGGTTCTTTATTGAACCTACTATTA
TAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAG
TCATCTGTGTCAAAGTATTTAAGACAGAGAGCGAAGCCGTAGAGCTTGCAAATG
ATACCCAGTGAGTTAACTGATTTGCGCAAGCCTGAAGAAAAAACTTCCATGCTTA
TATCACTCCCTTGACTTCCCTACCATTTTATTTTCAGCTATGGCTTGGCTGGTGGT
GTGATCTCTGATGATCTAGAGAGGTGTGAGCGCATTGCAAAGGTAGATTCAAGTT
GAACTCTGAACATGCACGTAAATTTTGAACGTGTGATGAGATCGAAGTATTTGGT
GAGCTGATAGGTAAACGTACAAACCCACACCTGCAGGTTATTCACTCGGGCAT
TGTTTGGATAAACTGCTCGCAACCGACTCTGGTTCAAGCTCCGTGGGGAGGGAA
AAGCGTAGTGGTTTTGGCCGGGAGCTAGGAGAATGGTGAATTGACTTCATGAGC
GCTATGGTTCACAAACTGCAGCTAGTATCCTCTGAACCATCTTTGCTGTTTTGCTT
TCAGGGGCCTCGAGAACTACCTGAGCGTGAAA

T .tauschii

CAGCAAGAAATGAAGGTGCTACAATTTTGCATGGGGGTGACCGACCAAAGGTAA
GCATTGATCACTACATTATATTACATGCTCTTGTCTCTCGGTTATTCATCCAAGCT
ACGTTTGTGAAATGGCAGCATCTCGGAAAAGGGTTCTTTATTGAACCTACTATTA
TAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAG
TCATCTGTGTCAAAGTATTTAAGACAGAGAGCGAAGCCGTAGAGCTTGCAAATG
ATACCCAGTGAGTTAACTGATTTGCGCAAGCCTGAAGAAAAAACTTCCATGCTTA
TATCACTCCCTTGACTTCCCTACCATTTTATTTTCAGCTATGGCTTGGCTGGTGGT
GTGATCTCTGATGATCTAGAGAGGTGTGAGCGCATTGCAAAGGTAGATTCAAGTT
GAACTCTGAACATGCACGTAAATTTTGAACGTGTGATGAGATCGAAGTATTTGGT
GAGCTGATAGGTAAACGTACAAACCCACACCTGCAGGTTATTCACTCGGGCAT
TGTTTGGATAAACTGCTCGCAACCGACTCTGGTTCAAGCTCCGTGGGGAGGGAA
AAGCGTAGTGGTTTTGGCCGGGAGCTAGGAGAATGGTGAATTGACTTCATGAGC
GCTATGGTTCACAAACTGCAGCTAGTATCCTCTGAACCATCTTTGCTGTTTTGCTT
TCAGGGGCCTCGAGAACTACCTGAGCGTGAAA

Hexaploid wheat variety Cadoux

A Genome

CAGCAAGAAGTGAAGGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGGTAA
GCATTGGTCATTACATTACATGCTCTTGTCTCTCGGTTATTCATCCAAGCTACGTT
TGTGAAATGGCAGCATCTCGGAAAAGGGTTCTTTATTGAACCTACAATTATAACA
GATGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGTCATCT
GTGTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAATGATACCC
AGTGAGTTACCTGATTGCGCAAGCCTGAAAAAAAAATGCCCGTGCTTATATCGTT
CCCTTGACTTCTCTGCCATTTTGTTTTCAGCTATGGCTTGGCTGGTGGTGTGATCT
CTGATGATCTACAGAGGTGTGAGCGCATTGCAAAGGTAGATTCAAGTGCGCCTG
TAAATTTGTAGCATGTGATGCAATCGAAGTATTCGAAATATTCGGTGAGCTGATA
GGTAAACGTACAATCCCACACCTGCAGGTTATTCACTCAGGCATTGTTTGGAT
AAACTGCTCGCAACCGACCCTTGTTCAAGCTCCGTGGGGAGGGAACAAGCGGAG
TGGTTTTGGCCGGGAGCTAGGAGAATGGTGAGTTACTTCATGAGCGCTATGGTTC
GCAAACTGCACTGCCCCTGAAAGTATCTTCTGAACTGTCTGCTGTTTTGCTTTCAG
GGGCCTCGAGAACTACTTGAGCGTGAAACAAGTCACCAGGTA

