Molecular studies of abiotic stress and abscisic acid mediated stress response pathway in barley

Atul Kamboj
Master of Science (Biotechnology)

This thesis is presented for the degree of Doctor of Philosophy
September 2014

Abstract

Salt and drought stresses are the two most important environmental stresses which limit plant growth and development. Salinity is currently affecting about 6 % of world's total land area and increased salinity may result in 50% loss of arable land by 2050. In Australia alone, about 5 % is currently affected by salinity and another 17 million ha (33 % of agricultural land) may be salinised or at risk by 2050. Drought stress also impedes production of crops. Australian crop yields may drop up to 30% by 2050 as a result of reduced rainfall. Thus, the major aim of the current study was to analyse the global expression of barley genes to acute salinity and drought stress.

In this work, the utility of deep cDNA sequencing (mRNA-seq) for barley transcriptome analysis is demonstrated as an alternative to microarrays despite the lack of a whole genome reference sequence [barley whole genome sequence has only recently been described; International Barley Genome Sequencing Consortium (IBSC)]. mRNA-seq was performed on acutely salt, drought and exogenous ABA stressed and unstressed leaf material of barley (*Hordeum vulgare* L. cv. Hindmarsh). The data yielded more than 100 million sequence tags which aligned to more than 20,000 transcripts in the IBSC database. Approximately 90% of the genes differentially expressed under salinity or drought were also differentially expressed in exogenous ABA stress. Genes such as chlorophyll a/b binding protein and MYB transcription factors were among the most differentially regulated during salt and drought stresses. Analysis of gene ontology showed "binding" as major molecular function in all three stress conditions. These differentially regulated genes can be the candidate tolerance genes that need to be further examined using reverse genetic approaches.

The phytohormone abscisic acid (ABA) is crucial for plant's adaptive response to salinity and drought stresses. ABA accumulates in plant cells to protect vegetative tissues and regulates developmental events. The major components of ABA mediated abiotic stress tolerance pathway include Pyrabactin resistance family (PYR/PYL/RCAR; soluble ABA receptors), protein phosphatase 2C subfamily A (PP2CA; negative regulators of ABA signalling cascade in the absence of ABA), SNF1 related protein kinase 2 (SnRK2; positive regulator of the cascade) and ABA transport related ABC transporters (ABCG25 and ABCG40). ABA signalling pathway is well

established in Arabidopsis, but is poorly defined in the major cereal crops, wheat and barley. Thus, the present study aimed to identify and characterise barley ABA mediated abiotic stress signalling related genes. Reciprocal BLAST search (RBH), an automated reciprocal BLAST pipeline tool was developed enabling the rapid identification of specific gene families of interest in related species, streamlining the collection of homologs prior to downstream molecular evolutionary analysis. The expressed barley ABA signalling related genes were initially identified using mRNA-seq and NCBI Unigene database, followed by further searches in the IBSC database. The identified sequences were analysed for certain functional motifs and key residues in the putative proteins, which resulted in 13 PP2CAs, 9 PP2CDs (PP2C subfamily D), 10 PYR/PYL/RCARs, 5 SnRK2 and 4 ABC transporters. The amino acid involved in the binding of PYR/PYL/RCAR and PP2C were strictly conserved among all sequences in rice and barley. The consensus for the characteristic motifs of these genes was generated for rice and barley. Of the above identified genes, only five barley PP2CAs, three PYR/PYL/RCARs, two SnRK2s (subfamily II and III) and one each of ABCG25 and ABCG40 showed significant differential expression, while the rest were only marginally affected. Putative PP2CAs, SnRK2s subfamily II, ABCG25 and ABCG40 were found to be up-regulated under all three abiotic stresses studied, whereas PP2CDs, PYR/PYL/RCARs and SnRK2s subfamily III were down-regulated.

Epigenetic factors, such as histone modification and DNA methylation, play a significant role in regulating gene transcription. The present study aimed to investigate the distribution of one such modification, the tri-methylation of lysine 4 of histone H3 (H3K4me3) that is involved in gene activation, in barley during abiotic stress. Chromatin immuno-precipitation and deep sequencing (ChIP-seq) identified 19,015 and 19,005 genomic regions exhibiting differential H3K4me3 under salt and drought stresses, respectively. The ChIP-seq dataset was then searched for putative genes with *a priori* annotations relating to ABA mediated signalling, which lead to identification of three PP2CA, four PP2CD, two PYR/PYL/RCAR, three SnRK2 and two ABC transporters type G. H3K4me3 was found at the 5'UTR, exon and intron regions of PP2CAs, whereas H3K4me3 was found at exon, intron and 3' regions of PP2CDs and SnRK2s. For PYR/PYL/RCARs, H3K4me3 was found at 5'UTR, exon, intron and 3'UTR regions. On the other hand, H3K4me3 was restricted only to exon and intron

regions of ABC transporters. Three PP2CAs, three PP2CDs, two PYR/PYL/RCARs, three SnRK2s and one ABCG40 were found to be significantly de-methylated at H3K4 during abiotic stress, whereas one each of PP2CDs and ABCG25 were significantly tri-methylated at H3K4.

This work also assessed the salt tolerance levels of sixteen barley varieties by key physiological assays including relative water content, levels of ABA and Na^+/K^+ ratio. The results showed a significant variation, ranging from highly salt tolerant to sensitive varieties. Further, the differential expression of key genes in the ABA-mediated abiotic stress tolerance pathway was analysed by quantitative real-time PCR of leaf RNA. Importantly, an arbitrary ranking method was developed and applied to both physiological and gene expression analysis. The tolerance ranking by gene expression closely correlated that by physiological indices. Thus, expression analysis of the ABA pathway can be used for rapid identification of potentially salt-tolerant barley varieties before undertaking physiological studies.

The work provides insights into genome-wide effects of salinity and drought stresses and is a new resource for the study of gene regulation in barley. The bioinformatics workflow may be applicable to other non-model plants to establish their transcriptomes and identify unique sequences. The in-depth expression and chromatin profiling of genes involved in ABA mediated abiotic stress signalling make it possible to identify candidates for future functional testing and development of tolerant plants through genetic modification or breeding.

Acknowledgements

Firstly, I would like to thank my principle supervisor, Prof. Mrinal Bhave for accepting me into her research group. Her unwavering zeal for research has been a great source of inspiration. I am also very indebted for all her support and guidance, the long hours spent on problem-solving discussions and always read every line of everything that I sent to her and gave her time selflessly. I benefitted immensely from your invaluable wisdom and experience. Thank you for being my teacher. To my co-supervisor, Dr Mark Ziemann, I am grateful for all your feed-backs, advice, kind words and constant encouragement. Thank you for all the late sittings, guidance in performing mRNA-seq, ChIP-seq and bioinformatics based analysis. To my co-supervisor, A/Prof Assam El-Osta, I am grateful for allowing me to use facilities at his lab and support the work.

I am also grateful to Swinburne University for awarding me the SUPRA scholarship, which allowed me to study without financial constraints. Thank You to Dr. Runyararo Memory Hove, a fellow PhD student then for sharing the mRNA-seq I work with me. Thank you to Soula Mougos, Ngan Nguyen, Chris Key, Dr Huimei Wu and Angela McKellar for their advice, and skilled technical assistance. I would also like to thank former and current PhD students, Dr Peter Gollan, Dr Abirami Ramalingam, Dr Rebecca Alfred, Dr Saifone Chuaboonmee, Dr Shanthi Joseph, Ms. Azadeh Niknejad, Ms. Gurvinder Kalra, Ms. Thi Hoang Yen Dang, Ms. Nadine Shagaghi, Mr. Avinash Karpe, Ms. Snehal Jadhav and Mr. Rohan Shah for the support, advice and encouragement.

To my, dearest mama, papa and Kanu (best parents and sister) thank you for all your encouragement, patience and ever-lasting support to constantly excel. Thank you for being outstandingly understanding throughout this journey. It really helped to keep me going. It was just not possible without you guys.

Publications arising from this work

Refereed journal articles

Ziemann M, Kamboj A, Hove R, Loveridge S, Bhave M (2013) Analysis of the barley leaf transcriptome under salinity stress using mRNA-seq. Acta Physiologiae Plantarum 35: 1915-1924.

Ziemann M, Kamboj A, Bhave M (2013) OrthoRBH: A streamlined pipeline for mining large gene family sequences in related species. Bioinformation 9: 267-269.

Kamboj A, Ziemann M, Bhave M (2014) Identification of salt tolerant barley varieties by a consolidated physiological and molecular approach. Under review

Conference presentations

Kamboj A, Ziemann M, Bhave M (2011) Identification of key components of the Abscisic acid signalling pathway in barley by next generation mRNA sequencing. 61th Australian Cereal Chemistry conference, Coolangatta, New South Wales (4-8 September 2011) (Talk).

Kamboj A, Ziemann M, Bhave M (2012) Use of next generation sequencing to profile the barley leaf transcriptome under salt stress. XXII International Congress on Sexual Plant Reproduction, Melbourne (13-17 February 2012) (Poster).

Declaration

I, Atul Kamboj, declare that the PhD thesis entitled 'Molecular studies of abiotic stress and abscisic acid mediated stress response pathway in barley' is no more than 100,000 words in length, exclusive of tables, figures, appendices, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma, and has not been previously published by another person. Except where otherwise indicated, this thesis is my own work.

Atul Kamboj
2014

Abbreviations

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

ABA	abscisic acid
At	*Arabidopsis thaliana*
BLAST	Basic Local Alignment Search Tool
BDT	Big Dye Terminator
bp	base pair (s)
cDNA	complimentary DNA
CDPK	Ca^{2+}-dependent protein kinase
CDS	coding sequence
C-terminal	carboxyl terminal
cv	cultivar
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleotide
EDTA	ethylenediaminetetraacetic acid
EST	expressed sequence tag
gDNA	genomic DNA
Hv	*Hordeum vulgare* (barley)
IBSC	International Barley Genome Sequencing Consortium
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kb	kilobase pairs
LB	Luria-Bertani media
mRNA	messenger RNA
MW	molecular weight
NCBI	National Centre for Biotechnology Information
NGS	next-generation sequencing
N-terminal	amino terminal
ORF	open reading frame
Os	*Oryza sativa* (rice)

PEG	polyethylene glycol
PCR	polymerase chain reaction
PM	plasma membrane
PP2C	Protein phosphatase 2C
PYR/PYL/RCAR	Pyrabactin resistance/PYR like/ regulator component of ABA receptor
RNA	ribonucleic acid
RNase	ribonuclease
Rpm	revolutions per minute
RT	room temperature
SNP	single nucleotide polymorphism
SnRK2	Sucrose non fermenting 1 related kinase 2
Ta	*Triticum aestivum* (common wheat)
TA	transcript assembly
TAE	tris acetate ethlenediaminetetraacetic acid buffer
TC	tentative consensus sequence
TIGR	The Institute for Genomic Research
TMH	transmembrane helical domain
UV	ultra-violet
X-gal	5-bromo-4-chloro-3-indolyl-b-D galactopyranoside

Table of contents

List of figures

List of tables

Chapter 1

General introduction and literature review

1.0 Introduction

This chapter presents a review of literature pertaining to the current study. A comprehensive discussion of the barley (*Hordeum vulgare*) industries will be provided, focussing on their importance as food and cash crops, in Australia and the world. Next, plant abscisic acid (ABA) and ABA mediated abiotic stress signalling genes will be reviewed in relation to their structure, subfamilies, function and regulation. The review provides updated information that allows the reader to appreciate the importance of discovering new ABA mediated abiotic stress signalling genes that can be incorporated into breeding programs for barley variety with more suited end-use. The chapter concludes by presenting the aims and objectives of the present study.

1.1 Barley

Barley and wheat are amongst the most important food crops of the world which belong to the grass family Poaceae, subfamily Pooideae and tribe Triticeae (Gaut 2002). The common ancestor of Triticeae has diverged from oats almost 25 million years ago (MYA), while barley and wheat split from the same ancestor 11-13 MYA (Figure 1.1; Salse and Feuillet 2007). Historically, humans harvested the wild forms of Triticeae species before the domestication and cultivation of these crops (Zohary and Hopf 2000).

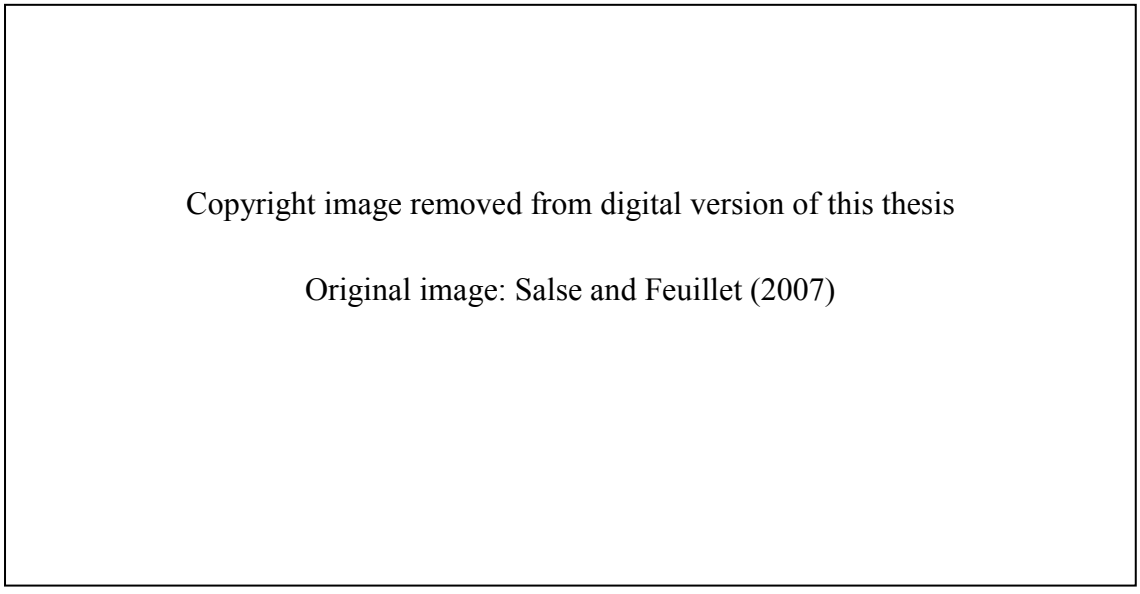


Figure 1.1 Phylogenetic relationships of monocotyledons
Evolutionary tree showing the location of barley and wheat. The Triticeae tribe is highlighted with a black rounded rectangles (Source: Salse and Feuillet 2007).

1.1.1 Origin and evolution of cultivated barley

Cultivated barley (*Hordeum vulgare*) is diploid (HH; 2n = 2x = 14). Archaeological and molecular evidence directs to domestication of barley about 10,000 years ago in the fertile crescent (Zohary and Hopf 2000), possibly through environmental or human selection (Bothmer *et al*. 2003) of wild *H. spontaneum*, developing tough, rachis, six-rowed spikes, less brittle and naked caryopsis (Salamini *et al*. 2002). Barley was initially used mostly for human consumption but over the years it has evolved majorly into a feed grain and for brewing beer (Newman and Newman 2008). Despite having undergone different evolutionary processes, genetic mapping (Devos and Gale 1997) and genomics studies (Ramakrishna *et al*. 2002) have shown that barley and wheat genomes have managed to maintain a significant conservation, to the extent that barley chromosomes can be substituted for wheat (Islam *et al*. 1981). This relatedness among both the crops allow the genetic, genomic or functional information obtained for barley to be comfortably extrapolated to its more complex cousin, wheat.

1.1.2 Importance of barley as crops

Barley is used for human consumption mostly in regions where other cereals are not well-adapted due to environmental conditions such as altitude, low rainfall or salinity. Wheat is generally preferred due to its superior sensory properties such as texture, aesthetic appearance and colour. It has great nutritional value, e.g., for cholesterol-lowering, colon health and blood sugar control (Newman and Newman 2008). Starch is the largest component of barley (57.0 - 65.2%), followed by fiber (15.6 - 20.2%), protein (12.5 - 15.4%), sugars (2.8 - 4.2%) and lipids (1.9 - 3.9%) (Aman and Newman 1986). Whole grain barley is a good source of the B-complex vitamins, especially vitamin B3, which is in four to five times higher amounts than in maize, oats and rye (Baik *et al*. 2011), and vitamin E (Kerckhoffs *et al*. 2002). It also has minerals such as potassium, calcium, phosphorus, iron, magnesium and zinc (Newman and Newman 2008). Globally, 70 % of the world barley production is utilised for animal feed and 30 % for malting purpose (FAO 2014).

1.1.3 Current barley genome sequencing projects

Bread wheat has a large genome size [17 Gigabases (Gb)], which is almost 6 times larger than that of maize and 40 times larger than rice (see below), whereas barley has a 5 Gb genome (Eversole *et al*. 2009). This complexity has created challenges in full sequencing of their genomes; however, recent advancements in genomic technology have made this aim almost fully attainable at present. *Arabidopsis thaliana* (150 Mb) was the first plant genome sequenced (The Arabidopsis Genome Initiative, 2000). Since then, other genomes such as rice (*Oryza sativa*; 389 Mb; International Rice Genome Sequencing Project 2005), grapevine (*Vitis vinifera*; 475 Mb; Travis 2008) and maize (*Zea mays*; 2.3 Gb; Schnable *et al*. 2009) have been sequenced also. Among the cereals, maize and rice are continually being improved partly due to information derived from their fully sequenced genomes, but barley and wheat lag behind. As a step towards filling this gap, the International Barley Genome Sequencing Consortium (IBSC) was established in 2006, with the target of a physical map and a complete high quality genome sequence (Schulte *et al*. 2009). The Roche 454 sequencing technology was initially tested on four BAC clones (Wicker *et al*. 2006) and subsequently employed for sequencing flow-sorted chromosome 1H (Mayer *et al*. 2009), then all others (Mayer *et al*. 2011). The data generated is publically available at the International Barley Genome Sequencing Consortium (http://barleygenome.org; last accessed June 2014) and the whole genome projected was released in October 2012 (http://prlog.org/10063090; last accessed June 2014). The release of the barley genome before wheat could be credited to its diploid nature, but it 'provides a faithful proxy' for wheat (Mayer *et al*. 2011) due to their phylogenetic relatedness mentioned above. These projects are thus significant steps towards gene discovery and selection and development of better varieties with increased yields and resistance to environmental stresses. The international wheat genome sequencing consortium (IWGSC) was established in 2005, with the aim of developing physical map, achieving sequence of the 21 bread wheat chromosomes and completing the reference sequence for each of these chromosomes. The first aim of sequencing 21 bread chromosomes has been achieved in July, 2014 (International wheat genome sequencing consortium 2014; http://www.wheatgenome.org/).

1.1.4 World barley production

Barley is grown on 47 million hectares (mha), likely due to being considered 'poor man's bread', but it continues to be cultivated since it is more adaptable and resilient in comparison to wheat (Zohary and Hopf 2000). It is more tolerant to drought, salinity and cold and can be cultivated at higher altitudes (e.g., >4,500 m on the Altipano of Peru and Bolivia) and latitudes (>65° north) and farther into deserts than other cereal crops (Ullrich 2011). Thus, a deeper understanding of these traits in barley could be important in further improvements of wheat through selective breeding and genetic engineering.

1.1.5 Barley production in Australia

In Australia, barley is one of the most important grain crops due to the quantity produced, the area cultivated and the revenue generated. Australian barley production has more than doubled in the period 1961 to 2011(0.98 t/ha in 1961 to 2.18 t/ha in 2011). In Australia, barley has annual production of 7 million tons per year (Barley Australia 2014; http://www.barleyaustralia.com.au/industry-information). Barley is grown in a narrow crescent stretching from Western Australia to southern Queensland. Australia is a leading exporter of barley, with 4.6 MT exported in 2011 (Barley Australia 2014) and the trade value of barley for the 2011 season was A$1.3 billion, (ABARES 2011). Thus, the economic value of barley in the Australian economy cannot be under-estimated. However, the production of this crop is constantly impeded by abiotic stresses such as drought and salinity (discussed below).

1.1.6 Abiotic stresses limiting the growth of barley

Among the abiotic stresses that limit growth of barley are salinity and drought. Salinity is a worldwide concern, currently affecting about 6% of world's total land area (Cramer *et al*. 2011) and increased salinity may lead up to 50% loss of arable land by 2050 (Wang *et al*. 2003). In Australia alone, about 2 million ha is currently affected by salinity and another 17 million ha (1/3 of agricultural land) is at risk of salinity by 2050 (Rengasamy 2010). Generally plants respond to salinity in two phases i.e. the osmotic (phase I) and ion-specific phase (phase II reviewed in Munns and Tester 2008). The osmotic phase is when water becomes less accessible to the roots, which leads to reduced leaf and root growth (Munns 1993). On the other hand, phase II starts when the

salt adds up to toxic levels in the leaves resulting, in dehydration and eventual death (Munns and Tester 2008). Salt stress has also been demonstrated to; (i) inhibit the activity of nucleic acid metabolism enzymes (Gomes-Filho *et al*. 2008), (ii) affect various processes involved in germination such as seed imbibition (reviewed in Wahid *et al*. 2010), (iii) reduce the availability of CO_2 resulting in a decline in photosynthesis and cell growth (Chaves *et al*. 2009) and (iv) increase the formation of reactive oxygen species (ROS) such as O_2, H_2O_2, O_2-, HO-, that cause oxidative damage to protein, DNA and lipids (reviewed in Gill and Tuteja 2010). Reduction in land available for cultivation due to salinity and increase in human population are major threats to sustainable agriculture. In order to maintain the food supply at current levels, global food production will need to increase by 38 % by 2025 and by 57 % by 2050 (Wild 2003). In summary, salinity stress threatens plant growth and crop yields.

Drought stress (water deficit) also impedes production of barley (Moffat 2002). Drought stress is detrimental to physiological processes such as photosynthesis, stomatal conductance, accumulation of lipids and gene expression (Rizhsky *et al*. 2004). The effects of water deficit on plants have been studied extensively and in summary it results in; (i) dehydration, (ii) stomatal closure leading to reduced CO_2 uptake (reviewed in Shinozaki and Yamaguchi-Shinozaki 2007), (iii) affects pollen development, thus compromising grain yield in cereals (Dolferus *et al*. 2011) and (iv) increased production of ROS affecting cellular structures. Australian barley yields may drop by 10-30 % by 2050 as a result of reduced rainfall (van Gool and Vernon 2006).

Noting the effects of the abiotic stresses on barley, it is of importance to develop tolerant cultivars. This relies mainly upon an understanding of their genetic and biochemical mechanisms of stress tolerance, an area of significant research focus globally. There is a significant amount of literature on numerous individual genes and interconnected regulatory pathways and response mechanisms in the plants, and the works of Wang *et al*. (2004), Langridge *et al*. (2006), Sreenivasulu *et al*. (2007), Munns and Tester (2008), Hirayama and Shinozaki (2010), Gill and Tuteja (2010), Sinha *et al*. (2011) and Atkinson and Urwin (2012), amongst others provide comprehensive reviews in this area. In the following sections, there is summary of some of the main features of plant responses to drought and salinity.

Exposure of plants to abiotic stresses lead to expression of signal transduction pathway genes, e.g., transcriptional factors (ABF/ABARE, CBF/DREB, WRKY families) (Nakashima *et al*. 2009a). The accumulation of compatible solutes, such as sugars (trehalose), sugar alcohols (mannitol), amino acids (proline) and amines (glycine betaine) has also been noted. Compatible solutes are essential for adjusting osmotic pressure and scavenging ROS (reviewed in Peleg *et al*. 2011). Other genes encode proteins that are involved in protection of the macromolecules and membranes under abiotic stress such as late embryogenesis abundant (LEA) proteins (Olvera-Carrillo *et al*. 2011), heat shock proteins (Hsps) (Wang *et al*. 2004). The increased production of ROS during abiotic stresses is initiated by the synthesis of ROS scavenging enzymes, such as catalases (CAT), superoxide dismutase (SOD) and other non-enzymatic antioxidant molecules such as ascorbic acid (Gill and Tuteja 2010). Phytohormones are also important in mediating a response to abiotic stresses. For example, the accumulation of ABA plays a role in stomatal closure, thus reducing water loss through transpiration (Cutler *et al*. 2010). Genes that are involved in water and ion uptake, such as aquaporins (reviewed in Tyerman *et al*. 1999), high-affinity potassium (K^+) transporters (HKT) (reviewed in Hauser and Horie 2010) and ion transporters (Blumwald 2000), also play significant roles in maintaining osmotic balances under dehydration stresses. The main focus of the present study is ABA mediated abiotic stress signalling genes. The next sections provide a comprehensive review of plant ABA and ABA mediated abiotic stress signalling genes.

1.2 Phytohormones

Plant hormones (Phytohormones) are a group of naturally occurring substances which regulate a large number of aspects of plant development and defensive responses. Historically phytohormonal signals were studied as individual pathways that mediated certain response to stress. Phytohormones have now been identified as acting in complex signalling networks, often with mutual effects referred as 'cross-talk', on plant's response to stress conditions such as drought, salt stress, wounding and pathogen attack. Considering the diverse structure and physiological functions, plant hormones are generally classified into several major classes i.e. auxins, ethylene, cytokinins, gibberellins, abscisic acid (ABA), salicylic acid, jasmonates, brassinosteroids and peptide hormones (Bai *et al.* 2010).

1.2.1 Abscisic acid

Abscisic acid (ABA) is present in different parts such as vacuoles, stomata etc. of the higher plants and is also produced by some algae and phytopathogenic fungi (Zeevaart and Creelman 1988). In the 1980s ABA was also found in mammalian brain (Chen *et al.* 1988) but it was concluded as having its origin from the plants in the animal diet and not been synthesized in the brain (Zeevaart and Creelman 1988). Research has now shown endogenous synthesis of ABA in humans. ABA is suggested to be an endogenous pro-inflammatory cytokine in human granulocytes (Bruzzone *et al.* 2007) and can stimulate the secretion of insulin in pancreatic beta cells (Bruzzone *et al.* 2008). Abscisic acid (ABA) plays a major role in regulating many of the agriculturally and physiologically critical aspects of plant development such as seed maturation and germination, which include the synthesis of seed storage proteins and lipids, the promotion of seed desiccation tolerance and dormancy (Leung and Giraudat 1998; Finkelstein *et al.* 2002). ABA also inhibits the phase transition from embryonic to germinative growth and from vegetative to reproductive growth (Finkelstein *et al.* 2002). A large amount of evidence proves that ABA play important role in responses including drought or osmotic induced stomatal closure, water tolerance induction and pathogen response (reviewed in Hetherington 2001; Kim *et al*. 2010). Moreover, ABA controls the expression of a large set of stress-responsive genes (explained in detail below) (Hoth *et al*. 2002; Nemhauser *et al*. 2006; Seki *et al*. 2002).

1.2.2 Discoveries in relation to abscisic acid

Abscisic acid (ABA) was detected in several independent investigations in late 1940s, and it became possible in 1960s to isolate and identify ABA by plant physiologists who were searching for endogenous growth inhibitors that prevented the growth of dormant plant tissues. In 1963, a group of scientists, F.T Addicott, H.R Carns and K Okhuma, were able to achieve the first crystallization of pure ABA from young cotton fruit and published the structure in 1965 (Ohkuma *et al.* 1965). ABA was first named as "abs*cisin* II" (Addicott and Lyon 1969) and was found to be a phytohormone affecting leaf abs*cis*sion and bud dormancy (Addicott and Carns 1983; Skriver and Mundy 1990; Lumba *et al.* 2010). In 1966 and 1967, the new name 'Abscisic acid' was agreed upon (Addicott and Lyon 1969). Since the discovery of the ABA structure, steady progress has been made in explaining its metabolism, synthesis and genes responsive to it. On

the other hand, early stage of the ABA perception by the plant cells was a puzzle until 2009, which was the turning point in ABA receptor protein research. In 2009, two research groups combined together upon the pyrabactin resistance (PYR)/pyrabactin like(PYRL)/ regulator component of ABA receptor(RCAR) family of soluble protein in Arabidopsis, which were proven to be a type of ABA binding receptor proteins (Ma *et al.* 2009; Park *et al.* 2009). The application of biochemical, chemical genetics and proteomics approaches have supplied the proof to obtain a clear view of the ABA binding, receptor, receptor complex formation and downstream signalling (Kline *et al.* 2010).

1.2.3 Biophysical and biochemical properties of abscisic acid

The international union of pure and applied chemistry (IUPAC) name for abscisic acid is [S-(Z, E)]-5-(1-Hydroxy-2, 6, 6 -trimethyl-4-oxo-2-cyclohexen- 1-yl)-3-methyl-2, 4-pentanedienoic acid (http://www.iupac.org/ ; last assessed June 2014).

Copyright image removed from digital version of this thesis

Original image: Addicott and Lyon (1969)

Figure 1.2 Chemical structure of abscisic acid (ABA) (Source Addicott and Lyon 1969)

The naturally occurring enantiomer of ABA has been proven to be (S)-(+)-abscisic acid. The substance $C_{15}H_{20}O_4$ (ABA) (Mr = 264.3 Da) is found to be soluble in many organic solvents. The melting point of ABA is 160°-161°C and it sublimes at 120 °C (Ohkuma *et al.* 1963). The half-life of ABA metabolism, calculated by time related disappearance curve was found to be 42 and 64 minutes for maize and *Commenlia* (dayflowers) respectively (Jia *et al.* 1996). The anion conductance of ABA is almost zero leading to no measurable interference of ABA transport with the electrical membrane potential and it is the only phytohormone which behaves ideally according to the anion trap mechanism for week acids (Hartung and Slovik 1991).

The molecular structure of ABA has a number of important features with respect to biological activity of plants. One of these is the side chain, which has two double bonds conjugated to the carboxylic acid. The configuration of the double bond next to the ring is *trans* and the one next to the acid group is *cis*. The biologically active 2-*cis*, 4-*trans* ABA is reversibly isomerized to the inactive 2-*trans*, 4-*tran*s ABA upon exposure to UV light. The active and inactive forms are easily identified by high performance liquid chromatography (HPLC) or gas chromatography (GC). Thus under high light conditions, the balance between the active and inactive ABA may shift to afford the significant quantities of the active form (Cutler *et al.* 2010). The biological activity of synthetic R-(-)- ABA and natural S-(+)-ABA has been of interest since the discovery of this plant hormone.

Copyright image removed from digital version of this thesis

Original image: Cutler *et al.* (2010)

Figure 1.3 Chemical structure of the natural S-(+)-ABA and synthetic R-(-)-ABA (Cutler *et al.* 2010).

The differences between the activities of synthetic and natural ABA on different physiological functions of plants were studied through comparison of different effects on plant processes:

Table 1.1 Activity of synthetic ABA on different physiological functions.

Function	Plant studied	Activity of R-(-)-ABA	Reference
stomatal closure	*Triticum aestivum*	weakly active	Walker-Simmons *et al.* 1992
seed germination	*Triticum aestivum*	comparable to S-(+)-ABA	
seed germination	*Arabidopsis thaliana*	comparable to S-(+)-ABA	Nambara *et al.* 2002
ABA regulated genes	*Arabidopsis thaliana*	regulate most of S-(+)-ABA controlled genes	
ABA responsive heterophylly genes	*Marsilea quadrifolia*	comparable to S-(+)-ABA	
morphogenic effects	*Marsilea quadrifolia*	stronger effect than S-(+) ABA	Lin *et al.* 2005

1.2.4 Regulation of ABA biosynthesis

Abscisic acid plays an important role during the different phases of the plant life cycle, including seed development and dormancy, stomatal closure and in responses to various stress conditions (Seiler *et al*. 2011). Many of these physiological processes are associated with the endogenous ABA levels and the diverse functions of ABA involve complex regulatory mechanisms that control its production, degradation, signal perception and transduction (Xiong and Zhu 2003). Thus an understanding of the mechanisms which regulate the plant ABA level is an important part of determining the action of ABA in plant physiological responses. The biosynthetic pathway of ABA has been largely defined in higher plants. Genetically tractable model plant, Arabidopsis has contributed much to the characterization of the main enzymes of the catabolic pathway (Wasilewska *et al*. 2008). The identification of genes encoding enzymes which are involved in ABA biosynthesis has disclosed details of the main ABA biosynthetic pathway (Seo and Koshiba 2002). As the rates of synthesis and breakdown determine the effective level of ABA *in situ*, both mechanisms should be studied in detail, although the rate of ABA breakdown remains largely unknown (Cutler and Krochko 1999). Mutants defective in ABA biosynthesis have been identified in a number of plant species such as Arabidopsis (Rock and Zeevaart 1991; Neuman *et al*.

2014), barley (Schwartz *et al.* 1997a), tobacco (Marin *et al.* 1996), maize (Tan *et al.* 1997), tomato (Burbidge *et al.* 1999; Taylor *et al.* 1988). Such mutants have been instrumental in revealing the pathway of ABA biosynthesis (Table 1.2).

Table 1.2 Mutants impaired in ABA biosynthesis.

Enzyme	**Mutant**	Plant	**Reference**
Zeaxanthin epoxidase	*aba1*	*Arabidopsis thaliana*	Rock and Zeevaart 1991
Zeaxanthin epoxidase	*aba2*	*Nicotiana plumbaginifolia*	Marin *et al.* 1996
9-*cis*-Epoxycarotenoid dioxygenase	*vpI4*	*Zea mays*	Tan *et al.* 1997
9-*cis*-Epoxycarotenoid dioxygenase	*notabilis*	*Lycopersicon esculentum*	Burbidge *et al.* 1999
Xanthoxin oxidase	*aba2*	*Arabidopsis thaliana*	Schwartz *et al.* 1997a
AB aldehyde oxidase	*aba3*	*Arabidopsis thaliana*	Schwartz *et al.* 1997a
AB aldehyde oxidase	*nar2a*	*Hordeum vulgare*	Schwartz *et al.* 1997a
AB aldehyde oxidase	*flacca*	*Lycopersicon esculentum*	Taylor *et al.* 1988
AB aldehyde oxidase	*aba1*	*Nicotiana plumbaginifolia*	Schwartz *et al.* 1997a
AB aldehyde oxidase	*sitiens*	*Lycopersicon esculentum*	Taylor *et al.* 1988
Neoxanthin synthase	*Nxd1*	*Nicotiana plumbaginifolia*	Neuman *et al*. 2014

1.2.5 ABA Biosynthetic pathway

Stage 1: Synthesis of carotenoid precursor

Direct and indirect routes are suggested for ABA biosynthesis. ABA is derived from C_{40} carotenoid by the indirect route (Taylor and Burden 1972) and from the C_{15} compound farnesyl pyrophosphate by the direct route (Zeevaart and Creelman 1988). The characterization of the ABA deficient mutants and mutated genes indicated that the indirect pathway is the main pathway and uses carotenoids for the synthesis of ABA. C_5 isopentenyl pyrophosphate (IPP) is the precursor for carotenoids. The IPP is usually synthesised from melvonic acid and 1-deoxy-D-xylulose-5-phosphate (DXP) in the cytosol and plastids respectively. This step is catalysed by the DXP synthase (DXS) enzyme. The IPP is converted to geranylgeranyl pyrophosphate (GGPP) a C_{20} product, which is then converted to a C_{40} carotenoid phytoene, using the enzyme phytoene

synthase (PSY; Figure 1.4). This is the first rate limiting step in carotenoid synthesis (Figure 1.4). Plant, algaeal and cyanobacterial PSY share conserved amino acid sequence region with bacterial phytoene synthase (CRTB) and with plant and mammalian squelene synthase. This is followed by the conversion of phytoene to ζ-carotene using phytoene desaturase (PDS) enzyme. The ζ-carotene is then converted to lycopene, followed by conversions to β-carotene and then to zeaxanthin, which is the first oxygenated carotenoid (Cunningham and Gantt 1998; Seo and Koshiba 2002; Danquah *et al*. 2013; Figure 1.4).

Stage 2: Specific ABA biosynthetic pathway

The first step which is more specific to the ABA biosynthetic pathway is the conversion of zeaxanthin to antheraxanthin and then to violoxanthin. The zeaxanthin epoxidase (ZEP) converts the zeaxanthin by the two step epoxidation to all-trans violaxanthin. ZEP was the first enzyme to be identified as the ABA biosynthetic enzyme in tobacco (Seo and Koshiba 2002). The ABA deficient mutants such as *aba2* in tobacco (Marin *et al.* 1996) and *Arabidopsis thaliana* (Schwartz *et al.* 1997a) were also found to be impaired in ZEP.

The all-trans violoxanthin is then converted to 9-*cis* violoxanthin or 9'-*cis* neoxanthin by the enzyme neoxanthin synthase, which is the latest to be identified as an ABA biosynthetic enzyme in Arabidopsis (North *et al*. 2007; Neuman *et al*. 2014) (Table 1.2). The 9-*cis*-violoxanthin or 9'-*cis*-neoxanthin is then converted to xanthoxin by oxidative cleavage, catalysed by the 9-*cis*-epoxycarotenoid dioxygenase (NCED). NCED comprises a gene family of several related genes. This gene was first isolated by characterization of the maize mutant *viviparous14* (*vp14*) (Tan *et al.* 1997). The vp14 protein specifically cleaves the 9-*cis* isomers of the epoxy xanthophylls such as 9-*cis*-violoxanthin and 9'-*cis*-neoxanthin (Schwartz *et al.* 1997b). NCED cDNA has been cloned from several species such as bean (*Phaseolus vulgaris*; Qin and Zeevaart 1999), cowpea (*Vigna unguiculate;* Iuchi *et al.* 2000), avacado (*Persea amaericana;* Chernys and Zeevaart 2000) and Arabidopsis (Neill *et al.* 1998).

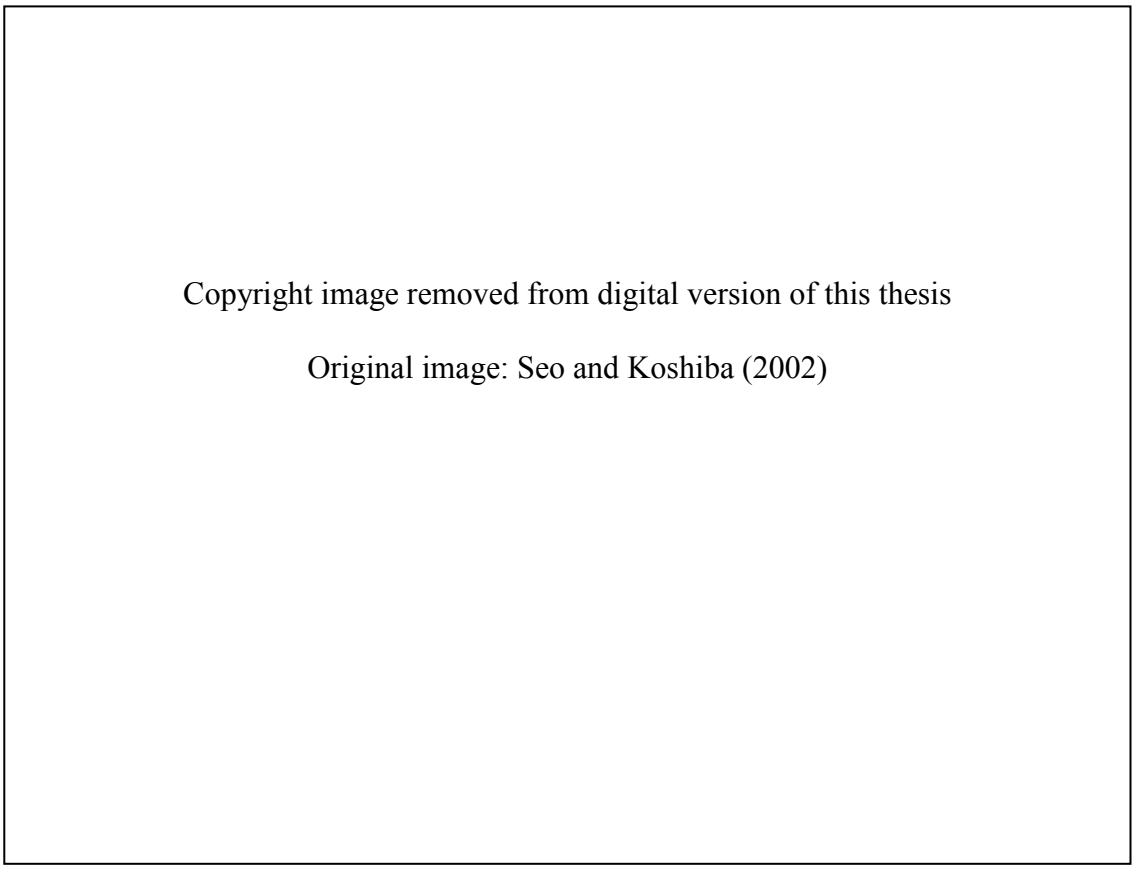


Figure 1.4 Biosynthetic pathway for ABA (Source Seo and Koshiba 2002).
(a) Carotenoid precursor synthesis in the early steps of ABA biosynthesis. ABA is synthesized from C_{40} carotenoids (phytoene, ζ-carotene, lycopene and β-carotene). Carotenoids are synthesized from a C_5compound, IPP. In plastids, IPP is synthesized via DXP from glyceraldehyde-3-phosphate and pyruvate. (b) Formation of epoxycarotenoid and its cleavage in plastid. The first step of this specific ABA synthetic pathway is the two-step epoxidation of zeaxanthin to form all-*trans*-violaxanthin catalyzed by ZEP. NCED catalyzes the oxidative cleavage of a 9-*cis* isomer of epoxycarotenoid such as 9-*cis*-violaxanthin and 9′-*cis*-neoxanthin to form xanthoxin. (c) Reactions in the cytosol for the formation of ABA. Three possible pathways are proposed. The first pathway via ABAld (1) is the most probable to function in plants, as shown by the characterization of *Arabidopsis* AO, which catalyzes the oxidation of ABAld. A member of SDR such as ABA2 in *Arabidopsis* converts xanthoxin to ABAld. The second pathway via xanthoxic acid (2) might also work. In this pathway, xanthoxin is first oxidized to xanthoxic acid by AO and then xanthoxic acid is converted to ABA, presumably by SDR. Pathway (3) via abscisic alcohol appears to be a shunt pathway but is important in mutants impaired in the oxidation of ABAld.

Stage 3: Xanthoxin to ABA

Three possible pathways have been suggested for the conversion of xanthoxin to ABA in the cytosol through abscisic aldehyde, xanthoxic acid or abscisic alcohol (Seo and Koshiba 2002) (Figure 1.4). The conversion of xanthoxin to abscisic aldehyde is catalysed by ABA2, which belongs to the short chain dehydrogenase/reductase (SDR) family (Rook *et al.* 2001). The Arabidopsis *aba2* mutant was found to be deficient in the synthesize of ABA from xanthoxin, whereas it could oxidize abscisic aldehyde to form by *in vitro* enzyme preparations (Schwartz *et al.* 1997a). ABA2 protein was found to be encoded by single gene in Arabidopsis, therefore loss of function of this gene

leads to severe ABA deficiency (Nambara *et al*. 2002). The abscisic aldehyde is then oxidized to carboxylic acid by abscisic aldehyde oxidase (AAO). Among four different Aradidopsis abscisic aldehyde oxidase (AAOs), AAO3 encode the enzyme (also called as aldehyde oxidase (AO)), which catalyses oxidation of abscisic aldehyde (Sekimoto *et al.* 1998; Seo *et al.* 2000a). Arabidopsis mutant with a defect in the *AAO3* gene shows wilty phenotype and ABA biosynthesis is hindered (Seo *et al.* 2000b). With the help of these studies, the pathway for conversion of xanthoxin to abscisic aldehyde, followed by its conversion to ABA, is supported (Seo and Koshiba 2002).

The second pathway for ABA synthesis through xanthoxic acid might also be functional (Milborrow 2001). Arabidopsis AO was found to oxidize xanthoxin in activity gel staining after native gel electrophoresis (Lee 1997), suggesting that the xanthoxin has been oxidized to xanthoxic acid by the AO. This indicates that two ABA biosynthetic pathways operate in an organ or development stage dependent manner (Seo and Koshiba 2002).

The third pathway for ABA biosynthesis through abscisic alcohol might be active in some mutants. Studies have shown that exogenously supplied abscisic aldehyde to *flacca* and *sitiens* mutant of tomato (with modified aldehyde oxidase; Table 1.2; Taylor *et al.* 1988) were reduced to abscisic alcohol, showing that abscisic aldehyde was first converted to abscisic alcohol and then oxidized to ABA (Rock and Zeevart. 1991).

1.2.6 ABA catabolism

ABA catabolism can be categorized via two pathways, hydroxylation and conjugation. (Nambara and Marion-Poll 2005). According to the ABA hydroxylation pathway, oxidation of one of the methyl group of the ring structure at three positions, C-7', C-8 and C-9' is the first step (Figure 1.5), of which C-8' is the primary (Cutler and Kruchko 1999; Danquah *et al*. 2013). Cytochrome P450 monooxygenase (CYP707A) is the enzyme involved in the catalysis of hydroxylation of ABA at C-8' to form unstable 8'-hydroxy ABA, which is later converted to phaseic acid (PA) by spontaneous isomerization (Kushiro *et al*. 2004; Saito *et al*. 2004). PA is then reduced to dihydrophaseic acid (DPA) by PA reductase (Gillard and Walton 1976). The oxidation

product of other methyl groups (C-7'and C-9') include 7'-hydroxy ABA, 9'-hydroxy ABA and neo-phaseic acid (Zhou *et al*. 2004) (Figure 1.5).

According to the conjugation pathway, carboxyl at C-1 and hydroxyl groups of ABA and its oxidative catabolites are the potential targets for conjugation with glucose leading to ABA inactivation by forming different conjugates (Danquah *et al*. 2013). Of these conjugates, ABA glucosyl ester (ABA-GE) is the most widespread, which is produced by glycosyltransferase. The gene encoding glycosyltransferase was first isolated from adzuki bean (*Vigna angularis*) and was named *AOG* (Xu *et al*. 2012a). The low membrane permeability of ABA-GE, makes it suitable for long-distance translocation and storage in vacuoles and apoplastic space (Jiang and Hartung 2008). ABA is found to be released from ABA-GE by β-glucosidases under drought condition. The enzymatic activity of β-glucosidases to catalyse the hydrolysis of ABA-GE to release ABA was first found in barley (Dietz *et al*. 2000). Two β-glucosidases (BG1 and BG2) were isolated in Arabidopsis (Lee *et al*. 2006; Xu *et al*. 2012a), which were found to be expressed in abiotic stress conditions such as dehydration and NaCl. Knockout mutants of these genes in Arabidopsis were hypersentive to abiotic stress, whereas over-expression lines were tolerant and contained more ABA (Xu *et al*. 2012a).

Figure 1.5 ABA catabolic pathways (Source Nambara and Marion-Poll 2005)
The 8′-hydroxylation is predominant pathway for ABA catabolism. Red and blue asterisks indicate active and less-active hydroxy groups for conjugation, respectively.

1.2.7 Self-regulation of ABA biosynthetic genes

The endogenous level of ABA is maintained by the balance between anabolism and catabolism. NCED has been proposed to be regulatory enzyme with respect to ABA biosynthesis as expression of NCED is co-related to endogenous ABA content and its overexpression can lead to significant ABA accumulation. ABA catabolic enzyme, ABA 8'hydroxylase is also expected to be an ABA regulatory enzyme (Nambara and Marion Poll 2005). Along with these two main enzymes, steps upstream of ABA metabolism are also responsible for determining the ABA level. Overexpression of genes encoding the regulatory enzyme for the carotenoid biosynthesis (phytoene synthase) and xanthophyll cycle, cause an enhanced accumulation of ABA in Arabidopsis seeds and seedlings (Lindgren *et al*. 2003; Frey *et al*. 1999). This suggests that the regulation of ABA metabolism is not only regulated by specific steps in ABA metabolism, but also by up-stream metabolism (Nambara and Marion-Poll 2005).

Genetic analysis of *sadI* (supersensitive to ABA and drought) mutant of Arabidopsis indicated that ABA biosynthesis is also controlled at the level of mRNA stability. *SAD1* locus encodes a peptide similar to Sm-like- small ribonucleoprotein (snRNP), expected to be involved in mRNA splicing, export and degradation may regulate the transcription rate of early signalling components. These components might be involved in the feedback circuit (Xiong and Zhu 2003). The *sad1* mutant showed reduction in levels of ABA and phaseic acid (PA) and it has also been demonstrated that SAD1 is a positive regulator of *AAO3* and *ABA3* genes. The role of SAD1 in regulation and metabolism of ABA is unclear, but the identification of RNA processing genes through genetic screens indicates that RNA processing is closely related to the regulation of ABA biosynthesis (Nambara and Marion Poll 2005). Arabidopsis ABI1, a member of the protein phosphatase 2C (explained below) is the negative regulator in ABA signalling. In *abi1* mutants, self-regulation loop of ABA biosynthesis was found to be partially impaired as in this mutant, ABA failed to activate the expression of NCED genes and significantly reduced transcript levels of ZEP and AAO were detected under exogenous ABA treatment. The involvement of ABI1 in the ABA regulation suggest that signalling for ABA biosynthesis is ABA dependent and there is a cross-link between the ABA biosynthesis and responsiveness pathway (Xiong and Zhu 2003).

1.3 ABA transporters for intercellular signalling

The physiological studies of ABA functions suggest that translocation and communication of this plant hormone between cells, organs and tissues plays an important role in whole plant physiological responses. ABA is mainly biosynthesized and metabolized in the vascular tissues whereas it acts in the distant guard cells leading to stomatal closure (Schachtman and Goodger 2008). Cell to cell ABA transport, which remained unclear, was clarified by the discovery of two specific plasma membrane bound ABA transporters, ATP binding cassette (ABC) named ABCGG25 and ABCG 40 (Kang *et al.* 2010; Kuromori *et al.* 2010).

The ATP-binding cassette (ABC) transporter gene *ABCG25* (AT1G71960) encodes a protein responsible for transport and response of ABA in Arabidopsis. *ABCG25* was found to be expressed mainly in the vascular tissues, where ABA is mainly synthesized. The Arabidopsis *abcg25* mutant was isolated by genetically screening for ABA sensitivity at the time of greening of cotyledons (Kuromori *et al.* 2010). The fluorescent protein fused ABCG25 is localized at the plasma membrane in plant cells. This ABC transporter was found to be conserved in many species from *E.coli* to humans and has been shown to transport different types of metabolites or signalling molecules (such as K^+ and protoporphyrin IX) including plant hormones in an ATP dependent pathway (Nagashima *et al.* 2008). The transport of ABA by ATP dependent ABCG25 was shown by the membrane vesicles derived from the *ABCG25* expressing insect cells. The *ABCG25* overexpression in plants leads to high temperature implying an influence on stomatal regulation. This suggested that ABCG25 plays an important role as an exporter of ABA through the plasma membrane and is involved in the inter-cellular ABA signalling pathway (Kuromori *et al.* 2010).

The other ABC transporter, AtABCG40 (AT1G15520) was found to function as an ABA importer in Arabidopsis cells (Kang *et al.* 2010). *ABCG40* was found to express in the leaves of young plantlets and also in primary and lateral roots, highest in the guard cells of leaves. The ABCG40::sGFP (green fluorescent protein) expression driven by native promoter in Arabidopsis guard cells, demonstrated the plasma membrane localization of ABCG40. The uptake of ABA in Arabidopsis was found to be increased in the cells expressing *ABCG40*, whereas the ABA uptake was decreased in the *abcg40*

mutants of Arabidopsis, which also showed slow closing of stomata in response to ABA and resulted in reduced stress tolerance. Moreover in response to exogenous ABA, the up regulation of ABA inducible genes is strongly delayed in *atabcg40* plants, which indicated that the ABCG40 plays an important role in the timely response to ABA (Kang *et al.* 2010).

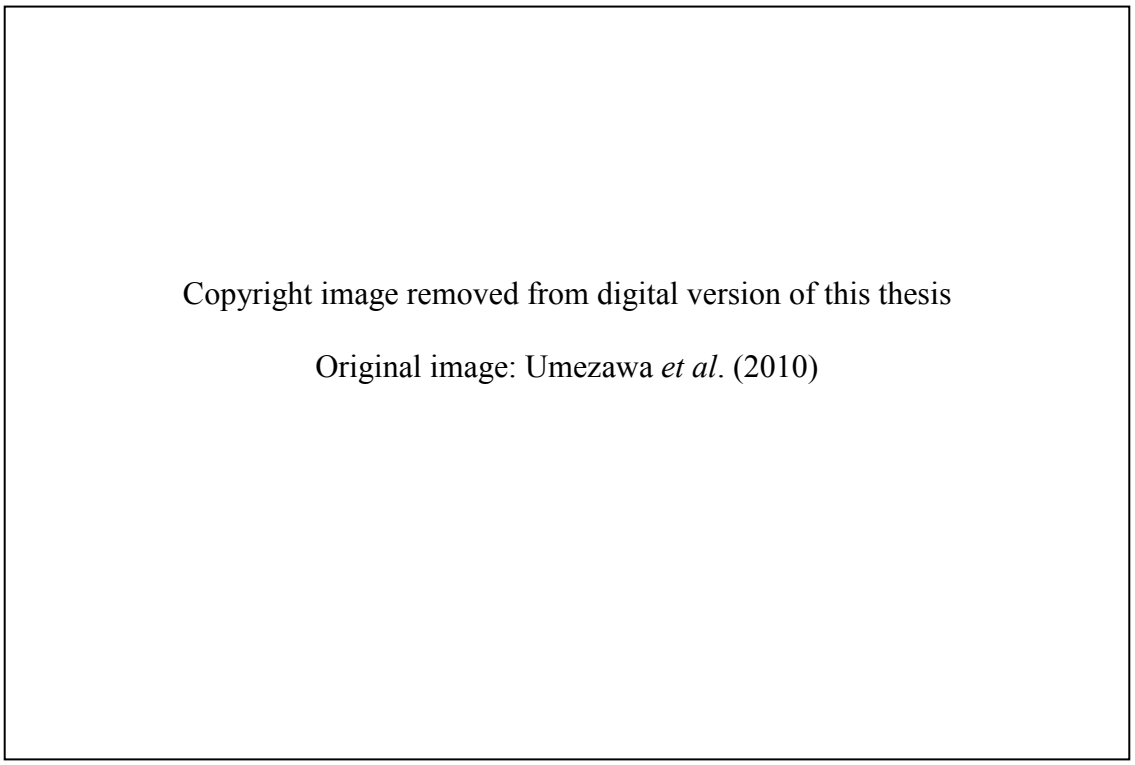


Figure 1.6 Hypothetical view of ABA intracellular transmission (Source Umezawa *et al*. 2010)
Arabidopsis leaf section showing two distinct cell types: vascular tissues, including vascular parenchyma cells, and guard cells on the leaf epidermis. AtABCG25 function in ABA efflux from ABA-biosynthesizing vascular cells, and ABA would diffuse into apoplastic areas. AtABCG40 function in ABA influx into guard cells to facilitate stomatal closure

A simple model can be suggested by the help of two reports: ABA is exported from ABA biosynthesizing cells to the apoplastic area, which is followed by the importing of the ABA from the apoplast into the guards cell (Kang *et al.* 2010; Kuromori *et al.* 2010). This model also explains for the ABA receptors (explained below in Section 1.6) are soluble and localized in the cytosol (Ma *et al.* 2009; Park *et al.* 2009). It also suggests the importance of an ABA transporter, which could deliver ABA in a regulated manner to start rapid and controlled response in different types of stress conditions (Kang *et al.* 2010).

ABC transporter proteins are characterized by the possession of one or two cytosolically oriented nucleotide binding folds (NBFs) or ATP binding cassette (ABC), which are linked to the multiple hydrophobic transmembrane domains (TMS) (Van Den Brûle and Smart 2002). The conserved region of TMS domain contains an ATP binding site consisting of Walker A box and Walker B box separated by approximately 120 amino acids and between the two boxes, a consensus sequence specific for ABC transporters (ABC signature motif) (Higgins 1992).

1.4 Regulatory proteins in abscisic acid signalling

Abscisic acid plays an important role in the regulation of many important aspects of plant development, such as synthesis of seed storage proteins and lipids, and promotion of seed desiccation, tolerance and dormancy. These processes involve regulatory proteins such as flowering time control proteins A, kinases (sucrose non fermenting 1 kinase 2) and transcription factors. ABA is also found to mediate physiological responses to environmental stress including drought or osmotica induced stomatal closure, the induction of tolerance to water, salt and pathogen response. These require receptors such as pyrabactin resistance (PYR/pyrabactin like (PYL)/regulator components of ABA receptor (RCAR, phosphatases (group A protein phosphatase 2A), kinases (SnRK2) and transcription factors (Figure 1.7) (Danquah *et al*. 2014).

Copyright image removed from digital version of this thesis

Original image: Nakashima and Yamaguchi-Shnozaki (2013)

Figure 1.7 Model of ABA pathways that control the transcription during seed maturation and under abiotic stress (Source Nakashima and Yamaguchi-Shnozaki 2013).

Regulatory factors that control ABA response have been identified by genetic, biochemical and cell biological studies. Genetic screens are used to study the irregular growth or gene expression responses to ABA. 'Reverse genetics' studies have been used for testing the functional role of particular gene; these studies usually employ screening of large mutagenized populations or a transgenic approach to over express or disrupt a target gene (McCallum *et al.* 2000; Finkelstein *et al.* 2002). Biochemical studies have led to the identification of wide variety of gene promoter elements, kinases, kinase inhibitors, phosphatases, phospholipases and transcription factors correlated with ABA response. Cell biological studies have analysed the action of candidates for the secondary messengers and signalling intermediates in regulating cellular response to ABA (Finkelstein *et al.* 2002). These ABA mediated regulatory factors are explained in detail below:

1.5 ABA mediated regulatory components involved during plant development

The ABA mediated regulatory components involved during plant development are flowering time control proteins A, kinases and other transcription factors. The kinases and transcription factors are also involved in ABA mediated abiotic stress signalling (explained in Section 1.6.5). This section includes the description of flowering time control proteins A.