B Genome

CAGCAAGAAGTGAAGGTGCTACAATTTTGCATGGGGGTGACCGACCAAAGGTAA
GCATTGATCACTACATTATATTACATGCTCTTGTCTCTCGGTTATTCATCCAAGCT
ACGTTTGTGAAATGGCAGCATCTCGGAAAAGGGTTCTTTATTGAACCTACTATTA
TAACAGACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAG
TCATCTGTGTCAAAGTATTTAAGACAGAGAGCGAAGCCGTAGAGCTTGCAAATG
ATACCCAGTGAGTTAACTGATTTGCGCAAGCCTGAAGAAAAAACTTCCATGCTTA
TATCACTCCCTTGACTTCCCTACCATTTTATTTTCAGCTATGGCTTGGCTGGTGGT
GTGATCTCTGATGATCTACAGAGGTGTGAGCGCATTGCAAAGGTAGATTCAAGTT
GAACTCTGAACATGCACGTAAATTTTGAATGTGTGATGAGATCGAAGTATTTGGT
GAGCTGATAGGTAAACGTACAAACCCACACCTGCAGGTTATTCACTCGGGCAT

TGTTTGGATAAACTGCTCGCAACCGACTCTGGTTCAAGCTCCGTGGGGAGGGAAC
AAGCGTAGTGGTTTTGGCCGGGAGCTAGGAGAATGGTGAGTTGACTTCATGAGC
GCTATGGTTCACAAACTGCAGCTAGTATCCTCTGAACCATCTTTGCTGTTTTGCTT
TCAGGGGCCTCGAGAACTACCTGAGCGTGAAACAAGTCACCAGGTA

D Genome

CAGCAAGAAGTGAAGGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGGTAA
GAATTGGTCATTACATTACGTGCTCTTGCCTCTCGGTTATTCATCCAAGCTACGTT
TGTGAAATGGCAGCATCTCGGAAAAGGGTTCTTTATTGAACCTACAATTATAACA
GACGTTAGCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGGCCAGTCATC
TGTGTCAAAGTATTCAAGACAGAGAGTGAAGCTGTAGAGCTTGCAAATGATACC
CAGTGAGTTACCTGATTTGCGCAAGCCTGAAAAAAAAGCCCGTGCTTATRTCGTT
CCCTTGACTTCYCTACCATTTTGTTTTCAGCTATGGCTTGGCTGGTGGTGTGATCT
CTGATGATCTACAGAGGTGTGAGCGCATTGCAAAGGTAGATTCAAGCACGCCCG
TAAATTTGTAGCGTGTGATGCAATTGAAGTATTCGAAATATTCAGTGAGCTGATA
GGTAAACGTACAATCCCACACCTGCAGGTTATTCACTCAGGCATTGTTTGGAT
AAACTGCTCGCAACCGACCCTGGTTCAAGCTCCGTGGGGAGGGAACAAGCGG
AGCGGTTTTGGCCGGGAGCTAGGAGAATGGTGAGTTACTTCACGAGCGCTATGG
TTTGCAAACTGCACTGCCCCTGGAAGTATCCTCTGAACTGTCTTTGCTGTTTTGCT
TTCAGGGGCCTCGAGAACTACTTGAGCGTGAAACAAGTCACCAGGTA

Appendix ii

*BADH*2 Genomic DNA sequences of hexaploid wheat and their progenitors

T. monococcum

TCTTAAACATTGTGACTGGATTAGGTCATGAAGCTGGCGCTCCTTTGTCGTCACA
CCCTGACGTCGACAAGGTACATATATTTGTCCTATACCTTTTTGTGATCTATCTAT
GGGTTATTCCTCACAATACATGGCTTATTTGATCTTTTAGGTTGCATTTACCGGGA
GCTATGCAACCGGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGGTTTG
TTCCACATGTCTCTATTTTTGAAAATTTTCACCAAATATCCCGTTTTTTCCCCTCTA
GATTTTTAGAATATTTACTAAGATACTATGAGCGCTCTCCTGATAGCAGACCATT
CTCACCTTTTGTGTTGCGTATTTTCTTGACAGCCTGTTACATTGGAACTTG

T. speltoides

GGTGTCTTAAACATTGTGACTGGATTAGGTCAGAAGCTGGCGCTCCTTTGTCGTC
ACACCCTGATGTCGACAAGGTACATATATTTGTCCTATATCTTTTTGTGATCCATC
TATGTCGATGGAGTATATTCCTCACAATACATGGCTTATTTGTGCTTTTAGGTTGC
ATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGCTGCAGCTCCTACA
GTCAAGGTTTGTTCCACATGTCTATATTTTTTTAATCTTTGTAGTAATGTATCACC
AAGTATCTGTTTTCCCCCTCCTTTAGATTTTTAGAATATTGACCAAGATACTACGA
GTGCTCTCCTGATAGCAGACCATTCTCACCTTTTGTGTTGCATCTTTTCTTGACAG
CCTGTTACATTGGAACTTG

T. tauschii

AGGACTTTTTCCACCAAGTTCCAATGTAACAGGCTGTCAAGAAAAGATGCAACA
CAAAAGGTGAGAATGGTCTGCTATCAGGAGAGCGCTCGTAGTATCTTGGTCAAT
ATTCTAAAAATCTAAAGGAGGGGAAAAACAGATACTTGGTGATACATTACTACA
AAGATTTAAAAAATAGAGACATGTGGAACAAACCTTGACTGTAGGAGCTGCAGC
AACCATAATCTTTTGACCGGTTGCATAGCTCCCGGTAAATGCAACCTAAAAGCAC
AAATAAGCCATGTATTGTGAGGAATATACTCCATCGACATAGATGGATCACAAA
AAGATATAGGACAAATATATGTACCTTGTCGACATCAGGGTGTGACGACAAAGG
AGCGCCAGCTTCATGACCTAATCCAGTCACAATGTTTAAGACACCT

A Genome

GGATAAGAAGACGAGATGTCGCACTGCAAATCTGACCGTTGGTCCAAAAGCACC
CAAATAGAGTCCACTCAACAGCTGTTTAAATAGAAGGATCAACAATCAAAACTA
ACTAAAGGTCACAAAAGGTAATTCACAGAATCATATTATTCAATCAAAATGTATG
TATCAACTAACTTGAGTGAATCACTGTCTAGAAACACTTAGATCGTAAAGGTTTC
CAGAGCTAAAGCATACTAATCATGTTTTAATAATCAGTTCAAAATAAGTGCCTTG
CTTGTCACCAAAAATTTGTGATGAAATAATTTCCTGTATGTTTGGGCCTCAAGCA
AGAAAACCCCTAGACTAGCCCAGCAAGGTTGTGCGTTTTGGAACAACCAGAATT
TACTATTTGGCAGCGACGTAGCCTACGATGGTTTCAAAAGTAGTTGCGTGAGCCG
AATCCGCTTGCACTTTTGGGCAAGGATCGCACACAAGAACAGCCCTTTGGTAACT
CATATGGATCTCTGAGGGGGCTTCCCTACAAAACGTAACTATCCTCCAAAAGCTA
GGATTAGTCTTTTTTTAGGGGAAGCTAGGATTAGTCAGATTAAACCAAATAGC