1.5.1 Flowering Time Control proteins A

The first ABA binding protein (ABAP1) was isolated from barley aleurone with anti-idiotypic ABA antibodies (AB2) (Razem *et al.* 2004). Flower Time Control Protein A (FCA) is the closest Arabidopsis homolog of the deduced amino acid sequence of ABAP1 (Macknight *et al.* 1997). FCA is a nuclear RNA binding protein which is specific to plants and promotes flowering by accumulation of the mRNA encoding Flowering Locus C (FLC) (Sheldon *et al.* 2000), MADS (Michaels and Amasino 1999) box transcription factor which is a key factor in the flowering time control. For functioning, FCA needs another protein, which is the RNA 3'end processing factor Flowering Locus Y (FY), which binds to its tryptophan-tryptophan (WW) protein interaction domain of FLC (Figure 1.8) (Simpson *et al.* 2003). On the attachment of ABA with high affinity to FCA, it reduces the interaction between the FLC and FY and leads to accumulation of full length FCA and FLC mRNA. The isolation of an ABA receptor as the mRNA stability component is appealing because many of the ABA

related loci are implicated in RNA metabolism (Razem *et al.* 2006). The expression of FCA is auto-regulated by promoting premature cleavage and polyadenylation in intron 3 of its own precursor mRNA and it requires the interaction with FY as well (Macknight *et al.* 2002). The RNA recognition motif in FCA is not present in the barley homologue ABAP1 protein of barley. One of these process is the negative auto-regulation of FCA expression by promoting the premature cleavage and polyadenylation in intron 3 of its precursor mRNA, which increases the premature mRNA encoding the major form of inactive, truncated protein FCAβ and thus decreases the mature FCA mRNA which encodes the active full length protein FCAγ. The second process is to down regulate the FLC expression through the direct action on FLC pre-mRNA or indirectly through a FLC regulator. The FCA-FY interaction, which controls the FLC expression, is the central point in flower regulation. This interaction is the point where ABA regulates the flowering of a plant (Bäurle and Dean 2006).

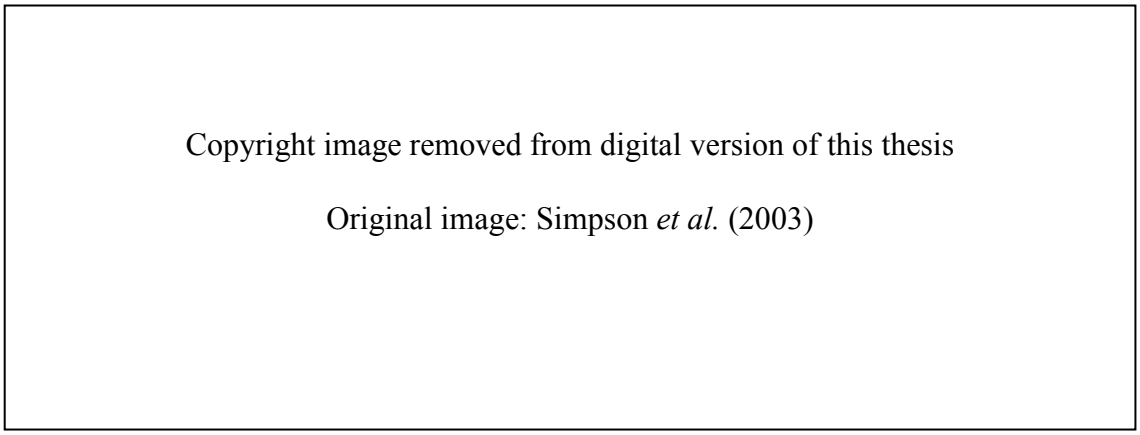


Figure 1.8 FCA and FY function model (Source Simpson *et al.* 2003)
FCA (red) binds target RNA through its two N-terminal RRMs and tethers the 3' end-processing machinery (pale blue) to this RNA via an interaction between the FCA WW domain and the PPLP domain of the 3'end-processing factor, FY (blue). One target of this interaction is FCA pre-mRNA (red). Auto-regulation of FCA expression presets the level of active FCA protein available to regulate the floral repressor, FLC (green), which is executed by FCA again interacting with FY.

The *in vitro* and *in vivo* assays have shown that the binding of ABA to FCA disrupts the FCA-FY interaction. The site for the binding of ABA at the FCA molecule is present near the C-terminus and protected but it does not include the FY binding WW domain. The ABA-FCA binding leads to the dissociation of the FCA-FY complex which abolishes the downstream signalling (Razem *et al.* 2006). The break in the downstream signalling includes two actions; one is a functional loss in auto- regulation of FCA expression with a decrease in the premature FCA mRNA, which encodes the inactive FCAβ and also as a consequence there is an increase in the mature FCA mRNA, which

encodes the active form FCAγ. Second is the increase in the expression of the central flowering repressor FLC. This biochemical pathway leads to the physiological consequence of significant delay in the flowering time (Razem *et al.* 2006). The regulation of the floral transition pathway by ABA mediating FCA was proved by implying the flowering mutants, *fca-1* and *fy-1*, along with the ABA biosynthetic mutants *aba1* and ABA signalling mutant *abi-2* in Arabidopsis. The ABA leads to the decrease in number of lateral roots slightly in the case of *fca-1* mutant whereas there is a considerable decrease noticed in the case of wild type plants. This suggests that the FCA might play a role in ABA-inhibited lateral root formation (Wang and Zhang 2008). On the other hand, *fca-1* seeds were seen not germinating in the presence of ABA and apertures of the *fca-1* guard cells significantly decreased after the application of ABA. It was thus proven experimentally that FCA does not play a role in other ABA physiological responses such as seed germination or stomatal closure (Razem *et al.* 2006).

The binding of the Arabidopsis FCA protein to ABA was also confirmed by the stereospecific studies. FCA binds to the naturally occurring, physiologically active (+)-ABA but not to the physiologically inactive (-)-ABA, which indicates that the ABA binding is stereo-specific (Razem *et al.* 2006). FCA binds to ABA at a ratio of roughly 0.72 mol of ABA per mol of protein, which suggests that the FCA has only one binding site for the ABA.

1.6 ABA mediated regulatory components involved during abiotic stresses

The ABA mediated regulatory components involved during abiotic stresses are pyrabactin resistance (PYR/pyrabactin like (PYL)/regulator components of ABA receptor (RCAR, phosphatases (group A protein phosphatase 2A), kinases (SnRK2) and transcription factors. These regulatory components are explained below.

1.6.1 Membrane bound receptors

Several experiments have suggested that the extracellular perception is important for ABA to achieve its function (Anderson *et al.* 1994). Thus plasma membrane localized ABA receptors can be major players for perceiving the extracellular ABA and

communicate the ABA signalling responses. G-protein coupled receptor (GPCR) is the major membrane bound ABA receptors and is explained below.

Several observations such as pharmacological evidence have suggested that G-protein coupled receptor (GPCR) might participate in ABA signal transduction in plants (Pandey *et al*. 2009). G protein mediated signal transduction process constitutes one of the most evolved effector signalling networks (Offermanns 2003). The principal components of this signalling pathway are hetrotrimeric G proteins and G protein coupled receptors (GPCRs). G proteins are comprised of Gα (encoded by GPA1), Gβ (encoded by AGB1) and Gγ (encoded by AGG1 and AGG2) subunits. The Gα subunit has got both the GTP-binding and GTPase activity sites, which makes it to act as bimodal molecular switch, typically with a GTP bound "on" mode and a GDP bound "off" mode. GPCRs are known to act as guanine nucleotide exchange factors (GEF) and a change in the GPCR conformation upon signal perception leads to exchange of GDP for GTP at the Gα subunit (Cabrera-Vera *et al.* 2003). The plasmon resonance spectroscopy and bimolecular fluorescence complementation studies have proven the physical interaction of GPCR and Gα. This promotes dissociation of the hetrotrimer G protein into free GTP- Gα and Gβγ dimers, both of these can interact with an array of downstream signalling elements. The intrinsic GTPase activity of Gα regenerates its GDP bound form, allowing it to reassociate with the Gβγ dimer and complete the cycle (Figure 1.9; Cabrera-Vera *et al.* 2003). The G protein cycle is also regulated by the accessory proteins, especially the GTPase accelerating proteins (GAPs) (Ross 2008) and the guanine nucleotide dissociation inhibitor (GDI) proteins which mainly inhibits the dissociation of GDP from Gα (Pandey *et al*. 2009).

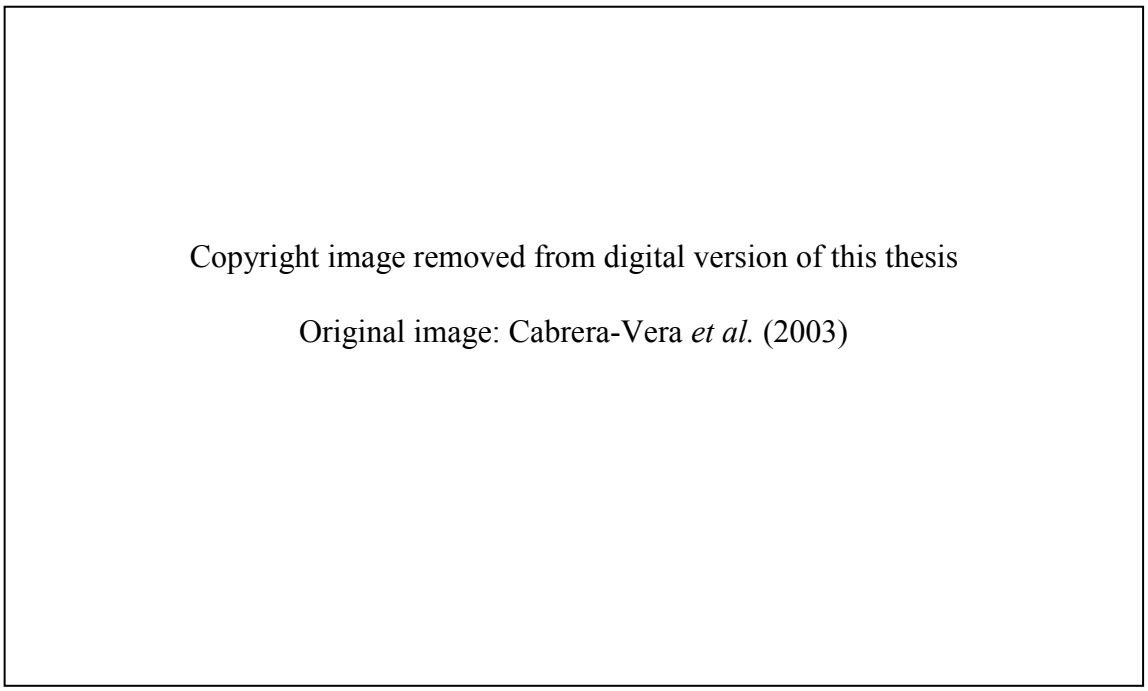


Figure 1.9 GPCR mediated G-Protein activation (Source Cabrera-Vera *et al.* 2003).
The interaction of an endogenous ABA (A) with cell surface receptor (R) facilitates the binding of the activate receptor (R*) with intracellular heterotrimeric G proteins. The R*-G protein binding promotes the exchange of GDP for GTP on the Gα-subunit. Gα-GTP then splits from Gβγ and R*. Both subunits are free to modulate the activity of a wide variety of intracellular effectors. Termination of the signal occurs when the γ-phosphate of GTP is removed by the intrinsic GTPase activity of the Gα-subunit, leaving GDP in the nucleotide binding pocket on Gα. Gα-GDP then reassociates with Gβγ and the cycle is complete. RGS proteins accelerate the intrinsic GTPase activity of Gα-subunits, thereby reduces the duration of signalling events.

Arabidopsis GTG1 and GTG2 have been identified and described as GPCR type G proteins, which exhibit specific GTP binding and intrinsic GTPase activity (Pandey *et al.* 2009). The role of Arabidopsis GTGs in ABA perception was confirmed by studies reporting the specific binding of ABA to purified recombinant GTG proteins and also by genetic studies (Pandey *et al*. 2009). This showed that *gtg1* and *gtg2* mutants of Arabidopsis were hyposentive to ABA in responses such as seed germination and primary root growth. Moreover expression of ABA responsive genes such as PP2C and SnRK2 (explained below) was reduced in *gtg1* and *gtg2* mutant (Pandey *et al*. 2009).

Soluble Receptors

In 1990s, microinjection studies and treatments with ABA analogs in Arabidopsis suggested that ABA may have both intracellular and extracellular sites of perception, and several proteins with the properties of intracellular ABA receptors have been described (Cutler *et al*. 2010). These include flowering time control proteins (included in plant development; explained in Section 1.5), putative ABA receptor and pyrabactin

resistance(PYR)/pyrabactin like (PYL)/regulatory component of ABA receptors (RCAR). The current data on these proteins is summarized below.

1.6.2 The putative ABA Receptor

A study of ABA binding proteins from *Vicia faba* led to the identification of a second ABA specific binding protein, which was likely involved in ABA-induced stomatal signalling (Zhang *et al.* 2002), hence named as abscisic acid receptor (ABAR) (Shen *et al.* 2006). The sequencing information has helped in the isolation of a complementary DNA fragment encoding the carboxy terminal half of about 770 amino acids of the putative H subunit (CHLH) of the magnesium protoporphyrin-IX chelatase (Mg chelatase) from the broad bean. CHLH plays an important role as the subunit of the Mg-chelatase in chlorophyll synthesis (Walker and Willows 1997) and plastid to nuclear retrograde signalling (Mochizuki *et al.* 2001). ABAR was found to show stereo specificity and affinity for ABA, underpinning the ABA receptor nature of this protein. The reduced expression or over expression of the ABAR gene has proven the effect of the ABA's response in stomatal movement, germination and gene expression (Hirayama and Shinozaki 2007). Mg chelatase is comprised of three subunits, namely CHLD, CHLI and CHLH. CHLH plays a central role and catalyses the insertion of Mg^{2+} into protoporphyrin IX (Proto) to form Mg-protoporphyrin IX (MgProto) (Figure 1.10). This is the first unique step of chlorophyll biosynthesis (Walker and Willows 1997). Using Arabidopsis *genome uncoupled 5* (*gun5*) mutant, which had a single amino acid Ala 990 to Val mutation, it was proven that CHLH is involved in the plastid to nucleus retrograde signalling by regulating the metabolism of the tetrapyrrole signal MgProto or by sensing the signal (Mochizuki *et al.* 2001).

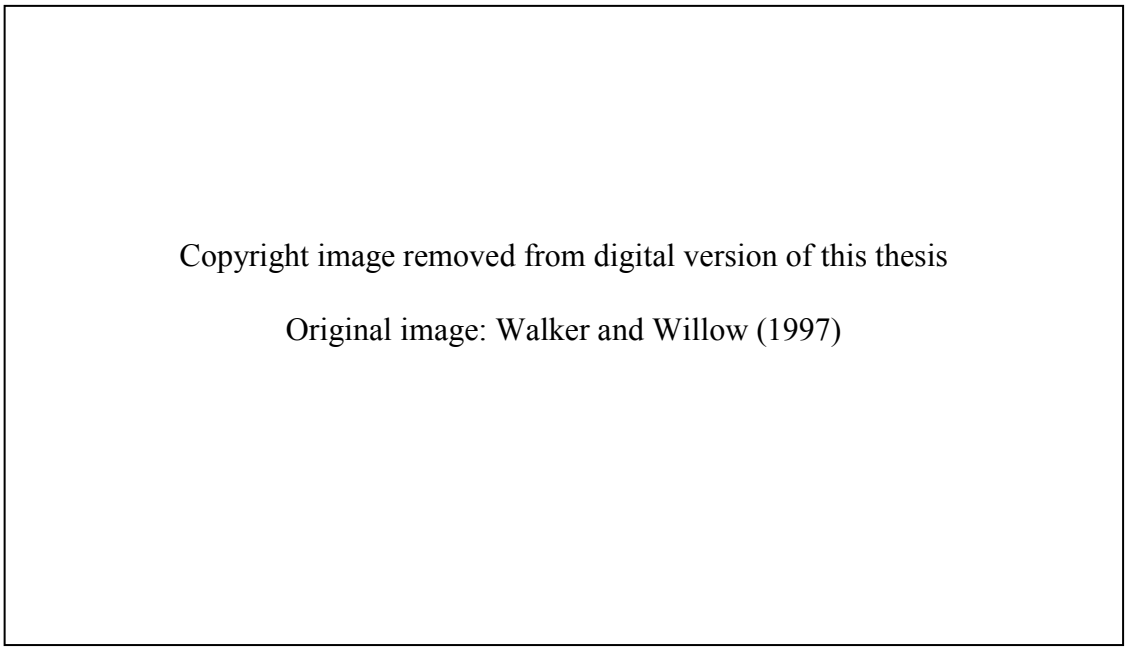


Figure 1.10 Reaction catalysed by Mg-chelatase (Source Walker and Willow 1997).

Initially, CHLH was found to be limited to green tissues only, however the data available at Genevestigator (http://www.genevestigator.ethz.ch) showed the presence of the Arabidopsis CHLH mRNA in seeds as well. It was then experimentally shown that ABAR was also expressed in non-green tissues including roots. Hence ABAR may function at the whole plant level (Shen *et al.* 2006).

The ABAR binds strerospecifically to (+) ABA and not to (-) ABA. When the guard cell protoplast was pre-treated with an antibody raised against ABAR, it was observed that the ABA–induced phospholipase D activity was decreased. This suggested that the ABAR may function as an ABA receptor (Kim 2007). The function of ABAR in ABA signalling was studied using the transgenic RNAi, antisense and overexpression lines. The underexpressing ABAR plants demonstrated significant ABA-insensitive phenotypes in seed germination, post-germination arrest of growth by ABA and ABA induced stomatal closure and inhibition of stomatal opening. On the other hand, the overexpressing ABA plants displayed ABA hypersensitive phenotypes and were found to be more resistant to dehydration from the leaves or whole plants (Shen *et al.* 2006). It was shown that the *abar-I* mutant seeds were deficient in lipid and mature protein bodies, which indicated possible distortion of late embryonic development (Finkelstein *et al.* 2002). These phenotypes were found comparable to the mutations in ABA-signalling genes such as ABI3, which has specific effects on seed maturation (Nambara

1992). It has been proven using transgenic plants that the ABAR is not involved in biosynthesis of ABA (Shen *et al.* 2006).

ABAR expression by RNAi was shown to decrease the levels of positive regulators of ABA signalling such as MYB2 (transcription factor; Abe *et al.* 2003), MYC2 (transcription factor; Abe *et al.* 2003), ABA insensitive 4 (ABI4; transcription factor; Finkelstein 1994), ABA insensitive 5 (ABI5; transcription factor; Finkelstein 1994) and open stomatata 1 (OST1; SnRK2; Mustilli *et al.* 2002). The ABAR also plays an important role in enhancing the negative regulators, ABA insensitive 1 (ABI1; PP2CA; Gosti *et al.* 1999), ABA insensitive 2 (ABI2; PP2CA; Gosti *et al.* 1999) and CIPKI5 (SnRK3; Guo *et al.* 2002) in leaves. The genes specific for the ABA signalling in seeds, such as *ABI3* (Nambara 1992), *ABI4* (Finkelstein 1994), *ABI5* (Finkelstein 1994) and their downstream genes such as *EM1* and *EM6* were found to be down regulated by ABAR. This was confirmed by expression level of these genes in the *gun 5* mutant of Arabidopsis, which is similar to the wild type *Arabidopsis thaliana* Columbia line (Finkelstein *et al.* 2002; Manfre *et al.* 2006). The *chh* is also a gun mutant but with a single nucleotide substitution at different sites, which resulted in single amino acid mutation Pro 642 to Leu (Mochizuki *et al.* 2001). The *chh* mutant of Arabidopsis was found to significantly decrease the ABA binding activity of the ABAR whereas the *gun 5* mutant failed to do this. This regulation of the ABA-signalling genes by the ABAR supported the hypothesis that ABAR is a positive regulator for ABA signalling (Shen *et al.* 2006).

1.6.3 Pyrabactin resistance (PYR)/ PYR like (PYL)/ Regulatory component of ABA receptor (RCAR)

Even after the isolation and characterization of ABA-receptor proteins such as GPCR and CHLH, their physiological and molecular relevance to well-known ABA-signalling proteins such as protein phosphatase 2C subgroup A (PP2CAs; negative regulators) and sucrose non fermenting 1 related kinase 2 (SnRK2s; positive regulators) (explained in Section 1.7.4) remained mysterious until 2009. Novel soluble ABA binding receptors, pyrabactin resistance (PYR)/ PYR like (PYL)/ regulatory component of ABA receptor (RCAR) were identified by two independent research groups using different approaches of chemical genetics and biochemistry (Ma *et al.* 2009; Park *et al.* 2009). The

molecular interactions of PYR/PYL/RCAR with already known ABA signalling factors have also been proven by these research groups. PYR/PYL/RCAR proteins belong to a family of soluble ligand binding proteins which are defined as the START domain superfamily (Cutler *et al.* 2010). It has also been named as *Bet v I* superfamily to recognise the original identification of the conserved domain from the major pollen allergen of the *Betula verrucosa* (white birch) (Cutler *et al.* 2010). The START-domain proteins are characterized by the conserved "helix grip" fold which forms the central hydrophobic ligand binding pocket (Iyer 2001).

Characterization of the synthetic ABA agonist pyrabactin showed that the Pyrabactin Resistance (PYR1) proteins (a member of PYR/PYL/RCAR) is a necessary component for the pyrabactin action *in vivo* (Park *et al.* 2009). Another member, PYL9 was shown to bind stereo-specifically to the natural (+) ABA, whereas the PYL5 can also bind to the unnatural (-) ABA stereoisomer (Santiago *et al.* 2009a). These observations suggested that many members of the protein family PYR/PYL/RCAR bind to (-) ABA and hence explained the pervasive bioactivity of (-) ABA (Cutler *et al.* 2010). The difference in selectivity for pyrabactin and the unnatural (-)-ABA in comparison to the natural (+)-ABA, suggest that the ligand binding pockets of the PYR proteins are likely to have variable residues involved selective receptor activation (Cutler *et al.* 2010).

There are 14 PYR/PYL/RCAR members in Arabidopsis, called as PYR1 and PYL1-PYL13. The triple (*pyr1;pyl1;pyl4*) and quadruple (*pyr1;pyl1;pyl2;pyl4*) mutant lines of Arabidopsis showed reduced sensitivity in germination and root growth as a response to the (+) ABA, whereas the control plants showed inhibited seed germination and root growth in response to ABA. The quadruple mutant also failed to show ABA induced stomatal closure (Park *et al.* 2009; Nishimura *et al.* 2010). Overexpression of the PYL5 in Arabidopsis confered drought resistance, proving the role of PYR/PYL/RCAR family in the stress response (Santiago *et al.* 2009a). The genetic analysis and structural studies together demonstrated that PYR/PYL/RCAR are ABA receptors that control many aspects of ABA signalling and physiology (Cutler *et al.* 2010).

Yeast two hybrid and other studies showed that both pyrabactin and ABA cause PYR1 to bind and inhibit the group A protein phosphatases (PP) 2Cs (PP2CAs), such as the ABA insensitive (ABI) 1 and ABI2 and hypersensitive to ABA 1 (HAB) 1 (Park *et al.* 2009; explained in Section 1.7.3). Other PYR/PYL/RCARs such as PYL2, PYL3 and PYL4 can interact with the protein 'hypersensitive to ABA 1' (HAB1; PP2CA) in presence of ABA (Cutler *et al*. 2010). Another member of receptor family, PYL9 binds to the (+) ABA and this interaction is increased ten times by the addition of ABI2, which suggests that PP2CAs play a role in stabilising the interaction between ABA and PYR/PYL/RCAR (Ma *et al.* 2009). Structural studies of different family members such as PYR1, PYL1 and PYL2 proteins, either alone or in the complex form with ABA and PP2CAs, have provided the physical explanation for the cooperative bonding.

Copyright image removed from digital version of this thesis

Original image: Umezawa *et al*. (2010)

Figure 1.11 Overview of ABA signalling and transport (Source Umezawa *et al.* 2010)

PYR/PYL/RCAR, PP2C and SnRK2 form a core signalling complex (yellow circle), which functions in at least two sites. One is the nucleus, in which the core complex directly regulates ABA-responsive gene expression by phosphorylation of AREB/ABF-type transcription factors. The other is the cytoplasm, and the core complex can access the plasma membrane and phosphorylate anion channels (SLAC1) or potassium channels (KAT1) to induce stomatal closure in response to ABA. Other substrates of SnRK2s have yet to be identified. The ABA transport system consists of two types of ABC transporter for influx or efflux. ABA movements are indicated by green lines and arrows, and major signalling pathways are indicated by red lines and arrows. Dotted lines indicate indirect or unconfirmed connections.

1.6.4 Protein phosphatases involved in ABA signalling

In the 1990s, genetic screen for the ABA-insensitive Arabidopsis mutants led to the identification of two genes, *ABA-Insensitive1* (*ABI1*) and *ABA-insenstive2* (*ABI2*), which encoded the group A protein phosphatases 2C (PP2CA) (Leung *et al.* 1994). The mutants *abi1-1* and *abi1-2* (explained below) have shown an overall ABA insensitivity in various tissues and developmental stages suggesting that the PP2C acts as the global regulator in ABA signalling (Umezawa *et al.* 2010).

The protein phosphatase enzymes are categorized with respect to their substrate specifcity into Ser/Thr, Tyr and dual specificity classes. The Ser/Thr phosphatases, known for their role in ABA-signalling, are further divided into PP1, PP2A and PP2C groups of enzyme. Of these, PP2A and PP2C have been proven to be involved in ABA signalling by genetic studies (Luan 2003). Pharmacological approaches also indicate the participation of Tyr and dual specificity phosphatases in ABA signalling (Cutler *et al.* 2010). Another phosphatase, inositol polyphosphate 1- phosphatase SAL1/FIERY1, which functions in the catabolic pathway of 1,4, 5 trisphosphate (IP3), also acts as the negative regulator for the ABA and stress signalling in Arabidopsis (Xiong *et al.* 2001).

PP2Cs are Mg^{2+}/Mn^{2+} dependent monomeric enzymes, which function as negative regulators of ABA signalling. This has been proven by number of genetic evidences and this function is conserved from moss to Arabidopsis (Komatsu *et al.* 2009). It has been suggested that PP2Cs site of action is nucleus, in spite of the localization in both the cytosol and nucleus (Moes *et al.* 2008). An order can be established in the functions of PP2Cs according to gene expression level, tissue expression pattern and the analysis of ABA responses in different combinations of *pp2c* knockout lines in Arabidopsis (Rubio *et al.* 2009).

There are 9 PP2CA members in Arabidopsis, called as ‘ABA insensitive 1’ (ABI1), ‘ABA insensitive 2’ (ABI2), ‘hypersensitive to ABA 1’ (HAB1), ‘hypersensitive to ABA 2’ (HAB2), ‘ABA hypersensitive germination 1’ (AHG1), ‘ABA hypersensitive germination 3’ (AHG3), AT5G59220, AT2G29380 and AT1G07430. At least six members of these (ABI1, ABI2, HAB1, HAB2, AHG1 and AHG3) have been identified to act as the negative regulators of ABA signalling pathway in Arabidopsis. The

negative regulation of ABA is organized by the PYR/PYL/RCAR family of receptors, which connects the ABA perception directly to the release of PP2C negative regulation system including sucrose non fermenting 1 kinase 2 (SnRK2; Ma *et al.* 2009; Park *et al.* 2009; explained in Section 1.7.4). PP2C regulation is also controlled by secondary messengers, such as phosphatidic acid, which binds directly to the PP2CAs (experimentally proven for ABI1), resulting in the decrease of phosphatase activity and binds to the plasma membrane and hence limiting the access to the nuclear factors (Li *et al.* 2009a). PP2CAs can also function as the central point which connects the different effectors of ABA perception involved in stress response. For example, ABI2 (PP2CA) interacts with a kinase from the SnRK3 family, named as the Salt Overly Sensitive (SOS) 2. The SOS2 plays an critical role in Arabidopsis salt tolerance by regulating the plasma membrane located Na^+/H^+ antiporter encoded by SOS1 (Ohta *et al.* 2003). It is possible that the ABA signalling through PYR/PYL/RCAR receptors can regulate the cation homeostasis through the SOS pathway. In the same way, interaction between the PP2C and the AKT2 K^+ channels might be involved in the connection of the ABA to control K^+ transport and membrane polarization under stress (Cherel *et al.* 2002).

Group A PP2Cs are redundantly functional at the molecular level, but they play different roles in different tissue and organs, as suggested by the tissue-specific expression patterns. The ABI1 (PP2CA from Arabidopsis) is expressed in tissues such as seeds and guard cells, whereas the AHG1 and AHG3 (ABA hypersensitive germination, Type of PP2CAs from Arabidopsis) are expressed mainly in seeds (Nishimura *et al.* 2007; Umezawa *et al.* 2009). The subcellular localization pattern of ABI1, AHG1 and AHG3 is nucleus (Umezawa *et al.* 2009), which support the results that ABI1 broadly regulates the different effects of ABA in tissues from seeds to guard cells (Leung *et al.* 1994). The AHG1 and AHG3 essentially regulate the functions in seeds, by mainly regulating the gene expression of ABA mediating signalling downstream genes in nucleus (Nishimura *et al.* 2007).

The mutations of abi1-1 and abi2-1 occur in the catalytic domain of PP2CA (which is involved in binding to PYR/PYL/RCAR), with a well conserved glycine being converted to aspartate. The mutants, *abi1-1* and *abi2-1* showed overall ABA insensitivity such unable to prevent seed germination or root growth in various tissues

and developmental stages, which suggest that PP2CAs act as the global regulators of ABA signalling. The same mutation in other members of PP2CA also induces strong ABA insensitivity (Yoshida *et al.* 2006).

PP2A is a holoenzyme composed of three subunits, i.e., subunit A, B and C. The catalytic subunit C complexes with the scaffolding subunit A to form an AC core enzyme, which binds to the regulatory subunit B. The Arabidopsis mutant *roots curling npa (rcn) 1* was affected in the scaffolding subunit A, showed reduction in levels of PP2A activity (Garbers *et al.* 1996). Particularly the *rcn1* shows ABA-insensitive stomatal response because of the lack of ABA active anion channels such as K^+ ion channels (Cutler *et al.* 2010). PP2A catalytic enzymes (PP2Ac) have five subunits (PP2A1-5) and the recessive mutation in the catalytic subunit 2 (PP2Ac-2) leads to an enhanced sensitivity to ABA in different processes. This suggest that the PP2Ac-2 act as the negative regulator of the ABA responses, whereas the RCN 1 seemed to play a positive role in ABA signalling regulation (Pernas *et al.* 2007).

1.6.5 Protein kinases involved in ABA signalling

The identification and characterization of PP2CAs indicate the importance of protein phosphorylation event as the fundamental mechanism by which the plants modulate the signal transduction events (Umezawa *et al.* 2010). The kinases which are involved in the ABA signalling include calcium independent, sucrose non-fermenting (SNF) 1 related kinase (SnRK) 2 enzyme and calcium dependent, SnRK3/CIPK and calcium dependent protein kinase (CDPK/CPK) enzymes. SNF1 related protein kinase 2 (SnRK2) was first noticed to be involved in the ABA signalling pathway in the case of wheat PKABA1, which was transcriptionally up regulated by the ABA, and in turn PKABA1 phosphorylates the transcription factor TaABF1 (ABA binding factor) and induced the suppression of GA-induced genes expressed in the cereal grains (Gómez-Cadenas *et al.* 1999). This was followed by the identification of an ABA activated serine-threonine protein kinase (AAPK), a member of the SnRK2 family from the fava beans, which was shown to play an important role in the ABA induced stomatal closure (Li *et al.* 2000). A dominant negative version of AAPK contributed to the stomata insensitivity to ABA induced closure, which was the result of eliminating the ABA activation of plasma membrane 'slow anion channels' (Li *et al.* 2000). Open Stomata 1

(OST1)/ SnRK2.6 was identified genetically in a screen using infrared thermography to identify plants having transpiration defects, was found as the Arabidopsis ortholog of fava beans AAPK (Mustilli *et al.* 2002). The disruption of AAPK and OST1 functions does not affect the regulation of the stomata by light and carbon dioxide, which suggest that these enzymes are involved specifically in ABA signalling (Cutler *et al.* 2010). Other ABA related kinases, SnRK2.2 and SnRK2.3 were found to be closely related to OST1 (Fujii *et al.* 2007).

SnRK2 is composed of a well conserved catalytic kinase domain and a comparatively diverse C-terminal domain (Yoshida *et al.* 2002). The C terminal domain is further divided in two subgroups, domain I and II (Yoshida *et al.* 2006). Domain I is found to be relatively similar in all the SnRK2 enzymes, whereas the domain II differs among the SnRK2 subgroups, particularly in the acidic patch region. The acidic patch is found to be rich in Asp in subclass II and III, whereas it is rich in Glu in the subclass I members (Kobayashi *et al.* 2004). Several studies have suggested that C-terminal region could be the regulatory domain of SnRK2 and has been shown experimentally that the deletion of domain II hinders the ABA dependent activation, which suggested that this domain is required for the ABA responsiveness (Yoshida *et al.* 2006). There are 10 SnRK2 members in Arabidopsis and rice, called as SnRK2.1-SnRK2.10 and they are categorised in three subclasses; I, II and III. SnRK2 Subclass I members are activated by osmotic stress but not by ABA, whereas subclass II and II members are activated by both osmotic and ABA stress. Subclass III members are strongly activated by ABA as compared to subclass II members (Cutler *et al*. 2010).

Double mutants such as *snrk2.2* and *snrk2.3* of Arabidopsis lacks only in ABA mediated stomatal movement, whereas the triple mutant (*snrk2.2, snrk2.3, snrk2.6*) lacks most ABA responses such as seed dormancy, root growth inhibition and reduced expression of ABA inducible genes (Fujii *et al.* 2007). This has proven that these three SnRK2s are global positive regulators of the ABA signalling (Nakashima *et al.* 2009b). Several membrane proteins have been isolated as SnRK2 substrates. One of these is a slow anion channel SLAC1, which plays a central role in guard cells (Negi *et al.* 2008). SLAC1 is activated by the phosphorylation by the SRK2E/OST1/SnRK2.6 which is in turn regulated by the PP2C (Lee *et al.* 2009). The other SnRK2 regulated protein is

KAT1, an inward-rectifying potassium channel. The KAT1 also plays an important role in the stomatal movement (Pilot *et al.* 2001).

SnRK subgroup 3 includes members such as SnRK3, calcium independent protein kinase (CIPK), which interact with the calcium binding proteins (SOS3, SCaBP, CBL). A few members of this family regulate the ABA signalling in a calcium dependent manner. The calcium binding protein ScaBP5/CBL1 of Arabidopsis together with its interacting protein kinase PKS3/CIPK15 function as the negative regulators of the seed germination and stomatal ABA response (Guo *et al.* 2002). The SCaBP/PKS3 might also perceive the temporary increase in the cytosolic calcium induced by ABA, which results in the suppression PKS3 repression on ABA signalling (Guo *et al.* 2002). CIPK3/PKS12 also plays a role of negative regulator in the case of ABA induced seed germination, whereas there is no loss of ABA induced stomatal closure in *cipk3* mutant. CIPK3 was shown to be regulating the cold and salt induced gene expression but not drought induced gene expression. *Cipk23* mutant has shown reduction in transpiration water loss and improved response to the ABA mediated stomatal closure and inhibition of stomatal opening, whereas the ABA induced seed germination remained unaffected (Cheong *et al.* 2007).

Another family of calcium dependent protein kinases, CDPK/CPK plays a positive role in the ABA signalling. These protein kinases are distinguished by the structural arrangement of the calmodulin-like regulatory domain located at the C-terminal (Hrabak *et al*. 2003). This type of CDPK/CPK has both the kinase and calcium sensor domains in a single polypeptide and hence could be directly activated by calcium. The CDPK protein family was first identified as ABA activated protein kinase from the constitutive expression of the CPK10/CDPK1 and CPK30/CDPK1a, which activated an ABA inducible promoter in maize leaf protoplast (Sheen 1998). The overexpression of the CPK32 resulted in the ABA hypersensitive inhibition of seed germination (Choi *et al*. 2005). These proteins were proven as the positive regulators of the ABA regulated stomatal aperture by the gene knockout mutation of *CPK3* and *CPK6.* These mutants have not shown any phenotypes in ABA response during seed germination and seedling growth (Mori *et al.* 2006).

1.6.6 *Cis*- acting transcription factors mediating ABA regulated gene expression

Analysis of the promoters for ABA inducible genes suggest that the ABA responsive gene expression requires multiple *cis* elements (Zhang *et al.* 2005). The biochemical and genetic studies were used to identify the regulatory elements responsible for mediating the ABA induced changes in the protein coding gene expression. The classes of regulatory sequences which act as conferring the ABA induction are G-BOX ABA Response Elements (ABARE), which are recognized by members of the basic leucine zipper (bZIP) transcription factor family. Additionally, many of ABA regulated genes contains the binding sites for transcription factors of the MYB and MYC families. (Cutler *et al.* 2010). All of these transcription factor families contain dozen of members, so to identify different factors which specifically regulate a particular gene requires genetic studies with mutants. Specific candidates have been chosen within each family on the basis of either in vitro binding to the conserved DNA sequences or their own ABA inducible expression (Cutler *et al.* 2010). There is a significant amount of literature on numerous individual transcription factors and their specifically regulated genes in the plants, and the works of Zhang *et al*. (2005), Cutler *et al*. (2010), Furihata *et al*. (2006), Lopez-Molina *et al*. (2000), Nambara *et al*. (2010), Bensmihen *et al*. (2002), Finkelstein *et al*. (2011) and Fujita *et al*. (2005), amongst others provide comprehensive reviews in this area.

1.7 Core signalling complex in ABA signalling

The signalling complex of ABA receptors (PYR/PYL/RCAR), group A protein phosphatase 2C (PP2CAs; negative regulator) and sucrose non fermenting 1 kinase 2 (SnRK2s; positive regulator) offers the double negative regulation system in ABA signalling (Umezawa *et al*. 2009). The double regulatory system is supported by the structural studies of PYR/PYL/RCAR (Melcher *et al.* 2009; Nishimura *et al.* 2009; Santiago *et al.* 2009b; Yin *et al.* 2009). The double negative regulatory system in ABA system is summarized below:

Under normal conditions

1. Under normal conditions (absence of ABA), PYR/PYL/RCARs does not interact with the PP2CAs, leaving PP2CAs in the active state.
2. PP2CAs negatively regulate SnRK2s by direct interactions and de-phosphorylation. This inactivates the SnRK2s
3. The dephosphorylated, inactive SnRK2s are unable to activate the downstream factors, such as the AREB/ABF bZIP-type transcription factor which in turn, fails to facilitate transcription of ABA responsive genes (Figure 1.12).

Under environmental stress

1. Once environmental cues or stresses induce ABA, the ABA bound PYR/PYL/RCARs interact with PP2CAs and inhibit its phosphatase activity.
2. The inactive PP2CAs fails to dephosphorylate the SnRK2s, which in turn remains active.
3. The active SnRK2s activate downstream factors by phsophorylation, such as the AREB/ABF bZIP-type transcription factor and which in turn facilitate transcription of ABA responsive genes (Figure 1.12).

The hypothesis of double negative regulation of ABA signalling by receptors, PP2CAs and SnRK2s is supported by the observation that SnRK2 activities are significantly reduced in the PYR/PYL quadruple mutant line of Arabidopsis (*pyr1;pyl1;pyl2;pyl4*; Park *et al.* 2009). This model has been classified as the core component system in ABA signalling, as all components regulate global response in plants (Umezawa et al. 2009). PYR/PYL/RCAR, PP2C and SnRK2 form a signalling complex is also referred to as the 'ABA signalosome' (Umezawa *et al*. 2010). The structural basis of PYR/PYL/RCAR-PP2CA-SnRK2 signalling complex is explained below:

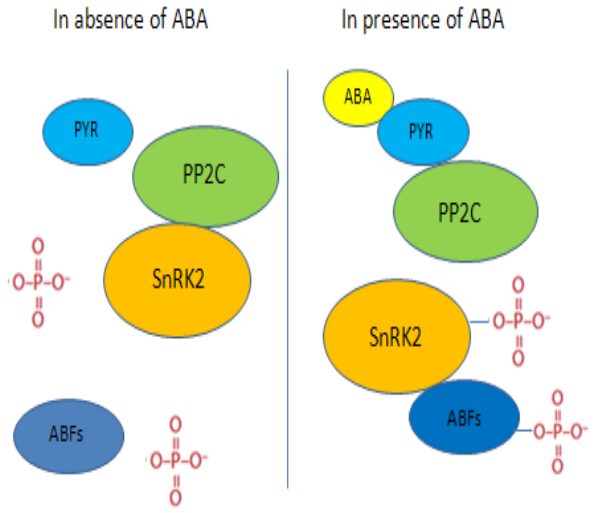


Figure 1.12 Summary of abscisic acid (ABA) signalling factors
PYR/PYL/RCAR, PP2C and SnRK2 form a signalling complex referred to as the 'ABA signalosome'. (A) Under normal conditions, PP2C negatively regulates SnRK2 by direct interactions and dephosphorylation of multiple residues of SnRK2. Once abiotic stresses or developmental cues up-regulate endogenous ABA, PYR/PYL/RCAR binds ABA and interacts with PP2C to inhibit protein phosphatase activity. In turn, SnRK2 is released from PP2C-dependent regulation and activated to phosphorylate downstream factors, such as the AREB/ABF bZIP-type transcription factor or membrane proteins involving ion channels.

1.7.1 Structural basis of ABA perception by PYR/PYL/RCAR

The binding of the ABA to the PYR/PYL/RCAR induces a conformational change in the two flanking loops of the ABA binding pocket, called as the gate and latch. The PYR/PYL/RCAR protein possesses an open confirmation of both loops. Binding of the ABA molecule to PYR/PYL/RCAR induces closure of the gate and latch loops. This consequently prepares the interaction surface and docking of PP2CAs onto the ABA bound PYR/PYL/RCAR receptor. The ABA perception by PYR/PYL/RCAR is summarized below (reviewed in Umezawa *et al.* 2010):

1. The PYR/PYL/RCAR has large, water filled internal cavity. ABA enters the cavity with the carboxyl group oriented towards the centre of the cavity. This structure explains the stereo-selectivity of the ABA.
2. The carboxyl group of ABA forms ionic bond with the side chain of lysine residue (PYL1K86, PYL2K64 and PYR1K59). The carboxyl group of ABA also form hydrogen bond network with the side chains of five polar residues (P 88, E 94, R 116, S 122, E 141 of Arabidopsis PYR1) through water molecules (Figure 1.13A).
3. Gate loop of PYR/PYL/RCAR connects the β strand 3 and β strand 4. The binding of ABA results in the proline residue (gate loop) (PYL1 P115, PYL P292 and

PYR1 P88) to move towards the cyclohexene (ABA) to close the gate on the cavity. The serine residue (gate loop) (PYL1 S112, PYL2 S89 and PYR1 S85) is flipped out of the cavity.

4. The latch loop connects the β strand 5 and β Strand 6. The imidazole ring of the histidine (latch loop; PYl1 H142, PYL2 H 119, PYR1 H115) moves into the cavity to form Van Der Waals bond with the cyclohexene of ABA, which induces conformational changes. The latch loop locks the closed gate loop by a hydrogen bond and Van Der Waals bond (Figure 1.13B). These loops of PYR/PYL/RCAR in the closed conformation then provide the surface for interaction with group A PP2Cs.

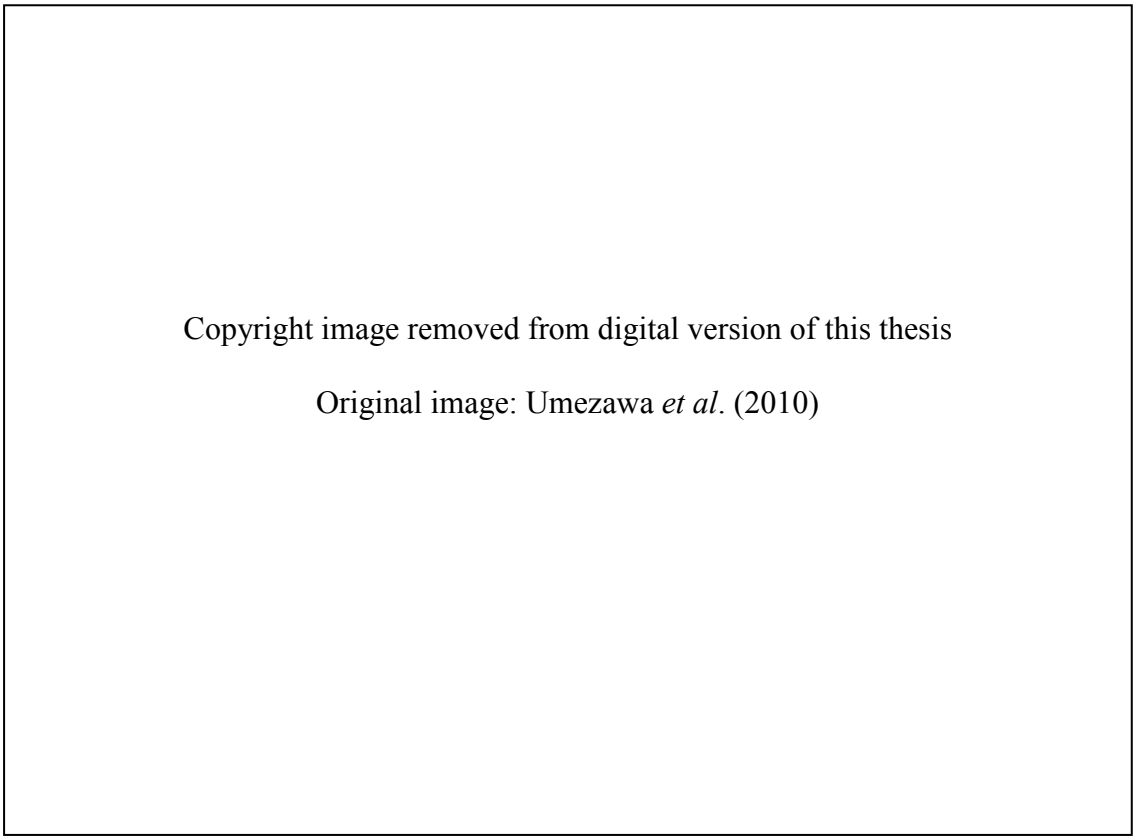


Figure 1.13 Structural analysis of PYR/PYL/RCAR receptor of ABA (Source Umezawa *et al*. 2010)

(A) Stereoselective ABA-binding mode of the PYR/PYL/RCAR proteins. A lysine residue (green) directly interacts with ABA, and polar residues (yellow) form a water-mediated hydrogen bond network with ABA. Water molecules and hydrogen bonds are shown by cyan spheres and dashed lines, respectively. Hydrophobic residues (blue) are localized around dimethyl and monomethyl groups of the cyclohexene moiety. For the structural formula of ABA in the inset, the cyclohexene moiety, pentadienoic acid moiety and hydroxyl group from the chiral carbon (shown by an asterisk) are colored blue, red and green, respectively.

1.7.2 ABA induced inhibition mechanism of PP2CAs by PYR/PYL/RCARs

A conserved proline of the gate forms a hydrogen bond with the single tryptophan present at the docking site of the PP2CA (Melcher *et al.* 2009; Yin *et al.* 2009). The binding of PP2CA to the closed form of the receptor is expected to lower the dissociation of ABA from PYR/PYL/RCAR, which in turn lowers the Kd. Structural studies provide another hypothesis, i.e., conversion of the PYR protein from dimeric to monomeric form in response to ABA binding. This appears as an important link in the signalling pathway, as the final PYR-PP2C complex formed is a heterodimer. The protoplast assay studies have proven that the entire PYR/PYL/RCAR family (except PYL13) is capable of activating this signalling mechanism in response to ABA (Fujii *et al.* 2009). The interaction of PYR/PYL/RCAR and PP2CAs is summarized below:

1. The serine residue of PYR/PYL/RCAR (PYL1 S112 and PYL2 S89) gets exposed on ABA binding (explained above). The serine residue forms hydrogen bond between its side chain and conserved catalytic residue glutamic acid of PP2CA (ABI1 E142 and HAB1 E203). The serine residue (PYR) also forms hydrogen bond with the glycine residue (ABI1 G180 and HAB1 G246) on the active site loop of PP2CA.
2. A tryptophan residue (ABI1 W300 and HAB1 W385) inserts its indole ring in between the closed gate and latch loops. This indole imine group (PP2CA) forms a hydrogen bond with the both loops (PYR) and the carbonyl group (ABA).
3. The gate loop is further locked down by the interaction of guanidinium group of arginine residue of the PP2CA (ABI1 R304 and HAB1 R389) with the conserved proline residue on the gate loop (PYL1 P115 and PYL2 P92) (Figure 1.13C).

1.7.3 PP2CA-SnRK2 complex (in absence of ABA)

PP2CAs has been proven to directly inactivate and dephosphorylates SnRK2s. SnRK2s phosphorylation sites have been determined using liquid chromatography-tandem Mass spectrtometry (LC-MS/MS; Umezawa *et al.* 2009). In response to ABA, SnRK2s are activated by phosphorylation of serine residue (SnRK2.6S175 of Arabidopsis) in its kinase activation loop. The same site is dephosphorylated by members of group A PP2C, resulting in the inactivation of SnRK2s. This suggests that SnRK2s are direct target of PP2CAs, providing clear insight into the PP2CA-dependent negative regulation of ABA signalling.

1.7.4 SnRK2 mediated transcription factors (in presence of ABA)

Analysis of the promoters for ABA inducible genes suggest that the ABA responsive gene expression requires multiple *cis* elements designated as ABA-responsive elements (ABRE) (Zhang *et al.* 2005). The classes of regulatory sequences which act as conferring the ABA induction are G-box ABA response elements (ABARE), which are recognized by members of the basic leucine zipper (bZIP) transcription factor family. Additionally, many of ABA regulated genes contains the binding sites for transcription factors of the MYB and MYC transcription factor families. (Cutler *et al.* 2010). All of these transcription factor families contain dozen of members, so to identify different factors which specifically regulate a particular gene requires genetic studies with mutants.

ABRE transcription factors require post translational modifications for their activation. Several studies have that ABRE are phosphorylated in response to ABA and this phosphorylation activates them for downstream action (Furihata *et al*. 2006). SnRK2s (SnRK2.2, SnRK2.3 and SnRK2.6 of Arabidopsis) has been proven to phosphorylate AREB transcription factors *in vitro*. Bimolecular fluorescence complementation (PiFC) has shown the co-localization and interaction of SnRK2.2, SnRK2.3 and SnRK2.6 with AREB in the Arabidopsis cell nucleui (Furihata *et al*. 2006).

1.8 ABA mediated abiotic stress signalling candidate genes in cereals

SnRK2 was the first ABA signalling related gene to be detected in cereals (wheat). There has been significant progress since in the characterisation of ABA signalling related genes in cereals (Table 1.3). Analysis of the MSU rice genomic annotation project database indicates 14 PYR/PYL/RCAR, 78 PP2C and 10 SnRK2 genes (Umezawa *et al*. 2010; Xue *et al*. 2008). Maize has 20 PYR/PYL/RCAR (Klingler *et al*. 2010), 68 PP2Cs (Zheng *et al*. 2010) and 11 SnRK2s (Huai *et al*. 2008). Seiler *et al*. (2014) showed nine PYR/PYL/RCAR, six PP2C and nine SnRK2 in barley. However, genome wide picture of ABA signalling related genes expressed in barley is still lacking and judging from the total number in rice, it is expected that there are more barley genes yet to be identified.

The entire gene families of ABA signalling related genes from wheat have not been characterised as yet. This may be due to the lack of its complete genomic sequence and the large allohexaploid genome, which makes gene analysis difficult. The first report of wheat SnRK2 was by Walker-Simmons *et al*. (1992) who demonstrated that SnRK2 are involved in ABA signalling pathway. So far three SnRK2s (Tian *et al*. 2013), 16 protein phosphatase (Lv *et al*. 2014) have been identified in wheat. However none of the PYR/PYL/RCAR have been characterised in wheat.

Table 1.3 ABA mediated abiotic stress signalling related genes in cereals.

Cereal	PYR/PYL/RCAR	PP2C	SnRK2	Reference
Rice	14	78	10	Umezawa *et al*. 2010; Xue *et al*. 2008
Maize	20	68	11	Klingler *et al*. 2010; Zheng *et al*. 2010; Huai *et al*. 2008
Barley	9	6	9	Seiler *et al*. 2014
Wheat	-	16	3	Tian *et al*. 2013; Lv *et al*. 2014

1.9 Responses of plant ABA mediated abiotic stress signalling related genes to abiotic stresses

Literature has demonstrated that ABA-signalling related candidate genes are regulated in response to abscisic acid (ABA) or environmental stimuli such as low temperature, drought, salinity and light (Xue *et al*. 2008). The present study focuses mainly on barley ABA signalling related genes; therefore, more emphasis will be placed on the responses of these genes to abiotic stresses that limit growth plants including wheat and barley, such as salinity, drought (detailed earlier in section 1.1.6).

Studies on the effect of abiotic stresses on the expression of ABA-signalling related candidate genes has yielded different responses ranging from down-regulation, up-regulation and no notable change in response to various stresses such as salinity, drought and nutrient deficiency (Xue *et al*. 2008; Seiler *et al*. 2014). The gene expression of PP2CAs during abiotic stress was shown to depend on the time course, intensity of the stress and specific developmental stages and tissue of the plant (Xue *et al*. 2008). In a study by Xue *et al*. (2008) of the PP2CAs in Arabidopsis, exposed to exogenous ABA (100μM), drought (150mM mannitol) and salt stress (150mM NaCl) resulted in the up-regulation of the genes except *AtPP2C7* (AT1G17550) under salt

stress condition, which showed no expression. An analysis of exogenous ABA and salt stress on the rice PP2CAs revealed up-regulation. In addition, OsPP2C49, OsPP2C09, OsPP2C68, OsPP2C06, OsPP2C50, OsPP2C08, OsPP2C51 and OsPP2C30 also showed up-regulation under drought stress. On the other hand gene OsPP2C53 showed no change in expression under drought stress, whereas another PP2C, OsPP2C37 showed no expression under drought stress (Xue *et al*. 2008). The expression of HvPP2C6, HvPP2C8 and HvPP2C9 was found to be up-regulated in the drought stress condition (Seiler *et al.* 2014).

Arabidopsis PYR/PYL/RCARs, *AtPYR1* (AT4G17870), *AtPYL1* (AT5G46790), *AtPYL2* (AT2G26040), *AtPYL8* (AT5G53160), *AtPYL5* (AT5G05440), *AtPYL6* (AT2G40330) and *AtPYL4* (AT2G38310) showed down regulation under salt stress condition (150mM NaCl). On the other hand three PYR/PYL/RCARs, *AtPYL3* (AT1G73000), *AtPYL9* (AT1G01360) and *AtPYL7* (AT4G01026) showed no change in expression, whereas no data was available for AT5G4587, AT5G45860, AT4G18620, and AT4G27920 (Winter *et al*. 2007). Barley PYR/PYL/RCAR, HvPYL3, HvPYL4, HvPYL5, HvPYL6, HvPYL7 and HvPYL8 were found to be down regulated under drought stress, whereas other PYR/PYL/RCAR has not shown any change in expression in the barley leaf (Seiler *et al*. 2014).

Of the SnRK2s in Arabidopsis, *AtSnRK2.6* (AT4G33950) and *AtSnRK2.7* (AT4G40010) showed up-regulation under salt stress condition (150 mM NaCl), whereas *AtSnRK2.1* (AT1G78290) showed a down-regulation. On the other hand *AtSnRK2.2* (AT3G50500) and *AtSnRK2.3* (AT5G66880) showed no change in expression under the same conditions (Winter *et al*. 2007). All of the SnRK2 genes in barley showed up-regulation under drought stress condition (Seiler *et al*. 2014).

It is thus evident that ABA signalling related genes are of significance in responding to abiotic stresses. However, while many studies have been conducted on PP2Cs, there is scant literature on PYR/PYL/RCARs. Further, despite the immense potential for crop improvement locked in ABA-signalling related genes, most studies using transgenics have been carried out for plants with fully-sequenced genomes such as Arabidopsis and rice, while wheat and barley (fully-sequenced genomes released in 2012) lag behind.

Therefore, there is a need to intensify the research on characterisation of ABA signalling related genes in these crops including details such as roles of specific isoforms, tissue-specific expression patterns and nature of responses (up or down regulations and associated effects). Further, abiotic stresses rarely occur in isolation in nature; thus study of these genes capable of responding to multiple stresses is essential for genetic screening or transgenics.

1.10 Gene expression profiling

Gene expression study is a fundamental step in screening for candidate genes to be employed in the development of tolerant crops in response to different environmental stresses. Over the years, transcript profiling has evolved from the use of hybridization-based methods such as microarrays to the high through put sequencing methods also called as next-generation sequencing (NGS) technology. Micro-arrays have been employed in different species such as rice (Walia *et al* 2005), barley (Ueda *et al*. 2004; Walia *et al*. 2007) and wheat (Kawaura *et al*. 2008). Despite their widespread use and significant results, microarray platforms have some limitations, such as limitations in the detection of genes with very low or high expression, or the specificity being limited by probe cross-hybridization, leading to problems in distinguishing isoforms (Kane *et al*. 2000). Lastly, microarrays generally require some knowledge of the target gene to be hybridized, thus limiting their use to annotated genes. Thus recent studies on global gene expression are moving towards mRNA-sequencing (mRNA-seq) by high throughput sequencing (next-generation) technologies.

mRNA-seq is a process wherein mRNA is reverse-transcribed to cDNA and converted into a library of short segments and sequenced. The first step in mRNA-seq is the isolation of total RNA from the plant samples of interest and then purification of mRNA from the total RNA using the poly-A tail (Figure 1.14). The mRNA is fragmented, then reverse-transcribed into cDNA library (200-300 nucleotide long). The cDNA is then ligated onto adapters, which is then loaded onto a flow-cell of high throughput sequencing machine and sequenced producing millions of short sequenced reads ranging in size from 36 to 400 nucleotides depending on the platform used (Martínez-Gómez *et al*. 2011). Subsequently, the short sequence reads are mapped to a reference genome of that species (if possible) using programs such as Burrows-Wheeler Aligner

(BWA) (Li and Durbin 2009) or if such a genome is lacking, the reads are assembled into contiguous sequences (contigs) using assembly programs such as ABySS (Simpson *et al*. 2009) prior to aligning them to the closest possible reference sequences (Wang *et al*. 2009a). The next phase in mRNA-seq is the quantification of differences in expression of transcripts between samples using software packages such as DESeq software (Anders and Huber 2010).

The platforms most widely used for mRNA-seq are the Illumina Genome Analysers (initially developed by Solexa), Roche/454 platform, Oxford nanopore, Pacific biosciences, Ion torrent and SOLiD platform from Life Technologies/Applied Biosystems and Nextseq 50, Miseq and Hiseq 2500 from the Illumina. When the first NGS platform was introduced in 2005 (Margulies *et al*. 2005), its application was limited to model plants with fully-sequenced genomes. However, recently the technology has been successfully employed for transcriptome sequencing of plants without such genomes. The surge of mRNA-seq is motivated by its advantages such as (i) greater sensitivity, making it possible to detect genes with low expression; (ii) cost-effectiveness; (iii) ability to profile genes without prior knowledge of sequence, allowing identification of novel transcripts; (iv) improved ability to distinguish between highly similar iso-forms (Wang *et al*. 2009a). In summary, mRNA-seq opens 'new playgrounds' for gene expression and epigenetic (explained below) analyses in plants lacking fully sequenced genomes (Bräutigam and Gowik 2010).