TAGCCAAGGCCAGCACAACAAAGCCGGGATAAGTTACTTGGGCTACAGCCCAGC
TTATCTAGGAACCAAATGTATCCTGCGTTTCATTATTCAGCTTGTATGTACCCTCC
ATAAAAAATAACGCCAGTTTGAACAAGCAGAACGTCTGGTGAAATTATAGACAT
TATTCTTTTTAATCATTTTAGAGGTTACTAATCCTAAGTTCTAACAATTATTTATC
ATGACAAGACCCTGACACCAAATGAAGGCTTTGGGGTGTGTAACACCATGCATT
TCGAATTGGAACAAACCTTGACTGGATAAATCTCAGCATGCAGTATGTACCTTTG
TCAATGTCGACATCATCAAATACTACAATAGGACTTTTTCCACCAAGTTCCAATG
TAACAGGCTGTCAAGAAAAGATGCAACACAAAAGGTGAGTTGGTCTGCTATCAG
GAGAGCGGTGTAGTATCTTGGTCAATATTTTAAAAACCTAAAGGAGGGGAAAAA
AGGATACTTGGTGATAAATTACTACAAAGATTAAAAGAATAGGGACATGTGGAA
CAAACCTTGACTGTAGGAGCTGCAGCAACCATAATCTTTTGACCGGTTGCATA
GCTCCCGGTAAATGCAACCTACAAGCTCAAATAAGCCATGTATTGTGAGGAATA
TACTCCATCGACATAGATGGATCACAAAAAGATATAGGACAAATATATGTACCT
TGTCGACGTCAGGGTGTGAGGACAAAGGAGCGCCAGCTTCATTACCTAATCCAG
TCACAATGTTTAAGACACCTGATG

B Genome

GGATAAGAAGACGAGATGTCGCACTGCAAATCTGACCGTTGGTCCAAAAGCACC
CAAATAGAGTCCACTCAACAGCTGTTTAAATAGAAGGATCAATAATCAAATTTA
ACTAAAGGTCGCAAAAGATAATTAACAGAATCATATTATTCAATCAAAATTTATG
TATCAACTAACTTGAGTGAATCACTGTCTAGAAACATTTAGATCGTAAAGGTTTC
CAGAGCTAAAGCAGACTACTCATGTTTTTATAATCAGTTCAAAATAAGTGCCTTG
CTTGTCACCAAAAAGCAGTGATGAAATAATTTCCGGTAGGTTTGGGCCTCAAGCA
AGAAAACCCCTAGACTAGCCCAGCAAGGTTGTGCATTTTGGAACAACCAGAATT
TACTCTTCGGCGGCAATGTAGCCTAGGATAGTTTCATAATTAGCTGCGTGAGTCG
AATCCGCTTGCACTTTTGGCTAAGAATCACACACCAGAACAGCTTTTTTGGTAAT
TCATATGGATCTCTGAGGGGGCTTCCCTACAAAACTTAACTATCCTCCAAAAGCT
AGGATTAGTCAGATTAAACCAAACAGCTAGCCAAGGCTAGCACAACAAAGCC
GGGATAAGTTACTTGGGCTCTAGCCCAGCTTAGCTAGGAACCAAATGTATCCTTC
GTTTCATTATTCAGCTTGTATGTACCCTCAATAAAAAATTACGCCAGTTCGAACA
AGCAGAAAGTCTGGTGAAATCATAGACATCATTCTTTTTAATCATTTTACAGGTT
GCTAATCCGAAGTTCTAACAATCATTTATCATGAAAAGCCCCAGACACCAAATA
AAGGCTTTGGGGTGTGTAACACCATGGTCAGTATTTAGTGATTCCACACAGGGCA
TTTTGAATTGGACAAGTCTCAGCATGCAGCAGCATGTACCTTTGTCAATGTCGAC
ATCATCAAATACTACAATAGGACTTTTGCCACCAAGTTCCAATGTAACAGGCTGT
CAAGAAAAGATGCAACACAAAAGGTGAGAATGGTCTGCTATCAGGAGAGCGCTC
GTAGTATCTTGGTCAATATTCTAAAAATCTAAAGGAGGGGGAAAACAGATACTT
GGTGATAAATTACTACAAAGATTTAAAAAATAGAGACATGTGGACAAACCTTGA
CTGTAGGAGCTGCAGCAACCATAATCTTTTGACCGGTTGCATAGCTCCCGGTA
TGCAACCTAAAAGCACAAATAAGCCATGTATTGTGAGGAATATACTCCATCGAC
AAATAGATGGATCACAAAAAGATATAGGACAAATATATGTACCTTGTCGACATC
AGGGTGTGACGACAAAGGAGCGCCAGCTTCATGACCTAATCCAGTCACAATGTT
TAAGACACCTGATG

D Genome

CATCAGGTGTCTTAAACATTGTGACTGGATTAGGTCATGAAGCTGGCGCTCCTTT
GTCGTCACACCCTGATGTCGACAAGGTACATATATTTGTCCAATATCTTTTTGTGA
TCCATCTATGTCGATGGAGTATATTCCTCACAATACATGGCTTATTTGTGCTTTTA
GGTTGCATTTACCGGGAGCTATGCAACCGGTCAAAAGATTATGGTTGCTGCAGCT
CCTACAGTCAAGGTTTGTTCCACATGTCTCTATTTTTTCAATCTTTGTAGTAATGT
ATCACCAAGTATCTGTTTTTCCCCTCCTTTAGATTTTTAGAATATTGACCAAGATA
CTACGAGCGCTCTCCTGATAGCAGACCATTCTCACCTTTTGTGTTGCATCTTTTCT
TGACAGCCTGTTACATTGGAACTTGGTGGCAAAAGTCCTATTGTAGTATTTGATG
ATGTCGACATTGACAAAGGTACATGCTGCTGCATGCTGAGACTTGTCCAATTC
GAAATGCCCTATGCGAAATCACTTCATACTGCATGGTGTTACACACCTCAAAGCC
TTCATTTGGTGTCTGGGGTTTGTCATGATAAATGATTGTTATTACTTAGGATTAGT
AACCTCTAAAATGATTAAAAAGAATGATGTCTATAATTTCACCAGACTTTCTGCT
TGTTCAAACTGGTGTAATTTTTTGTGGAGGGTACATACAAGCTGAATAACGAAAC
GAAGAATAGCCTTGGCTAGCTATTTGGTTTAATCTGACTAATCGTAGCTTTTGGA
GGATAGCTACTTTTTGTAGGGAAGCCCCCTCAAAGATCCATATGAGTTACCAAAG
AGTTGTTCTTGTGTGTGATTCTTGCCCAAAAGTGCAAGCGGATTCGGCTCACGCA
ACTACTTTAGAAACCATCCCGCCAAATAGTAAATTCTGGCTGTTCCAAAACGCAC
ACCCTTGCTTGGCTAGTCTAGGGGTTTTCTTGCTTGAGGCCCAAACATACCGGAA
ATTGTTTCATCACTAATTTTTGGTGACAAGCAAGGCAGTTATTTTGAACTGATT
ATTAAAACATGAGTAGTATGCTTTAGCTCTGGAAACCTTTACGATCTAAATGTTT
CTATACAGTGAATCACTCAAGTTAGTTGATACATACATTTTGATTGAATAATATG
ATTCTGTTAAGTACCTTTTGTGACCTTTAGTTTGATTTGATTGCTGATCCTTCTATT
TAAACAGCTGTTGAGTGGACTCTATTTGGGTGCTTTTGGACCAACGGTCAGATTT
GCAGTGCGACATCTCGTCTTCTTATCC

Appendix iii

cDNA sequences of BADH1 homologs of hexaploid wheat varieties in two time points (14DPA and 30DPA) and two tissue (seeds and leaves)

Banks14 DPA leaves

CAGCAAGAAGTGAAGGTGCTACAATTTTGCATGGTGGTGACCGACCAAA
GCATCTCGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGACGTTA
GCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGT
GTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAATGATAC
CCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTG
AGCGCATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCG
CAACCGACCCTGGTTCAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTT
TGGCCGGGAGCTAGGAGAATGGGGCCTCGAGAACTACTTGAGCGTGAAAC
AAGTCACCAGGTACA