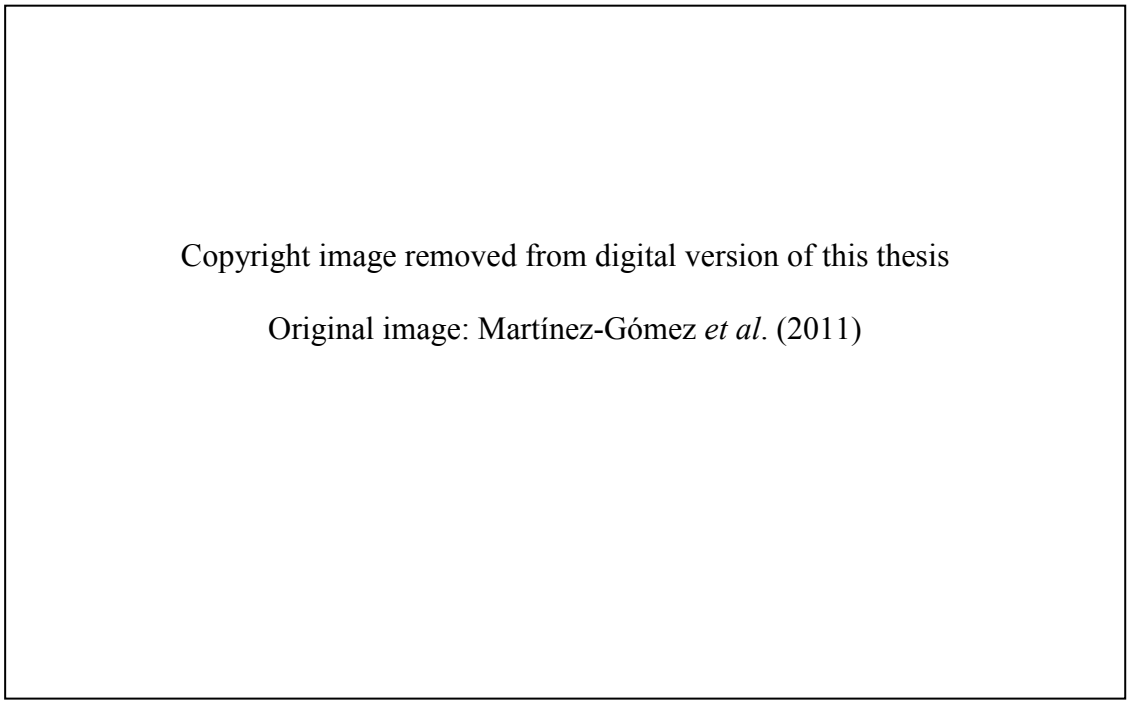


Figure 1.14 Summarised workflow of mRNA-seq using Illumina platform (Source Martínez-Gómez *et al*. 2011)

1.11 Epigenetic studies

Gene transcription is a complicated process that is regulated by both genetic and epigenetic factors. Epigenetic factors, such as histone modifications and DNA methylation, play a significant role in regulating transcription (Bernstein *et al*. 2007). With the widespread influence of these epigenetic factors on gene expression, it is expected to disrupt epigenetic factors at the time of stress or abiotic stress conditions (Young *et al*. 2011). To understand the role and influence of epigenetic factors is a central issue in understanding transcriptional regulation under abiotic stress conditions in plants.

1.11.1 Nucleosome structure and histone modifications

Since the discovery of nucleosome in 1974, several years of biochemical and biophysical studies were carried out to reveal crystal structure of nucleosome in 1997 (Luger *et al*. 1997). The nucleosome core consists of an octomer of small basic histone protein wrapped within 147 bp of DNA. The histone octomer has histone proteins dimerized through "histone fold" domains, which has three α-helices. Two H3:H4 dimers form central histone tetramer which is flanked on both sides by dimers of H2A and H2B. Each histone also contains N-terminal extension, called as "histone tails" and H2A and H2B shorter C-terminal extensions. This histone tails are mostly disordered in

the crystal structures are the sites of many post-translational modifications and contribute to the higher order of chromatin structures (Dutnall and Ramakrishnan 1997; Luger and Richmond 1998; Fletcher and Hansen 1996).

Posttranslational modifications of histone are known since 1960s (Allfrey *et al*. 1964), however their role in epigenetic control over nuclear processes began to be revealed in 1990s (Suganuma and Workman 2011). The identification of the first nuclear histone acetyltransferase and deacetylases, as transcriptional co-activator and co-repressor respectively, was a breakthrough in this regard (Brownell *et al*. 1996; Taunton *et al*. 1996). Since then histones have been found to be acetylated, methylated on lysine and arginine, ubiquitinated, 'sumolated', phosphorylated and ribosylated and also to undergo proline isomerization (reviewed in Suganurna and Workman 2011). Histone modifications may play a direct role in affecting chromatin structure or they may represent marks to be recognized by protein effectors. Some of the known histone modifications are described below:

Histone acetylation

Acetylation of histone lysine is regulated by opposing action of two families of enzymes, histone acetyltransferase (HAT) and histone deacetylases (HDAC; Xhemalce *et al*. 2011). The HAT utilizes acetyl CoA as co-factor and catalyses the transfer of an acetyl group to the ε-amino group of lysine side chains. HDAC enzymes oppose the effects of HAT and reverse lysine acetylation, which restores the positive charge of lysine. This stabilises the chromatin structure and is consistent with HDACs being transcriptional repressors.

Histone phosphorylation

Histone phosphorylation of histone takes place on serines, threonines and tyrosines predominantly but not exclusively in the N-terminal histone tails. The level of modifications is controlled by kinases and phosphatases that add or remove the modifications respectively (Oki *et al*. 2007). Most of the histone phosphorylation exists within the N-terminal tails. However, phosphorylation sites do occur in the core region as well for example H3Y41, which is deposited by the non-receptor tyrosine kinase JAK2 (Dawson *et al*. 2009).

Histone methylation

Histone methylation mainly occurs on the side chains of lysines and arginines. Unlike acetylation and phosphorylation, histone methylation does not alter the charge of the histone proteins. The added complexity for histone methylation is lysines may be mono-, di- or tri-methylated, whereas arginines may be mono-, symmetrically or asymmetrically di-methylated (Xhemalce *et al*. 2011; Ng *et al*. 2009; Lan and Shi 2009).

Several histone lysine methyltransferase (HKMT) have been identified since the discovery of SUV391 as HKMT, which targets H3K9 (Rea *et al*. 2000). The vast majority of these enzymes methylate lysine within the N-terminal tails. HKMT are specific enzymes as *Neurospora crassa* DIM5 specifically methylates H3K9 whereas SET7/9 methylates H3K4. HKMT enzymes also methylate the appropriate lysine to a specific degree i.e. mono-, di- and /or tri-methyl state. For example DIM5can trimethylate H3K9, whereas SET7/9 cam only mono-methylate H3K4.

The other histone modifications include deamination, β-N-acetylglucosamine, ADP ribosylation, Ubiquitylation and sumoylation, histone tail clipping and histone proline isomerization.

1.11.2 Modes of effect of histone modifications on gene regulation

Histone modifications exert their effect by two main mechanisms, (i) by influencing the overall structure of chromatin directly, (ii) by regulating the binding of effector molecules. However histone modifications are also essential for the DNA processes such as repair, replication and recombination.

Direct structural perturbation: Histone acetylation and phosphorylation reduce the positive charge of histones and lead to disruption of the electrostatic interaction between histones and DNA. This could lead to less compact chromatin structure, thereby facilitating DNA access by protein machineries. Acetylation occurs at number of histone tail lysines such as H3K9, H3K14, H3K18, H4K5, H4K8 and H4K12 (Kouzarides 2007; Figure 1.15). The large number of putative sites, indicate that in hyper actylated regions of genome, the histone tail charge can be effectively neutralized. This could have a great effect on chromatin structure (Kiefer *et al*. 2008). Multiple

histone acetylations are enriched at enhancer elements and promoters, where they are expected to facilitate transcription factor access (Wang *et al*. 2008).

Copyright image removed from digital version of this thesis

Original image: Bannister and Kouzarides (2011)

Figure 1.15 Major post-translational modifications (Source Bannister and Kouzarides 2011)
These post-translational modifications play an essential role in regulation o gene expression

On the other hand histone phosphorylation is site specific and there are fewer sites as compared to acetylated sites. H3S10 phosphorylation (H3S10ph) occurs genome wide in humans during mitosis and is associated with the genome becoming condensed (Wei *et al*. 1998). This seems contrary as the phosphate group adds negative charge to the histone tail, which is close to the negatively charged DNA backbone. The displacement of hetrochromatin protein (HP1) from hetrochromatin during metaphase by uniformly high levels of H3S10ph is essential for the detachment of chromosomes from interphase scaffolding, which in turn facilities chromosome remodelling essential for attachment of spindle fibres.

Regulating the binding of chromatin factors: Number of chromatin associated factors have been found to specifically interact with modified histones (Vermeulen *et al*. 2010; Bartke *et al*. 2010). There are multivalent proteins with specific domains which simultaneously recognize several modifications and other nucleosomal features (Bannister and Kouzaride 2011).

There distinct domain types recognizing lysine methylation have been found more than any other modification, reflecting the modifications relative importance (Bannister and Kouzaride 2011)., e.g., plant homeo domain (PHD) fingers and malignant brain tumor (MBT) domain. Among these domains, numerous domains can recognize same

methylated histone lysine. For example, H3K4me3, a mark associated with active transcription, is recognized by PHD finger within the ING family of proteins (ING1-5), bound by the tandem chromodomains within CHD1 and also by the tandem tudor domains within JMJD2A, a histone demthylase (Huang *et al*. 2006).

1.12 Summary of above literature and current research questions

The literature establishes that salinity and drought are the two most important environmental stress which limit plant growth and development. For plants to survive under such conditions, they must be able to sense and respond. Abscisic acid is a major phytohormone that regulates a wide range of plant traits and is especially important for adaptation to environmental conditions such as drought and salinity. The structural studies of PYR/PYL/RCAR support the double regulatory system of ABA signalling, which consists of four stages: ABA receptors (PYR/PYL/RCAR), protein phosphatases (PP2C), protein kinases (SnRK2) and their downstream targets. These factors make them crucial in fundamental life processes of plants such as photosynthesis, leaf abs*cis*sion, stomatal closure and fruit ripening. However, despite this significance, limited research has occurred on ABA-signalling related genes in cereals such as wheat and barley which are amongst the most important food sources for humans. Hence, there is need to further characterise the ABA signalling related genes from barley. Such research will allow the identification of potential genetic resources for crop improvement.

1.13 Aims and objectives

This will be addressed through the following specific aims:

1. To analyse the transcriptome of barley (*Hordeum vulgare* cv. Hindmarsh) leaf under control and salt, drought and exogenous ABA stresses using mRNA-seq.
2. To develop bioinformatics based tools to identify specific gene families of interest in the mRNA-seq data or any other database using orthologues from related species and extract the ABA mediated abiotic stress signalling related genes.
3. To investigate the expression modulation of the genes involved in ABA mediated signalling pathway under the above abiotic stresses.

4. To investigate the epigenetic mechanism of histone modifications in regulating ABA mediated abiotic stress signalling genes.
5. To explore salinity tolerance of selected barley varieties employing analyses of key physiological parameters.
6. To test the expression of selected genes involved in ABA mediated abiotic stress signalling in above varieties.

1.14 Strategy of study

The molecular study of ABA mediated abiotic stress signalling related genes will be undertaken with a range of analysis including:

(i) Transcriptome analysis of barley (*Hordeum vulgare* cv. Hindmarsh) leaf under control and salt, drought and exogenous ABA stresses using next-generation mRNA-seq technology.

(ii) Extraction of ABA mediated abiotic stress signalling related genes from mRNA-seq data and other databases using self-developed bioinformatics based tools.

(iii) Determining the expression modulation of the genes involved in ABA mediated signalling pathway under the above abiotic stresses using above mRNA-seq data. This investigation will be useful in determining the candidates for future functional testing and development of tolerant plants through genetic modification or breeding.

(iv) Investigating the distribution of one epigenetic modification, the tri-methylation of lysine 4 of histone H3 (H3K4me3) that is involved in gene activation, in barley during abiotic stress using ChIP-seq. This would provide a fundamental step in understanding the transcriptional regulation of these genes, which can be employed for development of tolerant crops.

(v) Determining the salinity tolerance of selected barley varieties employing analyses of key physiological parameters including relative water content, levels of the stress hormone abscisic acid (ABA) and Na^+/K^+ ratio. Further, analysing the differential expression of key genes in the ABA-regulated stress response pathway by quantitative real-time PCR of leaf RNA. This would help in determining, if expression analysis of the ABA pathway can be used for rapid identification of potentially salt-tolerant barley varieties before undertaking physiological studies.

Chapter 2

Materials and methods

Materials

2.1 Equipment and material

The instruments and apparatus used to carry out experiments are listed in Table 2.1 and the commercial kits and solutions used are listed in Table 2.2

Table 2.1 Equipment used in research

Equipment	Manufacturer	Purpose
MyCycler™	Bio-Rad, California, USA	PCR, Gradient PCR
Master cycler	Eppendorf, Hambug, Germany	PCR
MyiQ™ single-color real-time PCR detection system	Bio-Rad, California, USA	Real-time PCR
Genome Analyzer IIx	Illumina	Gene expression profiling
Electrophoresis power supply– EPS 301 Minnie Gel Unit	General Electric (GE), Buckinghamshire, UK (previously Amersham Biosciences)	Separation of nucleic acids
GeneQuant™ pro UV/Vis spectrophotometer	General Electric (GE), Buckinghamshire, UK (previously Amersham Biosciences)	RNA and DNA quantification
MCE®-202 MultiNA, microchip electrophoresis system	Shimadzu, Japan	RNA quantification for mRNA-seq
Plant growth cabinet	Thermoline, Coburg North, VIC, Australia	Growth of barley plants
Chemidoc XRS documentation station	Bio-Rad, California, USA	Visualisation and analysis of gel images
UV light transilluminator	UVP, USA	Visualisation of gels for gel purification
C3040 digital camera	Olympus, Tokyo, Japan	Capturing of UV exposed gel images
Finnpipette (0.5-10, 5-50, 20-200, and 100-1000 μL)	Thermo Electron, Madison, USA	Dispensing liquids

Table 2.2 Commercial kits and solutions used in study

Kit/Solution	Supplier	Purpose
Wizard® Genomic DNA Purification Kit, containing nuclei lysis, protein precipitation and DNA rehydration solutions	Promega, Madison, USA	Genomic DNA purification
RNase A (from bovine pancreas)	Sigma-Aldrich, St. Louis, USA	RNA digestion during genomic DNA purification
Wizard® Plus SV Minipreps DNA Purification System, containing cell resuspension solution, cell lysis solution, neutralization solution, column wash solution, alkaline	Promega, Madison, USA	Plasmid DNA purification
pGEM®-T Easy Vector System, containing pGEM®-T Easy Vector*, T4 DNA ligase,		Molecular Cloning
RQ1 RNase-free DNase I	Promega, Madison, USA	DNA digestion during RNA purification
RQ1 DNase 10× Reaction Buffer		
Perfectprep® Gel Cleanup Kit, containing binding buffer, wash buffer, elution buffer,	Eppendorf, Hamburg, Germany	Purification DNA from agarose gels
TRIsure™	Bioline, London, UK	Total RNA isolation
RNase Inhibitor		Inhibition of RNase activity

Bioscript™ Moloney Murine Leukaemia Virus Reverse Transcriptase		Reverse transcription
Biomix (2×)		PCR
dNTP set		cDNA synthesis
SensiFAST™ SYBR & Fluorescein Kit		Real-time PCR
Hyperladder™ I and V		Molecular weight markers for agarose gel electrophoresis
GeneRuler™ 50bp DNA ladder; GeneRuler™	Fermentas, Waltham, USA	
BDT (BigDye® Terminator) v3.1 Ready Mix	Applied Biosystems, Australia	DNA sequencing
Ethidium bromide (10 mg/mL)	Sigma-Aldrich, St. Louis, USA	Staining of agarose gels

*See Appendix I for pGEM®-T Easy Vector

2.2 Prepared solutions and materials

2.2.1 Sterilisation

Solutions were sterilised by autoclaving (121°C for 20 min), or filter sterilised through a 0.22 µM syringe filter (Millipore). All glassware and disposable plastic were also autoclaved as above. The sterilisation methods used in the preparation of buffers, media and solutions listed below are indicated with autoclaved or filter sterilised.

2.2.2 Buffers and solutions

All buffers and solutions in Table 2.3 and Table 2.4 were prepared using Milli-Q water (Millipore). The general use buffers and solutions listed below were prepared according to the instructions in Sambrook and Russell (2001).

Table 2.3 Composition of general buffers and solutions

Buffer/solution	Composition	Sterilization method
Agarose gel electrophoresis loading dye, 6×	30% (v/v) glycerol, 1 mg/mL xylene cyanol, 1 mg/mL bromophenol blue	autoclaved
Glacial acetic acid	0.01% (w/v)	autoclaved
TAE (Tris-acetate EDTA) buffer, 50×	M Tris base, 6.5 M EDTA disodium salt, pH 8.0	autoclaved
TB buffer	(10 mM Hepes, 15 mM $CaCl_2$, 250 mM KCl, pH 6.7)*, then solid $MnCl_2$ later added to final concentration of 55 mM	autoclaved

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

Table 2.4 Composition of buffers and solutions for sequencing

Buffer/solution	Composition	Sterilization method
BDT reaction buffer, 5×	400 mM Tris pH 9.0, 10 mM $MgCl_2$	autoclaved
$MgSO_4$ stock solution	0.2 mM in 70% ethanol	autoclaved

The following solution used for plant growth was prepared according to Hoagland and Arnon (1950).

Hoagland's solution (filter sterilised): 7 mM $Ca(NO_3)_2.4H_2O$, 5 mM KNO_3, 2 mM KH_2PO_4, 2 mM $MgSO_4.7H_2O$, 45 µM H_3BO_3, 9 µM $MnCl_2.4H_2O$, 0.7 µM $ZnSO_4.7H_2O$, 0.32 µM $CuSO_4.5H_2O$, 0.12 µM $NaMoO_4$, 28 µM FeEDTA in 1 M KOH.

2.2.3 DEPC treatment of solutions and materials to be used for RNA extractions

Only RNase-free certified disposable plasticware was used. Non-disposable plasticware, glassware and metal utensils were autoclaved twice. Electrophoresis tanks, gel casting trays and combs were soaked in 3% (v/v) H_2O_2 for 10 minutes, then thoroughly washed with RNase-free dH_2O (Sambrook and Russell 2001). Water to be used in for dilution was treated with diethylpyrocarbonate (DEPC). DEPC (1% (v/v) in dH_2O) was incubated at 37ºC overnight to inactivate RNase, and autoclaved to remove DEPC (Sambrook and Russell 2001). Milli-Q dH_2O was used for rinsing all equipment.

2.2.4 Media and solutions for microbial growth

The media and solutions used for culturing bacteria were prepared according to Sambrook and Russell (2001) and listed in Table 2.5.

Table 2.5 Solutions and media used for culturing bacteria

Solutions and media	Composition	Sterilization method
Ampicillin	20 mg/mL	Filter sterilization
IPTG (isopropyl-β-D-thiogalactopyranoside)	0.1 M	Filter sterilization
Luria broth (LB)	10 g/L tryptone, 5 g/L yeast extract, 5 g NaCl, 15 g/L agar (for plates only)	autoclaved
Nutrient agar (NA), Oxoid	23 g/L (for plates)	autoclaved
Super Optimal broth (SOB medium)	0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$	autoclaved
X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside)	5% (w/v) in dimethylformamide	filter sterilized
2×YT	16 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl, pH 7.0	autoclaved

2.2.5 Microbial strains

Escherichia coli JM109 (Promega), was used for general molecular cloning procedures. The 'competent' cells were used for transformation of recombinant DNA.

2.2.6 Plant material

Seeds of *Hordeum vulgare* listed below (Table 2.6) were kindly provided by the Australian Winter Cereals Collection (AWCC; Tamworth, NSW, Australia).

Table 2.6 Barley seeds used in this study

Cultivar/Varieties	Species	Use
Hindmarsh	*Hordeum vulgare*	mRNA-seq, Varietal analysis
Calmariout	*H. vulgare*	Varietal analysis
Mundah	*H. vulgare*	Varietal analysis
ELB14	*H. vulgare*	Varietal analysis
Golden promise	*H. vulgare*	Varietal analysis
Buloke	*H. vulgare*	Varietal analysis
Maythorpe	*H. vulgare*	Varietal analysis
Vlamingh	*H. vulgare*	Varietal analysis
Beecher	*H. vulgare*	Varietal analysis
Clipper	*H. vulgare*	Varietal analysis
Skiff	*H. vulgare*	Varietal analysis
CM72	*H. vulgare*	Varietal analysis
Numar	*H. vulgare*	Varietal analysis
Arivat	*H. vulgare*	Varietal analysis
Gairdner	*H. vulgare*	Varietal analysis
Franklin	*H. vulgare*	Varietal analysis

General Molecular Methods

2.3 Propagation of plants

2.3.1 Growth of plants for genomic DNA isolation

Hordeum vulgare L *cv.* Hindmarsh seeds were soaked *overnight* on filter papers pre-soaked in distilled H_2O. The number of plants to be grown was decided on the basis of using three biological replicates for each sample. Sprouted seeds were planted into seedling trays filled with potting mix consisting of 2 parts vermiculite (Bunnings, Australia) and 1 part perlite (Exfoliators, Australia). The pots were then maintained in a Thermoline growth cabinet with 12 hours (h)/day of light, 72% humidity and 25°C of temperature conditions, supplying daily with full-strength Hoagland's solution

(Hoagland and Arnon 1950). The plants were allowed to grow for 14 days. Leaves of each plant was harvested separately, snap-frozen in liquid nitrogen and stored at -80ºC for DNA extraction.

2.3.2 Growth of plants for total RNA isolation

The barley plants for total RNA isolation were germinated and grown under the same conditions as above. The plants were allowed to grow for 14 days for gene expression studies. Using plants of visually comparable height and vigour, salt, drought and exogenous ABA stresses were applied to separate plants by supplying 150 mM NaCl (Ozturk et al. 2002) or 20% PEG (Skribanek and Tomcsanyi 2008) or 100 µM ABA (Ishitani et al. 1995) in Hoagland's solution, respectively for 12 h, while 3 plants remained unstressed (controls). Leaves of each plant was harvested separately, snap-frozen in liquid nitrogen and stored at -80ºC for RNA extraction.

2.3.3 Growth of plants for physiological analysis

The barley plants for physiological analysis were germinated and grown under the same conditions as above. The plants were allowed to grow for 14 days. Using plants of visually comparable height and vigour, salt stress was applied to separate plants by supplying 150 mM NaCl (Ozturk et al. 2002) in Hoagland's solution, respectively for 12 h, while 3 plants remained unstressed (controls). Leaves and roots of each plant were harvested separately and used fresh for physiological analysis.

2.4 Isolation and quantification of nucleic acids

2.4.1 Extraction of plant genomic DNA (gDNA)

Plant genomic DNA (gDNA) was extracted using the Wizard® DNA Purification Kit (Promega, Australia), according to the supplier's protocol. Approximately 40 mg of the leaf tissue harvested above (Section 2.3.1) was ground to a fine powder in a in a microcentrifuge tube using a sterilized metal rod, mixed with 600 µL of nuclei lysis solution and incubated at 65 ºC for 15 minutes. 3 µL of RNase A (4 mg/mL) was added and the solution was incubated at 37 ºC for 15 minutes and allowed to cool for 5 minutes before proceeding. 200 µL of protein precipitation solution was added and the mixture was centrifuged at 14,000 *xg* for 5 minutes. The DNA-containing supernatant was then carefully mixed with 600 µL isopropanol and centrifuged as above. After

centrifugation, the DNA pellet was washed using 600 µL of 70% ethanol, centrifuged as above, dried and re-suspended in 100 µL DNA rehydration solution. The isolated gDNA was stored at -20 ºC.

2.4.2 Extraction of plant total RNA

RNA was extracted using the TRIsure reagent for tissue lysis (Bioline; http://www.bioline.com/h_au.asp) according to the supplier's protocol. Barley leaves (snap-frozen as above) were crushed in a microcentrifuge tube using a sterilized metal rod. Approximately 100 mg of the crushed tissue was mixed with 1 mL of TRIsure and incubated at room temperature (RT) for 5 minutes. 200 µL of 24:1 chloroform: isoamyl alcohol was added to the above suspension and the tube was shaken vigorously for 15 seconds and then incubated at RT for 3 minutes. The mixture was centrifuged at 4^0C at 14,000 *xg* for 15 minutes and the colourless supernatant was transferred to a sterile microcentrifuge tube containing 500 µL of isopropanol. The tube was held at RT for 10 minutes and then centrifuged at 4^0C at 14,000 *xg* for 10 minutes. The RNA pellet so formed was washed with 75% ethanol (made with DEPC-treated water) and centrifuged at 4^0C at 4,000 *xg* for 5 minutes. The pellet was air-dried and then dissolved in 33 µL of DEPC-treated water (Sambrook and Russell 2001).

An aliquot of the RNA was also separated on bioanalyser and/or agarose gels to check the quality and integrity (lack of degradation) of the preparation (Sections 2.9.2; 2.9.3). The total RNA of appropriate quality was then used for two different purposes:

(i) Next generation mRNA-sequencing (mRNA–Seq) experiments to compare the changes to the barley leaf transcriptomes under different conditions.

(ii) Semi-quantitative reverse transcriptase PCR (sqRT-PCR) and quantitative real-time PCR (qRT-PCR) to compare the expression changes of selected individual genes.

These experiments and other molecular methods are described below

2.5 Extraction of chromatin for chromatin immuno-precipitation

2.5.1 Crosslinking and nuclei isolation

The crosslinking of the chromatin and protein, nuclei isolation and ChIP procedure were performed according to Haring *et al*. (2007), with minor modifications. Briefly, fresh leaves of six barley plants together for each of three conditions were cut into small pieces and transferred to a 50 mL tubes. The tissue was submerged in 30 mL isolation buffer A (10 mM Tris pH 8, 400 mM sucrose, 10 mM Na-butyrate, 3 % w/v formaldehyde, 0.1 mM phenylmethylsulfonyl fluoride (PMSF), 5 mM beta-mercaptoethanol) and "clogged" with few pieces of polystyrene to prevent floating of the tissue. The tissue was then vacuum infiltrated for 10 minutes at room temperature (RT) to allow penetration of formaldehyde in the cells, which crosslinks the chromatin and protein to ensure preservation of chromatin structure during isolation and ChIP procedures. Filter material polystyrene was removed and 2.5 mL of 2 M glycine was added to each tube and mixed. It was again vacuum infiltrated for 5 minutes at RT to stop the crosslinking. The tissue was washed three times with water and then dried using a paper towel. The tissue was snap frozen in liquid nitrogen and crushed using a sterilized metal rod, then resuspended in 30 mL of isolation buffer B (10 mM Tris pH 8, 400 mM sucrose, 10 mM Na-butyrate, 0.1 mM PMSF, 5 mM beta-mercaptoethanol, proteinase inhibitors 1 µg/mL each of aprotinin, leupeptin and pepstatin) and incubated for 15 minutes at 4 ^{0}C with gentle shaking. The solution was filtered through 4 layers of Miracloth into a new ice-cold tube and centrifuged for 20 minutes at 3000 *xg* at 4 ^{0}C to collect pellet. After removing the supernatant the pellet was resuspended in 1 mL ice cold isolation buffer C (10 mM Tris pH 8, 250 mM sucrose, 10 mM Na-butyrate, 10 mM $MgCl_2$, 1 % v/v Triton ×100, 0.1 mM PMSF, 5 mM beta-mercaptoethanol, proteinase inhibitors 1 µg/mL each of aprotinin, leupeptin and pepstatin) and centrifuged at 12000 *xg* for 10 minutes at 4 ^{0}C. The supernatant was removed and was resuspended in 300 µL ice cold isolation buffer D (10 mM Tris pH 8, 1.7 M sucrose, 10 mM Na-butyrate, 2 mM $MgCl_2$, 0.15 % v/v Triton × 100, 0.1 mM PMSF, 5 mM beta-mercaptoethanol, proteinase inhibitors 1 µg/mL each of aprotinin, leupeptin and pepstatin). This was overlayed on 1500 µL of isolation buffer D in a new 2 mL tube and then centrifuged for 1 h at 16000 *xg* at 4 ^{0}C to purify the extracted chromatin.

2.5.2 Chromatin sonication

After removing supernatant, the nuclei pellet was resuspended in 320 µL ice cold nuclei lysis buffer (50 mM Tris pH 8, 10 mM EDTA, 0.4 % w/v SDS, 0.1 mM PMSF, proteinase inhibitors 1 µg/mL each of aprotinin, leupeptin and pepstatin). An aliquot of 10 µL representing un-sheared chromatin was stored on ice. Remaining chromatin suspension was sonicated with a pulse of 30 seconds at 3 µm amplitude. The tube was cooled in ice water during sonication and cool for 30 seconds (s) in between pulses. The insoluble fractions were pelleted by spinning for 5 minutes at 16000 *xg* at 4 ^{0}C. The clear supernatant containing the sheared chromatin was transferred to new tubes in 50 µL aliquots. The aliquots of sheared and un-sheared chromatin were reverse crosslinked by adding 140 µL TE (10 mM Tris pH 8, 1 mM EDTA), 5 µL 5 M NaCl and 5 µL 20 % SDS and incubated for 2 h at 65 ^{0}C to check the sonication efficiency by isolating DNA and running on agarose gel electrophoresis to ensure that chromatin is sonicated to length of 250 to 750 bp.

2.5.3 Chromatin pre-clearing

Forty microliter of Dynabeads A (Invitrogen; Cat No: 100.02D) washed twice in 40 µL dilution buffer (0.01 % SDS, 1.1 % Triton X-100, 1.2 mM EDTA, 16.7 mM Tris HCl pH 8.1, 167 mM NaCl, proteinase inhibitors 1 µg/ml) using the magnet to immobilise the magnetic beads. The beads were re-suspended in 40 µL of dilution buffer. The pre-clearing of chromatin (removal of cell lysates) was achieved by adding 450 µL of dilution buffer and 20 µL washed beads to chromatin aliquot (50 µL; as prepared above) and then rotated at 12 rpm for 2 h at 4 ^{0}C (labelled as "chromatin tube"). Pre-incubation of beads and the antibody was achieved by mixing 480 µL of dilution buffer, 20 µL washed beads and 0.5 µg of IgG antibody against peptide H3K4me3 in rabbit (Active motif; Cat no: 39159) and then rotated for 2 h at 4 ^{0}C (labelled as "antibody tube"). An aliquot of the chromatin solution (55 µL; without any beads) was stored on ice as control.

2.5.4 Chromatin immuno-precipitation

The "antibody tube" was placed on the magnet and supernatant (unbound beads and antibodies) was discarded. The "chromatin" tube was placed on the magnet and the supernatant (pre-cleared chromatin) was transferred to the "antibody tube". The tube

was then rotated overnight at 4 ^{0}C. The tube was briefly centrifuged for 10 s and then placed on magnet. The supernatant (immune-precipitated chromatin) was transferred to a fresh tube and washed once with 800 µL each of the following pre-chilled buffers; Low salt (20 mM Tris pH 8, 2 mM EDTA, 0.1 % w/v SDS, 1 % v/v Triton ×100, 150 mM NaCl), High salt (20 mM Tris pH 8, 2 mM EDTA, 0.1 % w/v SDS, 1 % v/v Triton × 100, 500 mM NaCl), LiCl (20 mM Tris pH 8.0, 1 mM EDTA, 1% v/v NP-40, 1% w/v Na-deoxycholate, 250 mM LiCl), TE by rotating for 5 minutes at 4 ^{0}C, followed by single wash with pre-chilled TE+0.01 % SDS at RT. After the last wash solution, the bead solution was centrifuged for 2 minutes at 200 *xg* at 4 ^{0}C and then transferred to new tube to reduce background. The tube was then placed on the magnet and the supernatant was then removed.

2.5.5 Elution of ChIP DNA

The beads were resuspended in 102 µL of ChIP elution buffer (20 mM Tris pH 8, 1 mM EDTA, 1 % v/v NP-40, 1 % w/v Na-deoxycholate, 250 mM LiCl) and 5 µL proteinase K (Sigma; Cat No: P4850-5mL) and then incubated for 10 minutes at 50 ^{0}C. The reverse-crosslinking between chromatin and proteins was performed by incubating at 65 ^{0}C for 2 h. The tube was briefly centrifuged for 10 s and then placed on magnet to transfer the supernatant (immuno-precipitated chromatin) to the tube with equal volume of isopropanol (~100 µL) and 5 volumes (~500 µL) of NTB binding buffer (from Nucleospin kit; Macherey-Nagel, Germany) to isolate DNA. The mixture was then transferred to the nucleospin column and centrifuged for 30 s at 14000 *xg*. The flow through was discarded and 500 µL of NT3 wash buffer was added, followed by spinning it again as before. After discarding the flow through, tube was centrifuged again for 1 minute. The spin column was then transferred to new tube and incubated for 2 minutes at RT to allow complete drying. Forty micro-litres of elution buffer from the kit was added, incubated for 1 minute at RT and then centrifuged at 14000 *xg* for 1 minute. The ChIP DNA was then quantified using Qubit (Invitrogen, Australia).

2.6 Next generation mRNA-sequencing experiment I (mRNA-seq I): analysis of barley leaf transcriptome under salt stress, for method development for un-sequenced genomes

2.6.1 mRNA-seq I data acquisition

Total RNA from two salt-stressed plants was pooled in equal quantities (0.5 µg) for mRNA-seq library preparation, to minimise any biological variation in transcriptomes, as per Mizuno *et al*. (2010) and Ando and Grumet 2010. Similarly, total RNA from two control plants was also pooled. The mRNA-seq methodology consists of four main stages: (A) library preparation, (B) cluster generation, (C) sequencing and (D) data analysis. This stage of the work was shared with Dr Runyararo Memory Hove (SUT), who was then a fellow PhD student. An overview is provided in Figure 2.1 and the key steps are described below, while the details of all reagents and steps are provided in Appendix II.

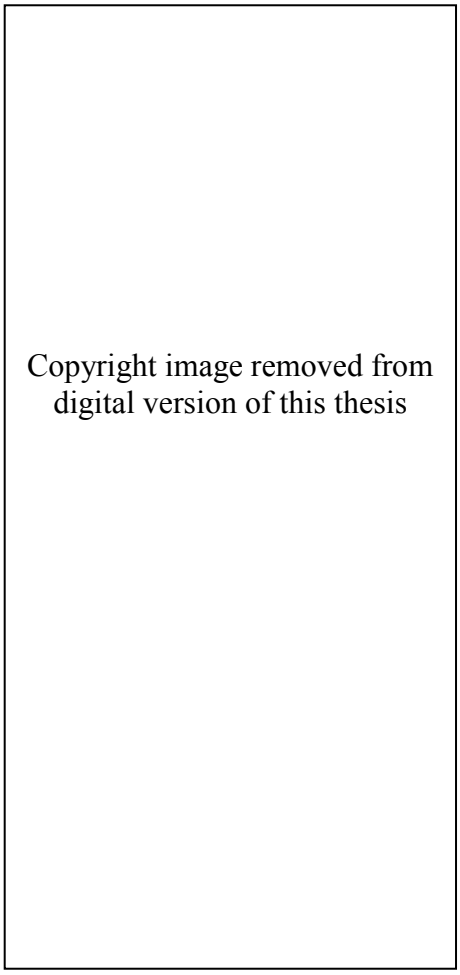


Figure 2.1 Summarised workflow of mRNA-seq using Illumina platform (http://grcf.jhmi.edu/hts/protocols/mRNA-seq_SamplePrep_1004898_D.pdf; September 2014

Library Preparation: The cDNA libraries were produced from control and salt stress samples as per the Illumina mRNA-seq library preparation protocol (http://grcf.jhmi.edu/hts/protocols/mRNA-seq_SamplePrep_1004898_D.pdf; September 2014) (Illumina, CA, USA) using the Illumina mRNA-seq sample prep kit (Illumina; RS-100-0801). The main steps are summarised below:

1. **Purification of mRNA from total RNA** - Separation of poly-A tailed mRNA from the above-purified DNase-treated total RNA (Section 2.9.1) was conducted using poly-T magnetic beads (Illumina).
2. **mRNA fragmentation** - Fragmentation of mRNA to small pieces was done using divalent cations and increased temperatures (94 ºC).
3. **Synthesis of first strand cDNA** - The synthesis of first strand cDNA using mRNA as template was carried out with the help of SuperScript™ II reverse transcriptase (200 U/µL; Invitrogen) and random primers.
4. **Synthesis of the second strand cDNA and purification of double-stranded cDNA** - This procedure involved the synthesis of the second strand cDNA by removal of the RNA template and then the synthesis of a DNA strand using the first strand cDNA as template using DNA polymerase I (10U/µL) and RNase H (5U/µL).
5. **End repair and adenylation of the 3´ end -** End repair involved the conversion of any overhangs created on the cDNA during the above steps to blunt ends using T4 DNA polymerase and Klenow DNA polymerase. The 3´adenylation step then entailed the addition of an 'A' nucleotide to both 3´ ends of the cDNA strands in preparation for ligation to the adapters, which have one 'T' overhang.
6. **Ligation of the barcode adapters -** This step involved the ligation of the PE adapters (5'-P-GATCGGAAGAGCGGTTCAGCAGGAATGCCGAG-3'; 5'-ACACTCTTTCCCTACACGACGCTCTTCCGATCT-3')1 to the cDNA fragments using T4 DNA ligase in preparation for hybridization (binding) to the flow-cells.
7. **Gel purification of the cDNA templates** - This step allowed the selection of a size range of the cDNAs for later enrichment. The ligation products were electrophoresed on a 2% agarose gel and the bands in the 200 bp size range were excised and purified using the QIAquick gel extraction kit (Qiagen).

8. **PCR enrichment of the purified cDNA templates** – In this procedure PCR reaction consisting of PCR primer PE 1.0 (5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGC TCTTCCGATCT 3′)1, PCR primer PE 2.0 (5' CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAAC CGCTCTTCCGATCT 3′)1 and Phusion DNA polymerase (2U/µL, Finnzymes) was employed to amplify the gel-purified cDNA templates, in order to create the final cDNA library for sequencing.

Cluster Generation: Cluster generation was done according to the Illumina user guide for cluster generation kit v4 (published November 2009) on the Illumina cluster Station. This procedure allowed the immobilization (binding) of the modified fragments from the cDNA library onto the flow cell. Cluster generation consists of five main steps, namely: (i) Denaturation of cDNA with NaOH and dilution to 15 pM; (ii) Cluster amplification with bridge PCR: hybridization of the template cDNA onto the oligo-nucleotide coated surface of the flow cell and amplification of the template DNA to form DNA clusters; (iii) Linearisation: removal of one of the two adapter molecules from the flow cell; (iv) Blocking: blocking of the 3′ OH end of the linearised clusters; (v) Hybridization: hybridization of the sequencing primers onto the linearised clusters.

Sequencing: Sequencing by synthesis was performed on the Illumina Genome Analyzer IIx (GAIIx) and according to the user guide for sequencing kit v4 using the sequencing-by-synthesis 76-cycle run (published August 2009).

Data Analysis: Image analysis, intensity scoring and base-calling was conducted using the Illumina Real-Time Analysis (v1.8) software (http://support.illumina.com/sequencing/sequencing_software/real-time_analysis_rta.ilmn), yielding >50 million 76 bp sequence reads altogether.

2.6.2 Bioinformatics analysis of mRNA-seq I results

This stage of the work was also shared with Dr Runyararo Hove SUT (as above). The mRNA-seq datasets were filtered for spurious reads using the Fastx Artifacts Filter and poor quality bases were removed from the 3' end with the Fastq Quality Trimmer using a threshold of Q30 (http://hannonlab.cshl.edu/fastx_toolkit/). Initially, the *Hordeum*

vulgare Unigene transcript sequences downloaded from the NCBI database, which consisted of 26,941 transcripts, including those annotated as 'complete CDS' and 'partial CDS' (http://www.ncbi.nlm.nih.gov/unigene; last accessed January 2013) were used as reference sequences. The 76 nucleotide (nt) cDNA reads were aligned using Burrows-Wheeler Aligner (BWA) (Li and Durbin 2009) with default settings. Unaligned reads were extracted with SAMtools software (Li *et al*. 2009b) and underwent de novo assembly using the 'Assembly By Short Sequence' (ABySS) software and contigs less than 100 bp were discarded and the remaining contigs were named 'novel' tentative consensus sequences (NTCs). To determine whether these NTCs represented novel sequences, they were BLASTN searched in the above Barley Unigene collection as well as in rice cDNA database (http://rice.plantbiology.msu.edu/). 'Novel' TCs were expected to find relatively higher % identity to the rice database compared to the barley Unigene database. After trial and error of various rice/barley BLAST % identity ratios (that turned out very stringent or too non-selective; explained in Section 3.2.3), a rice/barley BLAST % identity hit score ratio threshold of less than 2 was implemented and NTCs with a score less than 2 were discarded, leaving a set of 'novel' TCs (NTCs), which were subsequently appended to the Unigene reference. The workflow is given in Figure. 2.2

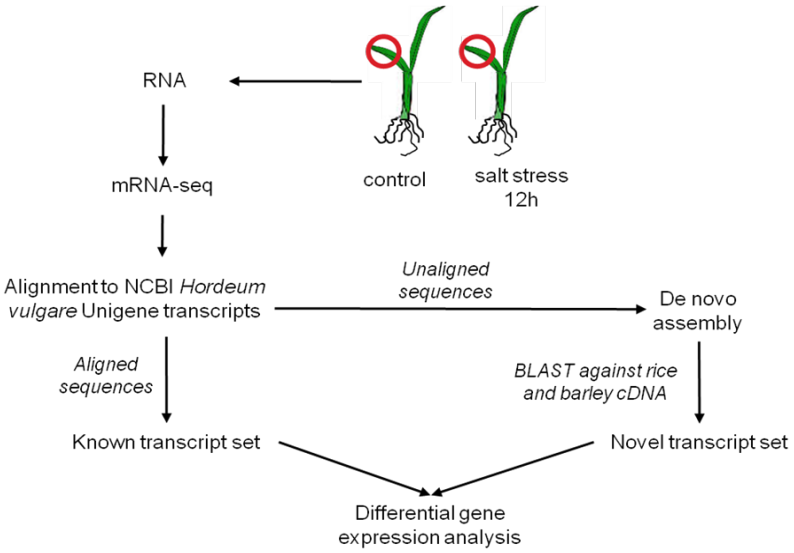


Figure 2.2 Study design, mRNA-seq was performed on RNA derived from control and salt stressed barley leaves.
Reads unaligned to current databases are assembled to discover novel sequences. These sequences are BLASTed to rice and barley databases to determine novelty. The final contig set is analysed for differential gene expression.

To perform differential gene expression (DGE) analysis, the Q30 quality-trimmed reads were aligned with BWA to the NCBI Unigene database. Counts for each transcript were extracted with SAMtools (idxstats feature) (http://samtools.sourceforge.net/; Li *et al*. 2009b) and then analysed by the DESeq software package (http://www-huber.embl.de/users/anders/DESeq/; Anders and Huber 2010), using the conservative ''blind'' method to estimate variance despite the lack of replicates. Transcripts with false discovery rate (Benjamini-Hochberg procedure) adjusted p values < 0.05, were considered significantly differentially expressed.

2.6.3 Gene ontology for mRNA-seq I

Gene ontology (GO) analysis was performed for up- and down-regulated sets of 200 genes (selected based on adjusted *p* value rank) by firstly mapping each barley gene to its closest BLASTN match in the rice cDNA database as above, using an e value threshold of <0.1. The rice locus name sets were then analysed with the agriGO Singular Enrichment Analysis tool (http://bioinfo.cau.edu.cn/agriGO/; Du et al. 2010), using the suggested rice whole-transcriptome background. Significance of the gene set enrichment was evaluated with Fisher's test using Yekutieli FDR adjustment, with a significance threshold set at 0.05.

2.7 Next generation mRNA-sequencing experiment II (mRNA-seq II): analysis of barley leaf transcriptomes under salt, drought and exogenous ABA stress conditions

2.7.1 mRNA-seq II data acquisition

Using the barley (*cv.* Hindmarsh) leaf total RNAs, eight mRNA-seq libraries were prepared: (i), two from total RNA of two separate plants treated with 150 mM NaCl for 12 h; (ii), two from total RNA of two separate plants treated with 20% PEG for 12 h; (iii) two from total RNA of two separate plants treated with 100 µM exogenous ABA for 12 h and (iv) likewise, two from total RNA of two control plants. mRNA-seq libraries were prepared as explained above. For cluster generation, 8 libraries prepared above were loaded on 4 lanes of flow cell. Data from two libraries on the same lane was differentiated using different barcode adaptors for each library loaded in the same lane (explained in step 6 of library preparation; Section 2.6.1). The sequencing by synthesis and data analysis was performed as explained above.

2.7.2 Bioinformatics analysis of mRNA-seq II results

mRNA-seq II datasets were filtered for spurious reads using the Fastx Artifacts Filter and poor quality bases were removed from the 3' end as above with the Fastq Quality Trimmer using a threshold of Q30 (http://hannonlab.cshl.edu/fastx_toolkit/). *Hordeum vulgare* transcript sequences downloaded from the International Barley Genome Sequencing Consortium database (IBSC; http://mips.helmholtz-muenchen.de/plant/barley/index.jsp; last accessed June 2014) were used as reference sequences. The 76 nt reads were aligned using the Burrows-Wheeler Aligner (BWA) (Li and Durbin 2009) using default settings. The differential gene expression (DGE) analysis was performed as explained above for mRNA-seq I dataset. The Gene ontology (GO) analysis was also performed for up- and down-regulated sets of 200 genes (selected based on adjusted *p* value rank) as described above.

2.8 Next generation ChIP-sequencing experiment

2.8.1 ChIP-seq data acquisition

Using the barley (*cv.* Hindmarsh) leaf total chrmatin, three ChIP-seq libraries were prepared: (i), one from total chromatin of six plants treated with 150 mM NaCl for 12 h; (ii), one from total chromatin of six separate plants treated with 20% PEG for 12 h and (iii) likewise, one from total chromatin of six control plants. ChIP-seq libraries were prepared as explained above. For cluster generation, three libraries prepared above were loaded on 4 lanes of flow cell. Data from two libraries on the same lane was differentiated using different barcode adaptors for each library loaded in the same lane (explained in step 6 of library preparation; Section 2.6.1). The sequencing by synthesis and data analysis was performed using Hi-seq 2500.

2.8.2 Bioinformatics analysis of ChIP-seq results

ChIP-seq datasets were filtered for spurious reads using the Fastx Artifacts Filter and poor quality bases were removed from the 3' end as above with the Fastq Quality Trimmer using a threshold of Q30 (http://hannonlab.cshl.edu/fastx_toolkit/). *Hordeum vulgare* transcript sequences downloaded from the International Barley Genome Sequencing Consortium database (IBSC; http://mips.helmholtz-muenchen.de/plant/barley/index.jsp; last accessed June 2014) were used as reference sequences. The 50 nt reads were aligned using the Burrows-Wheeler Aligner (BWA)

(Li and Durbin 2009) using default settings. The differential gene expression (DGE) analysis was performed as explained above for mRNA-seq dataset.

2.9 Molecular techniques applied for other gene expression analyses

2.9.1 DNase treatment of RNA

The total RNA extracted as above was treated with DNaseI (Promega, Australia) according to the manufacturer's instructions (http://au.promega.com/?origUrl=http%3a%2f%2fwww.promega.com%2f). Thirty microlitres of total RNA extracted (Section 2.4.2) was mixed with 10 units (U) of RQ1 RNase-free DNase I (Promega Australia), 5 µL of the supplied 10X reaction buffer (Promega where) and 2U of RNase inhibitor (Bioline where), making the final volume to 50 µL with DEPC-treated water. The mixture was incubated at 37°C for 30 minutes. The RNA was then repurified using the LiCl precipitation according to the Ambion technical Bulletin # 160 (http://www.ambion.com/techlib/tb/tb_160.html; last accessed June 2014). Briefly, 20 µL of 10 M LiCl (made in DEPC-treated water) was added to the DNase-treated RNA and the final volume made to 80 µL (with DEPC-treated water) to obtain a final concentration of 2.5 M LiCl. The mixture was kept at -20°C for 30 minutes and then centrifuged at 14,000 *xg* at 4°C for 15 minutes. The RNA pellet was washed twice using chilled 75% ethanol (made with DEPC-treated water) at 14,000 *xg* at 4°C for 15 minutes. The RNA pellet was air-dried at RT and then resuspended in 5 µL DEPC-treated water and stored at -80°C.

2.9.2 Qualitative analysis of RNA samples by Microchip electrophoresis

The quality of the purified total RNA from barley plants was determined for Next generation mRNA-sequencing (mRNA-seq) purposes using the MCE®-202 MultiNA, Microchip Electrophoresis System (Shimadzu, Japan). The purified total RNA from two plants was diluted 1:2 with RNA marker solution. The 28S and 18S bands were recognized based on the calibration curve from the RNA 6000 ladder (Applied Biosystems).

2.9.3 Qualitative analysis and approximate size determinations of DNA and RNA by agarose gel electrophoresis

The quality of RNA (before DNase treatment) or DNA (i.e. visible bands and lack of degradation) was assessed by agarose gel electrophoresis for gene expression analysis of selected individual genes. Agarose gels were typically prepared at 1% (w/v) concentration in 1X TAE buffer, with 0.5 µg/mL ethidium bromide added to the gel solution (Sambrook and Russell, 2001). Generally, 5 µL aliquots of RNA or DNA were mixed with 1 µL of 6X xylene cyanol loading dye. The DNA molecular weight markers used typically were GeneRuler™ 50bp DNA ladder (50 - 1,000 bp; Fermentas), Hyperladder™ I (200 - 1,037 bp; Bioline) or Hyperladder™ V (25 - 500 bp). Electrophoresis was accomplished using a Bio-Rad power pack at 80-100 V for 45-90 minutes. The gels were photographed on ultraviolet transilluminator using a Chemidoc XRS Documentation Station (Bio-Rad) and Quantity One software (Bio-Rad).

2.9.4 Spectrophotometric quantifications of RNA and DNA

The purified DNA or RNA was diluted to 1:100 with DEPC treated water to make the final volume 500 µL. Absorbance readings were recorded at 230 nm, 260 nm and 280 nm using a GeneQuant Pro spectrophotometer (GE Healthcare). The concentration was calculated on the basis of $1A_{260}$= 50 µg/mL of double stranded DNA and $1A_{260}$= 40 µg/mL of single stranded RNA (Sambrook and Russell, 2001). DNA or RNA to protein absorbance ratio (A_{260}/A_{280}) of 2.0 was considered desirable, as it indicates the pure DNA or RNA (Sambrook and Russell, 2001).

2.9.5 First strand complementary DNA (cDNA) synthesis

First strand complementary DNA (cDNA) was synthesised from total RNA of interest using the Bioscript MMLV reverse transcriptase (Bioline, Australia) according to the supplier's instructions. 1 or 2 µg of purified RNA (after DNase treatment) isolated and purified as above (Section 2.9.1) was mixed with 1 µL of oligo $d(T)_{18}$ (0.5 µg/µL; Invitrogen), making the final volume to 12 µL. The mixture was incubated at 70°C for 5 minutes and then chilled on ice. 10 U of RNase inhibitor, 40 mM dNTP, 1X Reaction buffer and 50 U of Bioscript MMLV reverse transcriptase (all from Bioline; http://www.bioline.com/h_au.asp) were added to the above prepared mixture. The final volume was made to 20 µL with DEPC-treated water. The mixture was incubated at

37°C for 1 hour to synthesize the first strand of cDNA. The reaction was terminated by heating at 70°C for 10 minutes and the resulting first strand cDNA preparation was stored at -20°C. The success and quality (lack of gDNA contamination) of the synthesised cDNA was assessed by reverse transcriptase PCR (RT-PCR) using intron-flanking actin primers (Table 2.7) and 1 µL cDNA as template.

2.10 Polymerase chain reaction (PCR)

2.10.1 Design and synthesis of primers

Primers for various polymerase chain reaction (PCR) (Mullis and Faloona 1987) were designed using Netprimer (http://www.premierbiosoft.com/netprimer/index.html; last accessed June 2014) with the following criteria: length between 15-25 bases; maximum of 5 °C difference in the annealing temperatures of the forward and reverse primers; % GC content of ~50% and minimal secondary structures such as hairpins and primer dimers (Rybicki 2005). All primers were synthesised commercially (Invitrogen™, Australia) and received as dried pellets. Primers were resuspended in dH_2O to a concentration of 0.1 µg/µL (approximately 10 µM depending on length and base composition) and stored at -20°C. The primer sequences for two housekeeping genes were designed as per Mohammadi *et al*. (2007) for α-tubilin and as per Hv.23088 (barley actin CDS sequence) for actin. The primers were designed for testing the expression patterns of selected barley genes in response to salinity, drought and exogenous ABA using sqRT-PCR and real-time PCR (Table 2.7). The other primers were designed for the amplification of ABA signalling related candidate genes (cDNA; Table 2.8).

Table 2.7 Primers used for sqRT-PCR and real-time PCR in barley

Primer name	Primer Sequence (5'-3')	Expected cDNA amplicon size (bp)	Annealing temp (^{0}C)
Hv.12354 α-tubilinF α -tubilinR	GGACCGTACGGGCAGATCT CACCAGACTGCCCAAACACA	72	59.9 59.7
Hv.23088 Hv-actinF Hv-actinR	TGAACCCAAAAGCCAACAGAG CACCATCACCAGAGTCGAGAAC	147	60.2 59.2
[1]Hv.469 Hv.469F1 Hv.469R1	CAACAAGCCCGACAGTGAAAT TGCACCCATCTCCGACAATA	62	60.0 59.3
[2]Hv.10251 Hv.10251F1 Hv.10251R1	GCTGCTTACAACACCATCTACAAA ATTCACATCTGGGCGGGC	118	59.3 61.0
[3]Hv.8888 Hv.8888F2 Hv.8888R2	CTACAGATATACACCCGCAGACAA ACCATGCAGCCCTCCAGC	122	59.1 61.2
[4]Hv.8276 Hv.8276F2 Hv.8276R2	GCGGATGTCGGCTACAACTG CACCACGAGGAGGCCGAA	81	61.3 62.1
[5]Hv.22598 Hv.22598F1 Hv.22598R1	GACCAGCAGGACATCACCAGA TGAAAGCAGGAGCGATGTAGAA	172	60.3 60.3
[6]Hv.20929 Hv.20929F1 Hv.20929R1	GTGGCGGCAAAATGGAGG GCGGCTGGTAGCTGGGG	139	61.8 61.9
[7]Hv.25312 Hv.25312F1 Hv.25312R1	CGCTCCTACTCCCACCTCCT CGCACCTCCCTCAGCTTGTC	152	60.6 62.8
PP2CA Hv.19158F Hv.19158R	ATCCCATCTCCTTCCGCC TGAGAAGCAATCCACAAACGC	206	59.21 61.12
PP2CA Hv.18841F Hv.18841R	CATCCAGGCGGCGGG GTCACGAACGGCTTCAGGTA	108	61.55 58.42
PYR/PYL Hv.15651F Hv.15651R	GTACAAGCGGTTCGTGCG GGACGGTGGTGACAGAGAGG	206	58.05 59.7
PYR/PYL Hv.9994F Hv.9994R	ACCTCCAGTCCCTCGCCC CCACCACCACGGATTTGATTA	182	61.64 60.54
SnRK2 Hv.1875F Hv.1875R	GGAAGCACCACCCCTCGC TGACCACACATCGGCAATCT	159	63.19 59.1
SnRK2 Hv.5014F Hv.5014R	AGTATGCCATCGTGACCTGAAG GGCTGCGAGTGAAGAACCG	114	59.28 61.35
ABCG25 Hv.77521F Hv.77521R	CGTGCTGCTGCTCTCCGA TTGACGCTGCCTCCCTCC	186	61.17 61.39
ABCG40 Hv.9836F Hv.9836R	AATGTCTCAGGTGGTGGTTGTAG TATGATTGTGCTGATACCCTAGTG	242	58.35 56.84
PP2CA AK251854F AK251854R	ATCCCATCTCCTTCCGCC TGAGAAGCAATCCACAAACGC	206	59.2 61.1

Primer name	Primer Sequence (5'-3')	Expected cDNA amplicon size (bp)	Annealing temp (^{0}C)
PP2CA MLOC_8131F MLOC_8131R	CATCCAGGCGGCGGG GTCACGAACGGCTTCAGGTA	108	61.5 58.4
PYR/PYL MLOC_71349F MLOC_71349R	GTACAAGCGGTTCGTGCG GGACGGTGGTGACAGAGAGG	206	58.0 59.7
PYR/PYL AK376521F AK376521R	ACCTCCAGTCCCTCGCCC CCACCACCACGGATTTGATTA	182	61.6 60.5
SnRK2 MLOC_3013F MLOC_3013R	GGAAGCACCACCCCTCGC TGACCACACATCGGCAATCT	159	63.1 59.1
SnRK2 MLOC_22145F MLOC_22145R	AGTATGCCATCGTGACCTGAAG GGCTGCGAGTGAAGAACCG	114	59.2 61.3
ABCG25 MLOC_62985F MLOC_62985R	CGTGCTGCTGCTCTCCGA TTGACGCTGCCTCCCTCC	186	61.1 61.3
ABCG40 MLOC_68581F MLOC_68581R	AATGTCTCAGGTGGTGGTTGTAG TATGATTGTGCTGATACCCTAGTG	242	58.3 56.8
Bowman_contig_222590F Bowman_contig_222590R	GCAGTTCAAGTCAGGCAGC CGTGAGAAGGTTGGAAAGGTGA	160	56.2 61.63
Bowman_contig_68397F Bowman_contig_68397R	CCGTCGGTGTTGAGACTGTTC GGTTCTTCAAATGTCGTAGTGGC	162	60.17 60.63

[1]Putative transketolase; [2]Indole-3-glycerol phosphate lyase; [3]DNA gyrase B subunit; [4]Glycosyl hydrolases, putative, expressed; [5]Ribulose bisphosphate carboxylase; [6]Serine carboxypeptidase II; [7]No annotation availabale

Table 2.8 Primers used for standard PCR in barley

Primer Name*	Primer sequence (5'-3')	Annealing temp (^{0}C)	Application
PP2CD NTC14333F1 NTC14333R1	GCGTCCACCTTCGTCGG ATCGCCGCTCTCACCAGTC	59.91 60.59	cDNA amplification
PYR/PYL NTC2524F1 NTC2524R1	GCCCAACCCAGCCATTC CCGCCCATCAATACTCTGTG	58.77 58.64	
ABCG40 NTC9836 F1 NTC9836R1	GTGGAGTACATGCCTGACGC ATGCTGTTGCCCTAAATGAGTT	58.51 58.52	
ABCG40 NTC4445F1 NTC4445R1	ATGTTTGCCATTACTTGTGTCCT CCGACGCTGCTCAACTGC	58.32 60.62	

2.10.2 PCR conditions

All the PCR reactions were carried out based on the principles described by Mullis and Faloona (1987). All the amplifications were generally carried out in 25 µL reactions containing 200 ng of gDNA template, 12.5 µL of 2× Biomix (Bioline; contains Taq polymerase, MgCl2, dNTPs) and 0.1 µg of each forward and reverse primer. Amplifications from cDNA or plasmids were performed using 20 ng cDNA or 10 ng

plasmid template. Negative controls (no template) were also included. Thermal cycling was done on the MyCycler Thermal Cycler (Bio-Rad) or Mastercycler (Eppendorf) using the cycling conditions in Table 2.9. The success of each PCR reaction was determined by subjecting 5 µL of each PCR product to agarose gel electrophoresis (Section 2.9.3)

Table 2.9 PCR thermal cycling reaction

Step	Temperature (^{0}C)	Time	Number of cycles
1. Initial denaturation	94	5 minutes	1
2. Denaturation	94	45 seconds	35
3. Annealing	Primer-specific	45 seconds	35
4. Extension	72	1 minute/kb	35
5. Final elongation	72	10 minutes	1
6. Hold	4	Hold	1

2.11 Individual gene expression analysis

2.11.1 Semi-quantitative reverse transcriptase-PCR (sqRT-PCR)

This technique was carried out for gene expression analysis. Three biological replicates, i.e., RNAs from three control and three stressed plants, were used to compare the changes in expression of selected genes. The sqRT-PCR reactions consisted of 1 µL of synthesised first strand cDNA as template, 12.5 µL of 2× Biomix (Bioline) and 0.1 µg of each primer (forward and reverse), made up to 25 µL with sterile MilliQ water. The actin and α-tubulin primers were used as housekeeping genes (Suprunova et al. 2004). The PCR thermal cycling conditions were identical to the typical PCR conditions (Table 2.9), except that each gene was amplified for 20, 25, 30, 35 cycles to determine the cycle number where the PCR is still in the exponential phase. 5 µL aliquots sqRT-PCR product was electrophoresed on agarose gels. The intensity of bands was recorded using the Chemidoc XRS Documentation Station (Bio-Rad), which uses Quantity One software (Bio-Rad). The differential expression of genes in RNAs from control and stressed plants was calculated using the following formula (Jang et al. 2004):

$$\text{Relative expression (Fold change)} = \frac{\text{Intensity (gene) (stress)}}{\text{Intensity(tubulin) (stress)}} \div \frac{\text{Intensity (gene) (control)}}{\text{Intensity(tubulin) (control)}}$$

2.11.2 Quantitative real time PCR (qRT-PCR) reaction condition

All amplifications were generally carried out in 20 µL containing 200 ng cDNA, 10 µL of 2 × SensiMixPlus (Bioline; contains SYBR Green I, fluorescein, heat-activated DNA Polymerase, dNTPs, $MgCl_2$, reaction buffer) and 0.5 µL (0.1 µg/µL) each of the forward and reverse primers. Each cDNA was analysed in duplicate. Reactions were analysed using 96-well plates and MyiQ™ real-time PCR detection system (Bio-Rad). Cycling conditions consisted of 95 ºC for 1 minute, followed by 40 cycles of 95 ºC for 15 seconds (s), 60 ºC for 30 s, and 72 ºC for 15 s. Data collection occurred during the annealing step. A melt curve was performed after each PCR by increasing the set-point temperature of 60 ºC by 0.5 ºC every 10 s, to assess the success and specificity of each PCR (single peaks only). Differential gene expression fold change (FC) was calculated using actin and α-tubulin as housekeeping controls that exhibit relatively constant expression (Supronova et al. 2004).