BobWhite 14 DPA leaves

CAGCAAGAAGTGAAGGTGCTACAATTTTGCATGGTGGTGACCGACCAAAG
CATCTCGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGACGTTAG
CACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGTG
TCAAAGTATTCAAGACAGAGAGTGAAGCTGTAGAGCTTGCAAATGATACC
CACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGA
GCGCATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGC
AACCGACCCTGGTTCAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTT
GGCCGGGAGCTAGGAGAATGGGGCCTCGAGAACTACTTGAGCGTGAAACA
AGTCACCAGG

Cadoux 14 DPA leaves

GGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGG
GTTCTTTATTGAACCTACAATTATAACAGACGTTAGCACATCAATGCAAA
TTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGTGTCAAAGTATTCAAG
ACAGAGAGTGAAGCTGTAGAGCTTGCAAATGATACCCACTATGGCTTGGC
TGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGCATTGCAAAGG
TTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTT
CAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGG
AGAATGGGGCCTCGAGAACTACTTGAGCGTGAAACA

NW51 14 DPA leaves

CAGCAAGAAGTGAAGGTGCTACAATTTTGCATGGTGGTGACCGACCAAAG
CATCTCGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGACGTTAG
CACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGTG
TCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAATGATACC
CACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGA
GCGCATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGC
AACCGACCCTGGTTCAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTT
GGCCGGGAGCTAGGAGAATGGGGCCTCGAGAACTACTTGAGCGTGAAACA

AGTCACCAGG

Banks 30 DPA leaves

TCAGCAAGAAGTGAAGGTGCTACAATTTTGCATGGTGGTGACCGACCAAA
GCATCTCGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGACGTTA
GCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGT
GTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAATGATAC
CCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTG
AGCGCATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCG
CAACCGACCCTGGTTCAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTT
TGGCCGGGAGCTAGGAGAATGGGGCCTCGAGAACTACTTGAGCGTGAAAC
AAGTCACCAGG

Bobwhite 30 DPA leaves

TCAGCAAGAAGTGAAGGTGCTACAATTTTGCATGGTGGTGACCGACCAAA
GCATCTCGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGACGTTA
GCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGT
GTCAAAGTATTCAAGACAGAGAGTGAAGCTGTAGAGCTTGCAAATGATAC
CCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTG
AGCGCATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCG
CAACCGACCCTGGTTCAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTT
TGGCCGGGAGCTAGGAGAATGGGGCCTCGAGAACTACTTGAGCGTGAAAC
A

Cadoux 30 DPA leaves

GGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGG
GTTCTTTATTGAACCTACAATTATAACAGACGTTAGCACATCAATGCAAA
TTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGTGTCAAAGTATTCAAG
ACAGAGAGTGAAGCTGTAGAGCTTGCAAATGATACCCACTATGGCTTGGC
TGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGCATTGCAAAGG
TTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACTCTGGTTC
AAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTAGG
AGAATGGGGCCTCGAGAACTACTTGAGCGTGAAACA

NW 51 30 DPA Leaves

AAGAAGTGAAGGTGCTCCAATTTTGCATGGTGGTGACCGACCAAAGCATC
TCGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGACGTTAGCACA
TCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGTGTCAA
AGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAATGATACCCACT
ATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGCA
TTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGA
CTCTGGTTCAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGG
GAGCTAGGAGAATGGGGCCTCGAGAACTACTTGAGCGTGAAACAAGTCAC
CAGG

Banks 14 DPA seeds

TCAGCAAGAAGTGAAGGTGCTACAATTTTGCATGGGGGTGACCGACCAAA
GCATCTCGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGACGTTA
GCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGT
GTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAATGATAC
CCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTG
AGCGCATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCG
CAACCGACTCTGGTTCAAGCTCCGTGGGGAGGGAACAAGCGTAGCGGTTT
TGGCCGGGAGCTAGGAGAATGGGGCCTCGAGAACTACTTGAGCGTGAAAC
A

BobWhite 14 DPA seeds

CAGCAAGAAGTGAAGGTGCTACAATTTTGCATGGGGGTGACCGACCAAAG
CATCTCGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGACGTTAG
CACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGTG
TCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAATGATACC
CACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGA
GCGCATTGCAAAGGTTATTCACTCGGGCATTGTTTGGATAAACTGCTCGC
AACCGACTCTGGTTCAAGCTCCGTGGGGAGGGAACAAGCGTAGCGGTTTT
GGCCGGGAGCTAGGAGAATGGGGCCTCGAGAACTACTTGAGCGTGAAACA
AGTCAC