2.12 Amplification ABA-mediated abiotic stress signalling related candidate genes

CDS sequences were amplified from the cDNA of barley *cv.* Hindmarsh by PCR using specific primers (Table 2.7), in 50 µL final volumes (Section 2.10.1). PCRs were performed using typical cycling conditions (Section 2.10.2), and at primer specific annealing temperature (Table 2.7). The PCR products (5 µL) were electrophoresed on 1.0 % (w/v) agarose gels to check the results.

2.13 Purification of DNA bands from agarose gel

Generally 45 µL of PCR product was mixed with 10 µL of 6X xylene cyanol loading dye, loaded onto a 1.0 % agarose gel and electrophoresed. DNA was purified using the Perfectprep® gel cleanup kit and its protocol (Eppendorf). In brief, DNA bands were excised, mixed with 3 equivalent volumes of binding buffer (e.g. 0.1 g gel in 0.3 mL binding buffer) and incubated at 50 ºC for 10 minutes to melt the gel. One volume of isopropanol was added, and the entire mixture was pipetted onto a spin column (provided with the kit). After centrifugation at 6,500 g for 1 minute to bind the DNA to the spin column, the column was washed with diluted wash buffer. To elute the bound DNA 30 µL of sterile Milli-Q water was added to the column, incubated for 10 minutes at RT and centrifuged at 6,500 g for 1 minute. The purified DNA was stored at -20 ºC.

2.14 Cloning of PCR products

Approximately 45 µL of PCR product was electrophoresed and gel-purified and an aliquot of the resulting purified DNA was ligated with the Promega pGEM-T Easy (Appendix I) cloning vector using the pGEM-T easy vector system I (Promega, Australia). Briefly, up to 3 µL of purified PCR product was mixed with 1 µL of pGEM-T Easy vector, 1 µL of 10X T4 DNA ligase buffer and 1 µL of T4 DNA ligase in a total volume of 10 µL. The ligation reactions were held at 4 ºC overnight and afterwards transformed into 100 µL of competent *E. coli* JM109 cells.

2.15 Transformation of chemically competent *E.coli* JM109cells

The chemically competent *E. coli* JM109 cells were prepared according to Inoue et al. (1990), with minor modifications. The cells from the stock culture were streaked onto a non-selective LB agar plate and grown overnight at 37 °C. A single colony inoculated into 10 mL LB medium and grown overnight at 37 °C with shaking (180 rpm). This culture was used to inoculate 500 mL SOB (0.5 % yeast extract, 2 % tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$) media and incubated at 18 °C with shaking until the cell density reached OD_{600}=0.4. The culture was held on ice for 10 minutes and the cells were transferred to 50 mL tubes, followed by centrifugation at 300 g for 15 minutes at 4 °C. The cells were resuspended in 80 mL ice-cold TB buffer (0.037 g EDTA, 1.36 g KCl. 0.27 g NaCl, 0.17 g phenylmethylsulfonyl fluoride, 0.12 g Tris-HCl in total volume of 100 mL), held on ice for 10 minutes and then centrifuged as above. The cells were washed again in 20 mL TB buffer. Dimethyl sulfoxide (DMSO) was then added to a final concentration of 7 % (v/v) with gentle mixing. The cells were held again on ice for 10 minutes, distributed into 100 µL aliquots, which were snap-frozen in liquid nitrogen and stored at -80 °C. Transformation was achieved by thawing the stored cells on ice, mixing them with the ligation reaction and incubating on ice for 1 hour (h). The cells were heat shocked at 42 °C for 3 minutes and then cooled on ice for 2 minutes, after which 500 µL of 2X YT media was added and cells were incubated at 37 ºC for 1 h. Aliquots of 100 µL of bacterial suspension were each spread onto LB/IPTG/X-gal/ampicillin agar plates, which were incubated at 37 ºC for 16 h. The pGEM®-T Easy plasmid contains the LacZ gene, encoding the α-peptide of β-galactosidase. Thus, blue white screening of non-recombinant (blue) and recombinant

(white) colonies was possible. Generally, at least six white colonies from each ligation were inoculated separately into 4 mL of luria broth (LB) containing 50 µg/µL of ampicillin, and grown overnight at 37 ºC on shaker (180 rpm). Plasmid DNA was purified from these bacterial cultures.

2.16 Plasmid purification

The plasmids were purified using the Wizard® *Plus* SV Minipreps DNA purification system according to manufacturer's instructions (Promega, Australia). Briefly, 3 mL of *E. coli* JM109 culture containing plasmids was pelleted by centrifugation at 6,700 g for 5 minutes. The pellets were resuspended in 250 µL of cell re-suspension solution, lysed with 250 µL of cell lysis solution, mixed with 10 µL of alkaline protease solution, and incubated at room temperature (RT) for 5 minutes. 350 µL of neutralization solution was then added and the lysate was centrifuged at 14,100 g for 10 minutes at RT. The supernatant was loaded onto a spin column assembly (provided with the kit) and centrifuged for 1 minute at 14,100 g to capture the plasmid DNA. The spin column was washed twice in wash buffer and plasmid DNA was eluted in 50 µL of distilled H_2O. The success of purification was assessed by electrophoresing 5 µL of plasmid DNA mixed with 1 µL of 6X xylene cyanol on a 1.0% (w/v) agarose gels. The concentration of each purified plasmid was determined using UV spectrophotometry. Purified plasmids were stored at -20 ºC.

2.17 Preparation of samples for DNA sequencing

Generally, DNA samples were prepared by mixing 300 ng of purified plasmid DNA, 3.2 pmol of a vector sequence based primer (either reverse primer SP6: 5‟-ATTTAGGTGACACTATAGAATAC -3‟; or forward primer T7: 5‟-GTAATACGACTCACTATAGGGC -3‟), 1.0 µL (for inserts 500-1,000 bp in length) of ABI BigDye Terminator (BDT) version 3.1 (Applied Biosystems; California, USA), and 3.5 µL of 5X BDT Dilution Buffer. The mixture was diluted to 20 µL with dH2O. The amount of BDT was increased to 1.5 µL for large inserts (≥1,000 bp), and reduced to 0.5 µL for small inserts (<500 bp). Cyclic sequencing reactions were conducted in a Bio-Rad MyCycler™, where the reactions were heated at 96ºC for 1 minute, then subjected to 35 cycles of 96ºC for 10 seconds, 50 ºC for 5 seconds and 60ºC for 4 minutes, as per the instructions of AGRF (Australian Genome Research Facility Ltd,

Melbourne, Australia; http://www.agrf.org.au/docstore/seq/How_to_Prepare_Your_Capillary_Separation_(CS)_Samples.pdf; accessed May 2014). The reactions were then precipitated using the magnesium sulphate clean-up protocol specified by AGRF. Briefly, the labelled products were mixed with 75 µL of 0.2 mM MgSO4 Stock Solution (in 70% ethanol) and incubated at RT for 15 minutes, followed by a 15 minute centrifugation at 14,100 g. The air-dried pellets were sent to AGRF (Melbourne) for capillary separation using a 3730xl DNA Analyzer (Applied Biosystems; California, USA).

2.18 Analysis of DNA sequencing chromatograms

Chromatograms were examined in the BioEdit Sequence Alignment Editor (Hall 1999; http://www.mbio.ncsu.edu/BioEdit/page2.html; last accessed May 2014) to check the quality of sequencing. Only clean chromatograms were used with minimal baseline noise and evenly spaced peaks, as well as satisfactory signal intensity (The University of Michigan, 'Interpretation of Sequencing Chromatograms', http://seqcore.brcf.med.umich.edu/doc/dnaseq/interpret.html; last accessed May 2013). The sequences were trimmed when necessary if separation between peaks was irregular or no longer evident. Plasmids were generally sequenced in both the forward and reverse directions and the data used to create a consensus sequence. Any ambiguous peaks were resolved by repeated sequencing.

2.19 Physiological analysis

2.19.1 Measurement of biomass, leaf water content and Na^+ and K^+ ion analysis

The barley plants were harvested as explained in Section 2.3.3. The shoot and root length was measured, then separated and the root and shoot fresh weight measured immediately. The roots were dried at 80 ºC for 48 h to record the dry weight. The shoots were then floated on water at room temperature for 48 h and turgor weight measured, followed by drying at 80 ºC for 48 h to record the dry weight. The relative water content of the shoots was determined as per Turner (1981) as RWC= (FW-DW)/ (TW-DW) (where FW= fresh weight, TW = turgor weight, DW = dry weight). The dried root and shoot tissues were then ashed at 550 ºC for 12 h. Five mL of 0.5 M HNO_3 was added to the whole ash and Na^+ and K^+ ion concentration determined by flame atomic absorption spectrometry (Varian, United States) using 9 standards ranging

from 0 to 25 µg/mL of [Na^+] or [K^+] ions, as per Munns et al. (2010). The relative water content stress tolerance index was calculated as RWCI = (RWC stressed plant / RWC control plant) ×100, Na^+ ion stress tolerance index as NaI= (Na^+ ion stressed plant / Na^+ control plant) ×100, K^+ ion stress tolerance index as KI = (K^+ ion stressed plant / K^+ control plant) ×100 and Na^+/K^+ ion stress tolerance index as Na/KI = (Na^+/K^+ ion ratio stressed plant / Na^+/K^+ ratio control plant) ×100 (Kausar et al. 2012).

2.19.2 Extraction and estimation of ABA from leaf tissue

The roots and shoots of the barley plants were dried at 80 ºC for 48 h. Eighty percent aqueous methanol was added to dried tissue of the plants and incubated at 4 ºC for 24 h in darkness. This was followed by addition of 100 % methanol and 10 mg/L butylated hydroxytoluene for 24 h and 6 h respectively at 4 ºC. The extracts were dried under nitrogen and the residue dissolved in 50 µL methanol and 200 µL of Tris-buffered saline (Popova et al. 1996). The enzyme-linked immunosorbent assay (ELISA) was performed in 96-well microtitre plates according to a protocol modified from Yang et al. (2001). The assay uses the rat monoclonal IgG2a antibody (cat no. Ab50594, Abcam, Australia) to S-*cis, trans* abscisic acid and rabbit polyclonal secondary polyclonal antibody to rat IgG (cat no. Ab6730 Abcam, Australia). Briefly, each well was coated with 200 µL primary antibody diluted in sodium bicarbonate buffer (1.5 g Na_2CO_3, 2.93 g $NaHCO_3$, 0.02 g NaN_3 per litre, pH 9.6) and the plate incubated overnight at 4 ºC. The coating solution was removed and plate was washed twice with 200 µL of wash buffer per well [phosphate buffered saline; PBS; 8.0 g NaCl, 0.2 g KCl, 1.44 g $Na_2HPO_4$and 0.24 g KH_2PO_4 per litre, pH 7.4 with Tween 20 (0.1 % v/v in final concentration)]. Any remaining protein binding sites were blocked by adding 200 µL of blocking buffer per well (2 % bovine serum albumin in PBS). The plate was then incubated for 1.5 h. Commercially available ABA ($C_{15}H_{20}O_4$; Sigma, Australia) was used to prepare standards ranging from 0 to 200 ng/mL in methanol to prepare a standard curve. The standard solutions and various dilutions of the ABA (extracted as above) (100 µL each) were added in duplicate to the plate and incubated for 1.5 h at 37 ºC. The plate was then washed twice with wash buffer, followed by addition of 100 µL of the diluted secondary antibody diluted in sodium bicarbonate buffer per well and incubation for 2 h. The plate was then washed twice with wash buffer. 100 µL of IgG-horseradish peroxidase (Cat# 554058, BD Biosciences Australia) was added to the plate

and then incubated at 30 ºC for 1 h. After washing with wash buffer, 100 µL of substrate solution (1.5 mg/mL O-phenylenediamine and 0.008 % H_2O_2) per well was then added. The reaction was stopped after 30 mins by adding 100 µL of 6 N H_2SO_4 and colour development was detected at 490 nm using a microplate reader (BioRad, Australia). ABA accumulation stress tolerance index (ABAI) was calculated as ABAI = ([ABA] Stressed plant / [ABA] Control plant) ×100.

2.20 Statistical analysis Except where otherwise stated, experiments were performed in triplicates. The data were presented as mean values ± standard deviation (mean ± SD) unless otherwise stated. The statistical significance was evaluated using one-sided Student's t test. A *p value* of <0.05 was considered statistically significant.

Bioinformatics Methods

2.21 Retrieval of rice and Arabidopsis PP2C, PYR/PYL/RCAR, SnRK2 (subfamily II and III) and ABA-transport related ABC transporter gene and protein sequences

The gene loci and putative protein accession numbers of the PP2C multigene family in Arabidopsis and rice, both were retrieved from Xue et al. (2008). The Arabidopsis accession numbers were then used to extract the sequences from the 'Gene' and 'Protein' sections of the Arabidopsis information resource (TAIR) database (http://www.arabidopsis.org/; last accessed June 2014). The rice loci were used to extract the corresponding gene and protein sequences from the MSU rice genome annotation project (http://rice.plantbiology.msu.edu/; last accessed June 2014) (originally The Institute for Genomic Research (TIGR) rice genome annotation project (http://rice.tigr.org; Ouyang et al. 2007)). Similarly, the Arabidopsis accession numbers previously reported for PYR/PYL/RCAR (Ma et al. 2009; Park et al. 2009), SnRK2 subfamily II and III (Yoshida et al. 2002; Hrabak et al. 2003) and ABA transport related ABC transporters (ABCG25 and ABCG40; Kuromori et al. 2010; Kang et al. 2010) were used to extract the sequences from the 'Gene' and 'Protein' sections of the TAIR database. The gene and putative protein sequences in rice were retrieved by use of BLASTn and BLASTx tool in the MSU rice genome annotation project using the reported Arabidopsis gene sequences.

The genomic location for all genes was identified through the 'gene structure' function in the MSU rice genome annotation project database (http://rice.plantbiology.msu.edu/; last accessed June 2014).

2.22 Method development for the identification of ABA-signalling related candidate genes from the mRNA-seq dataset of an organism without reference genome

The International Barley Genome Sequencing Consortium (IBSC) database was not available at the start of the project, and the purpose was to generate a transcriptome for non-model organisms. mRNA-seq I dataset developed from control and salt-treated leaf tissue of barley (*Hordeum vulgare*) *cv.* Hindmarsh (Section 2.6). The dataset had over 21 thousand known Unigene sequences and around four thousand novel tentative consensuses (NTC) sequences. A combination of methods were utilised to search this dataset for the PP2CA, PP2CD, PYR/PYL/RCAR, SnRK2 (subfamily II and III), ABCG25 and ABCG40 candidate genes, as follows:

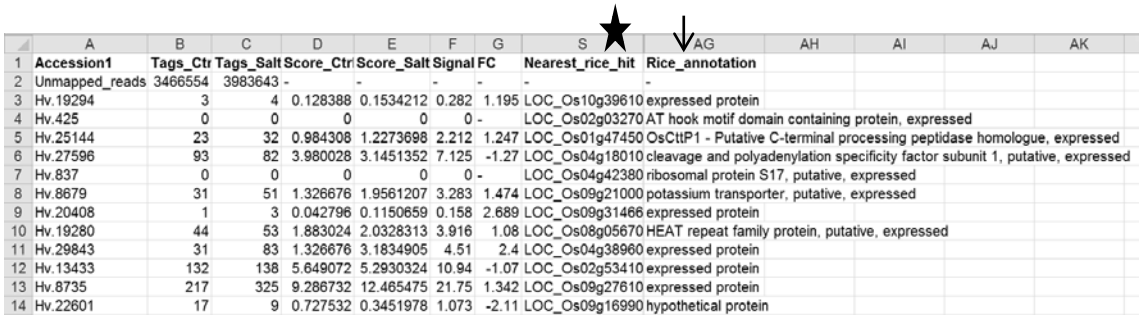
2.22.1 Method A: Keyword searches of the barley mRNA-seq dataset

Searches of the mRNA-seq I dataset were conducted using the terms such as 'pyrabactin resistance (PYR)/PYR-like (PYL)/regulatory components of ABA receptor (RCAR)', 'protein phosphatase', 'SnRK2'and 'ABC transporter' (of UNIX-like operating system family). The following command was used to sort the genes for a particular search and add into a specific file: grep – i 'A' NGS_data.txt > A_novel.txt (where A: search term specific to a gene family; NGS_data.txt: file having NGS mRNA-seq data; A_novel.txt: the file with data for a specific gene family (A); grep, -i, >: Ubuntu commands). The different .txt files were converted into .xlx format and the data merged into a single file. Likewise, to search for specific sub-families, the terms such as 'SnRK2 subfamily III', 'ABC transporters type G', 'ABCG25' etc. were used.

2.22.2 Method B: Search ABA related rice loci

In mRNA-seq I dataset every barley sequence had a orthologue rice locus listed against it. Screen shot of mRNA-seq datasheet is shown in Figure 2.3. The above extracted rice locus for ABA-signalling pathway related genes were searched in the barley

mRNA-seq I dataset. The barley sequences from the mRNA-seq I dataset, which are orthologue to selected rice sequences, were selected for further analysis.



	A	B	C	D	E	F	G	S	AG	AH	AI	AJ	AK
1	Accession1	Tags_Ctr	Tags_Salt	Score_Ctr	Score_Salt	Signal	FC	Nearest_rice_hit	Rice_annotation				
2	Unmapped_reads	3466554	3983643	-	-	-	-	-	-				
3	Hv.19294	3	4	0.128388	0.1534212	0.282	1.195	LOC_Os10g39610	expressed protein				
4	Hv.425	0	0	0	0	0	-	LOC_Os02g03270	AT hook motif domain containing protein, expressed				
5	Hv.25144	23	32	0.984308	1.2273698	2.212	1.247	LOC_Os01g47450	OsCttP1 - Putative C-terminal processing peptidase homologue, expressed				
6	Hv.27596	93	82	3.980028	3.1451352	7.125	-1.27	LOC_Os04g18010	cleavage and polyadenylation specificity factor subunit 1, putative, expressed				
7	Hv.837	0	0	0	0	0	-	LOC_Os04g42380	ribosomal protein S17, putative, expressed				
8	Hv.8679	31	51	1.326676	1.9561207	3.283	1.474	LOC_Os09g21000	potassium transporter, putative, expressed				
9	Hv.20408	1	3	0.042796	0.1150659	0.158	2.689	LOC_Os09g31466	expressed protein				
10	Hv.19280	44	53	1.883024	2.0328313	3.916	1.08	LOC_Os08g05670	HEAT repeat family protein, putative, expressed				
11	Hv.29843	31	83	1.326676	3.1834905	4.51	2.4	LOC_Os04g38960	expressed protein				
12	Hv.13433	132	138	5.649072	5.2930324	10.94	-1.07	LOC_Os02g53410	expressed protein				
13	Hv.8735	217	325	9.286732	12.465475	21.75	1.342	LOC_Os09g27610	expressed protein				
14	Hv.22601	17	9	0.727532	0.3451978	1.073	-2.11	LOC_Os09g16990	hypothetical protein				

Figure 2.3 Screenshots of the barley leaf mRNA-seq I dataset listing
The annotations column is shown using the arrow while the rice loci column is shown using the star.

2.22.3 Method C: Reciprocal BLAST hit (RBH) search

The barley mRNA-seq I dataset was searched using the ABA-signalling related protein sequences of rice (obtained as above) as the input data, for the RBH (Ziemann *et al*. 2013b). Briefly, a tBLASTn search was performed on each of the identified rice protein sequences against the barley mRNA-seq I experiment 76 nucleotide (nt) read data. The extracted mRNA-seq I reads were called candidate barley reads, and these were re-BLASTed (reciprocal BLAST) against the MSU rice genome annotation project database (http://rice.plantbiology.msu.edu/). The mRNA-seq I reads which extracted the same rice locus in reciprocal BLAST as the one used in input of initial BLAST were called 'RBH best match'. The RBH best match reads were then assembled to form contigs using CAP3. The assembled contigs were further BLASTed against mRNA-seq dataset to identify the 'known Unigene' and 'novel NTC' of interest in mRNA-seq I dataset.

2.23 Extraction of barley IBSC accession numbers for the ABA-signalling related candidate genes identified in the mRNA-seq I (Unigene and 'novel') dataset

With the release of barley sequencing database in 2012, the ABA signalling related sequences derived from mRNA-seq I as above were BLASTed in the IBSC BLAST tool (http://webblast.ipk-gatersleben.de/barley/viroBlast.php; last accessed June 2014). The IBSC accession numbers will be used further, for ease of comparisons. However, it should be noted that the fold change values derived from the original BLAST of 76 nt reads (mRNA-seq I) against NCBI Unigene dataset could be slightly different from those if the reads had been directly BLASTed against IBSC CDS database, as there is

difference in the length of sequences among IBSC and NCBI databases. It was not possible to do this, as the IBSC database was not available at the time, and the purpose was to generate a transcriptome for non-model organisms.

2.24 Identification of ABA-signalling related candidate genes from the IBSC CDS database

The complete list of PP2CA, PP2CD, PYR/PYL/RCAR, SnRK2 (subfamily II and III), ABCG25 and ABCG40 sequences in barley was extracted from barley IBSC CDS database by using the RBH method. Briefly, a tBLASTn search was performed on each of the identified rice protein sequences against the barley IBSC CDS database (http://mips.helmholtz-muenchen.de/plant/barley/index.jsp/; last accessed June 2014). The extracted barley CDS sequences were called candidate barley sequences, and these were re-BLASTed (reciprocal BLAST) against the MSU rice genome annotation project database (http://rice.plantbiology.msu.edu/). The barley CDS sequences which extracted the same rice locus in reciprocal BLAST as the one used in input of initial BLAST were called 'RBH best match' and identified for further analysis.

2.25 Analysis of important domains and residues in the putative proteins and grouping into sub-families

The CDS of the ABA-signalling related genes extracted from the barley IBSC CDS database were translated in all six open reading frames (ORF) using six frame translation function in BioEdit (http://www.mbio.ncsu.edu/BioEdit/BioEdit.html) and Generunner (http://www.generunner.net/). The correct reading frame was selected by conducting BLASTx of the MSU rice genome (http://rice.plantbiology.msu.edu/analyses_search_blast.shtml; last accessed June 2014) and NCBI database (http://blast.ncbi.nlm.nih.gov/Blast.cgi; last assessed June 2014). The translated sequences were then analysed for likely functional motifs and other characteristic features by aligning against the rice and Arabidopsis orthologues using ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalW22/). The % identity was also noted from ClustalW2. The sequences were assigned to respective subfamilies using the identities and relationships with rice and Arabidopsis orthologues.

2.26 Prediction molecular weight and isoelectric point of proteins

Molecular weight (MW) and isoelectric point (pI) of individual proteins were predicted from amino acid sequences in the ExPASy proteomics server (http://web.expasy.org/compute_pi/; last accessed June 2014).

2.27 Prediction of chromosomal location and gene exon-intron structure

The chromosomal location and genomic DNA of the barley sequences was derived by running a BLAST against IBSC barley database (http://webblast.ipk-gatersleben.de/barley/viroBlast.php; last assessed June 2014). The extracted genomic sequence along with CDS sequences were used to generate exon-intron diagram of barley genes using gene structure display server (http://gsds.cbi.pku.edu.cn/; last accessed June 2014).

2.28 Amino acid consensus generation

The weblogo browser (Schneider and Stephens 1990) (http://weblogo.threeplusone.com/create.cgi; last accessed June 2014) was utilised to generate sequence logos, which are graphical representations of amino acids at a particular position in a multiple alignment. The sequence logo consists of 'packs of symbols', one pack for each position in the sequence. The overall height of the pack indicates the sequence conservation at that position, while the height of symbol within a pack indicates the relative frequency of each amino at that position (http://weblogo.berkeley.edu/). The amino acid alignment for each conserved motif of rice and barley was entered separately into the submission form to generate the sequence logo, which displayed the relative frequency of amino acids.

2.29 Identification of barley ABA-signalling related candidate genes in mRNA-seq II dataset

mRNA-seq II dataset developed from control and salt, drought and exogenous ABA-treated leaf tissue of barley (*Hordeum vulgare*) *cv.* Hindmarsh, had over 20 thousand barley IBSC CDS sequences (Section 2.7). The ABA-signalling related candidate genes identified from the IBSC CDS database (Section 2.23) were searched in the mRNA-seq II dataset to determine the differential expression of above genes in abiotic stress condition.

Chapter 3

Analyses of the barley leaf transcriptomes using mRNA-seq

3.0 Abstract

Salinity and drought are major threats to crops in many parts of the world, and it is predicted to be a serious constraint to food security. However, understanding the impact of these stressors on plants is a major challenge due to the involvement of numerous genes and regulatory pathways. While transcriptomic analyses of barley (*Hordeum vulgare* L.) under salt and drought stresses have been reported with microarrays, there are no reports as yet of the use of mRNA-seq. We demonstrate the utility of mRNA-seq for transcriptome analysis of species without whole genome sequenced data by analysing cDNA libraries derived from acutely salt-stressed and unstressed leaf material of *H. vulgare* cv. Hindmarsh (barley whole genome sequence has only recently been described; International Barley Genome Sequencing Consortium (IBSC) *et al*. 2012). The data yielded more than 50 million sequence tags which aligned to 26,944 sequences in the Unigene reference database. To gain maximum information, we performed de novo assembly of unaligned reads and discovered more than 3,800 contigs, termed novel tentative consensus sequences, which are either new, or significant improvements on current databases. Another mRNA-seq study was employed to analyse the cDNA libraries derived from acutely salt, drought and exogenous stressed and unstressed leaf material of *H. vulgare* cv. Hindmarsh. The data yielded more than 100 million sequence tags which aligned to 20,537, 20,702 and 20,709 sequences in the IBSC CDS database for salt, drought and exogenous ABA stressed plants respectively. Around 90% of the genes showing differential expression for salinity or drought were also differentially expressed in exogenous ABA stressed plants. Genes such as chlorophyll a/b binding protein, aquaporins and MYB transcription factors were among the most differentially regulated during salt stress, whereas serine threonine kinase, chlorophyll a/b binding protein and MYB transcription factors were among the most differentially regulated during drought stress. Analysis of gene ontology has demonstrated that maximum number of genes showed binding as molecular function in all three stress conditions. The work provides comprehensive insights into genome-wide effects of salinity, drought and exogenous ABA application and is a new resource for the study of gene regulation in barley. Further, the bioinformatics workflow may be applicable to other non-model plants to establish their transcriptomes and identify unique sequences.

3.1 Introduction

Salt and drought stresses are the two most important environmental stresses which limit plant growth and development (explained in Section 1.1.6). A comprehensive understanding of the roles of individual genes, their transcripts including alternative splice forms, their protein products, as well as the 'sum' of all pathways that plants use to manage abiotic stresses need to grow for understanding the crop responses to drought and salinity. Despite being staple foods around the world, the elucidation of the complete genome sequence of wheat has been hindered by the complexity of its genome, while a draft barley whole genome sequence has only recently been described (International Barley Genome Sequencing Consortium *et al*. 2012).

The next-generation mRNA-seq, a high-throughput cDNA sequencing technology, is a powerful method for rapid characterisation of transcript sequences and gene expression levels in biological samples (explained in section 1.10). It is being applied widely in human genetics and medicine, but is still an emerging technology for plants. Marioni *et al*. (2008) critically evaluated gene expression profiling by RNA-seq by the Illumina platform to that by Affymetrix arrays from the same RNA samples, and concluded that RNA-seq was not only comparable in elucidating differentially expressed genes, but also had added capabilities of detecting transcripts with low-level expression, identifying sequence variants and new transcripts. Transcriptome analysis from short-read Illumina sequencing is now beginning to be carried out for crop species, e.g. rice (Mizuno *et al*. 2010; Zhang *et al*. 2010; whole genome sequenced in 2004), maize (Kakumanu *et al*. 2012; whole genome sequenced in 2009), cucumber (Guo *et al*. 2010; whole genome sequenced in 2009), grapewine (Zenoni *et al*. 2010; whole genome sequenced in 2007) which have the advantage of reference whole genome data, and also species such as chickpea (Hiremath *et al*. 2011; genome sequenced in 2013), and soybean (Severin *et al*. 2010; genome sequenced in 2010) and wheat (IBWSC 2014; genome not sequenced) without such information. By the FAO (2005) classification of salinity tolerance, both corn and soybean are moderately tolerant, wheat is tolerant, while barley is classified as 'highly tolerant'; hence, it may display important genetic attributes under salt challenge. The cultivar Hindmarsh was chosen for transcriptome analysis because it is among the most widely cultivated barley varieties in Australia (GRDC 2013; https://www.grdc.com.au/Research-and-

Development/NationalVarietyTrials/~/media/C9C378CA74BB45BDA89321333AB6FEB2.pdf; p27) and is particularly suited to regions of south-eastern Australia with lower rainfall (Modra Seeds Fact Sheet). This analysis aims to identify the genes which may confer resistance to salinity and drought stress and may thereby be candidates for future crop improvement where abiotic stress poses an increasing problem.

3.2 Results

Part A: Results of the next generation mRNA-sequencing experiment I (mRNA-seq I): analysis of barley leaf transcriptome under salt stress, for method development for un-sequenced genomes

This initial phase of work was undertaken in 2010. The main aim was to investigate utility of mRNA-seq for transcriptome analysis of species without whole genome sequenced data as the International Barley Genome Sequencing Consortium (IBSC) barley genome was not available at the time. mRNA-seq of control and acutely-stressed leaf tissue was conducted, as this is the site of photosynthesis and any stress effects have a direct relation to crop yield.

3.2.1 Results of total RNA extraction

The barley leaf mRNA-seq I transcriptome was developed using total RNA pooled from two control plants and total RNA pooled from two salt-stressed plants of barley cv. Hindmarsh. Firstly, total RNA was extracted from the single plant using the TRIsure reagent for tissue lysis. The method used for total RNA extraction is detailed in section 2.4.2. The quality and integrity of RNA was checked by running an aliquot of extracted RNA on MultiNA, microchip electrophoresis system (Section 2.9.2). The examples of electropherogram and gel images of total RNA used for mRNA-seq I (generated by bioanalyser) are shown in Figure 3.1 and 3.2. The total RNA was deemed suitable for further analysis by mRNA-seq if electropherogram showed two distinct ribosomal RNA peaks corresponding to both 18S and 28S for eurokaryotic RNA. The spectrophotometric quantification of total RNA extracted was also performed (as explained in section 2.9.4), yielding approximately 5 µg of total RNA per plant sample. This stage of the work was shared with Dr Runyararo Memory Hove (SUT), who was then a fellow PhD student.

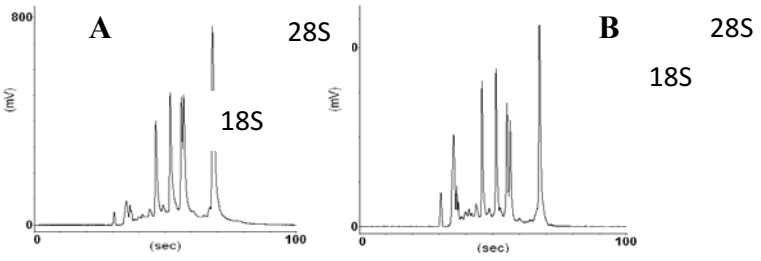


Figure 3.1 Example of electropherogram for quality test of total RNA used for mRNA-seq I

The electrophoergram were generated by the MultiNA, microchip electrophoresis system. A: total RNA from control plant; B: total RNA from salt stressed plant

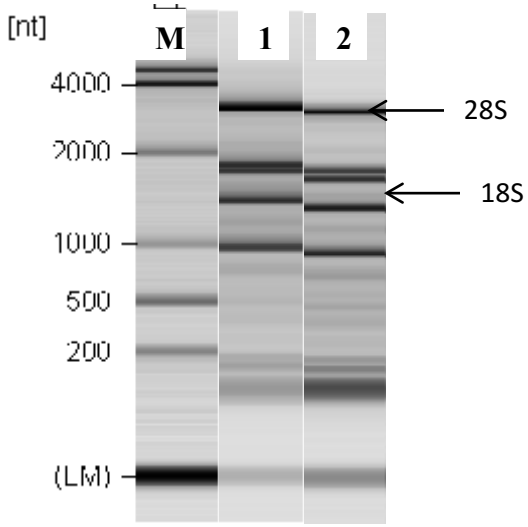


Figure 3.2 Example of gel images for quality test of total RNA used for mRNA-seq I

The gel images were generated by the MultiNA, microchip electrophoresis system. M: RNA 6000 ladder (Applied Biosystems) 1: total RNA from control plant; 2: total RNA from salt stressed plant

3.2.2 Generation of mRNA-seq I libraries and acquisition of the transcriptome dataset

Using the barley (cv. Hindmarsh) leaf total RNAs (section 3.2.1), two mRNA-seq libraries were prepared: (i), one from 0.5 µg each of total RNA pooled from two separate plants treated with 150 mM NaCl for 12 h (Seki et al. 2002), and (ii) likewise, total RNA pooled from two control plants. The pooling was done based on other studies (Mizuno *et al*. 2010; Ando and Grumet 2010), in order to minimize the effects of any biological variations between transcriptomes on the data and its interpretations. The libraries were each loaded on one lane of Illumina Genome Analyser IIx (section 2.6.1). The mRNA-seq results yielded 23,735,138, 76 nucleotide (nt) sequence reads from control and 26,685,685 from salt-treated library (Table 3.1). The reads were curated for any artifacts using the Fastx Artifacts Filter and any poor quality bases were removed from the 3′ end with a FastQ Quality Trimmer using a threshold of Q30 (http://hannonlab.cshl.edu/fastx_toolkit/) as described in sections 2.6.2. At the beginning of this study the barley genome was not sequenced, however the sequenced genome was released during the progress of this project (October 2012; The International Barley Genome Sequencing Consortium (IBSC), 2012). So theses reads were aligned against the NCBI Unigene database using Burrows-Wheeler Aligner (BWA; Li and Durbin 2009) with default settings. The remaining unaligned reads underwent filtering and then assembly using ABySS to discover 'novel' transcripts

(NTCs). This stage of the work was shared with Dr Runyararo Memory Hove (SUT), who was then a fellow PhD student.

By aligning 23.4 million and 26.1 million Q30 quality trimmed mRNA-seq reads from control and salt stressed plants to the NCBI Unigene database (reference genome) using Burrows-Wheeler Aligner (BWA) software (Li and Durbin 2009), it was found that significant proportion of reads 16.6 million from control and 16.4 million from salt-treated library (70.9 % control, 63.0 % salt) were mappable to the database (Table 3.1). From analysing this dataset, it was found that, of the 26,944 sequences present in the NCBI barley Unigene database, 21,336 transcripts were detected in the control and 21,574 in salt-stressed sample. The sequences which aligned successfully were designated as 'known UniGenes'. The selection criterion of signal strength ≥10 was applied to filter out the lowly expressed transcripts reducing the number of 'known UniGenes' to be 12,236 (datasheets can be provided on request).

Table 3.1 mRNA-seq I data yield from Genome Analyzer IIx sequencing.

	Control sample	Salt stress sample	Total
Original read length (nt)	76	76	76
Original number of reads	23,735,138	26,685,685	50,420,823
Number of reads after Q30 quality filtering (bp)	23,366,670	26,072,011	49,438,681
Sequence yield after Q30 quality filtering (bp)	1,701,749,555	1,863,246,963	3,564,996,518
Number of reads aligning to Unigene DB	16,573,963	16,430,198	33,004,161
% Reads aligned	70.93	63.02	66.76
Number of unaligned reads	6,792,707	9,641,813	16,434,520
Unaligned sequence (bp)	495,718,339	697,848,043	1,193,566,382

[a]: maximum read length can be obtained is 100bp using illumina genome analyser with low error rate; in this experiment it was set to 76 nt.[b]: total reads obtained from the salt stressed pants and the plant grown as control. [c]: Reads were aligned against barley Unigene database (http://www.ncbi.nlm.nih.gov/unigene; last accessed June 2014) using BWA algorithm, with default settings. Unaligned reads were defined in the form of tentative consensus using AbySS.

To discover 'novel' transcripts, the remaining 16,434,520 unaligned reads underwent a two-step de-novo assembly in ABySS (Table 3.1). A 18 assemblies of k-mer lengths (k-mer range 27, 29, etc. up to 63) generated 5,723,131 overlapping contigs. These contigs then underwent second phase of assembly, with the average contig length and N50 length reducing the number of transcripts to less than 50,000 (data not shown). The k55 assembly for downstream analyses was selected for this study, which yielded 39,707 contigs with length greater than 100bp. The average length of transcripts was 343.9 bp and the N50 length was 518bp (Table 3.2; Figure 3.3). The longest contig

(NTC21595; 13,710 bp) putatively encodes an auxin transport protein of the 'BIG-like' family, based on homology to a Brachypodium cDNA sequence.

Table 3.2 Results from the two-phase assembly using AbySS

Phase 1 of assembly	
Number of unmerged contigs from k27 to k63 assembly	5,723,131
Number of unmerged contigs C100 bp	954,420
Average length (bp)	235
N50 length (bp)	256
Longest contig (bp)	12,314
Phase 2 of assembly (k-mer = 55)	
Number of merged contigs	50,499
Number of merged contigs C100 bp	39,707
Average length (bp)	344
N50 length (bp)	518
Longest contig (bp)	13,710
Assembly size (bp)	13,696,077

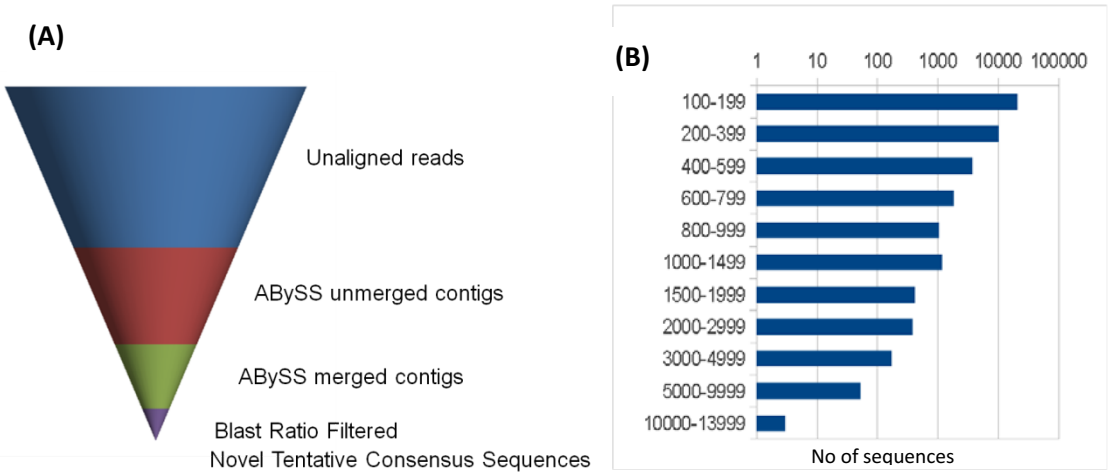


Figure 3.3 Process of identifying novel sequences from mRNA-seq data.
(A)The schematic consolidation of data by assembly of reads merging of overlapping contigs, filtering by BLAST ratio; (B) Length distribution for final assembly using k-mer of 55 was selected for further analysis.

3.2.3 Optimisation of selection criterion

To hone in on potentially novel sequences, a rice/barley BLASTN ratio was implemented as a potentially appropriate selection criterion with an idea to select the NTCs having strong rice and poor barley alignments. Three spread sheets were constructed from the original listing of NTCs after applying the rice/barley criterion of ≥3, ≥ 2, and < 2 respectively (data not shown). After applying the ≥3 for rice/barley

BLAST ratio as the selection criterion 2,607 NTCs were filtered in as potentially novel tentative consensus (NTCs), whereas the number of these sequence increased to 3,828 and 35,856 in the case of ≥ 2 and < 2 rice/barley BLAST ratio respectively (Table 3.3).

Table 3.3 Optimisation of rice/barley BLASTN ratio

Criterion	Rice/barley BLASTN ratio		
	<2	≥2	≥3
No of 'known' Unigenes	Known genes were not aligned to rice		
No of 'novel' TCs	35,856	3,828	2,607

The lists were sorted in the Microsoft excel in the ascending order with respect to the selection criterion and the numbers were then counted

Sequences were then randomly selected from all of the 3 spread sheets created for 'novel' sequences. The selected sequences were BLASTed in NCBI BLASTN tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi; last accessed, June 2014), using rice and barley as the selected organism in different BLASTN search (Table 3.4). Sequences with rice/barley BLASTN ratio less than 2 had high percent identity against the barley sequence as shown by NTC119230 (rice/barley ratio: 1.70; Table 3.4), which does not leave them as authentic 'novel' sequences anymore. The BLASTN results of sequences with the ratio of ≥ 2 and ≥ 3, both exhibited significant coverage and % identity to rice, whereas none of the sequences were identified as close match in barley. It was felt that perhaps ratio of ≥ 3 may be too stringent and some potentially unique sequences may be ignored. So rice/barley BLASTN ratio of ≥ 2 was decided to be adequate at this stage, which removed 90.3% of contigs, resulting in 3,828 potentially novel tentative consensus (NTCs). The selection criterion of signal strength ≥ 10 was further applied to filter out the lowly expressed transcripts reducing the final number of potential NTCs to be 2,144.

Table 3.4 BLASTN result of selected novel transcripts.

NTC No	Rice/barley BLAST ratio	Alignment with rice sequences in msu database		Alignment with barley sequences in Unigene database	
		% coverage	% identity	% coverage	% identity
NTC140024	3.198708	98%	86%	-	-
NTC360351	2.093985	90%	90%	-	-
NTC270220	1.993103	99%	93%	-	-
NTC165620	1.883495	62%	74%	-	-
NTC119230	1.704301	51%	70%	52%	100%

BLAST search did not found any sequence; BLASTN search was conducted using the BLAST tool in NCBI (http://BLASTt.ncbi.nlm.nih.gov/BLAST.cgi).

3.2.4 Analysis of differential gene expression

Quality trimmed reads from control and salt stressed plants were aligned to NCBI Unigene database, which led to decrease in number of unaligned reads from 16,434,520 to 13,505,967. DESeq was used to scan for differential gene expression between control and salinity stress in leaves using a negative binomial model. One hundred and ten transcripts were found to be significantly deregulated (FDR adjusted p value <0.05). From these, 48 transcripts showed up-regulation and 62 showed down-regulations. The top 20 differentially expressed transcripts (up and down-regulated) ranked by the fold change from the barley Unigene and NTC sets are shown in Table 3.5. The list of up-regulated genes includes a number of genes (or homologue) which have been shown previously to mediate osmotic/drought/salinity tolerance, such as lipoxygenase 2.1, cellulose synthase like protein, protein phosphatase 2C, dehydrin, calcium dependent protein kinase as well as those encoding membrane bound proteins such as a peptide transporter, two plasma membrane ATPases and a novel wall-associated receptor kinase. Down-regulated transcripts include those in the Jumonji, heat shock protein, Pumilio RNA binding and MYB transcription factor classes, as also several transcripts of unknown function. Few uncharacterised transcripts were also found to be differentially regulated such as Hv.5008 and Hv.29473 etc. These differentially regulated genes can be the candidate tolerance genes that need to be further examined using reverse genetic approaches.

Of the 26,941 transcripts present in the Unigene database, 26 sequences demonstrated an increase of ≥ 10 in differential expression, while 2,514 sequences exhibited a fold change of +1.5-10X. Over seven thousand transcripts showed differential expression ranging from -1.5 to +1.5X, whereas only 2,218 sequences had a decrease in expression ranging from -10 to -1.5X during salt stress condition. Thirty three transcripts exhibited a down-regulation of more than 10 fold (Figure 3.4A). Of the novel tentative consensus (NTC), only 1 sequence was up-regulated by >10 fold, while 267 showed an up-regulation of 1.5 to 10X. Only 903 NTCs exhibited fold change of -1.5 to 1.5X. Whereas 972 'NTCs' had a decrease in expression ranging from -1.5 to -1.5 X during salt stress condition, while only 1 sequence demonstrated a down-regulation of more than 10 fold (Figure 3.4B).

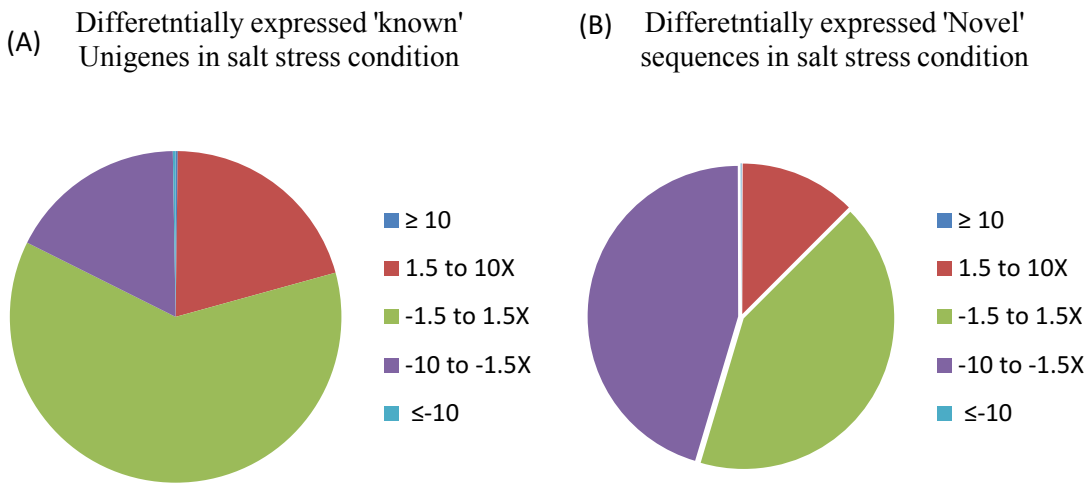


Figure 3.4 Differential expression of sequences in mRNA-seq I data salt as stress condition.

(A): 'known' Unigene sequences; (B); 'novel' TC sequences; The data was sorted using the Microsoft excel in an ascending order with respect to the fold change value (differential expression) and then graphs were drawn using the numbers.

Table 3.5 Top 20 up-regulated and down-regulated transcripts ranked by fold change (mRNA-seq I)

Accession	Reads control	Reads salt	Fold Change	Nearest Rice BLAST hit	Annotation
Hv.29838	20	827	39.40	LOC_Os07g36750	CSLF3-cellulose synthase like family F; beta1,3; 1,4glucansynthase, expressed
Hv.5008	125	3,701	29.86	–	Unclassified transcript
Hv.31363	4	224	55.72	LOC_Os12g37260	Lipoxygenase2.1, chloroplast precursor, putative, expressed
Hv.8934	29	640	21.11	LOC_Os04g40990	Malate synthase, glyoxysomal, putative, expressed
Hv.17368	44	843	18.38	LOC_Os05g46040	Protein phosphatase2C, putative, expressed
Hv.2654	9	261	27.86	LOC_Os07g44060	Halo acid dehalogenase-like hydrolase family protein, putative, expressed
Hv.29473	10	264	25.99	–	Unclassified transcript
Hv.3400	2	137	68.59	LOC_Os01g12580	Late embryogenesis abundant protein, putative, expressed
Hv.30848	3	149	48.50	LOC_Os03g48310	Plasma membrane ATPase, putative, expressed
Hv.32578	4	164	39.40	LOC_Os04g02000	Zinc finger family protein, putative, expressed
Hv.17120	40	603	14.93	LOC_Os10g41490	CAMK_CAMK_like.41—CAMK includes calcium/calmodulin dependent protein kinases, expressed
Hv.15443	41	566	13.93	LOC_Os04g47700	Expressed protein
NTC25482	58	754	13.00	LOC_Os02g42110	Wall-associated receptor kinase-like 22 precursor, putative, expressed
Hv.32190	6	172	27.86	LOC_Os03g19600	Retro transposon protein, putative, unclassified, expressed
Hv.5085	11	229	21.11	LOC_Os07g05365	PhotosystemII 10k Dapolypeptide, chloroplast precursor, putative, expressed
Hv.30861	2	104	51.98	LOC_Os03g48310	Plasma membrane ATPase, putative, expressed
Hv.10528	234	2,825	12.13	LOC_Os09g35880	B-box zinc finger family protein, putative, expressed
Hv.5729	1	56	27.86	LOC_Os06g38294	Peptide transporter PTR2, putative, expressed
NTC26185	3	230	16.00	LOC_Os09g25700	TsetseEP precursor, putative, expressed
Hv.12388	89	976	10.56	LOC_Os09g02180	Expressed protein
Hv.16656	5,229	25	-207.94	LOC_Os12g31000	Pumilio-family RNA binding repeat domain-containing protein, expressed
Hv.2383	1,413	10	-147.03	LOC_Os10g25060	Expressed protein
Hv.6975	798	5	-157.59	LOC_Os04g47140	Expressed protein
Hv.33010	382	2	-194.01	LOC_Os09g31380	jmjC domain-containing protein 5, putative, expressed
Hv.13882	771	10	-78.79	LOC_Os04g02880	Expressed protein
Hv.10251	950	14	-68.59	LOC_Os03g58300	Indole-3-glycerolphosphatelyase, chloroplast precursor, putative, expressed

Accession	Reads control	Reads salt	Fold Change	Nearest Rice BLAST hit	Annotation
Hv.34103	230	1	-238.86	LOC_Os03g08580	Expressed protein
Hv.37409	1,475	43	-34.30	LOC_Os04g57880	Heat shock protein DnaJ, putative, expressed
Hv.20312	1,873	57	-34.30	LOC_Os01g74020	MYB family transcription factor, putative, expressed
Hv.30597	712	22	-27.86	LOC_Os01g05060	Mitochondria lglycoprotein, putative, expressed
Hv.13356	234	3	-78.79	LOC_Os04g49450	MYB family transcription factor, putative, expressed
Hv.19411	4,919	160	-32.00	LOC_Os06g19444	CCT/B-box zinc finger protein, putative, expressed
Hv.9005	1,081	51	-21.11	LOC_Os03g55280	Semi aldehyde dehydrogenase, NAD binding domain-containing protein, putative, expressed
Hv.8557	391	14	-27.86	LOC_Os03g16780	Ankyrin repeat family protein, putative, expressed
Hv.20948	7,350	295	-24.25	LOC_Os05g37520	Expressed protein
Hv.19979	934	50	-18.38	LOC_Os07g42650	Expressed protein
Hv.19759	1,439	82	-18.38	LOC_Os02g40510	Response regulator receiver domain-containing protein, expressed
Hv.30983	2,304	135	-17.15	LOC_Os03g63910	PPR repeat domain-containing protein, putative, expressed
Hv.8625	214	5	-32.00	LOC_Os07g48050	Peroxidase precursor, putative, expressed
Hv.21993	270	11	-24.25	LOC_Os12g43600	RNA recognition motif containing protein, expressed

*Annotation were mined from the best BLASTN hits in the rice database

3.2.5 Gene ontology analysis

Gene ontology analysis was performed using the 200 differentially regulated barley sequences (ranked by *p* value). As few GO analysis tools exist as yet for barley, each barley sequence was mapped to its best BLAST hit in rice, and this list of rice locus was compared to the rice transcriptome-wide background using agriGO (Du *et al*. 2010). Of the 200 down regulated genes only 120 had GO annotation. These down-regulated genes did not show any significant enrichment. Of the 200 up-regulated genes only 114 genes have shown GO annotation. This list of up-regulated genes was significantly enriched for genes annotated with 'response to abiotic stimulus'. These genes were linked to response to osmotic stress, desiccation or water limitation (Table 3.6). Up and down regulated lists showed that the GO terms 'response to chemical stimulus' and 'response to abiotic stimulus' were over represented.

Table 3.6 Significant GO terms for up-regulated transcripts (mRNA-seq I)

GO term	Ontology	Description	Number in input list	Number in BG/Ref	p-value	FDR
GO:0009628	P	response to abiotic stimulus	5	60	1.3e-05	0.0039

The GO was performed using agriGO (http://bioinfo.cau.edu.cn/agriGO/analysis.php)

Putative ontologies were assigned to up-regulated barley sequences using the functional annotations. The ontologies covered three domains: cellular component, molecular function and biological process. On the basis of cellular component, the sequences were categorized in subsets such as 'cell part', 'cell' (around 25% each), 'organelle (around 8%) and 'macromolecular complex' (around 5%) (Figure 3.5A). In the domain molecular function, sequences were categorized in 5 different categories, such as 'catalytic activity' (around 60%) and 'binding' (around 50%) etc. (Figure 3.5B). This was followed by categorizing these transcripts in the domain 'biological processes'. Of the 7 different categories, most of the sequences belonged to 'metabolic process' (around 56%) and 'cellular process' (around 40%) (Figure 3.5C). The possible use of this study was to assign the location and role at the cellular level to the transcripts.

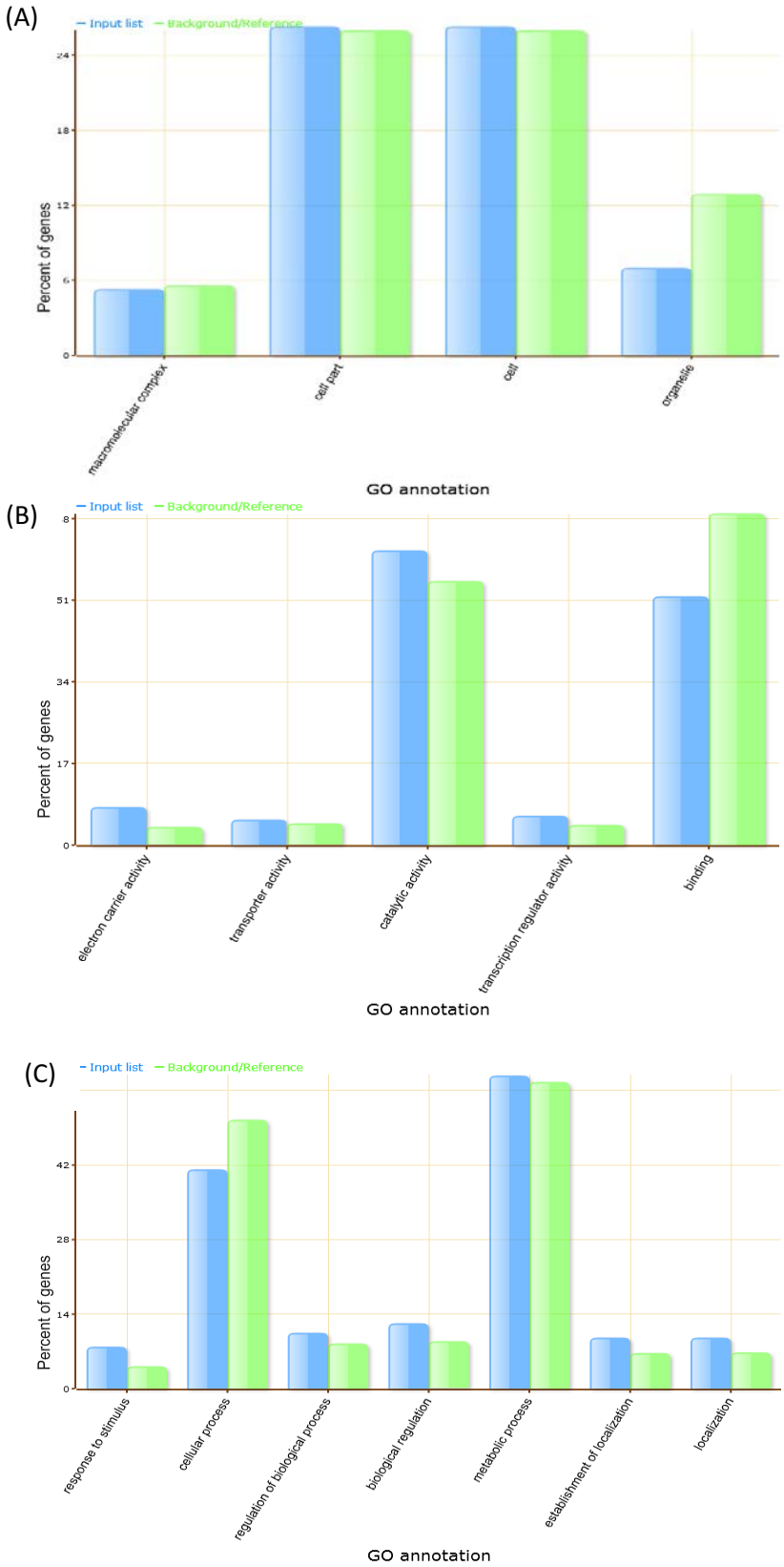


Figure 3.5 Gene ontology for up-regulated transcripts from mRNA-seq I with salt as stress condition

(A): Cellular component; (B) Molecular function; (C) Biological process; The data was sorted using the Microsoft excel in an descending order with respect to the *pvalue* and then graphs were drawn using the rice locus in agriGo (http://bioinfo.cau.edu.cn/agriGO/analysis.php).

3.2.6 Validation of differential expression from mRNA-seq I data using semi quantitative reverse transcriptase-PCR (sqRT-PCR)

To validate the mRNA-seq I data, the expression patterns of Indole-3-glycerol phosphate lyase (Hv.10251), RuBisCO activase (Hv.22598), serine carboxypeptidase (Hv.20929), Putative transketolase (Hv.469), Glycosyl hydrolases, putative (Hv.8276) and DNA gyrase (Hv.8888) were tested using semi- quantitative reverse-transcriptase (sqRT-PCR). Actin was used as a housekeeping control. The primers for these genes were designed to span introns, using the Unigene sequences as templates. It was also ensured that the primers met the following criteria; i) have minimal secondary structures, ii) comparable annealing temperatures between primer pairs and iii) GC content of approximately 50% (detailed in section 2.10.1).

The barley (cv. Hindmarsh) plants were grown under the same conditions as those used for mRNA-seq (detailed in section 2.3.2). After 14 days, three independent plants were treated with NaCl as done for mRNA-seq (150 mM NaCl for 12 h), while three others remained untreated (controls). Subsequently, the total RNA from the leaf samples of both the control and stressed plants was reverse transcribed into cDNA. Each cDNA preparation was tested for quality (lack of gDNA contamination) by amplification using intron-flanking actin primers (Table 2.7). The cDNA was used for further analysis if the PCR resulted in a single band at approximately 147 bp (Figure 3.6).

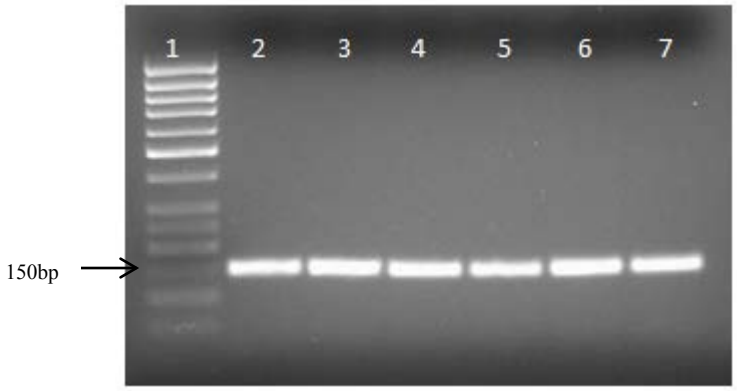


Figure 3.6 Example of a gel image showing quality cDNA
Lane 1: Generuler™ 50bp DNA ladder; Lanes 2 - 7: Actin (AK252278) PCR products from different cDNA preparations

After a quality check, the cDNA was used for sqRT-PCR of the six genes, whose success was noted by the amplification of bands at the expected sizes (Table 2.7) when electrophoresed on an agarose gel. The bands were then quantitated using Chemidoc XRS Documentation Station and Quantity One software (Bio-Rad). The fold changes of the genes were determined by comparing band intensities to those of actin (house-keeping gene). In general the sqRT-PCR fold changes were smaller than those of the mRNA-seq data (as expected, due to the sensitivity limitations of agarose gel based detections), but the direction (up/down regulation) held true (Figure 3.7; Appendix III), except for Hv.8276, likely due to its change being marginal and biological variability.

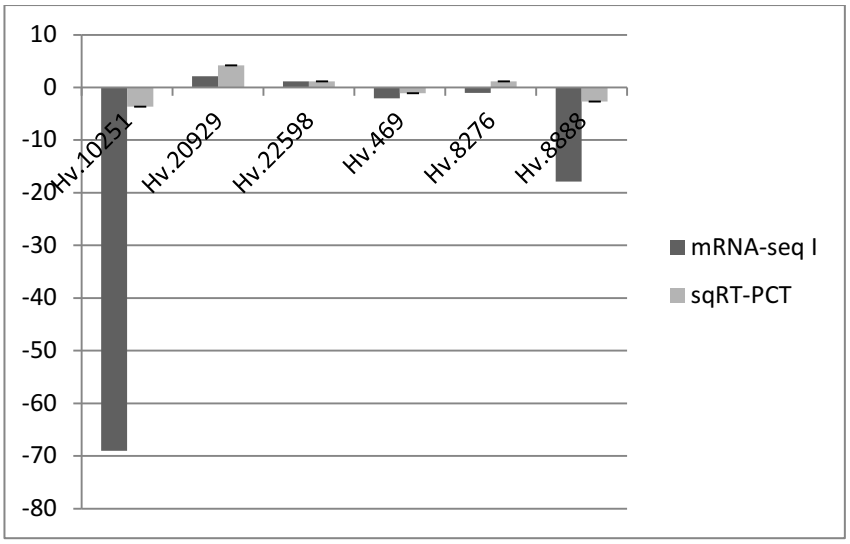


Figure 3.7 Validation of mRNA-seq expression data with sq-RT-PCR

Part B: Results of the Next generation mRNA-sequencing experiment II (mRNA-seq II): analysis of barley leaf transcriptome under salt, drought and exogenous ABA stress conditions

This phase of work was undertaken in 2012. The main aim was to identify the genes which may confer resistance to salinity and drought stress and may thereby be candidates for future crop improvement where abiotic stress poses an increasing problem.

3.2.7 Results of total RNA extraction

The barley leaf mRNA-seq II transcriptome was developed using RNA from two control plants and total RNA from two plants each of salt, drought and exogenous ABA stress conditions. Total RNA was extracted as explained above (Section 2.4.2). The quality and integrity of RNA was checked by running an aliquot of extracted RNA on MultiNA, microchip electrophoresis system (Section 2.9.2). The examples of electropherogram and gel images of total RNA used for mRNA-seq II (generated by bioanalyser) are shown in Figure 3.8 and 3.9. The total RNA was deemed suitable for further analysis by mRNA-seq if electropherogram showed two distinct ribosomal RNA peaks corresponding to both 18S and 28S for eurokaryotic RNA. The spectrophotometric quantification of total RNA extracted was also performed (as explained in section 2.9.4), yielding approximately 5 µg of total RNA per plant sample.

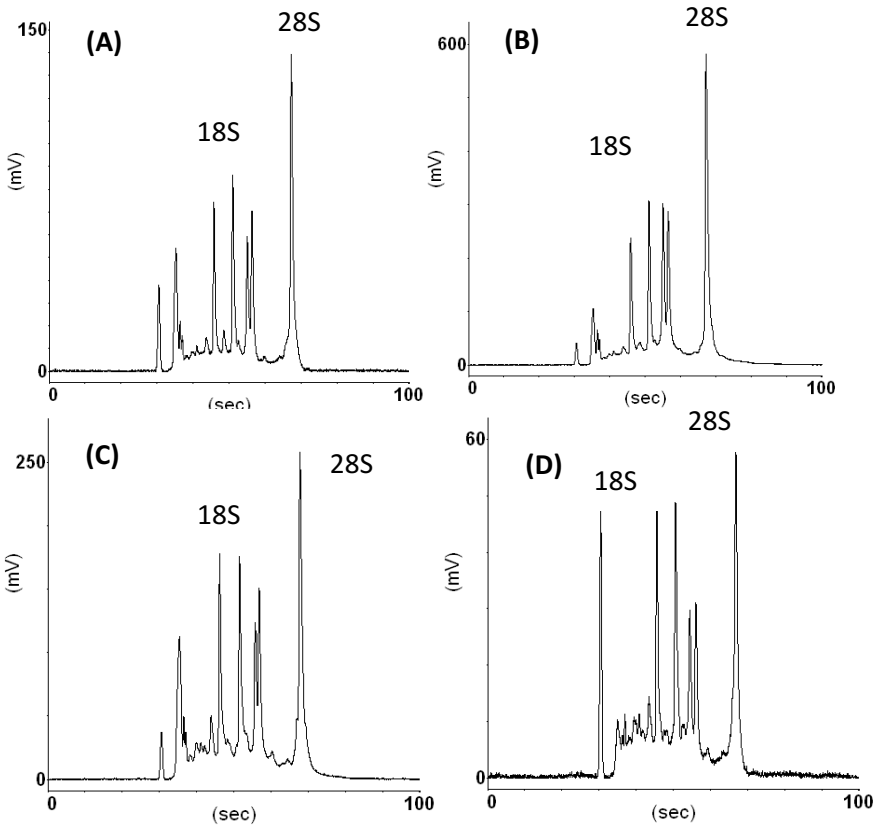


Figure 3.8 Example of electropherogram for quality test of total RNA used for mRNA-seq II

The electrophoergram were generated by the MultiNA, microchip electrophoresis system. A: total RNA from control plant; B: total RNA from salt stressed plant; C: total RNA from drought stressed plant; D: total RNA from exogenous ABA stressed plant

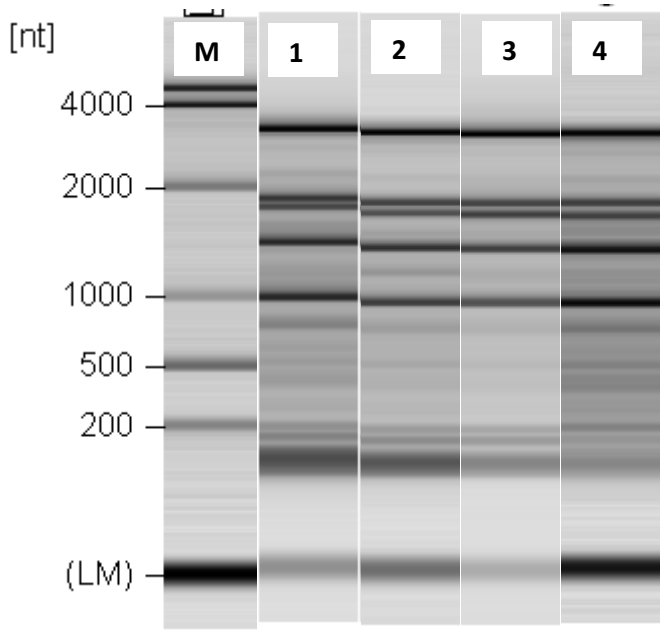


Figure 3.9 Example of gel images for quality test of total RNA used for mRNA-seq II

The gel images were generated by the MultiNA, microchip electrophoresis system. M: RNA 6000 ladder (Applied Biosystems) 1: total RNA from control plant; 2: total RNA from salt stressed plant

3.2.8 Generation of mRNA-seq II libraries and acquisition of the transcriptome dataset

Using the barley (cv. Hindmarsh) leaf total RNAs, eight mRNA-seq libraries were prepared: (i), two from total RNA of two separate plants treated with 150 mM NaCl for 12 h; (ii), two from total RNA of two separate plants treated with 20% PEG for 12 h; (iii) two from total RNA of two separate plants treated with 100 µM exogenous ABA for 12 h and (iv) likewise, two from total RNA of two control plants (Section 2.7.1). The mRNA-seq results yielded 27,299768, 76 nucleotide (nt) sequence reads from control and 25,289,768; 28,012,255; 28,789,122 from for salt, drought and exogenous ABA treated libraries respectively (Table 3.7). The reads were curated for any artifacts using the Fastx Artifacts Filter and any poor quality bases were removed from the 3′ end with a FastQ Quality Trimmer using a threshold of Q30 (http://hannonlab.cshl.edu/fastx_toolkit/) as described in Sections 2.7.2. Barley sequenced genome was released during the progress of this project (October 2012; The International Barley Genome Sequencing Consortium (IBSC), 2012). So theses reads were aligned against the IBSC CDS database using Burrows-Wheeler Aligner (BWA; Li and Durbin 2009) with default settings.

By aligning 21.2 million, 21.0 million, 23.5 million and 22.8 million Q30 quality trimmed mRNA-seq reads from control, salt, drought and exogenous ABA treated plants respectively to the IBSC CDS database using Burrows-Wheeler Aligner (BWA) software (Li and Durbin 2009), it was found that significant proportion of reads 11.9 million from control and 11.1 million, 15.2 million and 13.9 million from salt, drought and exogenous ABA-treated libraries respectively (56.2 % control, 53.06 % salt, 64.8 % drought, 61.1 % exogenous ABA) were mappable to the database (Table 3.7). From analysing this dataset, it was found that, of the 21,096 sequences present in the NCBI barley Unigene database, 20,537, 20,702 and 20,709 transcripts were detected in salt, control and exogenous ABA stressed samples, respectively. The selection criterion of signal strength ≥ 10 was applied to filter out the lowly expressed transcripts reducing the number to 9,956, 8,094 and 9,324 for salt, drought and exogenous ABA stressed plants, respectively (datasheets can be provided on request).

Table 3.7 mRNA-seq II data yield from Genome Analyzer IIx sequencing.

	Control sample	Salt stress sample	Drought stress sample	Exogenous ABA stress sample
Original read length (nt)	76	76	76	
Original number of reads	27,299,768	25,289,836	28,012,255	28,789,122
Number of reads after Q30 quality filtering (bp)	21,295,948	21,036,203	23,575,823	22,831,237
Sequence yield after Q30 quality filtering (bp)	1,461,735,864	1,396,337,866	1,547,060,573	1,505,553,743
Number of reads aligning to Unigene DB	11,975,748	11,162,917	15,296,031	13,955,050
% Reads aligned	56.23	53.06	64.88	61.12

[a]: maximum read length can be obtained is 100bp using illumina genome analyser with low error rate; in this experiment it was set to 76 nt.[b]: total reads obtained from the salt stressed pants and the plant grown as control. [c]: Reads were aligned against barley Unigene database (http://www.ncbi.nlm.nih.gov/unigene; last accessed June 2014) using BWA algorithm, with default settings. Unaligned reads were defined in the form of tentative consensus using AbySS.

3.2.9 Analysis of differential gene expression in salt stressed plant

Quality trimmed reads from control and salt stressed plants were aligned to barley CDS database, which led to identification of 20,537 differentially expressed transcripts. DESeq was used to scan for differential gene expression between control and salinity stress in leaves using a negative binomial model. Seven hundred and two transcripts were found to be significantly deregulated (FDR adjusted p value < 0.05). From these, 254 transcripts showed up-regulation and 448 showed down-regulations. The top 20 differentially expressed transcripts (up and down-regulated) ranked by the fold change from the barley transcripts are shown in Table 3.8. The list of up-regulated genes

includes a number of kinase such as serine/threonine kinase and OsWAK receptor like protein kinase which has been shown previously to mediate abiotic and abiotic stress tolerance as well as those encoding membrane bound proteins such as a peptide transporter. Down-regulated transcripts include those in the LTP family protein precursor, chlorophyll A/B binding protein and MYB transcription factor classes, as also several transcripts of unknown function.

Three transcripts demonstrated a differential expression of ≥ 10 in drought stressed condition, while 711 sequences exhibited a fold change of +1.5-10X (Figure 3.10A). Over seven thousand transcripts showed differential expression ranging from -1.5 to +1.5X, whereas only 1,316 sequences had a decrease in expression ranging from -10 to -1.5X during drought stress condition. One transcript exhibited a down-regulation of more than 10 fold.

Table 3.8 Top 20 up-regulated and down-regulated transcripts ranked by fold change in salt stressed plants (mRNA-seq II)

Accession No	Reads Control 1	Reads Control 2	Reads Salt 1	Reads Salt 2	Fold change	Signal Strength	Rice Locus	Annotation
MLOC_71749.1	11	0	1	551	31.86124	68.12164	LOC_Os04g51520.1	glycosyl hydrolases family 16, putative, expressed
MLOC_76394.3					16	32	LOC_Os08g37210.1	patatin, putative, expressed
MLOC_21281.1					16	16	LOC_Os09g37834.1	serine/threonine-protein kinase receptor precursor, putative,
MLOC_63089.10	46	11	22	800	10.60548	60.0137	LOC_Os03g18130.1	asparagine synthetase, putative, expressed
MLOC_56255.2	5	0	4	38	9.957915	10.73392	LOC_Os02g42150.1	OsWAK14 - OsWAK receptor-like protein kinase, expressed
MLOC_5353.1	11	0	22	247	9.295385	40.4198	LOC_Os08g31250.1	C1-like domain containing protein, expressed
MLOC_70041.1	14	2	4	241	8.834737	38.53464	LOC_Os01g24430.1	expressed protein
MLOC_64685.2	65	22	449	133	8.476263	142.8231	LOC_Os03g01820.1	fatty acid hydroxylase, putative, expressed
MLOC_73463.1	12	0	11	114	7.263838	20.15767	LOC_Os10g04730.1	TKL_IRAK_DUF26-la.6 DUF26 kinases have homology to DUF26 containing loci
MLOC_67287.1	2	0	63	6	7.165106	11.43984	LOC_Os10g23820.1	transferase family protein, putative, expressed
AK248931.1	10	7	25	92	7.003535	25.32099	LOC_Os08g28410.1	retinal pigment epithelial membrane protein, expressed
MLOC_22742.1	14	2	46	241	6.870131	18.97273	LOC_Os07g42570.1	dirigent, putative, expressed
AK359923	87	6	61	609	6.725575	64.20083	LOC_Os03g04570.1	peptide transporter PTR3-A, putative, expressed
MLOC_3110.1	6	0	0	70	6.667289	17.26035	LOC_Os11g32650.1	chalcone synthase, putative, expressed
MLOC_71895.4	257	5	209	1989	6.122031	163.2704	LOC_Os05g51570.1	vacuolar-processing enzyme precursor, putative, expressed
AK367001	18	4	7	186	5.974796	36.88702	LOC_Os01g24430.1	expressed protein
MLOC_72741.1	21	9	4	376	5.794151	25.2768	LOC_Os04g03164.1	expressed protein
MLOC_73077.1	81	10	157	540	5.057373	185.6711	LOC_Os01g71670.1	glycosyl hydrolases family 17, putative, expressed
MLOC_74627.1	1	0	25	7	5.029267	27.74707	LOC_Os04g01710.1	cysteine proteinase At4g11310 precursor, putative, expressed
MLOC_26558.1	36	4	9	149	4.528755	15.41173	LOC_Os01g71080.1	xylanase inhibitor, putative, expressed
AK361049	32	67	10	18	-5.07664	10.50669	LOC_Os02g05470	CCT motif family protein, expressed
MLOC_36627.1	12	80	8	15	-5.14619	12.44382	LOC_Os01g10800	pentatricopeptide, putative, expressed
AK353664	98	724	108	98	-5.27089	82.09979	LOC_Os10g30870	expressed protein
AK248244.1	103	718	58	12	-5.35497	64.70673	LOC_Os02g44320	LTPL113 - Protease inhibitor/seed storage/LTP family protein precursor
AK367396	49	283	43	33	-5.49354	31.33546	LOC_Os05g27100	expressed protein
MLOC_32563.1	3	460	34	9	-5.49578	55.7661	LOC_Os09g17740	chlorophyll A-B binding protein, putative, expressed
MLOC_52133.1	180	1723	208	168	-5.50529	158.5789	LOC_Os02g51790	ribosomal protein L29, putative, expressed
AK354532	21	170	27	3	-5.57	19.1755	LOC_Os11g05556	signal recognition particle 54 kDa protein, putative, expressed

Accession No	Reads Control 1	Reads Control 2	Reads Salt 1	Reads Salt 2	Fold change	Signal Strength	Rice Locus	Annotation
AK357437	32	184	24	24	-5.63621	21.77079	LOC_Os02g54710	expressed protein
AK357949	24	229	42	8	-5.90469	28.01227	LOC_Os12g01370	fatty acid desaturase, putative, expressed
AK358112	1	160	11	1	-5.916	16.62982	LOC_Os05g10370	acid phosphatase, putative, expressed
MLOC_52227.1	23	177	12	1	-5.9298	15.07369	LOC_Os10g18400	beta-galactosidase precursor, putative, expressed
MLOC_13104.1	1386	595	579	78	-5.97083	135.9318	LOC_Os07g37510	organic cation transporter-related, putative, expressed
MLOC_36886.1	46	87	10	11	-6.14649	23.49423	LOC_Os04g57720	OsRR6 type-A response regulator, expressed
MLOC_19326.1	148	342	137	66	-6.16296	43.1971	LOC_Os04g02880	expressed protein
AK367074	3613	19705	5311	183	-6.30427	1959.753	LOC_Os08g35760	Cupin domain containing protein, expressed
AK251420.1	123	985	97	154	-7.2993	152.4393	LOC_Os12g17910	T-complex protein, putative, expressed
MLOC_39958.1	13	236	19	1	-8.3347	18.17502	LOC_Os04g55159	LTPL125Protease inhibitor/seed storage/LTP family protein precursor
AK355398	110	1789	161	6	-8.46126	141.9257	LOC_Os04g55159	LTPL125 - Protease inhibitor/seed storage/LTP family protein precursor
MLOC_9957.3	80	2061	118	32	-11	250.8585	LOC_Os08g35760	Cupin domain containing protein, expressed

*Annotation were mined from the best BLASTN hits in the rice database

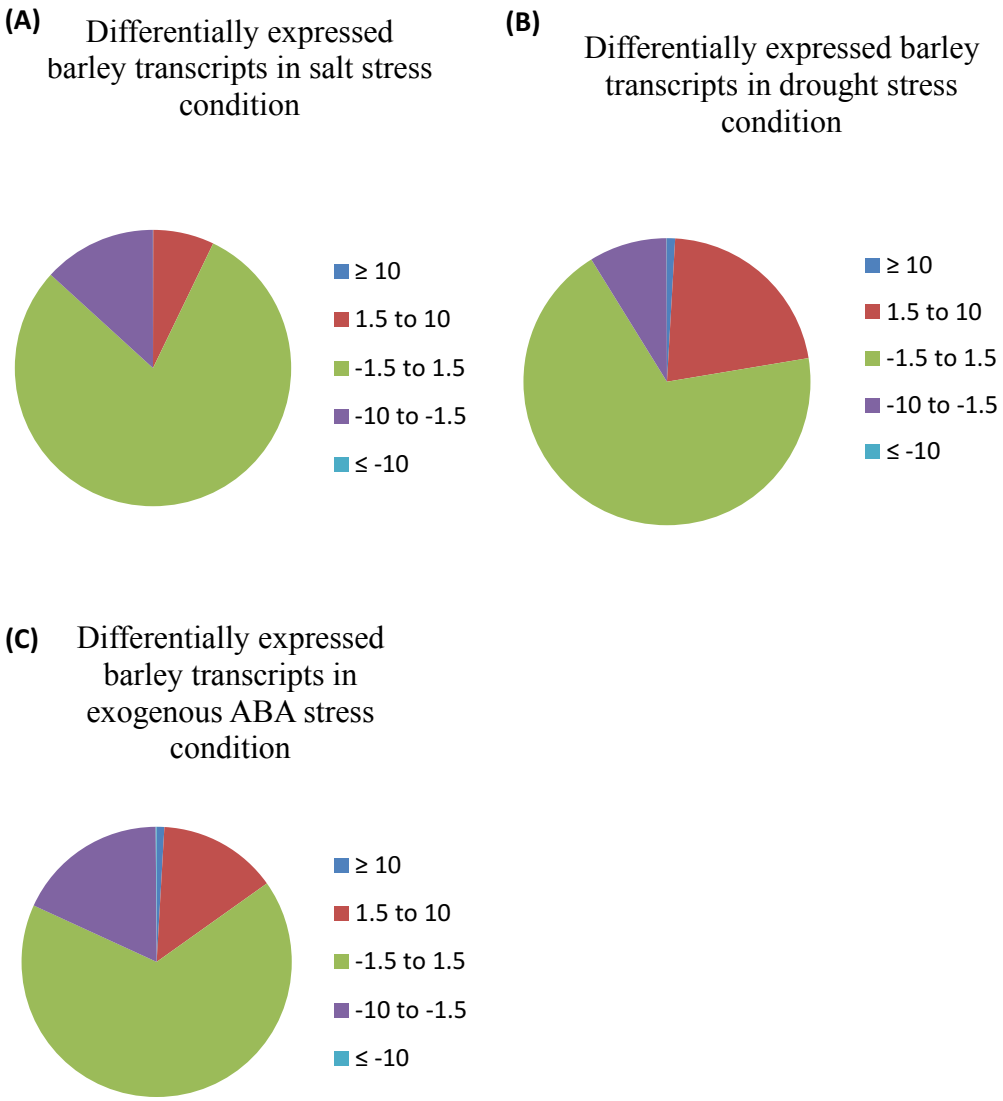


Figure 3.10 Differential expression of sequences from mRNA-seq II data salt, drought and exogenous ABA as stress conditions.

(A): Salt as stress condition; (B); Drought as stress condition; (C); Exogenous ABA as stress condition; The data was sorted using the Microsoft excel in an ascending order with respect to the fold change value (differential expression) and then graphs were drawn using the numbers.

3.2.10 Analysis of differential gene expression in drought stressed plant

Quality trimmed reads from control and drought stressed plants were aligned to barley CDS database, which led to identification of 20,702 differentially expressed transcripts. Three hundred and fifty four transcripts were found to be significantly deregulated (FDR adjusted *p* value <0.05). From these, 317 transcripts showed up-regulation and 37 showed down-regulations. The top 20 differentially expressed transcripts (up and down-regulated) ranked by the fold change from the barley transcripts are shown in Table 3.9. The list of up-regulated genes includes a number of genes (or homologue) which has been shown previously to mediate osmotic/drought/salinity tolerance through ABA dependent stress tolerance pathway, such as B-box Zinc finger family protein, MYB family transcription factor and chlorophyll A/B binding protein. Down-regulated transcripts include those in the stress responsive protein, universal stress protein domain and protein kinase. MYB family transcription factor was present in listing of both up and down regulated genes. But they are homologue to different rice sequences, which suggest that these transcription factors are different members of same family.

Seventy four transcripts demonstrated a differential expression of ≥ 10 in drought stressed condition, while 1,737 sequences exhibited a fold change of +1.5-10X (Figure 3.10B). Over five thousand transcripts showed differential expression ranging from -1.5 to +1.5X, whereas only 711 sequences had a decrease in expression ranging from -10 to -1.5X during drought stress condition. Three transcripts exhibited a down-regulation of more than 10 fold.

Table 3.9 Top 20 up-regulated and down-regulated transcripts ranked by fold change in drought stressed plants (mRNA-seq II)

Accession No	Reads Control 1	Reads Control 2	Reads Drought 1	Reads Drought 2	Fold change	Signal Strength	Rice Locus	Annotation
MLOC_50985.1	0	0	0	723	2706.735	34.1267	LOC_Os09g26670.1	expressed protein
MLOC_59741.1	0	0	8	523	1816.85	23.67179	LOC_Os07g45290.1	cytochrome P450 72A1, putative, expressed
MLOC_16719.1	0	0	0	399	1494.206	18.83951	LOC_Os05g50500.1	secretory protein, putative, expressed
MLOC_80104.1	0	0	0	253	947.8201	11.95124	LOC_Os02g07930.1	B-box zinc finger family protein, putative, expressed
MLOC_66196.1	0	0	0	255	873.0182	11.3318	LOC_Os04g40990.1	malate synthase, glyoxysomal, putative, expressed
MLOC_13672.1	1	0	0	713	504.0506	33.69802	LOC_Os01g17396.2	expressed protein
MLOC_33935.1	1	0	0	270	190.992	12.7965	LOC_Os05g50500.1	secretory protein, putative, expressed
MLOC_14118.2	15	17	17	6870	186.1487	326.9244	LOC_Os08g06110.2	MYB family transcription factor, putative, expressed
AK357075	5	16	22	2922	120.8039	140.3705	LOC_Os03g04100.1	expressed protein
MLOC_73689.1	2	0	0	268	104.672	12.74458	LOC_Os03g07870.1	expressed protein
AK359823	0	2	9	245	99.20954	12.16823	LOC_Os07g01710.1	phytosulfokine receptor precursor, putative, expressed
MLOC_19593.3	4	3	4	684	82.90145	32.88477	LOC_Os07g48570.1	expressed protein
MLOC_58632.3	0	3	1	294	79.56987	14.16048	LOC_Os04g08828.1	cytochrome P450, putative, expressed
AK374133	17	5	21	1354	53.88949	66.02567	LOC_Os09g35880.1	B-box zinc finger family protein, putative, expressed
AK360713	0	7	19	323	41.22592	16.69929	LOC_Os07g08160.1	early light-induced protein, chloroplast precursor, putative,
AK374463	6	1	1	338	40.86457	16.34717	LOC_Os02g03710.1	UP-9A, putative, expressed
AK373368	24	6	8	1279	37.09381	62.234	LOC_Os04g41560.2	B-box zinc finger family protein, putative, expressed
AK355455	2	3	17	203	36.6714	10.74374	LOC_Os01g51140.1	helix-loop-helix DNA-binding domain containing protein,
AK356734	17	9	28	1054	35.94189	52.51021	LOC_Os03g27019.1	expressed protein
AK248909.1	32	1182	2135	39268	29.70596	2045.607	LOC_Os01g52240.1	chlorophyll A-B binding protein, putative, expressed
MLOC_57283.1	136	27	33	25	-3.21568	10.66935	LOC_Os05g37130.4	glycosyl hydrolases family 17, putative, expressed
MLOC_42749.1	134	32	38	21	-3.21952	10.99596		
MLOC_71125.2	177	36	36	39	-3.2519	13.90717	LOC_Os05g25450.1	TKL_IRAK_CrRLK1L-1.3 - The CrRLK1L-1 subfamily has homology to the CrRLK1L homolog, expressed
AK371737	122	30	27	26	-3.28048	10.04381	LOC_Os01g53800.1	glutamate carboxypeptidase 2, putative, expressed
MLOC_66428.5	576	154	126	120	-3.40364	47.9733	LOC_Os05g02420.1	expressed protein
MLOC_17459.1	138	33	41	12	-3.68991	10.94883	LOC_Os05g50340.1	MYB family transcription factor, putative, expressed
AK369026	191	33	40	29	-3.71587	14.00277	LOC_Os01g07590.1	universal stress protein domain containing protein, putative,
MLOC_11418.1	643	123	140	93	-3.77048	47.92017		
MLOC_56286.2	315	117	75	45	-4.12534	27.99112	LOC_Os05g10940.1	metal cation transporter, putative, expressed

Accession No	Reads Control 1	Reads Control 2	Reads Drought 1	Reads Drought 2	Fold change	Signal Strength	Rice Locus	Annotation
MLOC_56250.1	6063	1735	1043	1034	-4.30916	488.8765	LOC_Os06g48180.1	glycosyl hydrolases family 16, putative, expressed
MLOC_78745.1	770	111	169	45	-4.72089	51.5681	LOC_Os03g22230.1	POEI47 Pollen Ole e I allergen and extensin family protein precursor
AK376331	124	72	40	4	-5.08837	12.79299	LOC_Os05g44140.1	non-symbiotic hemoglobin 2, putative, expressed
AK367109	163	18	26	14	-5.16624	10.29775	LOC_Os01g53920.2	receptor-like protein kinase 5 precursor, putative, expressed
MLOC_7230.1	170	13	14	22	-5.79933	10.02569		
AK368196	2923	282	389	48	-8.41399	167.2453	LOC_Os07g48500.1	stress responsive protein, putative, expressed
MLOC_2910.1	597	36	61	15	-9.53048	31.93507	LOC_Os03g22230.1	POEI47 Pollen Ole e I allergen and extensin family protein precursor
AK354735	159	144	20	16	-9.59734	19.14374	LOC_Os06g50310.1	E2F-related protein, putative, expressed
MLOC_60720.1	395	106	31	19	-11.4457	27.27887	LOC_Os07g12890.1	metal cation transporter, putative, expressed
MLOC_65218.1	802	498	33	82	-12.9481	77.17235	LOC_Os04g47360.1	OsPOP9 - Putative Prolyl Oligopeptidase homologue, expressed
MLOC_66582.1	192	130	10	18	-13.0851	19.16168	LOC_Os05g39560.1	metal cation transporter, putative, expressed

*Annotation were mined from the best BLASTN hits in the rice database

3.2.11 Analysis of differential gene expression in exogenous ABA stressed plant

Quality trimmed reads from control and drought stressed plants were aligned to barley CDS database, which led to identification of 20,709 differentially expressed transcripts. Five hundred and thirty seven transcripts were found to be significantly deregulated (FDR adjusted *p* value <0.05). From these, 330 transcripts showed up-regulation and 207 showed down-regulations. The top 20 differentially expressed transcripts (up and down-regulated) ranked by the fold change from the barley transcripts are shown in Table 3.10. Up-regulated transcripts include those in the stress signalling in plants such as DUF581 domain containing protein. The list of down-regulated genes includes a number of genes (or homologue) which has been shown previously to mediate stress tolerance through ABA dependent stress tolerance pathway, such as chlorophyll A/B binding protein and cyclase/dehydrase.

Eighty five transcripts demonstrated a differential expression of ≥ 10 in drought stressed condition, while 1,328 sequences exhibited a fold change of +1.5-10X (Figure 3.10C). Over six thousand transcripts showed differential expression ranging from -1.5 to +1.5X, whereas only 1,682 sequences had a decrease in expression ranging from -10 to -1.5X during drought stress condition. Ten transcripts exhibited a down-regulation of more than 10 fold.

Table 3.10 Top 20 up-regulated and down-regulated transcripts ranked by fold change in exogenous ABA stressed plants (mRNA-seq II)

Accession No	Reads Control 1	Reads Control 2	Reads ABA 1	Reads ABA 2	Fold change	Signal Strength	Rice Locus	Annotation
MLOC_71749.1	11	0	8	2097	133.3402	93.94522	LOC_Os04g51520.1	glycosyl hydrolases family 16, putative, expressed
MLOC_39585.1	0	2	331	72	129.6775	21.65971	LOC_Os02g33380.1	pectinesterase inhibitor domain containing protein, putative,
MLOC_77894.1	5	0	5	872	119.5829	39.21341	LOC_Os09g20240.1	DUF581 domain containing protein, expressed
MLOC_51393.1	2	0	3	291	94.6289	13.22339	LOC_Os01g13610.1	isoflavone reductase homolog IRL, putative, expressed
MLOC_3110.1	6	0	33	665	79.85011	31.6747	LOC_Os11g32650.1	chalcone synthase, putative, expressed
MLOC_22742.1	14	2	32	1452	65.00722	67.08881	LOC_Os07g42570.1	dirigent, putative, expressed
MLOC_5353.1	11	0	6	990	63.10087	44.81587	LOC_Os08g31250.1	C1-like domain containing protein, expressed
MLOC_13423.1	4	0	2	298	50.66088	13.57821	LOC_Os08g06100.1	O-methyltransferase, putative, expressed
AK249901.1	6	2	294	232	45.51688	27.21628	LOC_Os01g03340.1	BBTI4 - Bowman-Birk type bran trypsin inhibitor precursor,
MLOC_26558.1	36	4	119	2408	44.61776	115.6056	LOC_Os01g71080.1	xylanase inhibitor, putative, expressed
AK367914	13	1	11	799	40.48387	36.85493	LOC_Os05g33130.1	CHIT17 - Chitinase family protein precursor, expressed
AK371265	64	13	139	3755	35.80378	178.7323	LOC_Os12g43490.1	thaumatin, putative, expressed
AK354089	45	0	70	1969	32.02053	93.5389	LOC_Os01g71080.1	xylanase inhibitor, putative, expressed
MLOC_72965.1	21	3	69	1021	31.97068	50.53931	LOC_Os07g03710.1	SCP-like extracellular protein, expressed
MLOC_68610.1	34	0	25	1465	30.92498	68.1443	LOC_Os08g02230.1	FAD-binding and arabino-lactone oxidase domains containing protein, putative
MLOC_71895.4	257	5	932	10023	29.6578	509.6052	LOC_Os05g51570.1	vacuolar-processing enzyme precursor, putative, expressed
AK365104	75	1	117	3059	29.58575	146.1228	LOC_Os01g28450.1	SCP-like extracellular protein, expressed
MLOC_72741.1	21	9	21	1228	29.35678	57.60994	LOC_Os04g03164.1	expressed protein
MLOC_64685.2	65	22	3107	482	29.2156	197.9262	LOC_Os03g01820.1	fatty acid hydroxylase, putative, expressed
MLOC_35442.1	32	2	92	1198	26.77479	60.13094	LOC_Os07g13580.1	glucan endo-1,3-beta-glucosidase precursor, putative, expressed
MLOC_34273.1	28	216	22	29	-6.70727	24.40304	LOC_Os04g16734.1	maturase K, putative, expressed
AK357437	32	184	20	24	-6.87672	21.21748	LOC_Os02g54710.1	expressed protein
MLOC_56051.1	156	1931	343	38	-7.71165	212.9161	LOC_Os01g52240.1	chlorophyll A-B binding protein, putative, expressed
MLOC_945.1	347	5632	855	219	-7.84081	612.1337	LOC_Os01g52240.1	chlorophyll A-B binding protein, putative, expressed
MLOC_16300.1	4139	3574	957	406	-7.9705	635.4691	LOC_Os11g03290.1	nucleoside-triphosphatase, putative, expressed
MLOC_1250.1	54	75	18	4	-8.16527	11.03239	LOC_Os04g44510.1	GEM, putative, expressed
MLOC_13104.1	1386	595	323	16	-8.22613	150.2926	LOC_Os07g37510.1	organic cation transporter-related, putative, expressed
AK372159	95	135	35	4	-8.25374	19.96566	LOC_Os06g40500.1	expressed protein
AK356855	85	123	25	8	-8.8103	17.80919	LOC_Os06g40500.1	expressed protein

Accession No	Reads Control 1	Reads Control 2	Reads ABA 1	Reads ABA 2	Fold change	Signal Strength	Rice Locus	Annotation
MLOC_64100.2	64	178	13	24	-9.1497	21.92697	LOC_Os01g72330.1	OsRR4 type-A response regulator, expressed
AK248909.1	32	1182	147	32	-9.54049	123.773	LOC_Os01g52240.1	chlorophyll A-B binding protein, putative, expressed
MLOC_63724.2	111	91	28	0	-10.0686	16.15602	LOC_Os05g39990.1	expansin precursor, putative, expressed
MLOC_34262.1	172	690	51	69	-10.0975	80.71055	LOC_Os11g04720.2	OsRR10 type-A response regulator, expressed
MLOC_71348.1	231	77	18	20	-11.3389	21.59831	LOC_Os03g18600.1	cyclase/dehydrase family protein, putative, expressed
MLOC_50199.1	9	162	19	1	-11.8858	16.91374	LOC_Os06g12080.1	mTERF family protein, expressed
MLOC_10183.1	113	75	4	17	-12.4524	13.98394		
MLOC_71349.1	378	120	26	27	-13.17	34.44396	LOC_Os03g18600.1	cyclase/dehydrase family protein, putative, expressed
MLOC_44067.1	150	4470	356	71	-15.2327	456.7264	LOC_Os01g41710.1	chlorophyll A-B binding protein, putative, expressed
MLOC_7198.1	20	1249	81	20	-17.6521	125.3488	LOC_Os09g17740.1	chlorophyll A-B binding protein, putative, expressed
MLOC_32563.1	3	460	24	3	-23.9097	45.36855	LOC_Os09g17740.1	chlorophyll A-B binding protein, putative, expressed

*Annotation were mined from the best BLASTN hits in the rice database

3.2.12 Gene ontology analysis

Gene ontology analysis was performed using the 200 differentially regulated barley sequences from each stress category. As few GO analysis tools exist as yet for barley, each barley sequence was mapped to its best BLAST hit in rice, and this list of rice locus was compared to the rice transcriptome-wide background using agriGO (Du *et al*. 2010). Of the 200 down regulated genes in salt stress condition, only 121 had GO annotation. These down-regulated genes did not show any significant enrichment. Among the up-regulated genes only 107 genes exhibited GO annotation. This list of up-regulated genes was significantly enriched for genes annotated in 15 terms such as 'phosphotransferase', 'kinase' and 'protein modification' etc. Putative ontologies were assigned to the up-regulated barley sequences in salt stressed plants using the functional annotations. The ontologies covered three domains: cellular component, molecular function and biological process. On the basis of cellular component, the transcripts were categorized in subsets such as macromolecular complex, cell part, cell and organalle. Of these the category 'cell part' and 'cell' contains the highest amount of transcripts (almost 25% each), followed by categories organelle (around 9%) and macromolecular complex (around 5%) (Figure 3.11A). In the domain molecular function, sequences were categorized in 5 different categories, such as 'catalytic activity' (around 65%) and 'binding' (around 60%) etc. (Figure 3.11B). This was followed by categorizing these transcripts in the domain 'biological processes'. Of the 6 different categories, most of the sequences were found to belong to 'metabolic process' (around 60%) and 'cellular processes (around 45%) (Figure 3.11C). The possible use of this study was to assign the location and role at the cellular level to the transcripts.

Of the 200 down regulated genes in drought stress condition, only 123 had GO annotation. These down-regulated genes showed significant enrichment for 26 annotation terms such as 'ribonucleotide binding, 'kinases' and 'phosphorylation etc. Of the 200 up-regulated genes only 113 genes have shown GO annotation. This list of up-regulated genes was significantly enriched for genes annotated with 'oxidoreductase activity', 'iron, heme or tetrapyrrole binding' etc. In up-regulated genes, the categories 'cell part' and 'cell' contained the highest amount of transcripts (around 25% each), followed by 'organalle' (around 5%) (Figure 3.13A) on the basis of cellular component.

In the down regulated genes, 32% of transcripts were annotated as 'cell part' and 'cell' each, followed by around 18% as 'organelle' and 5% as 'macromolecular complex' (Figure 3.12A). In the domain molecular function, sequences were classified in 5 different categories for up and down regulated genes respectively. The up-regulated genes were categorised as 'catalytic activity' (68%), 'binding' (around 55%), 'electron carrier activity' (around 15%), 'transcription regulator' and 'antioxidant' (around 5% each) (Figure 3.13B), whereas the down-regulated genes were classified as 'binding' (72%), 'catalytic activity' (54%), 'transporter activity' (around 9%), 'electron carrier activity' (around 5%) and 'structural molecule activity' (around 3%) (Figure 3.12B). This was followed by categorizing these transcripts in the domain 'biological processes'. Of the 7 different categories, most of the sequences were found to belong to 'metabolic process' (around 56% for both up and down-regulated) and 'cellular process' (around 45% for up and 50% for down-regulated) (Figure 3.13C; 3.12C).

Of the 200 down regulated genes in exogenous ABA stress condition, only 109 had GO annotation. These down-regulated genes, which showed significant enrichment were widely distributed among 78 annotation terms such as 'membrane', 'transcription' etc. Among the 200 up-regulated genes only 100 genes have shown GO annotation with 7 significant GO terms such as 'hydrolase activity' and 'peptidase activity'. On the basis of cellular component, 25% of the transcripts were annotated as 'cell part and 'cell' each for up-regulated genes, whereas these GO terms formed around 37% each of the down-regulated genes (Figure 3.14A; 3.15A). In the domain molecular function, up regulated sequences were classified in 5 categories, whereas the down regulated genes in 6 with 'enzyme regulator activity' as an extra. Of these different categories, most of the sequences belonged to 'binding' (around 65% for up and 55% for down-regulated) and 'catalytic activity' (around 60% for up and 50% for down-regulated) (Figure 3.14B; 3.15B). In the domain biological processes, the up-regulated transcripts were annotated in 7 GO terms, whereas the down regulated transcripts showed two extra terms ('signalling' and 'signalling process). The highest number of transcripts were annotated as 'metabolic process' (around 60% for up and 56% for down-regulated) followed by 'cellular binding (around 50% for up and 55% for down-regulated) (Figure 3.14C; 3.14C).

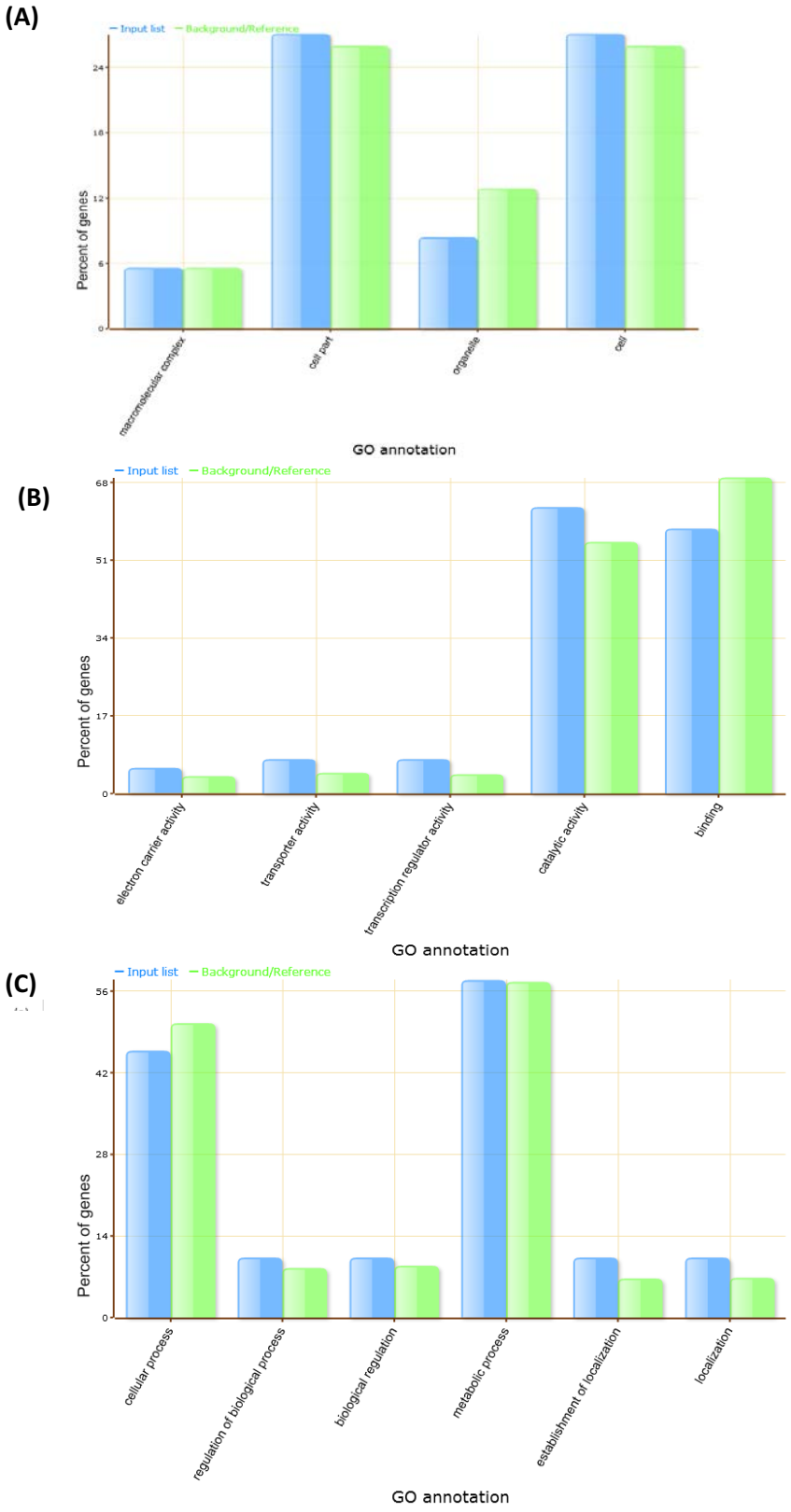


Figure 3.11 Gene ontology for up-regulated transcripts from mRNA-seq II (salt as stress condition)

(A): Cellular component; (B) Molecular function; (C) Biological process; The data was sorted using the Microsoft excel in an descending order with respect to the *pvalue* and then graphs were drawn using the rice locus in agriGo (http://bioinfo.cau.edu.cn/agriGO/analysis.php). Significant *p-value*s for Go are mentioned in Appendix IV

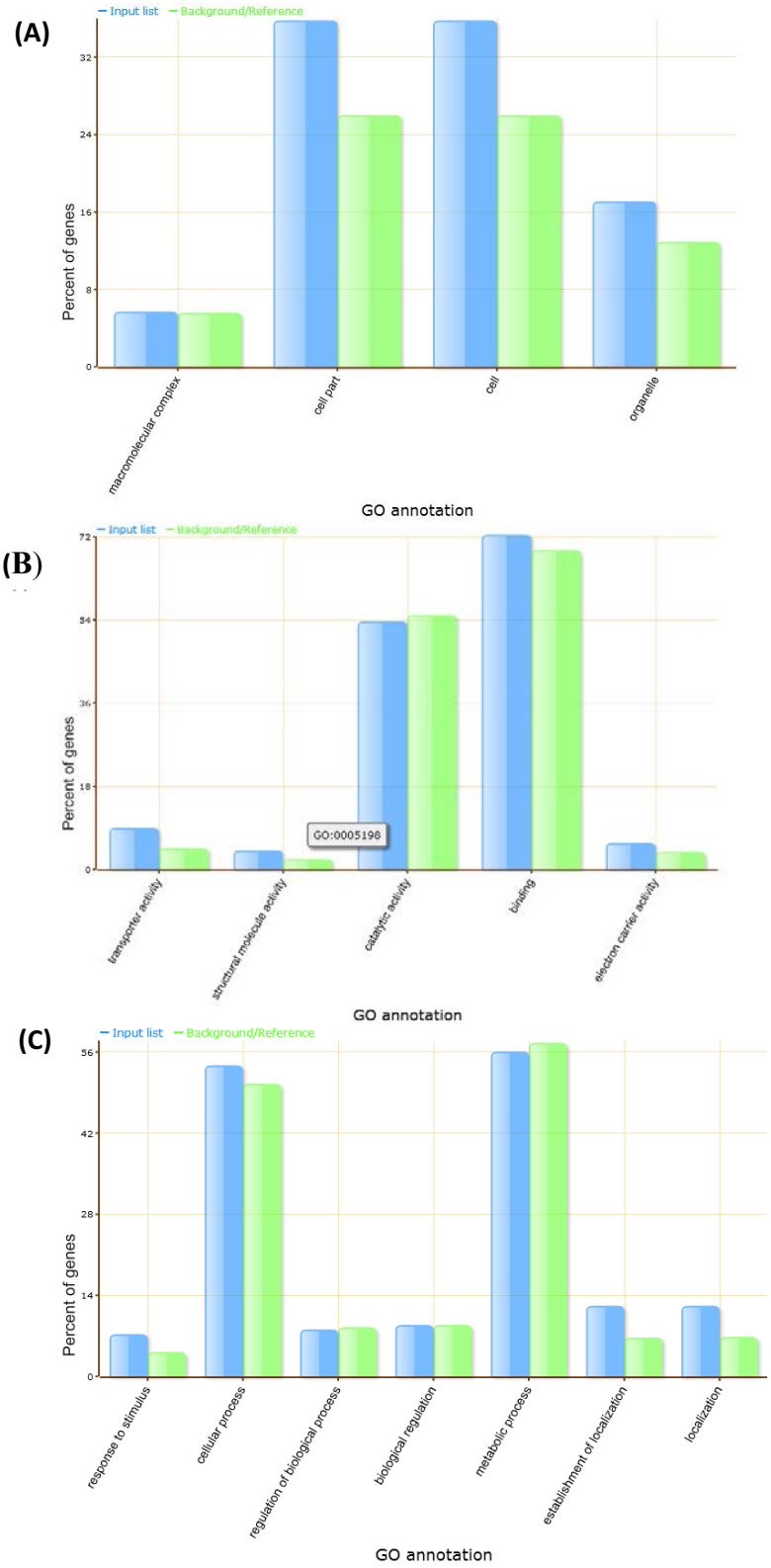


Figure 3.12 Gene ontology for down-regulated transcripts from mRNA-seq II (drought as stress condition)

(A): Cellular component; (B) Molecular function; (C) Biological process; The data was sorted using the Microsoft excel in an descending order with respect to the *pvalue* and then graphs were drawn using the rice locus in agriGo (http://bioinfo.cau.edu.cn/agriGO/analysis.php). Significant *p-value*s for Go are mentioned in Appendix V

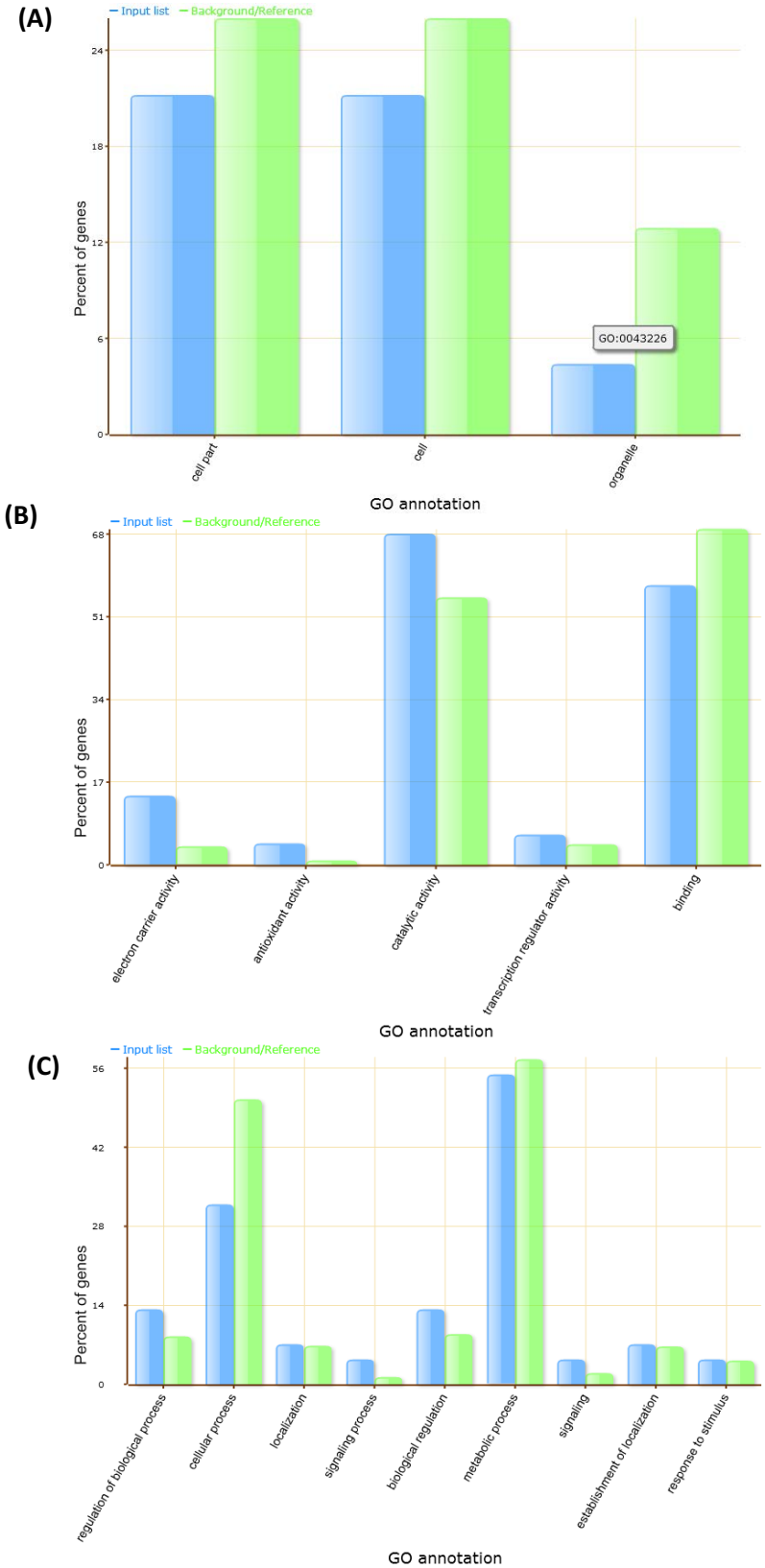


Figure 3.13 Gene ontology for up-regulated transcripts from mRNA-seq II (drought as stress condition)

(A): Cellular component; (B) Molecular function; (C) Biological process; The data was sorted using the Microsoft excel in an descending order with respect to the *pvalue* and then graphs were drawn using the rice locus in agriGo (http://bioinfo.cau.edu.cn/agriGO/analysis.php). Significant *p-value*s for Go are mentioned in Appendix VI

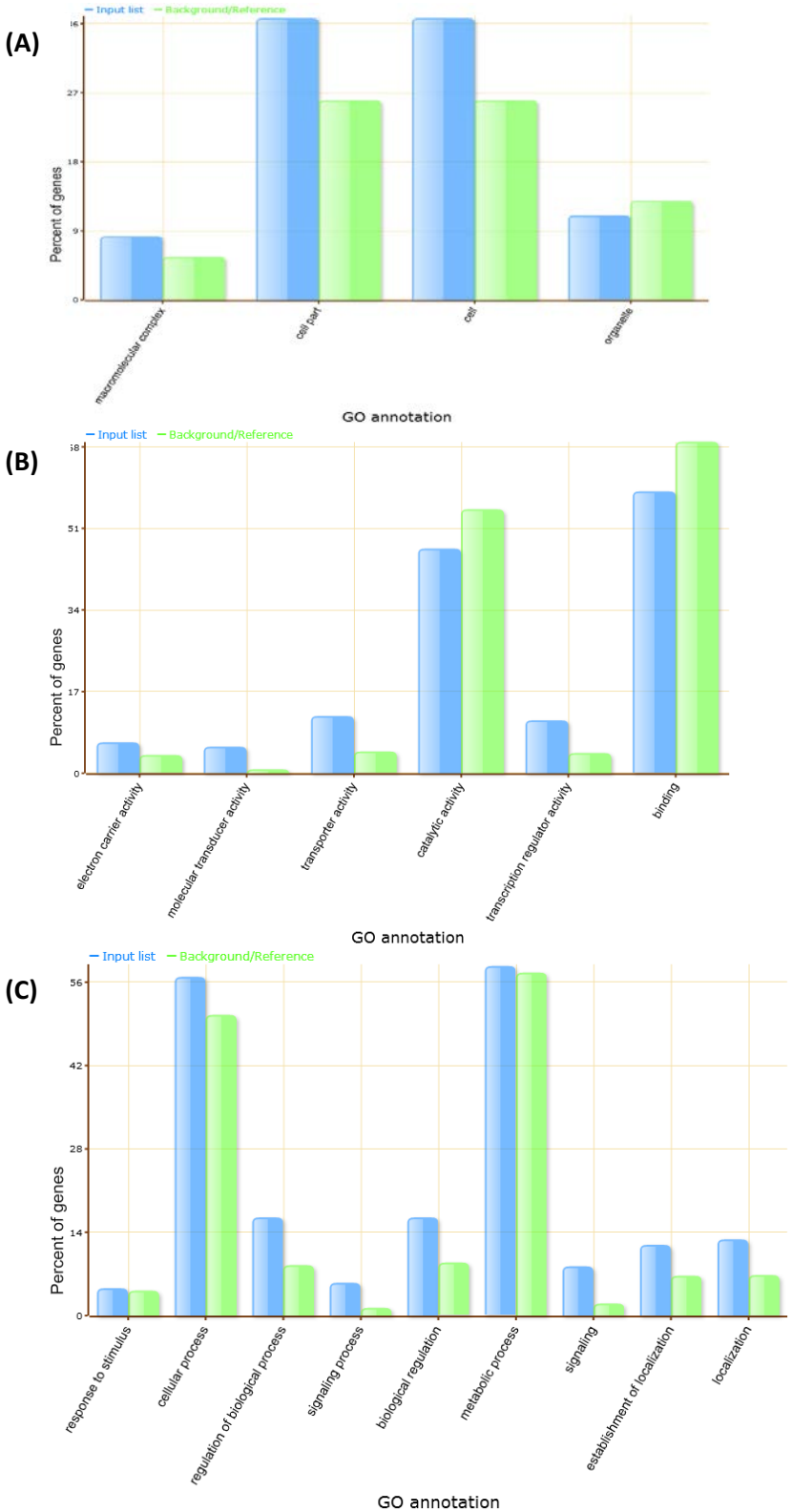


Figure 3.14 Gene ontology for down-regulated transcripts from mRNA-seq II (exogenous ABA as stress condition)

(A): Cellular component; (B) Molecular function; (C) Biological process; The data was sorted using the Microsoft excel in an descending order with respect to the *pvalue* and then graphs were drawn using the rice locus in agriGo (http://bioinfo.cau.edu.cn/agriGO/analysis.php). Significant *p-value*s for Go are mentioned in Appendix VII

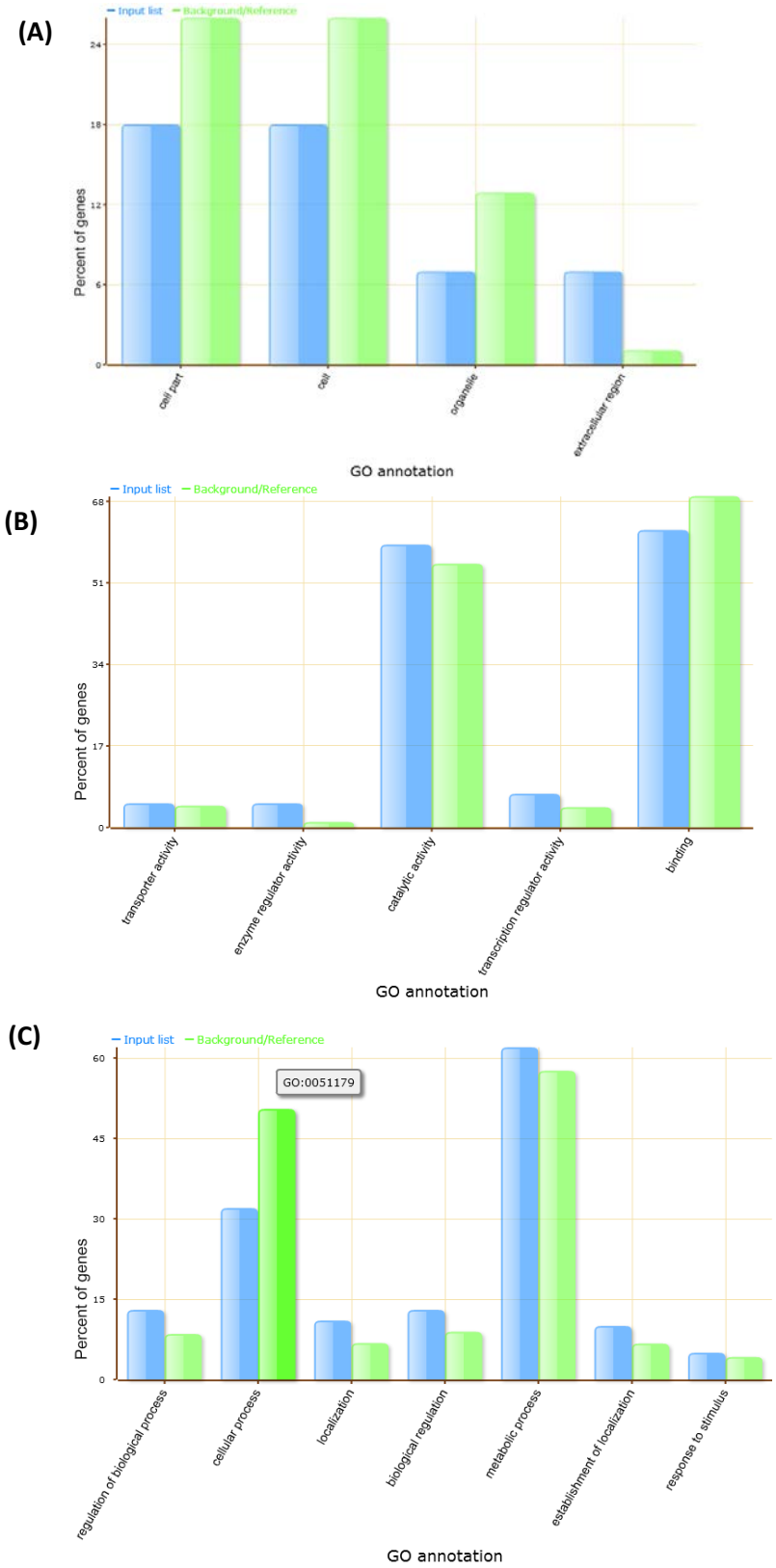


Figure 3.15 Gene ontology for up-regulated transcripts from mRNA-seq II (exogenous ABA as stress condition)

(A): Cellular component; (B) Molecular function; (C) Biological process; The data was sorted using the Microsoft excel in an descending order with respect to the *pvalue* and then graphs were drawn using the rice locus in agriGo (http://bioinfo.cau.edu.cn/agriGO/analysis.php). Significant *p-value*s for Go are mentioned in Appendix VIII

3.2.13 Validation of differential expression from mRNA-seq II using quantitative real time-PCR (qRT-PCR)

To validate the mRNA-seq II data, the expression patterns of HvPP2C06 (AK251854), HvPP2C09 (MLOC_8131), HvPYR1.1 (AK376521), HvPYR1.2 (AK71349), HvSnRK2.6 (AK22145), and HvABCG40 (MLOC_68581) were tested using quantitative real time PCR (sqRT-PCR). Actin was used as a housekeeping control as explained above 3.2.6.