Cadoux14dpaseeds

GGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCATCTCGGAAAAGG
GTTCTTTATTGAACCTACAATTATAACAGACGTTAGCACATCAATGCAAA
TTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGTGTCAAAGTATTCAAG
ACAGAGAGTGAAGCTGTAGAGCTTGCAAATGATACCCACTATGGCTTGGC
TGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCGCATTGCAAAGG
TTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAACCGACCCTGGTT
CAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGCCGGGAGCTA
GGAGAATGGGGCCTCGAGAACTACTTGAGCGTGAAACAAGTACCCAGG
TACA

NW51 14 DPA seeds

CCTGGTGACTTGTTTCACGCTCAGGTAGTTCTCGAGGCCCCATTCTCCTA
GCTCCCGGCCAAAACCGCTACGCTTGTTCCCTCCCCACGGAGCTTGAACC
AGAGTCGGTTGCGAGCAGTTTATCCAAACAATGCCCGAGTGAATAACCTT
TGCAATGCGCTCACACCTCTGTAGATCATCAGAGATCACACCACCAGCCA
AGCCATAGTGGGTATCATTTGCAAGCTCTACAGCTTCGCTCTCTGTCTTG
AATACTTTGACACAGATGACTGGTCCAAAGACTTCCTCTCTCCAAATTTG
CATTGATGTGCTAACGTCTGTTATAATAGTAGGTTCAATAAAGAACCCTT
TTCCGAGATGCTTTGGTCGGTCACCCCCATGCAAAATTGTAGCACCTTCA
CTTCTTG

Banks 30 DPA seeds

CAAGAAGTGAAGGTGCTACAATTTTGCATGGTGGTGACCGACCAAAGCAT
CTCGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGACGTTAGCAC
ATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGTGTCA
AAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAATGATACCCAC
TATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTGAGCG
CATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCGCAAC
CGACTCTGGTTCAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTTTGGC
CGGGAGCTAGGAGAATGGGGCCTCGAGAACTACTTGAGCGTGAAACA

BobWhite 30 DPA seeds

TCAGCAAGAAGTGAAGGTGCTACAATTTTGCATGGTGGTGACCGACCAAA
GCATCTCGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGACGTTA
GCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGT
GTCAAAGTATTCAAGACAGAGAGTGAAGCTGTAGAGCTTGCAAATGATAC
CCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTG
AGCGCATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCG
CAACCGACCCTGGTTCAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTT
TGGCCGGGAGCTAGGAGAATGGGGCCTCGAGAACTACTTGAGCGTGAAAC
AA

Cadoux 30 DPA seeds

TCAGCAAGAAGTGAAGGTGCTACAATTTTGCATGGTGGTGACCGACCAAA
GCATCTCGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGACGTTA
GCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGT
GTCAAAGTATTCAAGACAGAGAGTGAAGCTGTAGAGCTTGCAAATGATAC
CCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTG
AGCGCATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCG
CAACCGACTCTGGTTCAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTT
TGGCCGGGAGCTAGGAGAATGGGGCCTCGAGAACTACTTGAGCGTGAAAC
A

NW 51 30 DPA seeds

TCAGCAAGAAGTGAAGGTGCTACAATTTTGCATGGTGGTGACCGACCAAA
GCATCTCGGAAAAGGGTTCTTTATTGAACCTACAATTATAACAGACGTTA
GCACATCAATGCAAATTTGGAGAGAGGAAGTCTTTGGACCAGTCATCTGT
GTCAAAGTATTCAAGACAGAGAGCGAAGCTGTAGAGCTTGCAAATGATAC
CCACTATGGCTTGGCTGGTGGTGTGATCTCTGATGATCTACAGAGGTGTG
AGCGCATTGCAAAGGTTATTCACTCAGGCATTGTTTGGATAAACTGCTCG
CAACCGACCCTGGTTCAAGCTCCGTGGGGAGGGAACAAGCGGAGCGGTTT
TGGCCGGGAGCTAGGAGAATGGGGCCTCGAGAACTACTTGAGCGTGAAAC
AAGTCACCAGGTACA