The barley (cv. Hindmarsh) plants were grown under the same conditions as those used for mRNA-seq (detailed in Section 2.3.2). After 14 days, three independent plants were treated with NaCl (150 mM NaCl for 12 h), three with PEG (20 % for 12 h), three with exogenous ABA (100 µM for 12 h) as done for mRNA-seq II, while three others remained untreated (controls). Subsequently, the total RNA from the leaf samples of both the control and stressed plants was reverse transcribed into cDNA. Each cDNA preparation was tested for quality as explained above in Section 3.2.6.

After a quality check, the cDNA was used for qRT-PCR of the six genes, whose success was noted by the amplification of bands at the expected sizes (Table 2.7) when electrophoresed on an agarose gel. The bands were then quantitated using Chemidoc XRS Documentation Station and Quantity One software (Bio-Rad). The fold changes of the genes were determined by comparing band intensities to those of actin (house-keeping gene). In general the qRT-PCR fold changes were lower than those of the mRNA-seq data (Appendix IX) (as expected, due to the sensitivity limitations of agarose gel based detections), but the direction (up/down regulation) held true (Figure 3.16), except for HvPP2C06 and Hv ABCG40 under salinity stress, likely due to its change being marginal and biological variability.

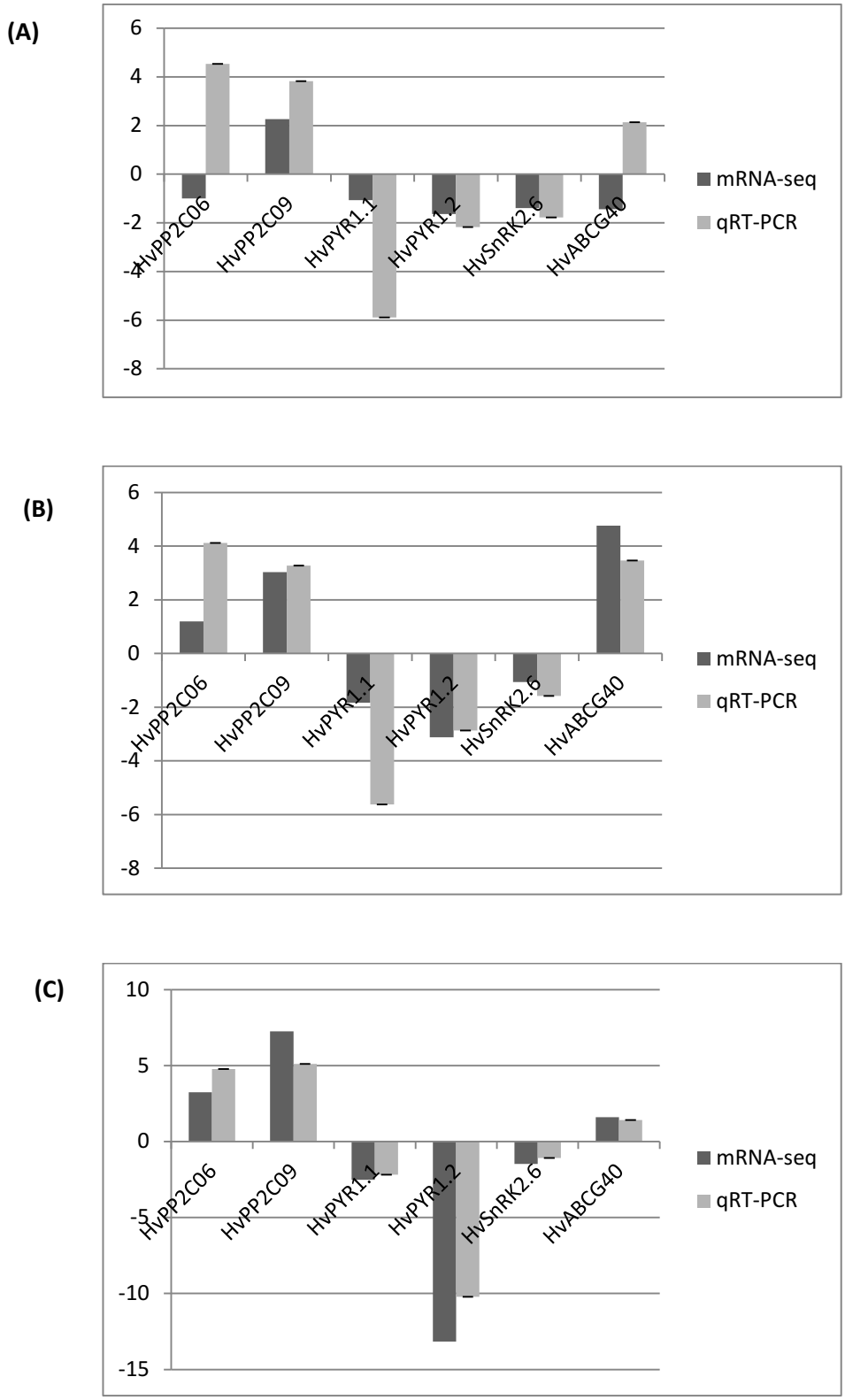


Figure 3.16 Validation of mRNA-seq expression data with sq-RT-PCR
(A): validation of salt stressed data; (B) validation of drought stressed data; (C): validation of exogenous ABA stressed data

3.3 Discussion

3.3.1 mRNA-seq I data analysis: salinity as stress condition

mRNA-seq facilitates rapid generation of large datasets for transcript identification and quantification, even in the absence of a reference genome. In this work barley leaf mRNA from salt stressed and control plants was loaded on two lanes of an Illumina Genome Analyser flow cell, which generated over 50 million 76 nucleotide reads. This amounted to 3.56 Gbp after quality trimming, whereas Genbank contained only 525,999 capillary-sequenced barley ESTs amounting to 272.6 Mbp (June 2014). This indicates that mRNA-seq has intense prospective for plant biology, as also indicated by recent studies on crop species such as rice (Mizuno *et al*. 2010; Zhang *et al*. 2010), soybean (Severin *et al*. 2010) and chickpea (Garg *et al*. 2011).

When plants are exposed to salinity in laboratory conditions a rapid temporary drop in growth is observed, which is followed by a gradual recovery to a new reduced rate of growth. The temporary effects are due to rapid and transient changes in plant water relations (Munns 2002). Salt specific effects can become visible after few days at high salinity. So ideally, to study the response of plants to acute stress, one should investigate expression profiles over a range of times, as done for microarray work (Seki *et al*. 2002) to study the dynamics of transcriptomes under acute salinity. Other investigators have undertaken transcriptome studies in acclimation models such as exposure to cyclic or continuous mild stress (Watkinson *et al*. 2003), which are relevant to study of plants exposed to chronic stress in the field. However, these stresses negatively impact developmental pathways, which then confound the comparison of control and stress samples due to growth retardation of stressed plants. Hence, in order to avoid the potential retardation obscuring the responses of plants to high salinity, researches often investigate the response to acute salinity stress. While it is preferred to have profiled many time points, the next-generation experiments are still too costly. As such, the 12 h time point at 150 mM NaCl (Ozturk et al. 2002) was selected as the one likely to produce the clearest effects including differential expressions (fold change) of several key transcription factor as well as response genes.

This work was aimed at two main points; one was to determine barley genes which were differentially expressed under acute salt stress and second to discover previously

unidentified ('novel') transcripts in the barley leaf. The application of stringent selection criterion of Rice/barley BLAST ratio (>=2) enabled the discovery of 2,144 'novel' TCs. Some of these identified sequences were new and others were more complete in size. Some of the differentially expressed genes identified in this experiment agree with the previous expression work using microarray (Ueda *et al*. 2004). Whereas the differentially expressed NTCs found in this experiment represent transcripts which were not detected using microarray. One example of such 'novel' sequence is wall associated receptor kinase-like 22 (2.0 kb) which has only a 450 bp BLAST hit in the barley Unigene databse, but a 1.2 kb homologue was found in rice and Brachypodium. A similar gene in Arabdidopsis, WALK4 is responsive to Na^{+}, $K^{+,}$ Ni^{2+} and Zn^{2+} cations (Hou *et al*. 2005).

Late embryogenesis abundant (LEA), which is a homologue of early responsive to dehydration stress (ERD4) is found to be strongly up-regulated by acute salinity in this mRNA-seq experiment. This gene has also been investigated in maize (Liu *et al*. 2009a), where it was found that this gene is not only induced upon salinity stress, but its over-expression in Arabidopsis leads to enhanced tolerance to drought and salinity. In barley, this gene has already been proven to convene tolerance to osmotic stresses (Xu *et al*. 1996). Other strongly up-regulated candidates for future functional work could include the chloroplast-localised lipoxygenase 2.1, a plasma membrane bound ATPase, as well as a protein phosphatase 2C. A highly expressed aquaporin was among the most decreased in expression, indicating water transport processes within the leaf could be strongly reduced under acute salt stress.

Comparison of this mRNA-seq dataset to previous array experiments (Walia *et al*. 2007) resulted a reasonable correlation of fold change, with some of the highest expressed transcripts showing contradictory fold change. Different lines of barley used, different regimens of salt stress and different analysis chemistries employed are among factors which could explain this disparity.

3.3.2 mRNA-seq II data analysis: salinity, drought and exogenous ABA as stress conditions

In this work barley leaf mRNA from salt, drought and exogenous ABA stressed and control plants was loaded on four lanes of an Illumina Genome Analyser flow cell, which generated over 100 million 76 nucleotide reads. As explained above, with the limitation of cost for carrying out next generation experiment, it was decided to use 12 h as the only time frame for all three stress conditions. This work was aimed to determine barley genes, which were differentially expressed under acute salt, drought and exogenous ABA stress conditions.

Membrane anchored receptor like protein kinases (RLK) homologue, known as Domain of unknown function 26 (DUF26) -receptor like kinase was strongly up-regulated by acute salinity in this dataset and has been investigated in Arabidopsis (Tanake *et al*. 2012) and rice (Nakashima *et al*. 2007), wherein this gene is not only induced upon salinity stress, but also in other abiotic stress such as cold, drought etc. Serine/threonine kinases and WAK-receptor like protein kinase are also found to regulate stress responsive genes in Arabidopsis (Gao *et al*. 2013), rice (Diedhiou *et al*. 2008) and barley (Cadenas *et al*. 1999). Other strongly up-regulated genes for future functional work could include the one known to be involved in biotic stress regulation such as xylanase inhibitor (in wheat; Sansen *et al*. 2004). A highly expressed protease inhibitor and chlorophyll a/b binding protein were among the most decreased in differential expression and supported the results in Arabidopsis (Xu *et al*. 2012b).

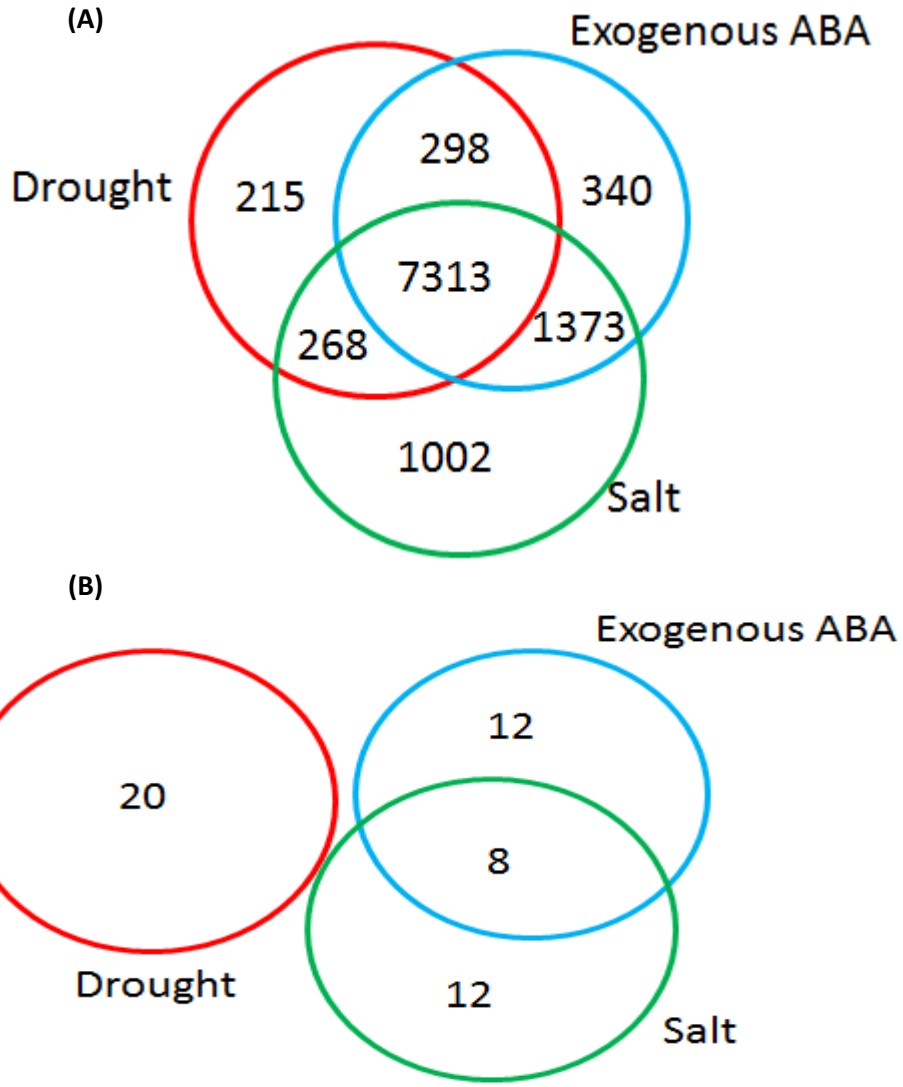
Cytochrome P450 superfamily of monooxygenases, mainly involved in the oxidation of organic substances was strongly up-regulated by acute drought stress in this dataset and has also been investigated in Arabidopsis (Narusaka *et al*. 2004), wheat (Houde *et al*. 2006) and barley (Talame *et al*. 2007), which showed similar results. Other strongly up-regulated genes for future work could include phytosulfokine receptor (PSKR), which enables phytosulfokine to be an important candidate in the signalling pathway for cellular dedifferentiation and proliferation in plants, as well as malate synthase, key enzyme in the glyoxylate cycle. The differential expression of PSKR was also supported data in wheat (Houde *et al*. 2006). A highly expressed glycosyl hydrolases was among the most down-regulated transcripts, indicating hydrolysis of glycosidic

bonds in cell wall polymers within the leaf could be strongly reduced under acute drought stress.

Pectinesterase inhibitor homologue, known as pectin methylesterase inhibitor are mainly involved in the inhibition of demethylesterification of highly heterogeneous polymers pectins. This enzyme was strongly up-regulated by exogenous ABA treatment in this dataset and has also been investigated in wheat (Hong *et al*. 2010), which showed similar results. This transcript was also found to be induced highly in salt (8.31) and drought (20.78) stress condition as supported by Hong *et al*. (2010) in wheat, but they demonstrated a signal strength of <10 in salt and drought stress condition (mRNA-seq), hence were screened out after the application of selection criterion. Other strongly up-regulated genes for future functional work could include DUF581 domain containing protein, which are suggested to be involved in the reallocation of the kinase to specific regions within the nucleus (Nietzsche *et al*. 2014). A highly expressed chlorophyll a/b binding protein was among the most decreased in differential expression and supported the results in Arabidopsis (Xu *et al*. 2012a). Of the top 20 down-regulated transcripts in exogenous ABA treatment, 7 belonged to the family of chlorophyll a/b binding proteins.

Based on the Venn diagram analysis, differences and cross talk of gene expression among salinity, drought and exogenous ABA stress conditions in barley were analysed. It was identified that 1002, 215 and 340 genes were differentially expressed only by salinity, drought and exogenous ABA stress treatment respectively by mRNA-seq analysis. Thirteen hundred and seventy three genes were differentially expressed by both salinity and exogenous ABA, 268 by salinity and drought. Similarly, 298 genes were differentially expressed by drought and exogenous ABA treatments as shown in Figure 3.17. Over seven thousand genes were found to demonstrate differential expression in all three stress conditions. Over 87% of the salinity and 94% of the drought-inducible or repressed genes were also expressed by exogenous ABA. Stress inducible genes have shown that four independent signal pathways function in the induction of genes in response to abiotic stress. Two are ABA dependent and two are ABA independent (Kasuga *et al*. 1999), which explains the induction of same genes during abiotic stress and exogenous ABA treatment. Around 94% of drought

deregulated genes were also found to be differentially expressed in salinity as physiological drought occurs during salinity stress as soluble salt limits the water uptake, thereby inducing drought stress condition (Carrow and Duncan 1998). These results in barley are consistent with previous observations in rice (Rabbani *et al*. 2003) and Arabidopsis (Shinozaki and Yamaguchi-Shinozaki, 2000; Seki *et al*. 2002). This also supports the greater crosstalk signalling pathways between salinity and ABA stress condition as compared to drought (Fig 3.17 b,c) and is also in line with Arabidopsis transcriptome data (Matsui et al. 2008).



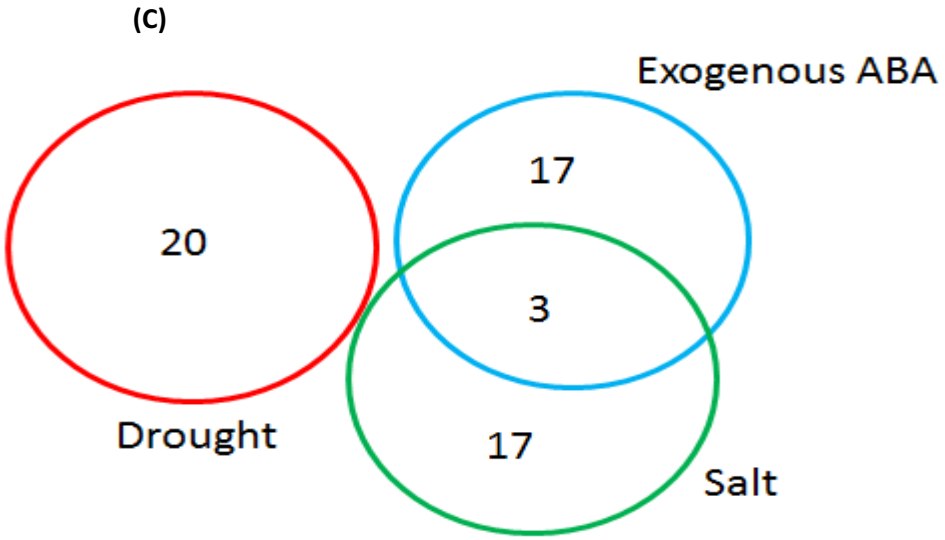


Figure 3.17 Comparison of transcripts showing differential expression in salt, drought and exogenous ABA stress conditions in mRNA-seq II
(A): Complete list of transcriptomes; (B) top 20 up-regulated; (C) top 20 down-regulated; The Venn diagram were drawn using the barley accession no in BioVenn (http://www.cmbi.ru.nl/cdd/biovenn/index.php).

Based on the Venn diagram analysis of the rice homologue of the barley transcriptome, differences and cross talk of gene expression under salinity stress condition in mRNA-seq I and II were analysed. It was identified that 3,737 and 559 genes to be differentially expressed by salinity only in mRNA-seq I and II respectively, whereas more than seven thousand genes were found to be differentially expressed in both the experiments (Figure 3.18). The data supports that for broad comparison between two samples run at different time, the primary determinants of gene expression changes result from biological differences, rather than the artefacts of platform (Yauk *et al*. 2004).

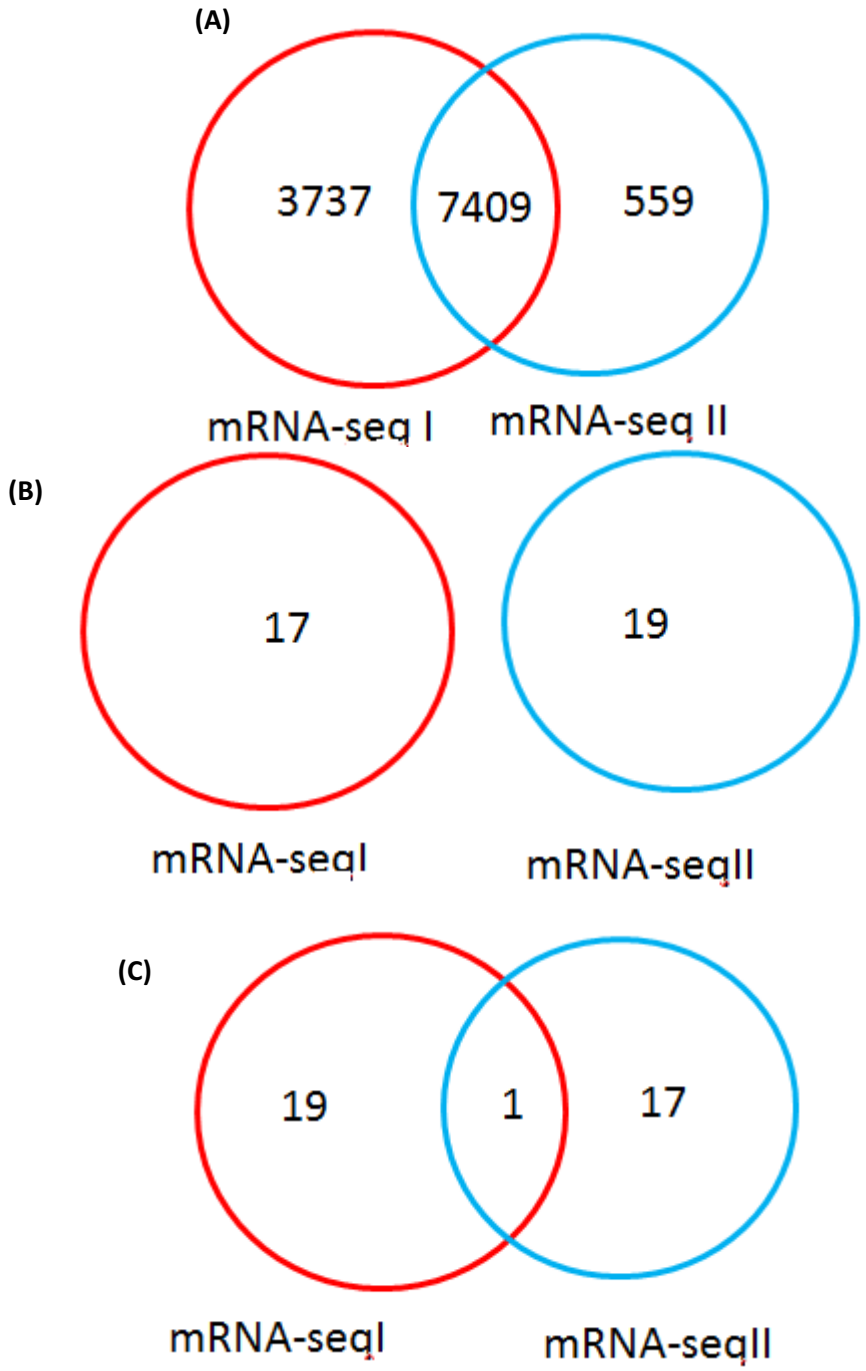


Figure 3.18 Comparison of transcripts showing differential expression in salt stressed condition from mRNA-seq I and II dataset.
(A): Complete list of transcriptomes; (B) top 20 up-regulated; (C) top 20 down-regulated; The Venn diagram were drawn using the barley accession no in BioVenn (http://www.cmbi.ru.nl/cdd/biovenn/index.php).

Investigation of the selected transcripts with sqRT-PCR and qRTPCR has shown general agreement between the techniques (Figure 3.7; 3.16). The direction of expression changes was largely consistent between the methods. Exceptions occurred when the differential expression detected were slight and the deviation were within the observed variation of the house keeping genes actin (+1.17; from mRNA-seq data). For

sequences having extreme differential expression, sqRT-PCR recorded small value as compared to the one calculated by mRNA-seq due to the already known limitations of sqRT-PCR which are saturation and detection range of ethidium bromide staining of gels and quantitation (Ziemann *et al*. 2013a).

The data demonstrates mRNA-seq is an excellent high-throughput methodology for gene expression which will be crucial in revealing the scale of variations in barley germ-plasm and accurate mapping of quantitative trait loci. As next-generation sequencing technologies and associated bioinformatics methods continue to improve, these will become more common place in plant biology and result in a comprehensive high-quality annotation of the barley and wheat genomes. Until then, this study provides a valuable dataset containing thousands of transcripts and a snapshot of differential expression due to abiotic stress. The outcomes serve as a useful reference for future hypothesis-driven studies in barley and the closely related and possibly the most important cereal, wheat. Phytohormone abscisic acid (ABA) mediated abiotic stress tolerance genes are crucial for plant's adaptive response to salinity, drought and other abiotic stresses. ABA signalling pathway is well established in Arabidopsis, but is poorly defined in the major cereal crops, wheat and barley. Thus, the next chapter sought to identify and characterize ABA mediated abiotic stress signalling related candidate genes in barley using mRNA-seq data and other *in vitro* searches.

Chapter 4

Identification and differential expression studies of key components of the abscisic acid mediated signalling pathway in barley under salt and drought stress conditions using mRNA-sequencing

4.0 Abstract

Abscisic Acid (ABA) regulates critical aspect of plant development as well as abiotic stress response. The major components of ABA dependent signalling pathway form a double negative regulatory system [ABA-| PYR-|PP2CA-|SnRK2]. The other important components in ABA signalling are the ABA transport related ABC transporters (ABCG25 and ABCG40). ABA signalling pathway is well established in Arabidopsis, but is poorly defined in the major crops, wheat and barley. Thus the present study aimed to identify and characterise barley ABA signalling related candidate genes in order to gain a picture of ABA signalling pathway in barley. The expressed barley ABA signalling related candidate genes were initially identified using next generation mRNA-sequencing (mRNA-seq) of barley and NCBI Unigene database, followed by further searches in the International Barley Genome Sequencing Consortium (IBSC) database. The identified sequences were analysed on the basis of certain functional motifs and key residues in the putative proteins, which resulted in 13 PP2CAs, 9 PP2CDs, 10 PYR/PYL/RCARs, 5 SnRK2 (subfamily II and III) and 4 ABA transport related ABC transporters. The amino acid involved in the binding of PYR/PYL/RCAR and PP2C were found to be strictly conserved among all members of PP2CA and PYR/PYL/RCAR families in rice and barley. Further, using mRNA-seq, the expression changes in ABA signalling related candidate genes in leaves in response to salinity, drought and exogenous ABA were conducted. Five barley PP2CAs, three PYR/PYL/RCARs, two SnRK2s (subfamily II and III) and one each of ABCG25 and ABCG40 expressed in the leaf showed differential expression (fold changes of $\geq +1.5$ or ≤ -1.5), while the rest were only marginally affected. PP2CAs, SnRK2s subfamily II, ABCG25 and ABCG40 were found to be up-regulated under all three abiotic stresses studied, whereas PP2CDs, PYR/PYL/RCARs and SnRK2s subfamily III were down-regulated. Studies on these genes and their expression under different stress conditions would provide a fundamental step in screening for candidate genes for development of tolerant crops through genetic modification or breeding.

4.1 Introduction

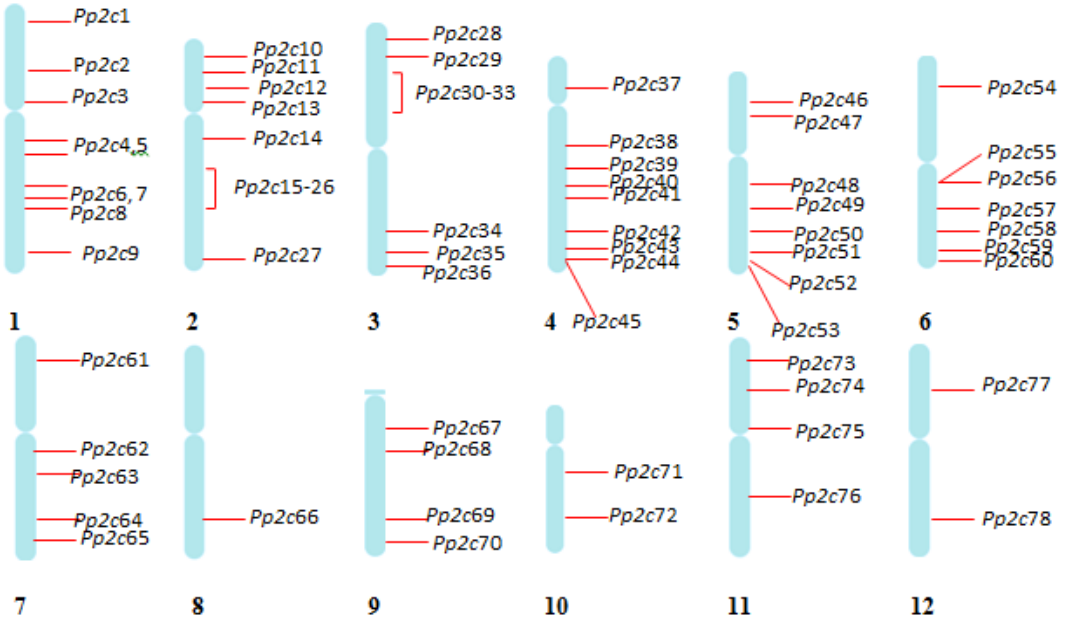
Abscisic acid (ABA) is a phytohormone crucial for plant's adaptive responses to abiotic stresses such as drought and salinity. ABA accumulates in plant cells and induces stomatal closure, leaf abscission and expression of many genes, the products of which may protect vegetative tissues (Umezawa *et al*. 2010; Cutler *et al*. 2010; Leung and Giraudat 1998; Finkelstein *et al.* 2002). The double-negative regulatory system of ABA signalling is comprised of the soluble receptors of ABA called the pyrabactin resistance (PYR)/PYR-like (PYL)/regulatory components of ABA receptor (RCAR)s (PYR/PYL/RCAR) and the enzymes of the classes protein phosphatase 2C (PP2C) and SNF1-related kinase 2 (SnRK2) (detailed in section 1.6). In summary, the PYR/PYL/RCARs bind to ABA and inhibit the inhibitory action of PP2Cs on SnRK2s. The SnRK2s are in turn involved in direct phosphorylation of ABA-responsive element binding protein (AREB)/ABA-responsive element binding factor (ABF)/basic leucine zipper (bZIP) proteins, which in turn regulate the expression of downstream response genes such as salt overly sensitive (SOS) and late embryogenesis abundant (LEA). The import of ABA across the plasma membrane occurs via the ATP binding cassette (ABC) transporter Type G protein called ABCG40, whereas ABCG25 acts as exporter of ABA (Umezawa *et al*. 2010). The PYR/PYL/RCAR protein family members have been identified in rice (Kim *et al*. 2012) and maize (Klingler *et al*. 2010). There has been little research as yet on the ABA regulated genes in the other major cereals, barley and wheat. Studies on these genes and their expression under different stress conditions would provide a fundamental step in screening for candidate genes for development of tolerant crops. Hence the aims of this study were to identify the key genes in the ABA-mediated abiotic stress response pathways in barley from the mRNA-seq databases developed in Chapter 3, and from the barley database by International Barley Genome Sequencing Consortium (IBSC), and explore their expression under salinity, drought and exogenous ABA stress by analysing the mRNA-seq datasets developed in Chapter 3 as well as other published resources such as rice mRNA-seq (Mizuno *et al*. 2010).

4.2 Results

4.2.1 ABA-mediated abiotic stress signalling related candidate genes and proteins in rice

Seventy-eight genes in rice putatively encoding the enzyme group protein phosphatase 2C (PP2C) have been reported by Xue *et al*. (2008). Their genomic loci obtained from the MSU rice genome annotation project database are shown in Figure 4.1. The isoelectric point (pI) and molecular weight (Mw) of the putative PP2C proteins were also extracted from Xue *et al*. (2008) (Appendix X). The protein sequences of Arabidopsis PYR/PYL/RCARs (Ma *et al*. 2009; Park *et al*. 2009), the SnRK2 subfamilies II and III (Yoshida *et al*. 2002), and the ABA transporters ABCG25 and ABCG40 (Kuromori *et al*. 2010; Kang *et al*. 2010) have been reported, but the MSU rice genome annotation project database locus information is not available in Umezawa *et al*. (2010), who provide the NCBI based rice locus. Hence the corresponding rice sequences were identified through BLASTx search on the MSU rice genome annotation project database using Arabidopsis sequences (Umezawa *et al*. 2010) as query. This resulted in the identification of 10 PYR/PYL/RCARs, six SnRK2s (subfamilies II and III) and two ABA-transporters, respectively (Figure 4.1; Appendix X). These sequences have been named on the basis of their closest orthologs in Arabidopsis. For example, the putative PYR/PYL/RCAR from rice that shares highest amino acid sequence identity with AtPYR1 is denoted as OsPYR1.

(A)



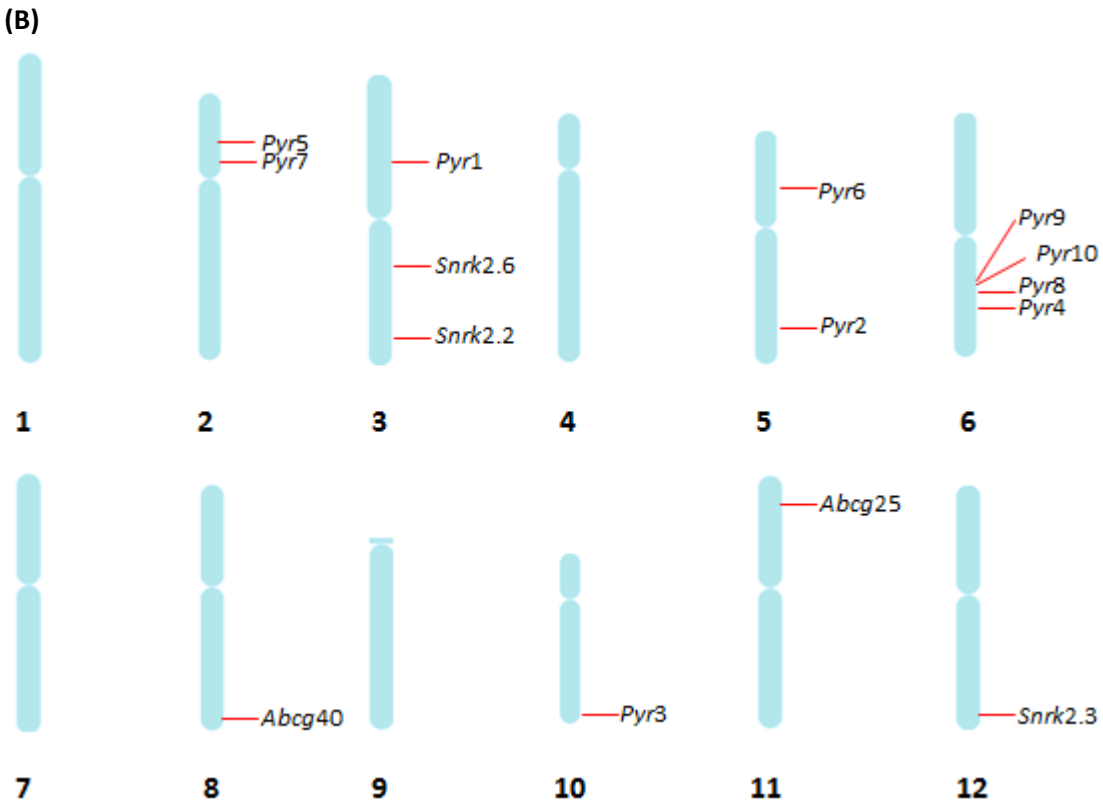


Figure 4.1 Genomic loci of ABA-signalling related candidate genes in *O. sativa*.
(A): Genomic loci of PP2C family; (B): genomic loci of PYR/PYL/RCAR, SnRK2 (subfamily II and III) and ABA-transport related ABC transporter families.

4.2.2 Identification of PP2CA and PP2CD candidate genes in the International Barley Genome Sequencing Consortium (IBSC) CDS database

Phylogenetically, 11 major PP2C subfamilies have been identified in rice (PP2C A, B, C, D, E, F1, F2, G, H, I, K) (Xue *et al*. 2008) and 13 in Arabidopsis (PP2CJ and PP2CL being additional) (Xue *et al*. 2008). PP2Cs are reported to have 11 functional motifs, some being specific to some subfamilies and others shared among several subfamilies (Xue *et al*. 2008). As the total number of putative PP2C loci in rice is 78, it was decided to focus only on identifying two subfamilies in barley, the PP2CA subfamily due to its primary role in ABA-mediated stress tolerance, and PP2CD due to it likely also being a positive regulator in ABA signalling (Xue *et al*. 2008).

The complete list of ABA-signalling related sequences of barley was first extracted from the International Barley Genome Sequencing Consortium (IBSC) database by the RBH method (Method C; Section 2.22.3). Briefly, this involved a tBLASTn search of each rice protein sequence (Xue *et al*. 2008) against the CDS database of the International Barley Genome Sequencing Consortium (http://mips.helmholtz-

muenchen.de/plant/barley/index.jsp/; last accessed May 2014), to find the candidate barley CDSs, which were re-BLASTed (reciprocal BLAST) against the MSU rice genome annotation project database (http://rice.plantbiology.msu.edu/). The barley CDSs (start codon to stop codon) which extracted the same rice locus in reciprocal BLAST as the one used in initial BLAST were called 'RBH best match'. This led to identification of 23 HvPP2CA and 11 HvPP2CD cDNAs from the barley IBSC CDS database (Table 4.1).

The confirmation of authenticity and subgrouping of these sequences was then conducted by analyses of the putative proteins. The 23 HvPP2CA and 11 HvPP2CD cDNAs were translated using six frame translation in BioEdit and the correct reading frame identified by BLASTx in MSU rice genome annotation project database and NCBI. The translated sequences were aligned with the relevant rice and Arabidopsis orthologs using ClustalW2 and the motifs searched for manually. The results showed that 13 PP2CAs and nine PP2CDs had the motifs and were probably functional (Table 4.1; examples in Figure 4.2; alignments given in Appendices XI, XII). Nine of the thirteen putative PP2CAs possessed all seven characteristic motifs of this subfamily (motif I, II, III, IV, VI, VII and VIII; Xue *et al*. 2008), while the other four (MLOC_61822.2, AK367207, MLOC_75421.1, AK358849) were partial. Of the nine putative PP2CDs, seven possessed all eight characteristic motifs for PP2CD (motif I, II, III, IV, V, VII, IX, XI; Xue *et al*. 2008), while three were partial (MLOC_51800, MLOC_16954, MLOC_4262.2). The other 12 sequences (10 PP2CA, 2 PP2CD) could not be assigned to any subfamily as they either lacked the motifs (due to premature stop codons) or exhibited only one or two motifs. The glycine (G) at position 10 of motif 4 of ABI1 and AHG3 (PP2CA), and arginine (R) at position 2 of motif 2, shown to form a bond with the active site of PYR/PYL/RCAR of Arabidopsis (Umezawa *et al*. 2010), were found to be strictly conserved in all PP2CAs in rice and barley as well. The other two PYR/PYL/RCAR binding residues of PP2CA (E142 and W300 of ABI1; Umezawa *et al*. 2010) were also completely conserved (Figure 4.2a).

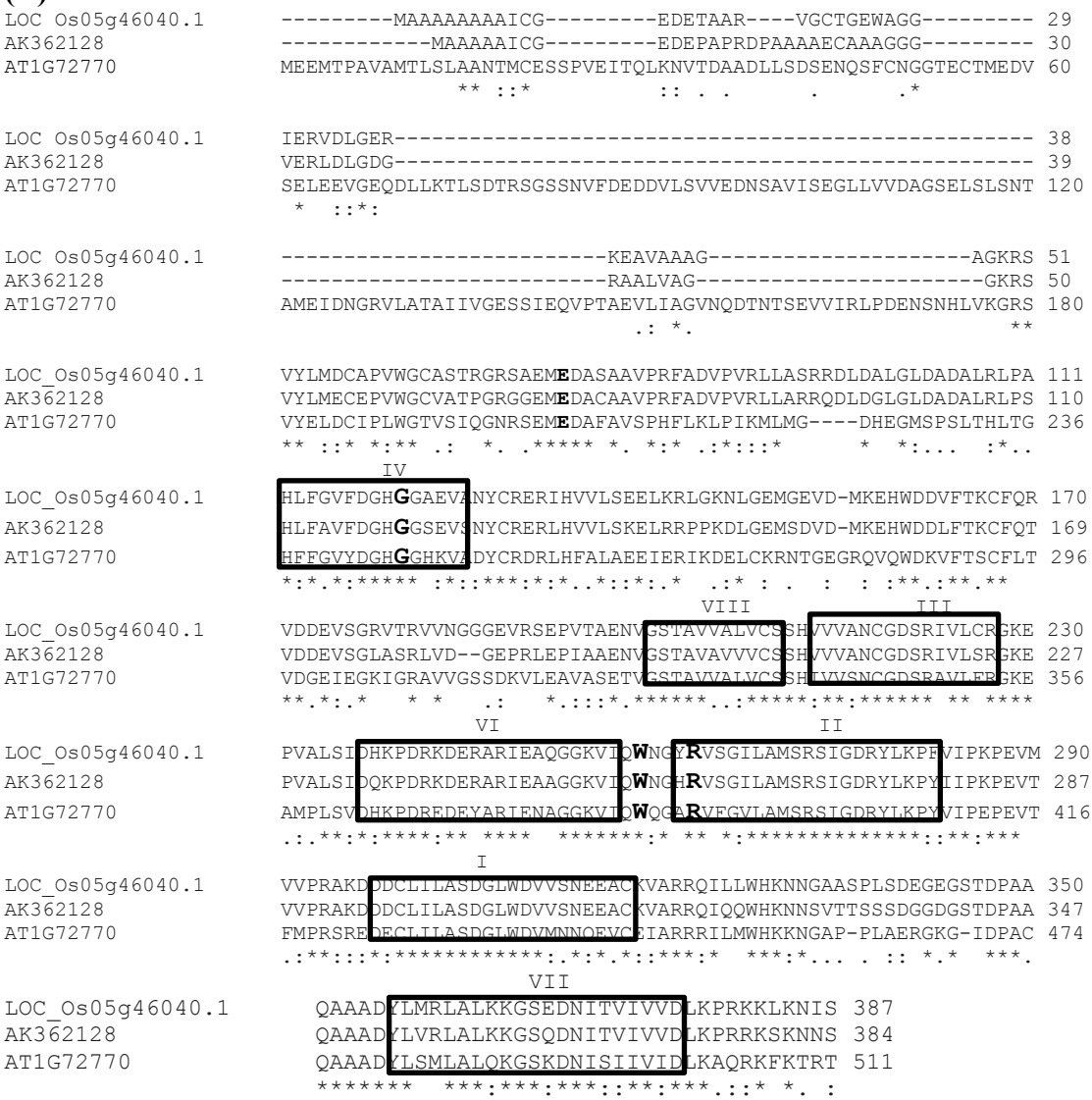
Table 4.1 Identification and motif analysis of PP2CAs and PP2CDs from IBSC barley database.

Barley	Length of barley sequence CDS(bp)/protein(aa)[a]	Mw(kDa)/pI[b]	Most closely related rice sequence (% identity between rice and barley CDS/protein sequences)[c]	Most closely related Arabidopsis sequence (% identity between Arabidopsis and barley CDS/protein sequences)[d]	PP2CA /PP2CD?[e]	Motifs[f]										
						I	II	III	IV	V	VI	VII	VIII	IX	X	XI
PP2CA																
AK251854	1431/477	49.68/4.94	LOC_Os01g40094 (82/83)	AT1G17550 (65/46)	YES	√	√	√	√	NA	√	√	√	NA	NA	NA
MLOC_72950.1	537/179	19.42/4.46	LOC_Os01g40094 (75/31)	AT1G17550 (73/27)	NO					NA				NA	NA	NA
MLOC_78417.1	705/235	24.39/4.72	LOC_Os01g40094 (69/32)	AT1G17550 (62/21)	NO					NA				NA	NA	NA
AK374059	1194/398	42.21/6.67	LOC_Os01g46760 (83/81)	AT2G29380 (57/44)	YES	√	√	√	√	NA	√	√	√	NA	NA	NA
MLOC_61822.2	468/156	16.92/5.25	LOC_Os01g46760 (84/85)	AT2G29380 (64/51)	YES	√	√	X	X	NA	√	√	X	NA	NA	NA
MLOC_8131.1	1179/393	41.86/5.78	LOC_Os01g62760 (82/70)	AT1G07430 (62/48)	YES	√	√	√	√	NA	√	√	√	NA	NA	NA
AK357955	1200/400	43.05/5.71	LOC_Os03g16170 (83/73)	AT1G07430 (61/50)	YES	√	√	√	√	NA	√	√	√	NA	NA	NA
AK367207	777/259	27.12/5.31	LOC_Os03g16170 (74/50)	AT1G07430 (64/43)	YES	√	√	√	X	NA	√	√	√	NA	NA	NA
AK369918	540/180	29.26/11.59	LOC_Os05g38290 (75/22)	AT3G62260 (67/15)	NO					NA				NA	NA	NA
MLOC_35206.1	315/105	11.46/5.03	LOC_Os05g38290 (71/42)	AT3G62260 (81/19)	NO					NA				NA	NA	NA
MLOC_69210.1	441/147	15.69/5.25	LOC_Os05g38290 (71/55)	AT3G62260 (70/27)	NO					NA				NA	NA	NA
MLOC_72576.1	519/173	18.13/5.93	LOC_Os05g38290 (71/23)	AT3G62260 (66/16)	NO					NA				NA	NA	NA
MLOC_75421.1	729/243	25.43/9.02	LOC_Os05g38290 (69/35)	AT1G07430 (61/27)	YES	X	X	√	√	NA	X	X	√	NA	NA	NA
AK362128	1152/384	41.10/5.43	LOC_Os05g46040 (80/78)	AT1G72770 (57/48)	YES	√	√	√	√	NA	√	√	√	NA	NA	NA
MLOC_44992.1	390/124	13.74/9.26	LOC_Os05g51510 (74/26)	AT4G08260 (60/18)	NO					NA				NA	NA	NA
MLOC_7586.1	1149/383	40.58/5.59	LOC_Os05g46040 (76/70)	AT1G72770 (60/47)	YES	√	√	√	√	NA	√	√	√	NA	NA	NA
AK376382	1185/395	41.58/6.32	LOC_Os05g49730 (80/70)	AT2G29380 (60/43)	YES	√	√	√	√	NA	√	√	√	NA	NA	NA
MLOC_15065.1	882/294	30.94/5.49	LOC_Os05g49730 (82/77)	AT2G29380 (59/50)	YES	√	√	√	√	NA	√	√	√	NA	NA	NA
MLOC_72575.1	699/221	23.03/6.16	LOC_Os05g49730 (63/38)	AT3G63340 (72/12)	NO					NA				NA	NA	NA
AK358849	777/259	26.98/4.83	LOC_Os05g51510 (77/67)	AT1G17550 (63/35)	YES	X	X	√	√	NA	X	X	√	NA	NA	NA
MLOC_71293.1	264/88	9.21/4.36	LOC_Os05g51510 (82/37)	AT4G08260 (69/25)	NO					NA				NA	NA	NA
AK367469	1047/349	37.10/6.42	LOC_Os09g15670 (82/79)	AT2G29380 (58/53)	YES	√	√	√	√	NA	√	√	√	NA	NA	NA
MLOC_9189.2	672/224	23.22/4.62	LOC_Os09g15670 (63/30)	AT5G06750 (60/17)	NO					NA				NA	NA	NA
PP2CD																
AK371581	1197/399	43.81/9.30	LOC_Os03g04430 (84/87)	AT4G38520 (70/67)	YES	√	√	√	√	√	NA	NA	√	√	NA	√
MLOC_67251.4	1197/399	43.25/9.61	LOC_Os03g10950 (77/67)	AT4G33920 (59/51)	YES	√	√	√	√	√	NA	NA	√	√	NA	√

Barley	Length of barley sequence CDS(bp)/protein(aa)[a]	Mw(kDa)/pI[b]	Most closely related rice sequence (% identity between rice and barley CDS/protein sequences)[c]	Most closely related Arabidopsis sequence (% identity between Arabidopsis and barley CDS/protein sequences)[d]	PP2CA /PP2CD?[e]	Motifs[f]										
						I	II	III	IV	V	VI	VII	VIII	IX	X	XI
AK364849	1140/380	41.47/9.33	LOC_Os03g55320 (84/82)	ATG33920 (83/58)	YES	√	√	√	√	√	NA	NA	√	√	NA	√
MLOC_16318.4	279/93	10.37/7.86	LOC_Os03g55320 (80/54)	AT4G33920 (80/47)	NO											
MLOC_26021.1	240/80	8.15/4.88	LOC_Os03g61690 (74/56)	AT3G17090 (75/27)	NO											
MLOC_51800	903/301	32.99/7.00	LOC_Os03g61690 (70/61)	AT3G17090 (66/47)	YES	√	√	√	√	X	NA	NA	√	X	NA	√
MLOC_71450.3	1170/390	42.85/6.71	LOC_Os04g49490 (88/89)	AT5G06750 (67/65)	YES	√	√	√	√	√	NA	NA	√	√	NA	√
MLOC_13716.1	1176/392	43.53/8.53	LOC_Os06g50380 (89/91)	AT4G38520 (69/70)	YES	√	√	√	√	√	NA	NA	√	√	NA	√
MLOC_61942.1	1176/392	43.32/8.23	LOC_Os10g39780 (88/88)	AT3G51370 (70/68)	YES	√	√	√	√	√	NA	NA	√	√	NA	√
MLOC_16954	675/225	23.99/8.55	LOC_Os12g39120 (82/80)	AT5G66080 (68/40)	YES	X	√	√	√	X	NA	NA	√	√	NA	√
MLOC_4262.2	504/168	19.12/9.36	LOC_Os12g39120 (80/75)	AT5G66080 (74/48)	YES	√	√	X	X	√	NA	NA	X	X	NA	X

[a]Length of barley CDS in base pairs (bp) and putative protein in amino acids (aa); [b] molecular weight (Mw) and isoelectric point (pI) of the putative proteins were predicted in the ExPASy Proteomics Server (http://web.expasy.org/compute_pi/); [c] most closely related rice sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW2; [d] most closely related Arabidopsis sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW2; [e]putative PP2CA or PP2CD was confirmed on basis of motif analysis; [f] sequence and position of domains were extracted from Xue *et al*. (2008); √: motif present; X: no motif found; NA: not applicable.