Appendix iv

cDNA sequences of BADH2 homologs of hexaploid wheat varieties in two time points (14DPA and 30DPA) and two tissue (seeds and leaves)

Banks14DPA leaves

TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGT
CCTCACACCCTGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACC
GGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGCCTGTTACATT
GGAACTTGGTGGAAAAAGT

BobWhite14 DPA leaves

TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGT
CCTCACACCCTGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACC
GGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGCCTGTTACATT
GGAACTTGGTGGAAAAAGT

Cadoux 14 DPA leaves

TTTTTCCACCAAGTTCCAATGTAACAGGCTTGACTGTAGGAGCTGCAGCA
ACCATAATCTTTTGACCGGTTGCATAGCTCCCGGTAAATGCAACCTTGTC
GACGTCAGGGTGTGAGGACAAAGGAGCGCCAGCTTCATTACCTAATCCA
GTCACAATGTTTAAGACACCTA

NW51 14 DPA leaves

TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGT
CCTCACACCCTGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACC
GGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGCCTGTTACATT
GGAACTTGGTGGAAAAAGT

Banks 30 DPA leaves

ACTTTTTCCACCAAGTTCCAATGTAACAGGCTTGACTGTAGGAGCTGCAG
CAACCATAATCTTTTGACCGGTTGCATAGCTCCCGGTAAATGCAACCTTG
TCGACGTCAGGGTGTGAGGACAAAGGAGCGCCAGCTTCATTACCTAATCC
AGTCACAATGTTTAAGACA

Bobwhite 30 DPA leaves

TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGT
CCTCACACCCTGAYGTSGACAAGGTTGCATTTACCGGGAGCTATGCAACC
GGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGCCTGTTACATT
GGAACTTGGTGGAAAAAGT

Cadoux 30 DPA leaves

GGTGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTT
GTCCTCACACCCTGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAA
CCGGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGCCTGTTACA
TTGGAACTTGGTGGAAAAA

NW51 30 DPA leaves

TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGT
CCTCACACCCTGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACC
GGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGCCTGTTACATT
GGAACTTGGTGGAAAAAGT

Seed
Banks 14 DPA seeds

TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGT
CCTCACACCCTGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACC
GGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGCCTGTTACATT
GGAACTTGGTGGAAAAAGT

BobWhite14 DPA seeds

GTGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTG
TCCTCACACCCTGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAAC
CGGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGCCTGTTACAT
TGGAACTTGGTGGAAAAAGT

Cadoux 14 DPA seeds

GTGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTG
TCCTCACACCCTGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAAC
CGGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGCCTGTTACAT
TGGAACTTGGTGGAAAAAGT

NW 51 14 DPA seeds

TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGT
CCTCACACCCTGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACC
GGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGCCTGTTACATT
GGAACTTGGTGGAAAAAGT

Banks 30 DPA seeds

TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGT
CCTCACACCCTGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACC
GGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGCCTGTTACATT
GGAACTTGGTGGAAAAAGT

BobWhite 30 DPA seeds

TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGT
CCTCACACCCTGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACC
GGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGCCTGTTACATT
GGAACTTGGTGGAAAAAGT

Cadoux 30 DPA seeds

TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGT
CCTCACACCCTGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACC
GGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGCCTGTTACATT
GGAACTTGGTGGAAAAAGTCCTAA

NW 51 30 DPA seeds

TGTCTTAAACATTGTGACTGGATTAGGTAATGAAGCTGGCGCTCCTTTGT
CCTCACACCCTGACGTCGACAAGGTTGCATTTACCGGGAGCTATGCAACC
GGTCAAAAGATTATGGTTGCTGCAGCTCCTACAGTCAAGCCTGTTACATT
GGAACTTGGTGGAAAAAGT