(A)


LOC Os05g46040.1 ---------MAAAAAAAAICG---------EDETAAR----VGCTGEWAGG--------- 29
AK362128 ----------MAAAAAICG---------EDEPAPRDPAAAAECAAAGGG--------- 30
AT1G72770 MEEMTPAVAMTLSLAANTMCESSPVEITQLKNVTDAADLLSDSENQSFCNGGTECTMEDV 60
LOC Os05g46040.1 IERVDLGER---------- 38
AK362128 VERLDLGDG---------- 39
AT1G72770 SELEEVGEQDLLKTLSDTRSGSSNVFDEDDVLSVVEDNSAVISEGLLVVDAGSELSLSNT 120
LOC Os05g46040.1 ----------KEAVAAAG----------AGKRS 51
AK362128 ----------RAALVAG----------GKRS 50
AT1G72770 AMEIDNGRVLATAIIVGESSIEQVPTAEVLIAGVNQDTNTSEVVIRLPDENSNHLVKGRS 180
LOC Os05g46040.1 VYLMDCAPVWGCASTRGRSAEMEDASAAVPRFADVPVRLLASRRDLDALGLDADALRLPA 111
AK362128 VYLMECEPVWGCVATPGRGGEMEDACAAVPRFADVPVRLLARRQDLDGLGLDADALRLPS 110
AT1G72770 VYELDCIPLWGTVSIQGNRSEMEDAFAVSPHFLKLPIKMLMG----DHEGMSPSLTHLTG 236
IV
LOC Os05g46040.1 HLFGVFDGHGGAEVANYCRERIHVVLSEELKRLGKNLGEMGEVD-MKEHWDDVFTKCFQR 170
AK362128 HLFAVFDGHGGSEVSNYCRERLHVVLSKELRRPPKDLGEMSDVD-MKEHWDDLFTKCFQT 169
AT1G72770 HFFGVYDGHGGHKVADYCRDRLHFALAEEIERIKDELCKRNTGEGRQVQWDKVFTSCFLT 296
VIII III
LOC Os05g46040.1 VDDEVSGRVTRVVNGGGEVRSEPVTAENVGSTAVVALVCSGHVVVANCGDSRIVLCRGKE 230
AK362128 VDDEVSGLASRLVD--GEPRLEPIAAENVGSTAVAVVVCSGHVVVANCGDSRIVLSRGKE 227
AT1G72770 VDGEIEGKIGRAVVGSSDKVLEAVASETVGSTAVVALVCSGHLVVSNCGDSRAVLFRGKE 356
VI II
LOC Os05g46040.1 PVALSIQHKPDRKDERARIEAQGGKVIQWNGYRVSGILAMSRSIGDRYLKPFVIPKPEVM 290
AK362128 PVALSIQQKPDRKDERARIEAAGGKVIQWNGHRVSGILAMSRSIGDRYLKPYIIPKPEVT 287
AT1G72770 AMPLSVQHKPDREDEYARIENAGGKVIQWQGARVFGVLAMSRSIGDRYLKPYVIPEPEVT 416
I
LOC Os05g46040.1 VVPRAKDQDCLILASDGLWDVVSNEEACKVARRQILLWHKNNGAASPLSDEGEGSTDPAA 350
AK362128 VVPRAKDQDCLILASDGLWDVVSNEEACKVARRQIQQWHKNNSVTTSSSDGGDGSTDPAA 347
AT1G72770 FMPRSREQECLILASDGLWDVMNNQEVCGIARRRILMWHKKNGAP-PLAERGKG-IDPAC 474
VII
LOC Os05g46040.1 QAAADKLMRLALKKGSEDNITVIVVDLKPRKKLKNIS 387
AK362128 QAAADKLVRLALKKGSQDNITVIVVDLKPRRKSKNNS 384
AT1G72770 QAAADKLSMLALQKGSKDNISIIVIDLKAQRKFKTRT 511

(B)


IX
AK371581 -MAAVMDYFRSCWGARSRSGRRSKKGSDAAGRQDGLLWYKDAGQAATGEFSMAVVQANNL 59
LOC Os03g04430.1 MLAAVMDYFRSCWGPRSPAGHRVR-GSDVAGRQDGLLWYKDAGQLVTGEFSMAVVQANNL 59
AT4G38520 MLSGLMNFLNACLWPRS--DQQARSASDSGGRQEGLLWFRDSGQHVFGDFSMAVVQANSL 58
IV
AK371581 LEDQSQVESGSLSMADPGPQGTFVGVYDGHGGPETSRFINDNMFHHLRRFATEHKCMSAD 119
LOC Os03g04430.1 LEDQSQVESGALSMAEPGPQGTFIGVYDGHGGPETARFINDHMFHHLRRFATEHKCMSTD 119
AT4G38520 LEDQSQLESGSLSSHDSGPFGTFVGVYDGHGGPETSRFINDHMFHHLKRFTAEQQCMSSE 118
VIII III
AK371581 VIRKAFQATEDGFLSVVSKEWSVKPQIAAVGSCCLVGVICSGTLYIANAGDSRAVLGRLV 179
LOC Os03g04430.1 VIRKAFQATEEGFLSLVSKQWSLKPQIAAVGSCCLVGVICSGTLYVANLGDSRAVLGRFV 179
AT4G38520 VIKKAFQATEEGFLSIVTNQFQTRPQIATVGSCCLVSVICDGKLYVANAGDSRAVLGQVM 178
XI II
AK371581 KATGQVVAMQLSAEHNACYEEVRQELQSSHPHDPQIVVLKHNVWRVKGLIQISRSIGDVY 239
LOC Os03g04430.1 KSTGEVVATQLSSEHNACYEEVRQELQASHPDDPQIVVLKHNVWRVKGLIQISRSIGDVY 239
AT4G38520 RVTGEAHATQLSAEHNASIESVRRELQALHPDHPDIVVLKHNVWRVKGLIQVSRSIGDVY 238
I
AK371581 LKREYNRTPLHSKFRLRETFKKPILSSEPAIAVHQIQPSDQFVIFASDGLWEHLSNQEA 299
LOC Os03g04430.1 LKREYNREPLHSKFRLRETFKRPILSSEPAIAVHQIQPNDHFVIFASDGLWEHLSNQEA 299
AT4G38520 LKRSFNREPLYAKFRLRSPFSKPLLSAEPAITVHTLEPHDQFIICASDGLWEHMSNQEA 298

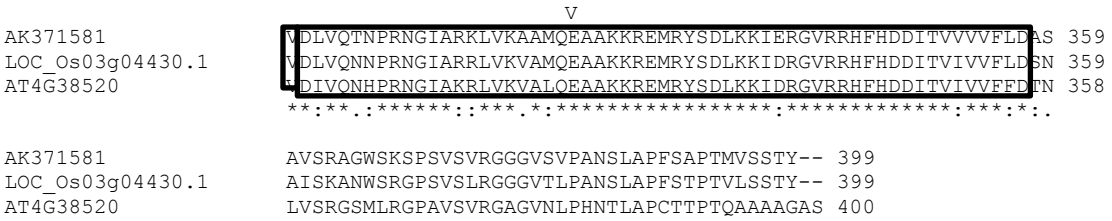


Figure 4.2 Example of alignments of amino acid sequences of the putative barley PP2CAs and PP2CDs with those in rice and Arabidopsis

(A) Alignment of barley PP2CA AK362128 with rice PP2CA encoded by LOC_Os05g46040 and Arabidopsis PP2CA encoded by AT1G72770;
(B) Alignment of barley PP2CD AK371581 with rice PP2CD encoded by LOC_Os03g04430 and Arabidopsis PP2CD encoded by AT4g38520.
Asterisk indicate the conservation of residue in all the three species and dashes indicate the gaps or absence of sequence data; alignments were made using CLUSTALW2 (http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=clustalw2; last accessed May 2014) and then bioedit was used to align the conserved functional motifs; box: conserved functional motifs; roman numbers: indicate the specific motifs; bold indicate amino acid involved in binding of PP2CA to PYR/PYL/RCAR.; rice sequences were derived from the MSU rice genomic annotation project (http://rice.plantbiology.msu.edu/; last accessed May 2014) and Arabidopsis sequences were derived from TAIR (http://www.arabidopsis.org/index.jsp); motif sequences and the positions were derived from Xue *et al*. (2008).

The gDNA contigs for bowmen and Morex lines and chromosomal location were extracted by BLASTn in barley IBSC genomic database. Comparison of gDNA sequences and cDNAs using the gene structure display server resulted in potential intron-exon structure as explained below (Section 4.2.6). Searches in the IBSC genomic database also showed the chromosomal distribution of HvPP2CAs and PP2CDs. The HvPP2CAs were located on all chromosomes except 6, and the HvPP2CDs were on 1H, 2H, 4H, 5H and 7H (Table 4.5). The results of interspecies sequence comparisons are detailed below.

4.2.3 Identification of PYR/PYL/RCAR candidate genes in the IBSC CDS database

The carboxyl group of ABA forms an ionic bond with the side chain of a lysine residue of PYR/PYL/RCAR in Arabidopsis (K 59 of PYR1) and a water-mediated hydrogen bond network with side chains of five polar residues (P 88, E 94, R 116, S 122, E 141 of Arabidopsis PYR1). Thirteen other residues directly involved in ABA binding are F 61, V 83, L 87, A 89, I 110, H 115, L 117, Y 120, F 159, A 160, V 163, V 164 and N 167 of Arabidopsis PYR1 (Yin *et al*. 2009). The complete list of barley PYR/PYL/RCARs were first extracted from barley (*Hordeum vulgare*) IBSC CDS database by using the RBH method (Method C, as explained above). This led to 12 PYR/PYL/RCAR deduced cDNA sequences reported in the barley IBSC CDS database (Table 4.2). The confirmation of authenticity of these sequences was then conducted by was then

conducted by analyses of the putative proteins. The cDNAs of the 12 PYR/PYL/RCARs were translated using six frame translations in BioEdit and the correct reading frame identified by using BLASTx in MSU rice genome annotation project database and NCBI. The translated sequences were aligned with the relevant rice and Arabidopsis orthologs using ClustalW2 and amino acids involved in ABA binding searched for manually. The results showed that 10 of the 12 PYR/PYL/RCARs had such amino acids and were probably functional (Table 4.2; examples in Figure 4.3; alignments given in Appendix XIII). Eight of the ten PYR/PYL/RCARs possessed all 19 characteristic amino acids (14 directly involved in ABA binding and five co-ordinate ABA through water molecule; Yin *et al*. 2009), while the other two (MLOC_3912.1 and MLOC_49654.1) were partial. The other two deduced sequences could not be classified as PYR/PYL/RCAR as they either lacked these characteristic amino acids (due to premature stop codons) or exhibited only one or two amino acids. The residues involved in the direct binding of PYR/PYL/RCAR to ABA at positions 1, 2, 3, 4, 7, 8, 9, 10, 13 and 14 were strictly conserved as compared to the residues at positions 5, 6, 11 and 12. The residues that co-ordinate ABA through the water molecule were also found to be strictly conserved at all positions 1 to 5**.**

Table 4.2 Identification and motif analysis of PYR/PYL/RCARs from IBSC barley database.

Barley	Length of barley sequence CDS(bp)/protein(aa)[a]	Mw(kDa)/pI[b]	Most closely related rice sequence (% identity between rice and barley CDS/protein sequences)[c]	Most closely related Arabidopsis sequence (% identity between Arabidopsis and barley CDS/protein sequences)[d]	PYR?[e]	Residues predicted to be directly involved in ABA binding[f]														Residues that coordinate ABA through water				
						1	2	3	4	5	6	7	8	9	10	11	12	13	14	1	2	3	4	5
MLOC_39291.1	597/199	21.71/5.23	LOC_Os02g13330(86/80)	AT2G26040(64/61)	Y	K	F	V	L	A	V	H	L	Y	F	T	V	V	N	P	E	R	S	E
RICE						.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.	.	.
ARA						.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.	.	.
MLOC_46394	591/197	22.36/5.92	LOC_Os06g33690(74/63)	AT5G53160(64/53)	Y	K	F	V	L	A	F	H	L	Y	F	V	V	P	N	P	E	P	S	E
RICE						.	.	.	.	.	.	.	.	.	.	I	.	.	.	.	.	M	.	.
ARA						.	.	.	.	.	I	.	.	.	.	.	.	.	.	.	.	R	.	.
AK376521	666/222	23.20/7.71	LOC_Os03g18600(81/72)	AT5G05440(62/54)	Y	K	F	V	L	A	V	H	L	Y	F	V	I	V	N	P	E	R	S	E
RICE						.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ARA						.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MLOC_71349.1	672/224	23.38/6.38	LOC_Os03g18600(88/85)	AT5G05440(60/52)	Y	K	F	V	L	A	V	H	L	Y	F	I	I	V	N	P	E	R	S	E
RICE						.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.	.	.
ARA						.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.	.	.
MLOC_3912.1	495/165	12.97/9.18	LOC_Os05g12260(80/70)	AT5G53160(68/44)	Y	K	F	V	L	A	F	H	L	X	X	X	X	X	X	P	E	R	X	X
RICE						.	.	.	.	.	.	.	.	X	X	X	X	X	X	.	.	.	X	X
ARA						.	.	.	.	.	I	.	.	X	X	X	X	X	X	.	.	.	X	X
MLOC_60739.1	489/163	17.57/7.20	LOC_Os05g39580(92/91)	AT5G05440(69/65)	Y	K	F	V	L	A	V	H	L	Y	F	V	I	V	N	P	E	R	S	E
RICE						.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ARA						.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MLOC_65591.1	618/206	22.00/9.19	LOC_Os05g39580(84/81)	AT5G05440(67/55)	Y	K	F	V	L	G	I	H	L	Y	F	V	I	V	N	P	E	R	S	E
RICE						.	.	.	.	A	V	.	.	.	.	.	.	.	.	.	.	.	.	.
ARA						.	.	.	.	A	V	.	.	.	.	.	.	.	.	.	.	.	.	.
MLOC_1585.1	516/172	17.73/9.40	LOC_Os06g36670(62/60)	AT2G26040(50/50)	N																			
MLOC_49654.1	348/116	12.58/5.21	LOC_Os06g36670(84/66)	AT2G26040(67/50)	Y	K	F	V	L	A	I	H	L	Y	X	X	X	X	X	P	E	C	S	X
RICE						.	.	.	.	.	V	.	.	.	X	X	X	X	X	.	.	R	.	X
ARA						.	.	.	.	.	V	.	.	.	X	X	X	X	X	.	.	R	.	X
MLOC_72289.1	588/196	21.26/6.31	LOC_Os06g36670(90/85)	AT2G26040(61/57)	Y	K	F	V	L	A	V	H	L	Y	F	T	V	V	N	P	E	R	S	E
RICE						.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ARA						.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.	.	.

Barley	Length of barley sequence CDS(bp)/protein(aa)[a]	Mw(kDa)/pI[b]	Most closely related rice sequence (% identity between rice and barley CDS/protein sequences)[c]	Most closely related Arabidopsis sequence (% identity between Arabidopsis and barley CDS/protein sequences)[d]	PYR?[e]	Residues predicted to be directly involved in ABA binding[f]													Residues that coordinate ABA through water				
						1	2	3	4	5	6	7	8	9	10	11	12	13	14	1	2	3	4	5
MLOC_80832.1	489/163	18.18/7.72	LOC_Os06g36670(61/42)	AT2G26040(54/30)	N																			
AK363238	636/212	22.67/4.96	LOC_Os10g42280(85/77)	AT4G17870(58/58)	Y	K	F	V	L	A	I	H	L	Y	F	A	V	V	N	P	E	R	S	E
RICE						.	.	.	.	.	.	.	.	.	.	.	I	.	.	.	.	.	.	.
ARA						.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

[a]Length of barley CDS in base pairs (bp) and putative protein in amino acids (aa); [b] molecular weight (Mw) and isoelectric point (pI) of the putative proteins were predicted in the ExPASy Proteomics Server (http://web.expasy.org/compute_pi/); [c] most closely related rice sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW2; [d] most closely related Arabidopsis sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW; [e] putative PYR/PYL/RCAR was confirmed on positioning of various residues involved in binding of ABA; [f] positioning of various residues was extracted from Yin *et al*. (2009); X: residue not found; The conserved amino acid residue in rice and Arabidopsis are shown with dots, otherwise the amino acid mentioned in rice and Arabidopsis rows.

```
LOC_Os02g13330.1      MEPHMERALREAVASEAERRELEGVVRAHHTFPAAERAAGPGRRPTCTSLVAQRVDAPLA 60
MLOC_39291.1          MEHHMESALRQGLT-EPERRELEGVVEEHHTFPG--RASG-----TCTSLVTQRVQAPLA 52
AT2G26040             ---MSSSPAVKGLT-DEEQKTLEPVIKTYHQFEP---DPT-----TCTSLITQRIHAPAS 48
                           . .  :.:: : *:: ** *:. :* *      .      *****::**:.** :

LOC_Os02g13330.1      AVWPIVRGFANPQRYKHFIKSCELAAGDGATVGSVREVAVVSGLPASTSTERLEILDDDR 120
MLOC_39291.1          AVWDIVRGFANPQRYKHFIKSCALAAGDGATVGSVREVTVVSGLPASTSTERLEILDDDR 112
AT2G26040             VVWPLIRRFDNPERYKHFVKRCRLISGDGD-VGSVREVTVISGLPASTSTERLEFVDDDH 107
                      .** ::* * **:*****:* * * :***  *******:*:*************::***:

LOC_Os02g13330.1      HVLSFRVVGGDHRLRNYRSVTSVTEFSSPSSPPRPYCVVVESYVVDVPEGNTEEDTRMFT 180
MLOC_39291.1          HILSFCVVGGEHRLRNYRSVTSVTEFTDQPSGP-SYCVVVESYVVDVPEGNTEEDTRMFT 171
AT2G26040             RVLSFRVVGGEHRLKNYKSVTSVNEFLNQDSGK-VYTVVLESYTVDIPEGNTEEDTKMFV 166
                      ::*** ****:***:**:*****.** .  *    * **:***.**:**********:**.

LOC_Os02g13330.1      DTVVKLNLQKLAAVATS-SSPPAAGNHH 207
MLOC_39291.1          DTVVKLNLQKLAAIATTTSSPPPLDGQS 199
AT2G26040             DTVVKLNLQKLGVAATS---APMHDDE- 190
```

Figure 4.3 Example of alignment of putative amino acid sequences of putative barley PYR/PYL/RCARs with rice and Arabidopsis

Alignment of barley PYR/PYL/RCAR MLOC_39291.1 with rice PYR/PYL/RCAR encoded by LOC_Os02g13330.1 and Arabidopsis PYR/PYL/RCAR encoded by AT2G26040.
Asterisk indicate the conservation of residue in all the three species and dashes indicate the gaps or absence of sequence data; alignments were made using CLUSTALW2 (http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=clustalw2; last accessed May 2014) and then bioedit was used to align the conserved functional motifs; bold indicate residues that are directly involved in ABA binding and grey shading indicate residues that form hydrogen bond with ABA through water; rice sequences were derived from the MSU rice genomic annotation project (http://rice.plantbiology.msu.edu/; last accessed May 2014) and Arabidopsis sequences were derived from TAIR (http://www.arabidopsis.org/index.jsp); positioning of amino acid residues were derived from Yin *et al*. (2009).

The gDNA contigs for barley PYR/PYL/RCAR in bowmen and Morex lines and chromosomal location were extracted by BLASTn in barley IBSC genomic database. Comparison of gDNA sequences and cDNAs using the gene structure display server resulted in potential intron-exon structure as explained below (Section 4.2.6; Table 4.5). Searches in the IBSC genomic database also showed that HvPYR/PYL/RCARs were located 1H, 3H, 4H and 7H (Table 4.5). The results of interspecies sequence comparisons are detailed below.

4.2.4 Identification of SnRK2 (subfamily II and III) candidate genes in the International Barley Genome Sequencing Consortium (IBSC) CDS database

SnRK2s have been studied in a number of plant species, and exhibit conserved features such as (i) a serine residue (Ser175 in Arabidopsis SnRK2.6), the phosphorylation of which has been associated with activation of SnRK2 (ii) an ABA-responsive box only found in strongly ABA responsive kinases (C terminal Domain II), and (iii) a SnRK2-conserved box conserved in all kinases (C terminal Domain I) (Belin *et al*. 2006). The C-terminal domain I was found only in SnRK2 subfamily III members, whereas C-

terminal domain II was found in both SnRK2 subfamily II and III. The complete list of barley SnRK2s (Subfamily II and III) were first extracted from barley (*Hordeum vulgare*) IBSC CDS database by using the RBH method (Method C, as explained above). This led to seven SnRK2s (Subfamily II and III) deduced cDNA sequences reported in the barley IBSC CDS database (Table 4.3). The confirmation of authenticity and subgrouping of these sequences was then conducted by analyses of the putative proteins. The cDNAs of the seven SnRK2s (Subfamily II and III) were translated using six frame translations in BioEdit and the correct reading frame identified by using BLASTx in MSU rice genome annotation project database and NCBI. The translated sequences were aligned with the relevant rice and Arabidopsis orthologs using ClustalW2 and characteristic domains searched for manually. The results showed that six of the seven SnRK2s (Subfamily II and III) had such domains and were probably functional (Table 4.3; examples in Figure 4.4; alignments given in Appendix XIV). MLOC_22145.2, AK374298 and MLOC_3013.1 appeared to belong to subfamily III as they possessed the ABA-responsive box, whereas AK372880, AK251684 and MLOC_69212 likely belonged to subfamily II (as they lacked this box). All of the identified SnRK2s (subfamily II and III) were complete sequences, which were located on the chromosomes 1H, 2H, 4H and 5H (Table 4.5). SnRK2 (subfamily II and III) gDNA contigs of bowmen and Morex lines of barley were extracted by BLASTn in barley IBSC genomic database (Table 4.5). The results of interspecies sequence comparisons are detailed below.

Table 4.3 Identification and motif analysis of SnRK2s (subfamily II and III) from IBSC barley database.

Barley	Length of barley sequence CDS(bp)/protein(aa)[a]	Mw(kDa)/pI[b]	Most closely related rice sequence (% identity between rice and barley CDS/protein sequences)[c]	Most closely related Arabidopsis sequence (% identity between Arabidopsis and barley CDS/protein sequences)[d]	Predicted SnRK2 subtype[e]	SnRK2?[f]	Presence of Ser phosphorylated for the activation of kinase[g]	ABA responsive box[g]	SnRK2 conserved box[g]
MLOC_22145.2	1083/361	40.64/4.80	LOC_Os03g41460(90/96)	AT5G66880(73/80)	III	YES	√	√	√
MLOC_36654.6	726/242	27.13/4.84	LOC_Os03g41460(74/58)	AT5G66880(70/50)	III	NO			
AK374298	1098/366	41.53/4.86	LOC_Os03g55600(90/95)	AT3G50500(72/77)	III	YES	√	√	√
MLOC_3013.1	1071/357	40.15/4.94	LOC_Os12g39630 (87/86)	AT4g33950 (72/78)	III	YES	√	√	√
AK372880	1026/342	38.82/5.75	LOC_Os03g27280 (87/90)	AT4G40010 (69/70)	II	YES	√	NA	√
MLOC_69212	1023/341	38.59/5.45	LOC_Os07g42940 (87/91)	AT1G78290(71/72)	II	YES	√	NA	√
AK251684	564/188	21.29/8.48	LOC_Os10g41490 (87/95)	AT1G78290(65/69)	II	YES	√	NA	√

[a]Length of barley CDS in base pairs (bp) and putative protein in amino acids (aa); [b] molecular weight (Mw) and isoelectric point (pI) of the putative proteins were predicted in the ExPASy Proteomics Server (http://web.expasy.org/compute_pi/); c: most closely related rice sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW2; d: most closely related Arabidopsis sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW2; e: SnRK2 subfamily II or III was confirmed on positioning of ABA responsive box and SnRK2 conserved box; f: serine residue involved in the activation of SnRK2s g: ABA responsive box, only found in strongly ABA responsive kinases (C terminal Domain II). SnRK2 conserved box: conserved in all kinases (C terminal Domain I) (Belin *et al*. 2006); √: domain/box present; X: no domain/box found; NA: not applicable.

```
LOC_Os03g41460.1      MDRAALTVGPGMDMPIMHDGDRYELVRDIGSGNFGVARLMRSRADGQLVAVKYIERGDKI 60
MLOC_22145.2          MDRAALTVGPGMDMPIMHDGDRYELVKDIGSGNFGVARLMRNRADGQLVAVKYIERGEKI 60
AT5G66880             MDRAPVTTGP-LDMPIMHDSDRYDFVKDIGSGNFGVARLMRDKLTKELVAVKYIERGDKI 59
                      ****.:*.** :*******.***::*:***************.:   :***********:**

LOC_Os03g41460.1      DENVQREIINHRSLRHPNIIRFKEVILTPTHLAIVMEYASGGELFERICNAGRFSEDEAR 120
MLOC_22145.2          DENVQREIINHRSLRHPNIIRFKEVILTPTHLAIVMEYASGGELFERICNAGRFSEDEAR 120
AT5G66880             DENVQREIINHRSLRHPNIVRFKEVILTPTHLAIIMEYASGGELYERICNAGRFSEDEAR 119
                      ******************:**************:*********:***************

LOC_Os03g41460.1      FFFQQLISGVSYCHSMQVCHRDLKLENTLLDGSTAPRLKICDFGYSKSSVLHSQPKSTVG 180
MLOC_22145.2          FFFQQLLSGVSYCHSMQVCHRDLKLENTLLDGSTAPRLKICDFGYSKSSVLHSQPKSTVG 180
AT5G66880             FFFQQLLSGVSYCHSMQICHRDLKLENTLLDGSPAPRLKICDFGYSKSSVLHSQPKSTVG 179
                      ******:**********:***************.************************

LOC_Os03g41460.1      TPAYIAPEVLLKKEYDGKIADVWSCGVTLYVMLVGAYPFEDPDEPKNFRKTIQRILGVQY 240
MLOC_22145.2          TPAYIAPEVLLKKEYDGKIADVWSCGVTLYVMLVGAYPFEDPDEPKNFRKTIQRILSVQY 240
AT5G66880             TPAYIAPEVLLRQEYDGKIADVWSCGVTLYVMLVGAYPFEDPEEPRDYRKTIQRILSVKY 239
                      **********::*******************************:**:::********.*:*

LOC_Os03g41460.1      SIPDYVHISPECRDLIARIFVANPATRISIPEIRNHPWFLKNLPADLMDDSKMSSQYEEP 300
MLOC_22145.2          SIPDYVHISSECRDLIAKIFVGNPATRITIPEIRNHPWFLKNLPADLVDDSTMSSQYEEP 300
AT5G66880             SIPDDIRISPECCHLISRIFVADPATRISIPEIKTHSWFLKNLPADLMNESNTGSQFQEP 299
                      **** ::**.** .**::***.:*****:****:.*.**********:::*. .**::**

LOC_Os03g41460.1      EQPMQSMDEIMQILAEATIPAAGSGGINQFLNDGLDLDDDMEDLDSDPDLDVESSGEIVY 360
MLOC_22145.2          EQPMQSMDEIMQILAEATIPAAGSR-INQFLNDGLDLDDDMDDLDSDADLDVESSGEIVY 359
AT5G66880             EQPMQSLDTIMQIISEATIPAVRNRCLDDFMTDNLDLDDDMDDFDSESEIDIDSSGEIVY 359
                      ******:* ****::******. .  :::*:.*.*******:*:**:.::*::*******

LOC_Os03g41460.1      AM 362
MLOC_22145.2          AM 361
AT5G66880             AT 361
```

Figure 4.4 Example of alignment of putative amino acid sequences of putative barley SnRK2s with rice and Arabidopsis

Alignment of barley SnRK2 MLOC_221451.2 with rice SnRK2 encoded by LOC_Os03g41460.1 and Arabidopsis SnRK2 encoded by AT5G66880.

Asterisk indicate the conservation of residue in all the three species and dashes indicate the gaps or absence of sequence data; alignments were made using CLUSTALW2 (http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=clustalw2; last accessed May 2014) and then bioedit was used to align the conserved functional motifs; box: conserved functional motifs; bold: indicate residues important in binding to PP2C; rice sequences were derived from the MSU rice genomic annotation project (http://rice.plantbiology.msu.edu/; last accessed May 2014) and Arabidopsis sequences were derived from TAIR (http://www.arabidopsis.org/index.jsp); motif sequences and the positions were derived from Belin *et al*. (2006).

4.2.5 Identification of ABA-transport related ABC transporter candidate genes in the International Barley Genome Sequencing Consortium (IBSC) CDS database

The putative ABA-related ABC transporter sequences (ABCG25 and ABCG40) were analysed for the conserved motifs of the Nucleotide binding domain (NBD). (i) Walker Box A and Walker Box B separated by approximately 180-220 amino acids are ATP/GTP binding motif (ii) the ABC transporter signature motif located in between the two Walker Boxes (Bairoch 1992). Barley ABCG25 and ABCG40 were first extracted from barley (*Hordeum vulgare*) IBSC CDS database by using the RBH method (Method C, as explained above). This led to six ABA-transport related ABC transporter deduced cDNA sequences reported in the barley IBSC CDS database (Table 4.4). The confirmation of authenticity of these sequences was then conducted by analyses of the

putative proteins. The cDNAs of the six ABA-transport related ABC transporters were translated using six frame translations in BioEdit and the correct reading frame identified by using BLASTx in MSU rice genome annotation project database and NCBI. The translated sequences were aligned with the relevant rice and Arabidopsis orthologs using ClustalW2 and characteristic motifs searched for manually. The results showed that four of the six ABA-transport related ABC transporters had such domains and were probably functional (Table 4.4; examples in Figure 4.5; alignments given in Appendix XV). MLOC_11284.1 and MLOC_64058.1 were found to encode the region between the characteristic motifs and C-terminal region of putative ABA related ABC transporter sequence without any characteristic motifs respectively; hence were not classified as probably functional ABC transporter, whereas the other four sequences had the ABC transporter characteristic features.

Table 4.4 Identification and motif analysis of ABCG25 and ABCG40 from IBSC barley database.

Barley	Length of barley sequence CDS(bp)/protein(aa)[a]	Mw(kDa)/pI[b]	Most closely related rice sequence (% identity between rice and barley CDS/protein sequences)[c]	Most closely related Arabidopsis sequence (% identity between Arabidopsis and barley CDS/protein sequences)[d]	ABC transporter ?[e]	Motifs[f]		
						Walker Box A	Walker Box B	Signature motif
MLOC_11283.1	1935/645	73.39/7.98	LOC_Os08g43120(82/82)	ABCG40(67/55)	YES	√	√	X
MLOC_11284.1	1311/437	49.79/7.11	LOC_Os08g43120(87/67)	ABCG40(77/45)	No			
MLOC_54794.6	3783/1261	143.01/8.93	LOC_Os08g43120(84/84)	ABCG40(64/52)	YES	√	√	√
MLOC_64058.1	249/83	-	-	-	NO			
MLOC_68581.1	4044/1348	153.85/8.09	LOC_Os08g43120(83/79)	ABCG40(62/48)	YES	√	√	√
MLOC_62985.1	1923/641	68.02./9.07	LOC_Os11g07600(83/83)	ABCG25(59/52)	YES	√	X	X

[a]Length of barley CDS in base pairs (bp) and putative protein in amino acids (aa); [b] molecular weight (Mw) and isoelectric point (pI) of the putative proteins were predicted in the ExPASy Proteomics Server (http://web.expasy.org/compute_pi/); [c] most closely related rice sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW2; [d] most closely related Arabidopsis sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW2; [e] putative ABC transporter was confirmed on positioning of ABA responsive box and SnRK2 conserved box; [f] Sequence and positioning of motifs were extracted from van den Brule and Smart (2002); √: box present; X: no box found; NA: not applicable.

The gDNA contigs for barley ABCG25 and ABCG40 in bowmen and Morex lines and chromosomal location were extracted by BLASTn in barley IBSC genomic database. Comparison of gDNA sequences and cDNAs using the gene structure display server resulted in potential intron-exon structure as explained below (Section 4.2.6; Table 4.5). Searches in the IBSC genomic database also showed that Hv.ABCG25 and Hv.ABCG40 were located on the chromosomes 3H, 4H and 7H (Table 4.5). The results of interspecies sequence comparisons are detailed below.

```
LOC_Os08g43120.1      ------------------------------------------------------MAAA 4
MLOC_54794.6          ------------------------------------------------------------
ABCG40                MEGTSFHQASNSMRRNSSVWKKDSGREIFSRSSREEDDEEALRWAALEKLPTFDRLRKGI 60


LOC_Os08g43120.1      VELTGDGGTTAETRWLSP-------------PLTHDDNRGFLQMLREKKERLGVGAAKVE 51
MLOC_54794.6          ------------------------------------------------------------
ABCG40                LTASHAGGPINEIDIQKLGFQDTKKLLERLIKVGDDEHEKLLWKLKKRIDRVGIDLPTIE 120


LOC_Os08g43120.1      VRLEKLTVEADVRVGRRAVPTLLNCAINAAQELAACAHMCTTRKKPMKIINEATGTIRPS 111
MLOC_54794.6          ------------------------------------MCTTRKKPIKIINGASGTIRPS 22
ABCG40                VRFDHLKVEAEVHVGGRALPTFVNFISNFADKFLNTLHLVPNRKKKFTILNDVSGIVKPG 180
                                                          : ..*** :.*:* .:* ::*.
                              Walker Box A
LOC_Os08g43120.1      RMTLLLGAPGSGKTTLLKALAGKLDSSLKMKGKVTYNGEEVNSSTPQYLHAYVSQYDLHH 171
MLOC_54794.6          RMTLLLGAPGSGKTTFLKALAGKLDSSLKLKGKVMYNGEEVKSSTPQYLHAYISQYDLHH 82
ABCG40                RMALLLGPPSSGKTTLLLALAGKLDQELKQTGRVTYNGHGMNEFVPQRTAAYIGQNDVHI 240
                      **:****.*.*****:* *******..** .*:* ***. ::. .**   **:.* *:*


LOC_Os08g43120.1      AEMTLSS----LKKLICT-------GHNLWRRKQPYNKLYY----------------QAI 204
MLOC_54794.6          AEMTVRETIDFSSKMLGTNNEFEMLGEAIRRKKGVINKVDQDLDSFIKAT-TFGEGSNLT 141
ABCG40                GEMTVRETFAYAARFQGVGSRYDMLTELARREKEANIKPDPDIDIFMKAMSTAGEKTNVM 300
                      .***: .      ::  .        .   *.*    *                    :
                                    Signature motif Walker box B
LOC_Os08g43120.1      KIECMQILGLSECADTLVGDEMRRGISGGQKKRATIGEMLVGLAFCFFMDDISTGLDSST 264
MLOC_54794.6          TNYIIKILGLSECADTMVGDELRRGISGGQKKRATIGEMLVGLAFCFFMDDISTGLDSST 201
ABCG40                TDYILKILGLEVCADTMVGDDMLRGISGGQKKRVTTGEMLVGPSKALFMDEISTGLDSST 360
                      .   ::****. ****:***:: **********.* ****** :*.:***:*********


LOC_Os08g43120.1      TFEIMKFLQQMAHLMDLTMVISLLQPPPETLELFDDIILLCEGQIVYHGPRENATDFFET 324
MLOC_54794.6          TYEIVKFLQQMAHLMDLTMVISLLQPPPETLELFDDIILLCEGQIIYHGPRENATNFFEI 261
ABCG40                TYQIVNSLRNYVHIFNGTALISLLQPAPETFNLFDDIILIAEGEIIYEGPRDHVVEFFET 420
                      *::*:: *:: .*::: * :******.***::*******:.**:*:*.***::..:***


LOC_Os08g43120.1      MGFKCPSRKNVADFLQEVTSKMDQKQYWIGNANKYQYHSIEKFAESFRTSYLPRLVENDH 384
MLOC_54794.6          MGFKCPSRKNVADFLQEVTSKMDQKQYWIGDENKYQYRPIEKFAESFRSSYLPRFAKDDL 321
ABCG40                MGFKCPPRKGVADFLQEVTSKKDQMQYWARRDEPYRFIRVREFAEAFQSFHVGRRIGDEL 480
                      ******.**.*********** ** ***    : *::  :.:***:*:: :: *   ::


LOC_Os08g43120.1      FESTNAGKSK-EVKTSTSRMISSWNIFKACFSREVLLLKRNSPVHIFKTIQITVLALVIS 443
MLOC_54794.6          CRTNNTGKSK-EIITSATRRISRWNIFKACFSREVLLLKRNSPLHIFKTVQITVMALVIS 380
ABCG40                ALPFDKTKSHPAALTTKKYGVGIKELVKTSFSREYLLMKRNSFVYYFKFGQLLVMAFLTM 540
                        . :  **:    *: .  :.  ::.*:.**** **:**** :: **  *: *:*::


LOC_Os08g43120.1      TLFLRTNMRHDTVLDANKYMGALFMAVVIVNFNGMTEIAMTIKRLPIFYKQREILALPGW 503
MLOC_54794.6          TIFLRTNMNHKTVLDANKYMGSLFMAVVIVNFNGMTEIAMTIKRLPTFYKQRELLALPGW 440
ABCG40                TLFFRTEMQKKTEVDGSLYTGALFFILMMLMFNGMSELSMTIAKLPVFYKQRDLLFYPAW 600
                      *:*:**:*.:.* :*.. * *:**: :::: ****:*::*** :** *****::*  *.*


LOC_Os08g43120.1      ALLSSVFLLSLPISFVETGLWTGLTYYVIGYAPSFVRFIQHFVVLFAMHQMSMSLYRFLA 563
MLOC_54794.6          ALLSSVFLISLPMSLLETGLWTSLTYYVIGYAPSFLRFIQQFLVLFAMHQMSMGLYRFLA 500
ABCG40                VYSLPPWLLKIPISFMEAALTTFITYYVIGFDPNVGRLFKQYILLVLMNQMASALFKMVA 660
                      .   . :*:.:*:*::*:.* * :******: *.. *::::::::*. *:**: .*::::*


LOC_Os08g43120.1      AIGRTQVMANMLGTAALIAIYILGGFVISKDNLQPWLRWGYWTSPFTYAQNAVALNEFLD 623
MLOC_54794.6          AIGRTQVMANMLGTAALIAIYIFGGFVISKDNLQPWLQWGYWTSPFTYAQNAVALNEFLD 560
ABCG40                ALGRNMIVANTFGAFAMLVFFALGGVVLSRDDIKKWWIWGYWISPIMYGQNAILANEFFG 720
```

```
                    *:**. ::** :*: *::.:: :**.*:*:*::: *  **** **: *.***:  ***:.

LOC_Os08g43120.1    DRWATEFHFANANTVGETILKVRGLLTEWHWYWICVSILFGFSLVFNILSIFALQYMRSP 683
MLOC_54794.6        ERWAIEFHYANAKTVGEAILKIRGLLMEWHWYWICVGILFGFSLVFNILTIFALEFMKSP 620
ABCG40              HSWSRAVENS-SETLGVTFLKSRGFLPHAYWYWIGTGALLGFVVLFNFGFTLALTFLNSL 779
                    . *:  .. : ::*:* ::** **:* . :**** .. *:** ::**:   :** ::.*

LOC_Os08g43120.1    HKHQVNINATKVKVDYNSQ------IVGNGTASTDQVILPFQPLSLVFDHINYFVDMPKE 737
MLOC_54794.6        HKHQVNIDSTKTKTECKKQ------KVGTGNASTGQVVLPFQPLSLVFDHINYFVDMPKE 674
ABCG40              GKPQAVIAEEPASDETELQSARSEGVVEAGANKKRGMVLPFEPHSITFDNVVYSVDMPQE 839
                     * *. *    .. : : *       *  *  ..  ::***:* *:.**:: * ****:*
                                                Walker Box A
LOC_Os08g43120.1    MTKYGVTDKKLQLLQDVSGAFRPGVLTALMGITGAGKTTLLDVLAGRKTGGYIEGTVKIA 797
MLOC_54794.6        MMKYGVTEKKLQLLQDVSGVFRPGVLTALMGVTGAGKTTLLDVLAGRKTGGYIEGTIRIA 734
ABCG40              MIEQGTQEDRLVLLKGVNGAFRPGVLTALMGVSGAGKTTLMDVLAGRKTGGYIDGNITIS 899
                    * : *. :.:* **:.*.*.***********::*******:************:*.: *:

LOC_Os08g43120.1    GYPKKQETFSRISGYCEQSDIHSPNLTVYESLQFSAWLRLPSNVKSHQRNMFIDEVMDLV 857
MLOC_54794.6        GYPKKQDTFSRISGYCEQSDIHSPNLTVHESLQFSAWLRLPSNVNSRQRDMFIDEVMDLV 794
ABCG40              GYPKNQQTFARISGYCEQTDIHSPHVTVYESLVYSAWLRLPKEVDKNKRKIFIEEVMELV 959
                    ****:*:**:********:*****::**:*** :*******.:*...:*.:**:***:**
                              Walker Box B
LOC_Os08g43120.1    ELTGLKNAMVGLAGATGLSAEQRKRLTIAVELVASPSIIFMDEPTTGLDARAAAIVMRTV 917
MLOC_54794.6        ELTGLKNAMVGIAGATGLSAEQRKRLTIAVELVASPSIIFMDEPTTGLDARAAAIVMRTV 854
ABCG40              ELTPLRQALVGLPGESGLSTEQRKRLTIAVELVANPSIIFMDEPTSGLDARAAAIVMRTV 1019
                    *** *::*:**:.* :***:**************.**********:**************

LOC_Os08g43120.1    RKTVDTGRTVVCTIHQPSIEIFESFDELLLMKRGGQLIYSGSLGPLSSNMIKYFEAIPGV 977
MLOC_54794.6        RKTVDTGRTVVCTIHQPSIGIFESFDELLLMKRGGQIIYSGPLGPLSSNMIKYFEAIPGV 914
ABCG40              RNTVDTGRTVVCTIHQPSIDIFEAFDELFLLKRGGEEIYVGPLGHESTHLINYFESIQGI 1079
                    *:***************** ***:****:*:****: ** *.**  *:::*:***:* *:

LOC_Os08g43120.1    PRIKEGQNPAAWMLDISSRTAEYEIGVDYAEIYQRSSLYWENRQLIDDLGKPEPNTEDLH 1037
MLOC_54794.6        PRIKEGQNPAAWVLDISSHITEYEIGVDYAEIYRSSSLYRENMLLIDELGQPAPNTEDLH 974
ABCG40              NKITEGYNPATWMLEVSTTSQEAALGVDFAQVYKNSELYKRNKELIKELSQPAPGSKDLY 1139
                     :*.** ***:*:*::*:   *  :***:*::*: *.** .*  **.:*.:* *.::**:

LOC_Os08g43120.1    FPPKYWQDFRAQCMACLWKQNCAYWKNSEHNVVRFINTFAVSIMFGIVFWKIGSTIKDEQ 1097
MLOC_54794.6        FPPGYWQNFRAQCMACLWKQRCAYWKNSEHNVVRFLNTFAVSIMFGIVFWKIGSIIKGQQ 1034
ABCG40              FPTQYSQSFLTQCMASLWKQHWSYWRNPPYTAVRFLFTIGIALMFGTMFWDLGGKTKTRQ 1199
                    **. * *.* :****.****. :**:*. :..***: *:.:::*** :**.:*.  * .*

LOC_Os08g43120.1    DVFNILGVVYGSALFLGFMNCSILQPVVGMERVVLYREKAAGMYSTMAYAIAQVAVELPY 1157
MLOC_54794.6        DVFNILGVVYGSALFLGFMNCSILQPVVTMERVVLYREKAAGMYCTLAYAIAQMAIELPY 1094
ABCG40              DLSNAMGSMYTAVLFLGLQNAASVQPVVNVERTVFYREQAAGMYSAMPYAFAQVFIEIPY 1259
                    *: * :* :* :.****: *.: :**** :**.*:***:*****.::.**:**: :*:**

LOC_Os08g43120.1    MFVQVFIFSAIVYPMIGFQMTATKFFWFALYMVLSFLYYTLYGMMTVALTPNIEIAAGLS 1217
MLOC_54794.6        MLVQVLIFASIVYPMIGFEMTAVKFFWFVLYMVLSFMYYTLYGMMTVALTPNLEIAAGLS 1154
ABCG40              VLVQAIVYGLIVYAMIGFEWTAVKFFWYLFFMYGSFLTFTFYGMMAVAMTPNHHIASVVS 1319
                    ::**.:::. ***.****: **.****: ::*  **: :*:****:**:*** .**: :*

LOC_Os08g43120.1    FLIFIFWNVFSGFIIGRQMIPVWWRWVYWANPAAWTVYGLMFSQLGDRTELIQVPGQPEQ 1277
MLOC_54794.6        FLIFIFWNVFSGFIIGRELIPIWWRWVYWANPAAWTVYGLMFSQLGDRTEIIHVPGQPDQ 1214
ABCG40              SAFYGIWNLFSGFLIPRPSMPVWWEWYYWLCPVAWTLYGLIASQFGDITEPMADSN---M 1376
                      :: :**:****:* *  :*:**.* **  *.***:***: **:** ** :  ..

LOC_Os08g43120.1    TVKEFLEGYLGLQDRYFNLVTSLHVAIIALFTFLFFLSIKHLKFQRR 1324
MLOC_54794.6        TVREFLEGYLGLENRYFNLVTCLHLAIIALFAFLFFIFIKHLKFQRR 1261
ABCG40              SVKQFIREFYGYREGFLGVVAAMNVIFPLLFAVIFAIGIKSFNFQKR 1423
                    :*::*:. : * .: ::.:*:.::: :  **:.:* : ** ::**:*
```

Figure 4.5 Example of alignment of putative amino acid sequences of putative barley ABC transporters with rice and Arabidopsis

Alignment of barley ABC transporter MLOC_54794.6 with rice ABC transporter encoded by LOC_Os08g43120.1 and Arabidopsis ABC transporter encoded by ABCG40.

Asterisk indicate the conservation of residue in all the three species and dashes indicate the gaps or absence of sequence data; alignments were made using CLUSTALW2 (http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=clustalw2; last accessed May 2014) and then bioedit was used to align the conserved functional motifs; bold: indicate ABC signature motif; box: indicate Walker box A and B; rice sequences were derived from the MSU rice genomic annotation project (http://rice.plantbiology.msu.edu/; last accessed May 2014) and Arabidopsis sequences were derived from TAIR (http://www.arabidopsis.org/index.jsp); **E.** motif sequence and the position were derived from Brule and Smart (2002).

Table 4.5 Physical mapping of ABA signalling related candidate genes.

Barley	Bowmen Contig	Morex Contig	Chromosome
PP2CA			
AK251854	Bowman_contig_900084	Morex_contig_37338	3HL
AK374059	Bowman_contig_65864 Bowman_contig_2516	Morex_contig_50301 Morex_contig_45906	3HL
MLOC_61822.2	Bowman_contig_65864	Morex_contig_45906	3HL
MLOC_8131.1	Bowman_contig_125801	Morex_contig_141242	3HL
AK357955	Bowman_contig_11688	Morex_contig_40533	4HL
AK367207	Bowman_contig_877761 Bowman_contig_863384	Morex_contig_61910 Morex_contig_56082	2HS
MLOC_75421.1	Bowman_contig_18587	Morex_contig_67887	7HS
AK362128	Bowman_contig_1981261	Morex_contig_158601	1H
MLOC_44992.1	Bowman_contig_222530	Morex_contig_275926	3HS
MLOC_7586.1	Bowman_contig_1987710	Morex_contig_139582	3HL
AK376382	Bowman_contig_23560	Morex_contig_156974	1H
MLOC_15065.1	Bowman_contig_23560	Morex_contig_156974	1H
AK358849	Bowman_contig_21265	Morex_contig_161087	1H
AK367469	Bowman_contig_1993612	Morex_contig_7812	5H
PP2CD			
AK371581	Bowman_contig_65251	Morex_contig_5695	4HL
MLOC_67251.4	Bowman_contig_126656 Bowman_contig_29423	Morex_contig_52783	4HL
AK364849	Bowman_contig_219091 Bowman_contig_654	Morex_contig_57315 Morex_contig_130636	5HL
MLOC_51800	Bowman_contig_1987951	Morex_contig_37090	5HL
MLOC_71450.3	Bowman_contig_69404	Morex_contig_59847	2HL
MLOC_13716.1	Bowman_contig_857885	Morex_contig_1566336	7HL
MLOC_61942.1	Bowman_contig_857936	Morex_contig_46025	1H
MLOC_16954	Bowman_contig_61582	Morex_contig_1575045	5HS
MLOC_4262.2	Bowman_contig_61582	Morex_contig_134917	5HS
PYR/PYL/RCAR			
MLOC_39291.1	Bowman_contig_863466	Morex_contig_2552923	3HS
MLOC_46394	Bowman_contig_1981872	Morex_contig_289557	7HL
AK376521	Bowman_contig_143	Morex_contig_1649329	4HL
MLOC_71349.1	Bowman_contig_1993341	Morex_contig_59669	4HL
MLOC_3912.1	Bowman_contig_971337 Bowman_contig_66869	Morex_contig_40322 Morex_contig_134370	3HL
MLOC_60739.1	Bowman_contig_291349	Morex_contig_44721	1H
MLOC_65591.1	Bowman_contig_17747 Bowman_contig_13001	Morex_contig_50387	3HL
MLOC_49654.1	Bowman_contig_1109674	Morex_contig_345720	3HS
MLOC_72289.1	-	Morex_contig_61377	7HL
AK363238	Bowman_contig_10577	Morex_contig_65876	1H
SnRK2			
MLOC_22145.2	Bowman_contig_74470	Morex_contig_160302	4HS
AK374298	Bowman_contig_1985329	Morex_contig_160473	5HL
MLOC_3013.1	Bowman_contig_12040	Morex_contig_127028	5HS
AK372880	Bowman_contig_113325	Morex_contig_1561710	2HL
MLOC_69212	Bowman_contig_144141	Morex_contig_5609	2HS
AK251684	Bowman_contig_12821	Morex_contig_135022	1H
ABC Transporter			
MLOC_11283.1	Bowman_contig_142310	Morex_contig_1560750	3HL
MLOC_54794.6	Bowman_contig_845065	Morex_contig_39409	7HS
MLOC_68581.1	Bowman_contig_1981866	Morex_contig_54987	4HL
MLOC_62985.1	Bowman_contig_863423	Morex_contig_47158	4HS

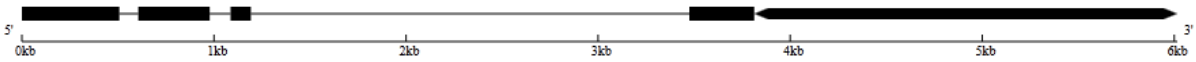
*Data obtained from IPK Gatersleben barley BLAST server (http://webblast.ipk-gatersleben.de/barley/)

4.2.6 Intron-exon structures of barley ABA-signalling related candidate genes

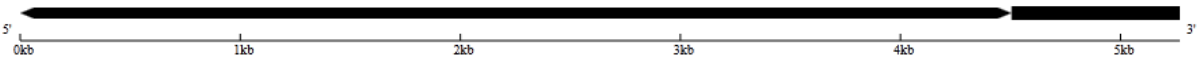
Comparison of the barley gDNA and cDNA ABA-related sequences (both extracted from IBSC) using the gene structure display server resulted in potential intron-exon structures (http://gsds.cbi.pku.edu.cn/; last accessed May 2014). Unlike Arabidopsis PP2CAs, all of which have 3-4 exons (Yoshida *et al*. 2006), HvPP2CAs had one to four exons, as noticed in rice PP2CAs (Xue *et al*. 2008) (Figure 4.6). All HvPP2CAs had the same number of exons as in rice, except AK357955, MLOC_74521 and AK358849, which were shorter (due to missing exon(s), and the latter two were missing some motifs (Table 4.1). HvPP2CDs had 2-4 exons, as found in rice (Xue *et al*. 2008). Six of the HvPP2CDs had the same number of exons as in their rice orthologues, while AK364849, MLOC_16954 and MLOC_4262 were shorter by 1, 2 and 2 exons, respectively. The latter two lacked the motifs at C and N termini, respectively. HvSnRK2s (subfamily II and III) showed 8-9 exons, whereas the Arabidopsis and rice orthologues demonstrated 6-10 and 7-9 exons respectively (Saha *et al*. 2013). Three of the SnRK2s (subfamily II and III) had the same number of exons as their orthologues in Arabidopsis, rice and maize (Huai *et al*. 2008), while MLOC22145.2 and AK372880 were shorter by 1 exon. Barley PYR/PYL/RCAR and ABCG25 sequences exhibited 1-3 and 2 exons respectively, similar to its orthologue in Arabidopsis (Guzman *et al*. 2012; Kuromori *et al*. 2010). ABCG40 was found to have 9-14 exons in barley (Figure 4.6). Analysis of the corresponding genomic sequences indicated that all introns displayed the standard GT/AG splice junctions, except MLOC_67251 (PP2CD) which contained GC/AG in its second introns.

PP2CA

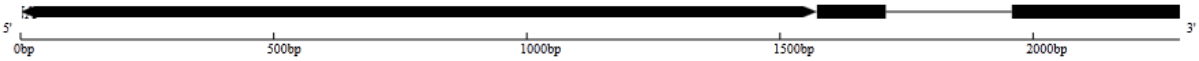
AK251854.1



AK374059

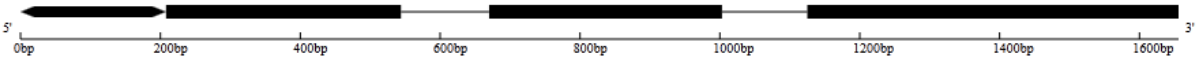


MLOC_61822.2

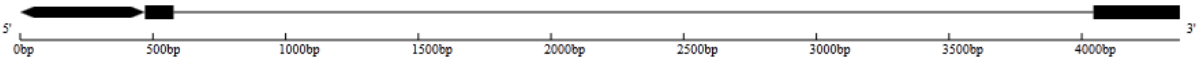


MLOC_8131.1

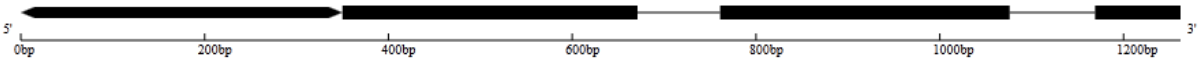
AK357955



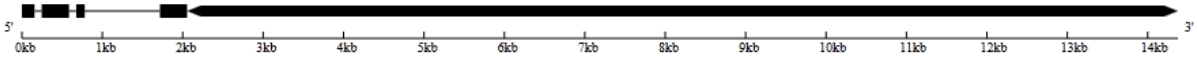
AK367207



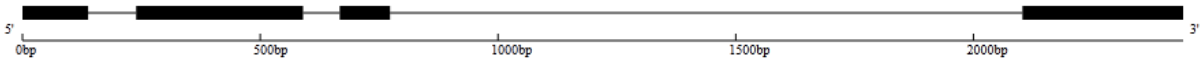
MLOC_75421.1



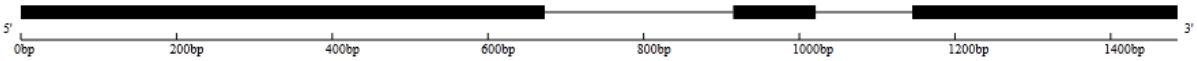
AK362128



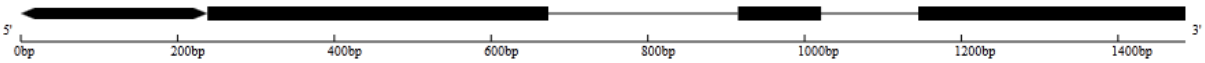
MLOC_7586.1



AK376382

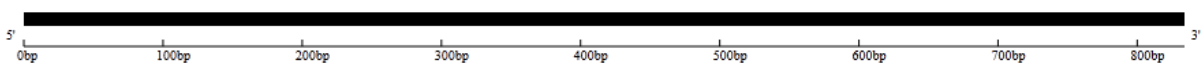


MLOC_15065.1



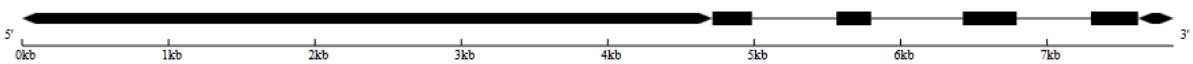
AK358849

AK367469



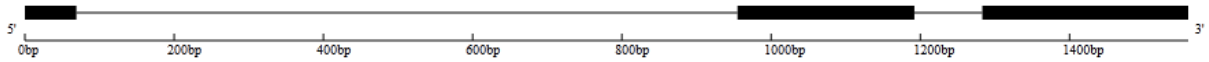
PP2CD

AK371581

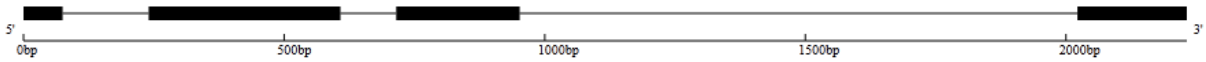


MLOC_67251.4

AK364849

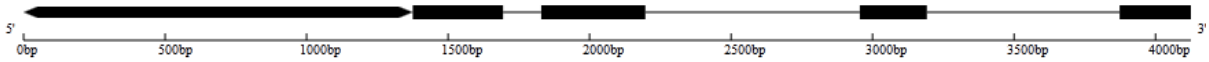


MLOC_51800

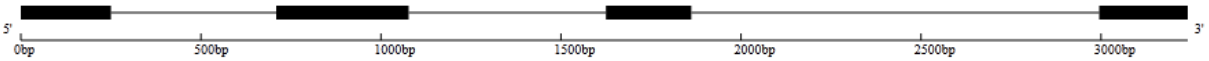


MLOC_71450.3

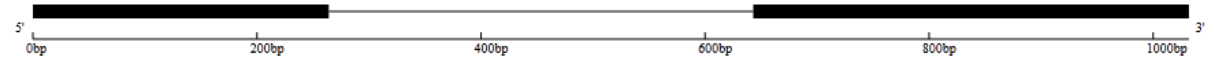
MLOC_13716.1



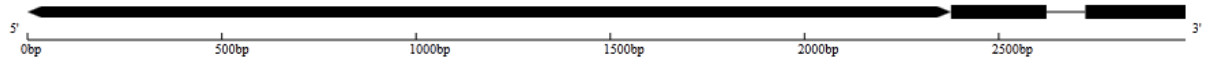
MLOC_61942.1



MLOC_16954

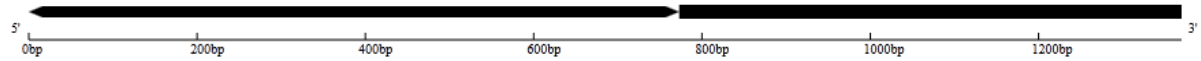


MLOC_4262.2

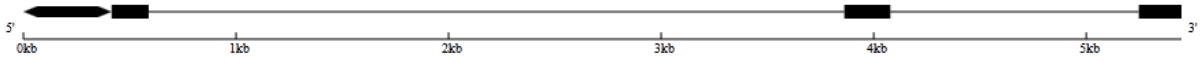


PYR/PYL/RCAR

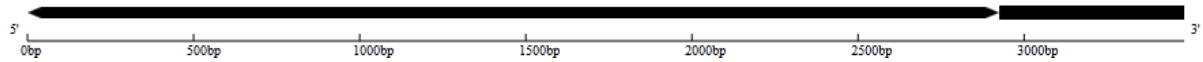
MLOC_39291.1



MLOC_46394

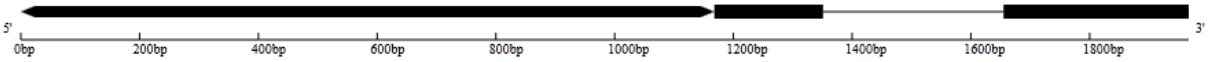


AK376521

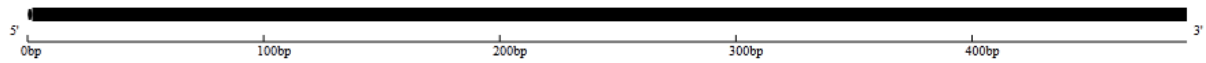


MLOC_71349.1

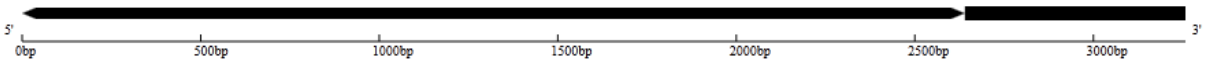
MLOC_3912.1



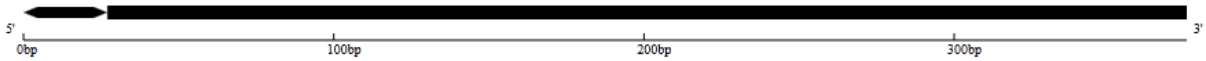
MLOC_60739.1



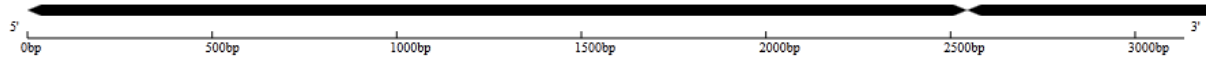
MLOC_65591.1



MLOC_49654.1



MLOC_72289.1



AK363238

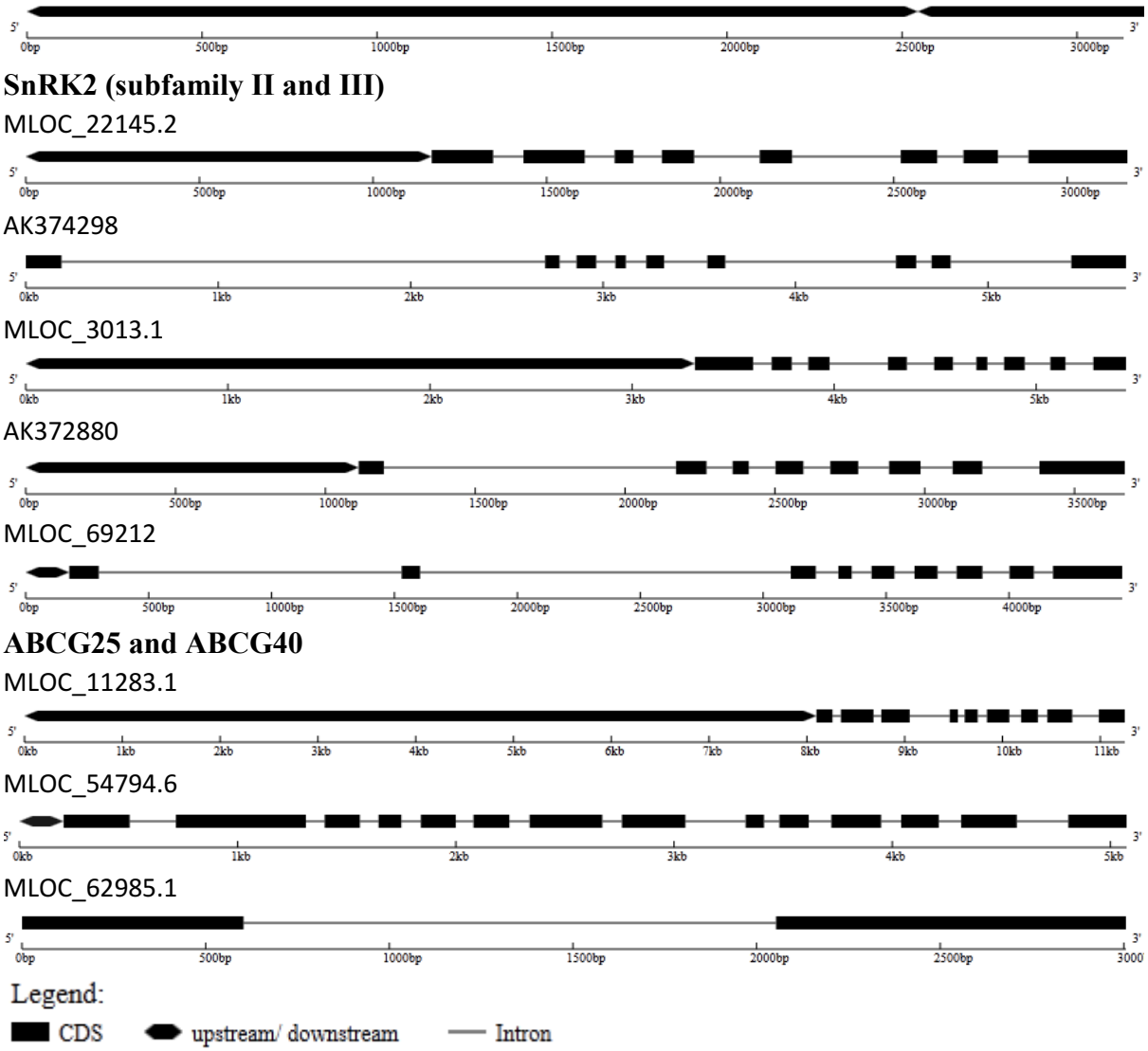


Figure 4.6 Schematic of the intron/exon junctions of the barley ABA signalling related candidate gene.

The putative exons are shown as rectangles, the putative introns are shown as black lines and un-translated regions are shown as double sided wedge. These structures were generated using gene structure display server (http://gsds.cbi.pku.edu.cn/; last accessed May 2014).

4.2.7 Comparison of HvPP2CA and HvPP2CD cDNAs and putative proteins to those in rice and Arabidopsis

The ABA-related barley CDS (IBSC) BLASTed against MSU rice genomic annotation project database and TAIR to extract the closest orthologues in rice and Arabidopsis respectively. The result of reciprocal BLAST search (Method C; as explained above) was also considered to determine closest rice orthologues of these barley sequences. The barley CDS and protein sequences were then aligned with their respective rice and Arabidopsis orthologues using ClustalW2. Alignment of the CDS and putative proteins of the 13 HvPP2CAs using ClustalW2 showed higher % identity to rice orthologues (69

to 84 % DNA identity) compared to Arabidopsis (57 to 81 % DNA identity) (Table 4.1), as would be expected due to them being members of the grass family. The rice AHG3 (LOC_Os04g08560, Pizzio *et al*. 2013) and Arabidopsis ABI1 (AT5G57050; Miyazono *et al*. 2009) and ABI3 (AT1G72770; Melcher *et al*. 2009) have been functionally proven to bind to the cytoplasmic ABA receptors PYL4, PYL1 and PYL2, respectively. Of these, AT1G72770 (1536 bp, 511 aa) had two orthologues in barley (MLOC_7586.1, AK362128). MLOC_7586.1 (1149 CDS; 383aa) and AK362128 (1152bp CDS; 384 aa) showed 47% and 48% protein identity to it, respectively (Table 4.1; Figure 4.2a). The difference in the length of AT1G72770 (Arabidopsis) to its rice (LOC_Os_05g46040.1; 1164 bp; 387 aa) and barley orthologues is due to presence of 236 amino acids on N terminal of first characteristic motif (motif IV) in AT1G72770, whereas the rice and barley orthologues have only 111 and 105 amino acids respectively (Table 2a) and these observations are in line with previous reports for Arabidopsis and rice sequences (Xue *et al*. 2008). The amino acid residues involved in the binding of PP2CA to PYR/PYL/RCAR (Section 4.2.2) were found to be strictly conserved in these PP2CAs of all three plants. However, barley orthologue for LOC_Os04g08560 and AT5G57050 were not found. MLOC_61822.2 (468bp CDS; 156 aa), a partial PP2CA showed the highest % DNA (84 %) and protein (85 %) identity to its rice orthologue (LOC_Os01g46760; 1212 bp, 403 aa) as compared to other PP2CAs. This sequence could be of importance with respect to ABA response and need to be Sanger sequenced to extract its complete sequence.

PP2CD MLOC_13716.1 (1176 bp, 392aa) showed highest % protein identity with it rice (LOC_Os06g50380; 1179 bp, 392 aa; 91 %) and Arabidopsis (AT4G38520; 1203bp, 400 aa; 70 %) ortholgues and also highest % DNA identity with its rice orthologue (89 %) compared to other barley sequences (Table 4.1; Figure 4.2b). None of the PP2CDs have the amino acid residues required for binding to PYR/PYL/RCAR, which indicates that PP2CDs play an important role in plant's stress response but through a pathway other than regulated by ABA. Xue *et al*. (2008) have previously reported the consensus sequences for all 11 motifs in Arabidopsis; however the consensus for rice is unclear. Therefore, the rice and barley putative amino acid alignments of PP2CAs and PP2CDs were utilised to develop a consensus sequence for each of the 11 motifs using weblogo (http://weblogo.threeplusone.com/create.cgi;

Schneider and Stephens 1990; last accessed May 2014). The results indicated that Motif 1 was highly conserved, with 90.4% identity between barley and rice consensus, while Motif 7 was least conserved (19 %) (Table 4.6).

Table 4.6 Characteristics of conserved motifs in the putative rice and barley PP2CA, PP2CD, SnRK2 and ABC transporter proteins

Motif	Motif length (aa)	No of motifs aligned	% identity	Plant	Sequence of motif
PP2Cs					
Motif 1	21	20	90.4	Rice	DEFLILASDGLWDVM(SA)NQEAV
		19		Barley	DEFLILASDGLWDV(LV)SN(EQ)EAV
Motif 2	21	20	33.7	Rice	WRVKGVLAMS(RA)AF(GR)DR(YH)LKQW
		20		Barley	WRVKGVLA(MV)(ST)R(ASG)(IF)GDR(YH)LK(PQR)W
Motif 3	15	20	66.6	Rice	LY(VA)AN(AV)GDSRAVLCR
		20		Barley	LYV(AS)N(AV)GDSRAVL(CS)R
Motif 4	15	20	46.6	Rice	-FVAVYDGHGGPEAA
		19		Barley	(FL)(FV)(GA)V(FY)DGHGGPE(ATV)A-
Motif 5	57	10	70.1	Rice	DIV—(NS)PR-G-A-RLVKAAL-EAA(KR)EMRY-DLK-I(ED)RGVRRHFHDDITV(IV)VVFLD
		7		Barley	-(LR)V--(NS)PR-GIARRLV(RK)(AS)A—EA(AT)(KR)KREMRY-D---IERGVRRHFHDDITVVVL(YF)LD
Motif 6	21	10	42.8	Rice	DHK(PVS)NR(ES)DERERIE(AV)AGGY(AV)I
		11		Barley	DH(KR)P(NDS)R(ES)DERERIEAAGG-V(IV)
Motif 7	21	9	19.0	Rice	LT(EK)L(AV)ALSR(GS)SDNI(TS)-(ILV)VVQL
		11		Barley	L(TSV)(EK)(LM)A-SRGS-DNI(TS)V(ILV)VVQL
Motif 8	11	20	27.2	Rice	GST(AC)VV(AV)(VL)I--
		20		Barley	GST(AC)(VC)VA(VL)I--
Motif 9	41	10	68.3	Rice	---DGLLW—(DE)LG-H(AV)G(ED)(FY)SMAV(AV)QAN—LEDQ-QV—(GS)P
		7		Barley	---DGLLW—(DE)L(GR)-HAAG(ED)FS(MF)AVVQAN(NE)-LEDQ-QVE-GP
Motif 10*	41	NA	NA	Rice	NA
		NA		Barley	NA
Motif 11	29	10	65.5	Rice	AE(QR)L(ST)-(DE)HN(AV)—EEVR(QR)EL-(AE)LHPDDSQI
		7		Barley	AE(QR)LS-(DE)HN(AV)—EEVR-E(LV)-SQHPDDPQI
SnRK2					
C-terminal Domain I	16	5		Rice	Q-MDEIMQILAEATI-
		5		Barley	QS(ML)DEIMQI—EATI-
C-terminal Domain II	28	3		Rice	Insufficient number of sequences
		3		Barley	Insufficient number of sequences

Motif	Motif length (aa)	No of motifs aligned	% identity	Plant	Sequence of motif
ABC transporters					
Walker box A	9			Rice	GATG(AG)G-TT
				Barley	GATGA(AG)-TT
Walker box B				Rice	NA
				Barley	Insufficient number of sequences
ABC signature				Rice	NA
				Barley	Insufficient number of sequences

The Weblogo Browser (http://weblogo.threeplusone.com/create.cgi) was utilised to generate sequence logos. * Motif 10 is absent in PP2CAs and PP2CDs (Xue *et al*. 2008).

4.2.8 Comparison of HvPYR/PYL/RCAR cDNAs and putative proteins to those in rice and Arabidopsis

Ten barley PYR/PYL/RCARs obtained from barley IBSC CDS datas*et al*igned with rice and Arabidopsis orthologues as obtained above using ClustalW2. Similar to PP2CAs, PYR/PYL/RCARs showed higher % DNA identity compared to % protein identity for both rice and Arabidopsis. The rice PYL4 (LOC_Os02g13330; Pizzio *et al*. 2013) and Arabidopsis PYL1 (AT5G46790; Miyazono *et al*. 2009) and PYL2 (AT2G26040; Melcher *et al*. 2009) have been functionally proven to bind to the PP2CA AHG3, ABI1 and ABI3, respectively. AT2G26040 (PYL2; 573 bp, 190 aa) had been identified as the closest orthologue of three barley PYR/PYL/RCARs (MLOC_39291.1, MLOC_49654.1 and MLOC_72289.1) by the BLAST tool in NCBI. Of these barley sequences, MLOC_39291.1 was later classified as PYL4 instead of PYL2 on the basis of its closest rice orthologue (LOC_Os02g13330). MLOC_39291.1 (597 bp CDS) showed LOC_Os02g13330 (624 bp) and AT2G26040 (573 bp) as the closest orthologues, with 86 % and 64 % DNA identity, respectively. The 199 aa putative protein of MLOC_39291.1 was 80 % and 61 % identical to the 207 aa rice and 190 aa Arabidopsis proteins. MLOC_49654.1 (348 bp CDS; 116 aa) showed 84 % and 67 % DNA identity and 66 % and 50 % protein sequence identity with LOC_Os06g36670 (624 bp, 207 aa) and AT2G26040 (573 bp, 190 aa), whereas MLOC_72289.1 (588 bp CDS; 196 aa) showed 90 % and 61 % DNA identity and 85 % and 57 % protein identity with LOC_Os06g36670 and AT2G26040 respectively (Table 4.2; Figure 4.3). The amino acid residues involved in the binding of PYR/PYL/RCAR to ABA (Section 4.2.3) were found to be strictly conserved in these PYR/PYL/RCAs of all three plants. However, barley orthologue for AT5G46790 was not found. Partial PYR/PYL/RCAR, MLOC_60739 (489 bp, 163 aa) demonstrated highest % DNA (92 % with rice and 68 % with Arabidopsis) and % protein (91 % with rice and 65 % with Arabidopsis) identity to its rice (LOC_Os05g39580; 654 bp, 217 aa) and Arabidopsis (AT5G05440; 612 bp, 203 aa) orthologues compared to other PYR/PYL/RCARs. This sequence could be of importance with respect to ABA response and need to be Sanger sequenced to extract its complete sequence.

4.2.9 Comparison of HvSnRK2 (subfamily II and III) cDNAs and putative proteins to those in rice and Arabidopsis

The closest rice and Arabidopsis orthologue for the five barley SnRK2s (subfamily II and III) obtained from were extracted by BLAST in MSU rice genomic annotation project and TAIR database respectively. Unlike PP2CAs and PYR/PYL/RCARs, SnRK2s (subfamily II and III) showed higher % protein identity compared to % DNA identity, except MLOC_3013 which showed higher % DNA identity (87 %) with rice (LOC_Os12g39630) compared to % protein identity (86 %). Rice SnRK2.6 (LOC_Os03g41460) has been functionally proven to bind to the PP2CA HAB1 (Zhou *et al*. 2004). LOC_Os03g41460 was identified as the closest orthologue of barley MLOC_22145.2, which also demonstrated highest % identity to rice and Arabidopsis compared to other barley SnRK2s (Table 4.3; Figure 4.4). MLOC_22145.2 (1083 bp, 361 aa) showed 90 % and 73 % DNA and 96 % and 80% protein identity to its rice (LOC_Os03g41460; 1089 bp, 362 aa) and Arabidopsis (AT5G66880; 1086 bp, 361 aa) orthologues compared. The rice and barley putative amino acid alignments of SnRK2s (subfamily II and III) were utilised to develop a consensus sequence for both the domains using weblogo (http://weblogo.threeplusone.com/create.cgi; Schneider and Stephens 1990; last accessed May 2014). The results indicated that SnRK2 conserved box was conserved, with 68.7% identity between barley and rice consensus, but was unable to generate consensus for ABA responsive box because of not having enough sequences (Table 4.6).

4.2.10 Comparison of HvABCG25 and HvABCG40 cDNAs and putative proteins to those in rice and Arabidopsis

The closest rice and Arabidopsis orthologues for four ABA-transport related ABC transporter (ABCG25 and ABCG40), were extracted by BLAST in MSU rice genomic annotation project and TAIR database respectively. Unlike other AB-related gene families studied, ABC transporters exhibited similar % DNA and protein identity with respective rice orthologues, except MLOC_68581.1 which has shown higher % DNA identity. ABCG25 and ABCG40 in Arabidopsis have been functionally proven to be involved in transport of ABA (Kang *et al*. 2010; Kuromori *et al*. 2010). The barley orthologue of ABCG25 was identified as MLOC_62985.1 (1923bp CDS; 641aa), which showed 83% and 58% DNA identity and 83% and 52% protein identity with

LOC_Os11g07600 (1839 bp, 612 aa) and AT1G71960 (ABCG25; 1989 bp, 662 aa) respectively (Table 4.4; Figure 4.8). MLOC_11283 (1935 bp, 645 aa) MLOC_54794.6 (3783 bp 1261 aa) and MLOC_68581.1 (4044bp CDS; 1348 aa) were identified as orthologue of ABCG40. MLOC_11283 (1935 bp, 645 aa) demonstrated highest % DNA (67 %) and % protein (55 %) identity to its Arabidopsis orthologue (AT1G15520; ABCG40). MLOC_54794.6 (3783 bp 1261 aa) demonstrated highest % DNA (84 %) and % protein (84 %) identity to its rice orthologue (LOC_Os08g43120) compared to other barley sequences. MLOC_68581.1 (4044bp CDS; 1348 aa), the longest ABA-related ABC transporter showed 83% and 62% DNA identity 79% and 48% protein sequence identity with LOC_Os08g43120 (3975 bp, 1324 aa) and AT1G15520 (ABCG40; 4272 bp, 1423 aa), which is also the least % protein identical to its rice and Arabidopsis orthologues as compared to other sequences of the same subfamily studied. Bairoch (1992) has previously reported the consensus sequences for Walker box A, Walker box B and ABC signature motif in Arabidopsis. Therefore, the rice and barley putative amino acid alignments of ABCG25 and ABCG40 were utilised to develop a consensus sequence for each of these boxes using weblogo (http://weblogo.threeplusone.com/create.cgi; Schneider and Stephens 1990; last accessed May 2014). The Walker box A was found to be highly conserved (78%) and a consensus could be generated for it by weblogo, whereas Walker box B and ABC signature motif were not generated due to having insufficient number of sequences (Table 4.6).

Identification of ABA-signalling related candidate transcripts and analysis of their differential expression under stress conditions by mRNA-seq

As detailed in Chapter 3, the barley leaf mRNA-seq I transcriptome was developed using total RNA pooled from two control plants and total RNA pooled from two salt-stressed plants of barley cv. Hindmarsh. The mRNA-seq I reads from these libraries were initially aligned against the NCBI Unigene database (as the International Barley Genome Sequencing Consortium (IBSC) barley genome was not available at the time) and the unaligned reads were assembled *de novo* to identify potentially novel transcripts (NTCs). This collectively led to mRNA-seq I dataset. The full sequence data set is available at Sequence Read Archive in Genbank (accession number SRA062960; http://www.ncbi.nlm.nih.gov/sra?term=SRA062960). The NCBI Unigenes and NTCs

of the key ABA-signalling gene families were later BLASTed against the barley genome when it became available, using the IBSC BLAST tool (http://webblast.ipk-gatersleben.de/barley/viroBlast.php). The IBSC accession numbers have been added to respective results tables (see below) for ease of comparisons.

The second mRNA-seq experiment analysed leaf RNAs from control, salt-stressed (150 mM NaCl for 12 h), drought stressed (20 % PEG for 12 h), and exogenous ABA stressed plants (100 µM ABA for 12 h) (Section 2.7). As detailed in Chapter 3 (Section 3.2.8), the mRNA-seq of these libraries led to 27.3, 25.3, 28.0 and 28.8 million 76 nt reads, respectively. These reads were aligned against the IBSC CDS database using burrows wheeler aligner (BWA) to develop the mRNA-seq II dataset with the IBSC accession numbers (Kamboj *et al*. unpublished).

Both of these datasets were analysed further in this chapter for identifying the leaf transcripts of the key ABA signalling pathway genes, as well as comparing their differential expression under the three abiotic stress conditions. The differential expression was assessed by fold changes (FC), calculated as normalised salt reads/normalised control reads, and FC of ≥ +1.5 or ≤ -1.5 was considered notable as per convention (Smyth 2004; Zhou *et al*. 2010). It should be noted that for mRNA-seq I, the FCs are derived from the original BLAST of the 76 nt reads against NCBI Unigenes and may not be directly comparable to the FCs in mRNA-seq II, where the reads were directly BLASTed against IBSC CDS database, as there is some difference in the length of some of the reported sequences between IBSC and NCBI. The transcript identifications and differential expressions for the four main gene families are detailed below.

4.2.11 Candidate barley PP2CA and PP2CD transcripts and their differential expression under stress conditions

The mRNA-seq I dataset were searched by terms such as ‘protein phosphatase’ using ‘Ubuntu’, leading to 17 NCBI Unigenes and 13 NTCs for PP2Cs. Another search by the RBH method developed in this project (detailed in Section 2.22.3) identified 4 NCBI Unigenes for PP2CA, and three Unigenes for PP2CDs and 1 NTC for PP2CD (Table 4.7). The collective list filtered down to 11 Unigenes and 5 NTCs, after

applying the criteria of (i) signal strength (≥10) to the NCBI Unigenes, and (ii) signal strength (≥10) as well as rice/barley BLAST ratio (≥2) to NTCs (as detailed in Section 3.2.3) (Table 4.7). The translation products of the 16 sequences were aligned with the relevant rice and Arabidopsis orthologs using ClustalW2 (data not shown) and the PP2C motifs searched for manually, as above. The results showed that three Unigenes belonged to subfamily PP2CA, three Unigenes and 1 NTC belonged to PP2CD (Table 4.8), and the rest had insufficient information for grouping. To confirm NTC14333, primers were designed to anneal at start and end of contig. These primers were used as pair to amplify the contig from cDNA constructed of extracted RNA from barley leaf. Due to the possibility of multiple sequence types, PCR products were cloned into pGEM®-T Easy (Promega, Australia), using *E. coli* JM109 as the host and then sequenced. The sequencing results (data not shown; summary in Appendix XVI) confirmed it to encode a PP2CD. These Unigenes and NTC were later BLASTed against the barley genome, and the IBSC accession numbers (the NTC being MLOC_16954) have been added to Tables 4.7 and 4.8 for convenience of further comparisons. The FC for the three PP2CAs ranged from +1.22 to +4.60, with only Hv.19158 (FC 4.60) and Hv.18841 (FC 3.71) showing significant up-regulation (≥ +1.5) (Table 4.7). Of the 4 PP2CDs, only Hv.20807 (FC -1.75) and NTC14333 (FC 2.13) showed notable changes. As explained above, these FCs were derived from BLAST of the mRNA-seq I reads against NCBI Unigenes and could vary somewhat from the FCs of mRNA-seq II transcripts, which were analysed by direct BLAST to IBSC CDS database.

Table 4.7 PP2C transcripts identified from the mRNA-seq I dataset.

Barley Unigene, or NTC	Signal strength[a]	Rice/ barley BLAST ratio[b]	Fold change[c]	Method of identifi-cation[d]	Barley Accession No (IBSC)[f]	Motif Analysis (Pass/ Fail/NA?)[g]
PP2CA						
Hv.11499	43.12	NA	1.22	A, C	AK357955	Pass
Hv.19158	100.94	NA	4.60	A, C	AK251854	Pass
Hv.18841	111.97	NA	3.71	A, C	AK370347/MLOC_8131	Pass
Hv.9794	3.91	NA	4.37	C	AK374059	NA
Hv.13561	5.49	NA	11.83	A	AK367469	NA
Hv.31236	0.08	NA	-1.11	A	AK376382/ AK364195/ MLOC_15065	NA
Hv.33774	4.83	NA	-1.57	A	AK358849	NA
Hv.33382	0.62	NA	6.27	A	AK367207	NA
Hv.34576	0.30	NA	-	A	AK369918	NA
PP2CD						
Hv.20807	139.94	NA	-1.75	A,C	AK356867/AK364849	Pass
Hv.11437	65.12	NA	1.41	A,C	AK371581	Pass
Hv.9421	61.68	NA	1.41	A, C	AK362082/AK364555/ MLOC_61942	Pass
NTC14333	11.82	3.4	2.13	A, C	MLOC_16954	Pass
NTC37557	0.41	4.6	-1.67	A	MLOC_71450	NA
Unclassifiable						
Hv.3594	27.20	NA	1.43	A	AK363794	Fail
Hv.8804	9.69	NA	-3.01	A	AK249474	NA
Hv.14888	72.53	NA	1.02	A	AK361878	Fail
Hv.2811	135.95	NA	1.36	A	AK375108	Fail
Hv.13120	63.00	NA	1.38	A	MLOC_63900	Fail
Hv.17813	28.34	NA	-1.21	A	MLOC_79155	No
NTC14842	5.97	2.7	-2.62	A	MLOC_36023/MLOC_31 577	NA
NTC21810	2.98	2.6	-1.42	A	MLOC_76799	NA
NTC25618	27.3	12.3	1.31	A	MLOC_21732	Fail
NTC28545	5.97	5.3	-2.05	A	MLOC_13677	NA
NTC33701	4.27	3.5	-2.59	A	-	NA
NTC35511	2.31	2.4	-2.35	A	-	NA
NTC42633	13.66	15.2	1.14	A	MLOC_60206	Fail
NTC43575	10.61	29.3	-1.63	A	MLOC_77804	Fail
NTC43622	4.77	2.5	1.32	A	MLOC_58433	NA
NTC5067	1.68	3.1	-1.93	A	MLOC_55971	NA
NTC7152	39.80	2.6	1.58	A	MLOC_37399	Fail

[a]Signal strength indicates the expression of a sequence in barley cv. Hindmarsh from mRNA-seq I dataset, the formula used for calculating signal strength being log (normalised control + salt reads, 2); [b]rice/barley BLAST ratio is comparison of % identity in BLAST result of transcript against the already identified rice sequence from MSU rice genome annotation project database and the barley sequence in NCBI unigene database; [c]fold change is the relative expression of a transcript under salt stress as compared to control conditions, calculated as normalised salt reads/normalised control reads; [d]methods used to retrieve the transcripts from mRNA-seq I data A: keyword searches of the barley mRNA-seq dataset B: rice locus search (not applied in this case) C: RBH search (see Methods); [e]the selection criteria were signal strength (≥10) for known barley Unigene sequences, and signal strength (≥10) and rice/barley BLAST ratio (≥2) for NTCs, sequences meeting the criteria are shown in grey; [f]IBSC accession number derived by BLASTing in the IBSC BLAST tool; [g] putative PP2CA or PP2CD was confirmed on basis of motif analysis; NA: Not applicable (motif anlaysis not performed as failed selection criteria).

Table 4.8 Identification and motif analysis of PP2CAs and PP2CDs from mRNA-seq I.

Barley NTC/Unigene number from Genbank	Length of barley sequence transcript(bp)/ CDS(bp)/protein(aa)[a]	Mw(kDa)/ pI[b]	Signal strength[c]	Fold change[d]	Most closely related rice sequence (% identity between rice and barley CDS/protein sequences)[e]	Most closely related Arabidopsis sequence (% identity between Arabidopsis and barley CDS/protein sequences)[f]	Predicted PP2C subtype[g]	Motifs[h]											Barley Accession No (IBSC)[f]
								I	II	III	IV	V	VI	VII	VIII	IX	X	XI	
Hv.11499	1572/1206/401	43.14/5.71	43.124	1.22	LOC_Os03g16170 (83/72)	AT1G07430 (59/49)	PP2CA	√	√	√	√	x	√	√	√	x	x	x	AK357955
Hv.19158	2116/1476/491	51.07/4.84	100.94	4.60	LOC_Os01g40094 (83/83)	AT1G17550 (57/44)	PP2CA	√	√	√	√	x	√	√	√	x	x	x	AK251854
Hv.18841	1704/1182/393	43.09/6.07	111.97	3.71	LOC_Os01g62760 (77/68)	AT1G07430 (58/46)	PP2CA	√	√	√	√	x	√	√	√	x	x	x	AK370347/MLOC_8131
Hv.9421	1947/1182/393	43.38/8.26	61.68	1.41	LOC_Os10g39780 (87/88)	AT4G38520 (61/68)	PP2CD	√	√	√	√	√	x	x	√	√	x	√	AK362082/AK364555/MLOC_61942
Hv.20807	1467/1143/380	41.47/9.33	139.94	-1.75	LOC_Os03g55320 (84/82)	AT4G33920 (60/58)	PP2CD	√	√	√	√	√	x	x	√	√	x	√	AK356867/AK364849
Hv.11437	2159/1203/400	43.91/9.30	65.129	1.41	LOC_Os03g04430 (84/86)	AT4G38520 (62/66)	PP2CD	√	√	√	√	√	x	x	√	√	x	√	AK371581
NTC14333	868/868/289	31.92/9.34	11.821	2.13	LOC_Os12g39120 (82/83)	AT4G33920 (68/56)	PP2CD	√	√	√	√	√	x	x	√	√	x	√	MLOC_16954

[a]Length of barley transcripts extracted from mRNA-seq data and CDS in base pairs (bp) and putative protein in amino acids (aa); [b] molecular weight (Mw) and isoelectric point (pI) of the putative proteins were predicted in the ExPASy Proteomics Server (http://web.expasy.org/compute_pi/); [c]Signal strength indicates the expression of a sequence in barley cv. Hindmarsh from mRNA-seq I dataset, the formula used for calculating signal strength being log (normalised control + salt reads, 2); [d]fold change is the relative expression of a transcript under salt stress as compared to control conditions, calculated as normalised salt reads/normalised control reads [e] most closely related rice sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW2; [f] most closely related Arabidopsis sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW2; [g]putative PP2CA or PP2CD was confirmed on basis of motif analysis; [h] sequence and position of domains were extracted from Xue *et al*. (2008); √: motif present; X: no motif found; NA: not applicable.

Of the 13 putative PP2CAs identified from barley IBSC genome that had the characteristic motifs of PP2Cs (Table 4.1), only six were found in the mRNA-seq II dataset (Table 4.9). These filtered down to three each (all shared) in salt, drought and exogenous ABA stress, and two more under ABA, which met the signal strength criterion (≥10) (Table 4.9). Their FCs varied from strong responses (FC +2.74 to +16.38) in ABA-treated samples to moderate responses under drought (FC +1.19 to +3.03) and salt (FC -1.00 to +2.26) stresses (Table 4.9). One PP2CA (MLOC_8131) showed significant up-regulation (FC +2.26 to +7.26) in all three stresses, as also noted in mRNA-seq I. All five PP2CAs with signal strength >10 under ABA also showed significant up-regulation. All nine PP2CDs from barley IBSC genome that showed the characteristic domains (Table 4.2) were identified in mRNA-seq II dataset. The same four sequences with signal strength ≥10 were noted under each stress, but none showed significant expression modulation (Table 4.9). Only two PP2CAs (AK251854 and MLOC_8131) met the signal strength criterion (≥10) under salt stress condition in both mRNA-seq I and mRNA-seq II datasets. MLOC_8131 showed significant up-regulation in both datasets, whereas AK251854 showed significant up-regulation (FC 4.60) in mRNA-seq I. As explained above none of the PP2CDs, which met the signal strength criterion showed significant differential expression in mRNA-seq II dataset, whereas two PP2CDs (MLOC_61942 and MLOC_16954) demonstrated significant fold change in mRNA-seq I.

Table 4.9 Differential expression of ABA-signalling related genes identified from mRNA-seq I and II datasets.

Gene	Drought		ABA		Salt (mRNA-seq II)		Salt (mRNA-seq I)	
	Fold Change[a]	Signal Strength[b]	Fold Change[a]	Signal Strength[b]	Fold Change[a]	Signal Strength[b]	Fold Change[a]	Signal Strength[b]
PP2CA								
AK251854	1.19	14.7	3.24	36.14	-1.00	17.19	4.60	100.94
MLOC_8131.1	3.03	11.51	7.26	30.06	2.26	12.05	3.71	111.97
AK357955	2.74	8.39	2.74	10.61	-1.75	4.37	1.22	43.12
AK362128	1.94	18.97	3.96	39.44	1.20	18.23	-	
MLOC_7586.1	1.68	2.52	16.38	20.86	3.68	5.71	-	
AK367469	2.30	0.19	20.15	1.85	-	-	11.83	5.49
AK358849							-1.57	4.83
AK374059							4.37	3.91
AK367207							6.27	0.62
AK369918							-	0.30
MLOC_15065/ AK376382							-1.11	0.08
PP2CD								
AK371581	1.40	21.89	1.33	27.54	-1.13	22.09	1.41	65.12
MLOC_67251.4	-5.16	2.86	-1.06	5.92	-1.13	6.04	-	-
AK364849	-1.40	15.04	-1.08	21.33	-1.19	20.85	-1.75	139.94
MLOC_51800	2.00	8.21	-1.25	6.33	-1.91	5.45	-	-
MLOC_71450.3	1.69	3.35	-4.97	1.86	-5.04	1.96	-1.67	4.6
MLOC_13716.1	1.00	42.95	-1.44	45.94	-1.23	50.37	-	-
MLOC_61942.1	1.07	17.84	1.08	22.65	1.21	24.56	1.41	61.68
MLOC_16954	-1.29	2.47	-1.17	3.23	-1.38	3.13	2.13	11.82
MLOC_4262.2	1.08	3.21	1.51	4.93	1.44	4.87	-	-
PYR/PYL/RCAR								
AK376521	-1.84	48.08	-2.51	55.34	-1.07	78.16	-2.08	13.68
MLOC_71349.1	-3.12	33.95	-13.16	34.44	-1.64	54.23	-6.13	13.68
MLOC_3912.1	-1.23	9.91	1.26	15.58	1.55	18.27		
MLOC_49654.1	-1.42	1.81	-1.22	2.50	1.03	2.84		
AK363238	-1.33	23.70	1.02	35.12	1.21	38.76	-1.85	24.51
MLOC_39291.1							-1.37	1.18
MLOC_65591.1							-	0.17
SnRK2								
MLOC_69212.1(III)	1.06	58.96	-1.61	59.25	-1.14	69.85	1.67	233.1
AK251684.1(II)	2.49	0.59	11.79	2.93	9.63	3.38	14.78	29.73
AK372880(III)	1.11	0.81	17.46	8.72	7.76	4.32	1.15	21.16
MLOC_22145.2(II)	-1.06	15.96	-1.47	17.92	-1.39	18.46	-1.55	22.47
AK374298(II)	1.30	28.22	1.00	31.23	1.01	31.85	1.24	81.59
MLOC_3013.(II)1	-1.66	67.74	1.18	116.90	1.26	124.78	-1.57	185.11
ABC Transporter								
MLOC_68581.1(G40)	4.76	19.61	1.61	12.50	-1.44	8.17	-	-
MLOC_62985.1(G25)	-1.38	2.75	1.36	4.88	-1.59	3.37	2.07	20.0

[a]Fold change is the relative expression of a transcript in a stress condition (150mM NaCl, 20% PEG or 100 μM ABA) as compared to controlled condition from mRNA-seq. The formula used is normalised stress/normalised control reads; [b] Signal strength indicates the expression of a particular sequence in barley variety Hindmarsh from mRNA-seq dataset. The formula used for calculating signal strength is log (normalised control + stress reads, 2).

4.2.12 Candidate barley PYR/PYL/RCAR transcripts and their differential expression under stress conditions

The barley PYR/PYL/RCARs extraction from mRNA-seq I dataset by keyword searches did not fetch any sequences. The possible reason could be that these receptors were discovered in Arabidopsis only in 2009 (Park *et al*. 2009), so the databases may not have had these annotations. Searches for PYR/PYL/RCAR rice loci (Method B) in mRNA-seq I dataset resulted in three NCBI Unigenes and one NTC, and the RBH identified seven NCBI Unigenes and one NTC. The total of nine PYR/PYL/RCARs filtered to four NCBI Unigenes and one NTC after applying the respective selection criteria (Table 4.10). Alignment of their putative protein sequences with the rice and Arabidopsis orthologs (data not shown) showed that three Unigenes and one NTC had the characteristic residues that directly bind to and co-ordinate ABA through a water molecule (Table 4.11). NTC2524, primers were designed to anneal at start and end of contig. These primers were used as pair to amplify the contig from cDNA constructed of extracted RNA from barley leaf. Due to the possibility of multiple sequence types, PCR products were cloned into pGEM®-T Easy (Promega, Australia), using *E. coli* JM109 as the host and then sequenced (data not shown; summary in Appendix XVI) and showed only 50% identity with the de-novo assembled contig, suggesting some assembly errors. It was later found to correspond to IBSC MLOC_45380.1. The IBSC accession numbers have been added to Tables 4.9 and 4.11. All showed significant down-regulation except NTC2524 (Table 4.9).

Table 4.10 PYR/PYL/RCAR, SnRK2 (subfamily II and III) and ABA-related ABC transporter transcripts identified from the mRNA-seq I dataset.

Barley NTC/Unigene	Signal strength[a]	Rice/barley BLAST ratio[b]	Fold change[c]	Method of selection[d]	Barley Accession No (IBSC)[f]	Motif Analysis (Pass/ Fail/NA?)
PYR/PYL/RCAR						
Hv.15651	13.68	NA	-6.13	B, C	AK361631/MLOC_71349.1	Pass
Hv.9994	61.50	NA	-2.08	B, C	AK376521	Pass
Hv.12427	24.51	NA	-1.85	B, C	AK363238	Pass
NTC2524	45.25	4.69	1.01	C	MLOC_45380.1	Pass
NTC30477	3.67	5.68	-1.45	B	MLOC_74341.1	NA
Hv.1590	1.18	NA	-1.37	C	MLOC_39291.1	NA
Hv.37400	0.17	NA	-	C	MLOC_65591.1	NA
Hv.12683	57.17	NA	1.39	C	-	Fail
Hv.25663	8.19	NA	-1.63	C	AK363238/ MLOC_14439.1	NA
SnRK2 Subtype II and III						
Hv.1875 (III)	185.11	NA	-1.57	B, C	AK355634/MLOC_3013.1	Pass
Hv.2096 (III)	81.59	NA	1.24	B, C	AK374298	Pass
Hv.5014 (III)	22.47	NA	-1.55	B, C	AK366496/AK363699/MLOC_22145.2	Pass
Hv.6253 (II)	233.1	NA	1.67	B, C	AK366400/AK374249/MLOC_69212.1	Pass
Hv.17120 (II)	29.73	NA	14.78	B, C	AK251684.1	Pass
Hv.21359 (II)	21.16	NA	1.15	B	AK372880	Pass
Hv.11385 (III)	110.31	NA	-1.40	C	MLOC_62759.3	Fail
Hv.20961 (III)	357.47	NA	1.64	B	AK36293	Fail
Hv.34453 (III)	10.22	NA	-2.92	C	AK370199	Fail
Hv.2008 (III)	129.31	NA	-2.08	C	AK362030	Fail
Hv.7083 (III)	9.97	NA	1.06	C	AK365481	NA
Hv.22672 (II)	26.23	NA	-1.44	B, C	MLOC_11948.5/MLOC_17371.1	Fail
NTC21223 (II)	38.67	2.01	-1.30	C	MLOC_63787.3	Fail
ABC transporter						
Hv.7752 (G25)	20.00	NA	2.07	B, C	AK357645/MLOC_62985.1	Pass
NTC4445(G40)	15.13	17.8	2.27	B, C	MLOC_68581.1	Pass
NTC9836(G40)	24.38	7.9	2.53	B, C	MLOC_68581.1	Pass
Hv.15015(G40)	11.25	NA	1.89	B	MLOC_68581.1	Fail
NTC2926(G40)	2.06	2.6	2.01	B	MLOC_68581.1	NA
NTC3321(G40)	0.23	2.2	1.79	B	MLOC_68581.1	NA
NTC8909(G40)	0.48	2.2	1.25	B	MLOC_68581.1	NA

[a]Signal strength indicates the expression of a sequence in barley cv. Hindmarsh from mRNA-seq I dataset, the formula used for calculating signal strength being log (normalised control + salt reads, 2); [b]rice/barley BLAST ratio is comparison of % identity in BLAST result of transcript against the already identified rice sequence from MSU rice genome annotation project database and the barley sequence in NCBI unigene database; [c]fold change is the relative expression of a transcript under salt stress as compared to control conditions, calculated as normalised salt reads/normalised control reads; [d]methods used to retrieve the transcripts from mRNA-seq I data A: keyword searches of the barley mRNA-seq dataset B: rice locus search (not applied in this case) C: RBH search (see Methods); [e]the selection criteria were signal strength (≥10) for known barley Unigene sequences, and signal strength (≥10) and rice/barley BLAST ratio (≥2) for NTCs, sequences meeting the criteria are shown in grey; [f]IBSC accession number derived by BLASTing in the IBSC BLAST tool; [g] putative sequences were confirmed on basis of motif analysis; NA: Not applicable (motif anlaysis not performed as failed selection criteria).

Table 4.11 Identification and motif analysis of PYR/PYL/RCAR from mRNA-seq I.

Barley NTC/Unigene number from Genbank	Length of barley sequence transcript(bp)/CDS(bp)/protein(aa)[a]	Mw(kDa)/pI[b]	Signal strength[c]	Fold change[d]	Most closely related rice sequence (% identity between rice and barley CDS/protein sequences)[e]	Most closely related Arabidopsis sequence (% identity between Arabidopsis and barley CDS/protein sequences)[f]	Residues predicted to be directly involved in ABA binding													Residues that coordinate ABA through water					Barley Accession No (IBSC)[f]
							1	2	3	4	5	6	7	8	9	10	11	12	13	14	1	2	3	4	5	
Hv.15651	849/522/173	18.37/6.05	13.68	-6.13	LOC_Os03g18600 (70/91)	AT5G05440 (53/60)	K	F	V	L	A	V	H	L	Y	F	I	I	V	N	P	E	R	S	E	AK361631/ MLOC_71349.1
Hv.9994	1135/750/249	26.00/8.84	61.50	-2.08	LOC_Os03g18600 (83/69)	AT5G05440 (64/54)	K	F	V	L	A	V	H	L	Y	F	V	I	V	N	P	E	R	S	E	AK376521
Hv.12427	1119/609/202	22.67/4.96	24.51	-1.85	LOC_Os10g42280 (88/77)	AT4G17870 (57/61)	K	F	V	L	A	I	H	L	Y	F	A	V	V	N	P	E	R	S	E	AK363238
NTC2524	589/588/195	20.96/10.07	45.25	1.00	LOC_Os05g12260 (66/67)	AT5G53160 (57/45)	K	I	V	L	A	F	H	L	Y	-	-	-	-	-	P	E	R	S	-	MLOC_3912.1

[a]Length of barley transcripts extracted from mRNA-seq data and CDS in base pairs (bp) and putative protein in amino acids (aa); [b] molecular weight (Mw) and isoelectric point (pI) of the putative proteins were predicted in the ExPASy Proteomics Server (http://web.expasy.org/compute_pi/); [c]Signal strength indicates the expression of a sequence in barley cv. Hindmarsh from mRNA-seq I dataset, the formula used for calculating signal strength being log (normalised control + salt reads, 2); [d]fold change is the relative expression of a transcript under salt stress as compared to control conditions, calculated as normalised salt reads/normalised control reads [e] most closely related rice sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW2; [f] most closely related Arabidopsis sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW2; [g] putative PYR/PYL/RCAR was confirmed on positioning of various residues involved in binding of ABA; f: Positioning of various residues was extracted from Yin *et al*. (2009); -: residue not found

Searches of the barley IBSC genome database (Section 4.2.3) had led to 10 PYR/PYL/RCARs which had the characteristic amino acids and were probably functional (Table 4.2). Of these, only five were found expressed in mRNA-seq II dataset (Table 4.9), which filtered down to three in drought and four each in salt and exogenous ABA stress with signal strength ≥10. Their FCs varied from -3.12 to -1.33 in drought, -13.16 to 1.02 in ABA and -1.64 to 1.55 in salt stress (Table 4.9). AK376521 and MLOC_71349.1 showed significant down-regulation in drought and exogenous ABA stress, whereas MLOC_71349.1 exhibited down-regulation and MLOC_3912.1 exhibited up-regulation in salt stress. Only three PYR/PYL/RCARs (AK376521, MLOC_71349.1 and AK363238) met the signal strength criterion (≥10) under salt stress condition in both mRNA-seq I and mRNA-seq II datasets. MLOC_71349.1 showed significant down-regulation in both datasets, whereas AK376521 and AK363238 showed significant down-regulation in mRNA-seq I only.

4.2.13 Candidate barley SnRK2 subfamily II and III transcripts and their differential expression under stress conditions

Keyword searches of mRNA-seq I resulted in >100 SnRK2s (NCBI Unigenes and NTCs together) and were not analysed further due to a high number of probable non-specific transcripts. Method B (search by rice loci) led to four NCBI Unigenes each for SnRK2 subfamilies II and III. Eleven SnRK2 (subfamily II and III) transcripts (10 Unigenes, 1 NTC) were identified by RBH, of which five were identified for the first time and six were amongst the 8 identified by Method B. Thus a total of 13 putative barley SnRK2s (subfamily II and III; 12 ‘known’ Unigene and 1 ‘novel’) were identified The sequences filtered down to 11 NCBI Unigenes and one NTC after applying the selection criteria (Table 4.10). Alignments of their putative proteins with the rice and Arabidopsis orthologs (data not shown) showed that Hv.1875, Hv.2096 and Hv.5014 belonged to subfamily III as they possessed the ABA-responsive box, whereas Hv.6253, Hv.17120 and Hv.21359 lacked this box and likely belonged to subfamily II (Belin *et al*. 2006) (Table 4.12) while the rest could not be confirmed as SnRK2s. Similar to rice, barley also has three members of SnRK2 subfamily III, whereas only two members have been identified in Arabidopsis (Table 4.12). The IBSC accession numbers have been added to Tables 4.9 and 4.12. Of the three members of subfamily II, only two (Hv.1875 and Hv.5014) showed significant down regulation, whereas two

of three members of subfamily II (Hv.6253 and Hv.17120) showed significant up-regulation (Table 4.9).

Table 4.12 Identification and motif analysis of SnRK2 (subfamily II and III)s from mRNA-seq I

Genbank Unigene/ NTC	Barley transcript (bp)/ CDS(bp)/ protein(aa) [a]	Mw (kDa)/ pI[b]	Signal strength[c]	Fold change[d]	Most closely related rice sequence (% identity between rice and barley CDS/protein sequences)[e]	Most closely related Arabidopsis sequence (% identity between Arabidopsis and barley CDS/protein)[f]	Predicted SnRK2 subtype[g]	Serine (Ser175 in Arabidopsis SnRK2.6) phosphorylated for activation of kinase[h]	ABA responsive box	SnRK2 conserved box	Barley Accession No (IBSC)[f]
Hv.1875	1322/1056/351	39.44/ 4.92	185.11	-1.57	LOC_Os12g39630 (87/85)	AT4G33950 (72/77)	III	√	√	√	AK355634/ MLOC_3013.1
Hv.2096	1591/1089/362	41.19/ 4.86	81.59	1.24	LOC_Os03g55600 (91/94)	AT3G50500 (73/77)	III	√	√	√	AK374298
Hv.5014	2243/1089/362	42.52/ 4.26	22.47	-1.55	LOC_Os03g41460 (90/81)	AT5G66880 (67/68)	III	√	√	√	AK366496/ AK363699/ MLOC_22145.2
Hv.6253	1447/1023/341	38.59/ 5.45	233.1	1.67	LOC_Os07g42940 (87/91)	AT1G78290 (62/72)	II	√	NA	√	AK366400/ AK374249/ MLOC_69212.1
Hv.17120	1408/1020/339	39.42/ 5.85	29.73	14.78	LOC_Os10g41490 (86/88)	AT4G40010(65/61)	II	√	NA	√	AK251684.1
Hv.21359	1329/1029/342	38.82/ 5.75	21.16	1.15	LOC_Os03g27280 (87/91)	AT4G40010(70/68)	II	√	NA	√	AK372880

[a]Length of barley transcripts extracted from mRNA-seq data and CDS in base pairs (bp) and putative protein in amino acids (aa); [b] molecular weight (Mw) and isoelectric point (pI) of the putative proteins were predicted in the ExPASy Proteomics Server (http://web.expasy.org/compute_pi/); [c]Signal strength indicates the expression of a sequence in barley cv. Hindmarsh from mRNA-seq I dataset, the formula used for calculating signal strength being log (normalised control + salt reads, 2); [d]fold change is the relative expression of a transcript under salt stress as compared to control conditions, calculated as normalised salt reads/normalised control reads [e] most closely related rice sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW2; [f] most closely related Arabidopsis sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW2; [g] putative SnRK2 subtype II or III was confirmed on positioning of ABA responsive box and SnRK2 conserved box; [h] serine residue involved in the activation of SnRK2s; ABA responsive box, only found in strongly ABA responsive kinases (C terminal Domain II). SnRK2 conserved box: conserved in all kinases (C terminal Domain I) (Belin *et al*. 2006); √: domain/box present; X: no domain/box found; NA: not applicable

Seven SnRK2s altogether were identified in the barley IBSC CDS database, of which only six had the characteristic features (Table 4.3). Five of these were identified in mRNA-seq II, but filtered down to four (@ signal strength ≥10), common to all stresses (Table 4.9). Only MLOC_3013.1 (subfamily III) demonstrated significant down-regulation under drought (Table 4.9). Under exogenous ABA stress, only MLOC_69212 (subfamily II) showed significant down-regulation (FC -1.61), whereas none of the sequence showed significant differential expression in salt stressed plants (Table 4.9). Only four SnRK2s (subfamily II and III) met the signal strength criterion (≥10) under salt stress condition in both mRNA-seq I and mRNA-seq II datasets. As explained above none of these sequences demonstrated significant differential expression in salt stress condition of mRNA-seq II, whereas two sequences of each subfamily showed significant differential expression in mRNA-seq I (salt stress condition). AK372880 showed signal strength of ≥10 only in mRNA-seq I, but non-significant fold change (1.15) in mRNA-seq II.

4.2.14 Candidate barley ABCG25 and ABCG40 ABC transporter transcripts and their differential expression under stress conditions

Keyword searches of mRNA-seq I resulted in over 100 NCBI Unigenes and NTCs altogether for putative ABA-related ABC transporters, which was too great a number to manually analyse for conserved features. Method B (searches by rice loci) led to 2 Unigenes and 5 NTCs orthologous to rice ABCG25 and ABCG40 together. Three ABA-related ABC transporters were identified by RBH. The sequences filtered down to four after application of above explained selection criteria (Table 4.10). Their translated proteins were aligned with the relevant rice and Arabidopsis orthologs (data not shown) and the characteristic domains searched. The Hv.15015 encoded a partial ABA transporter without any characteristic motifs; hence it was excluded. The other three exhibited the characteristic residues for the ABA-related ABC transporter ABCG25 and ABCG40 (Table 4.13). The two NTCs were 98% identical and both likely encoded G40 However, the amplified, cloned and Sanger-sequenced NTC4445 was 100% identical to the assembled contig, while the Sanger sequence of NTC9836 was 93% identical to its contig and its BLAST led to another type of ABC transporter not known to be involved in ABA transport (data not shown; summary in Appendix XVI). There is only one reported sequence each for the ABCG25 and ABCG40 in rice

and Arabidopsis, hence it appears that NTC9836 and NTC4445 may be the same sequence, assembled differently in the *de-novo* assembly of the mRNA-seq I reads. The IBSC accession numbers of these sequences have been added to Table 4.9 and 4.13. All demonstrated significant up-regulation under salt stress (Table 4.9).

Table 4.13 Identification and motif analysis of ABA-transport related ABC transporters from mRNA-seq I.

Barley NTC/Unigene number from Genbank	Barley transcript(bp)/ CDS(bp)/ protein(aa)	Mw (kDa)/pI[b]	Signal strength[c]	Fold change[d]	Most closely related rice sequence (% identity between rice and barley CDS/protein sequences)[e]	Most closely related Arabidopsis sequence (% identity between Arabidopsis and barley CDS/protein sequences)[f]	Predicted ABC trans-porter subtype[g]	Walker Box A	Walker Box B	ABC signature motif	Barley Accession No (IBSC)
Hv.7752	2479/1845/614	65.56/9.00	20.00	2.07	LOC_Os11g07600 (83/83)	AT1G71960 (57/52)	G25	√	X	X	MLOC_68581.1
NTC4445	1599/1602/533	60.69/8.75	15.13	2.27	LOC_Os08g43120 (71/73)	AT1G15520 (72/48)	G40	√	√	√	MLOC_68581.1
NTC9836[a]	1790/1790/596	67.19/6.50	24.38	2.53	LOC_Os08g43120 (68/83)	AT1G15520 (72/54)	G40	√	√	X	MLOC_68581.1

Length of barley transcripts extracted from mRNA-seq data and CDS in base pairs (bp) and putative protein in amino acids (aa); [b] molecular weight (Mw) and isoelectric point (pI) of the putative proteins were predicted in the ExPASy Proteomics Server (http://web.expasy.org/compute_pi/); [c]Signal strength indicates the expression of a sequence in barley cv. Hindmarsh from mRNA-seq I dataset, the formula used for calculating signal strength being log (normalised control + salt reads, 2); [d]fold change is the relative expression of a transcript under salt stress as compared to control conditions, calculated as normalised salt reads/normalised control reads [e] most closely related rice sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW2; [f] most closely related Arabidopsis sequence on the basis of positioning of conserved motifs and percent identity was calculated using ClustalW2; [g] putative ABC transporter was confirmed on positioning of ABA responsive box and SnRK2 conserved box; sequence and positioning of motifs were extracted from Brule and Smart (2002); √: box present; X: no box found; NA: not applicable.

Six ABA-related ABC transporters (ABCG25 and ABCG40) were found in the barley IBSC CDS database, of which only four had these characteristic features (Table 4.4). Of these, only two were identified in mRNA-seq II dataset. Only one (MLOC_68581.1) had signal strength ≥ 10 in drought and ABA stress and also demonstrated significant up-regulation, and none showed signal strength ≥ 10 under salt stress (Table 4.9). None of the ABCG25 and ABCG40 demonstrated signal strength of ≥ 10 under salt stress condition in mRNA-seq II datasets, whereas both the members of ABC transporter family not only met signal strength criterion but also showed significant up-regulation in mRNA-seq I.

4.2.15 Analysis of reported mRNA-seq data for differential expression studies of ABA-signalling related genes of rice under salt stress

mRNA-seq technology is a powerful tool for gene expression profiling and has been used successfully in model plant Arabidopsis (Weber *et al*. 2007) and several cereal species including soyabean (Fan *et al*. 2013), wheat (Gillies *et al*. 2012), rice (Mizuno *et al*. 2010), maize (Li *et al*. 2010a) and sorghum (Dugas *et al*. 2011). There are few studies published utilizing NGS to study abiotic stress responses in plants, which include studies on salt stress in rice, soyabean (Mizuno *et al*. 2010; Fan *et al*. 2013), cold stress in perennial ryegrass and meadow fescue (Tamura and Yonemaru 2010) and drought stress in sorghum, white mustard and soyabean (Dugas *et al*. 2011; Dong *et al*. 2012).

It was not possible to compare the data in this study with all other available transcriptomes as it required extraction and characteristic motif analysis of ABA signalling related CDS sequences for all these plants. So the data in this study was only compared to the sequenced transcriptome of rice (Mizuno *et al*. 2010), as the rice ABA signalling related CDS sequences have already been identified and analysed for their characteristic motifs (Table 4.1, 4.2, 4.3 and 4.4). According to Mizuno *et al* (2010), 7 day old rice (*Oryza sativa* L. 'Nipponbare') plants were stressed for 1 hour by providing 150 mM NaCl solution. These stressed plants along with control plants were used for mRNA-seq generating 36 nucleotide reads. These reads were then analysed against the Rice Annotation Project database (RAP-DB: http://rapdb.dna.affrc.go.jp/).

RAP-DB locus number for ABA-signalling related candidate genes were extracted from Umezawa *et al* (2010). These locus numbers were then searched in the supplemental data table (S2; RPKM of RAP2 annotated genes) from Mizuno *et al* (2010) to extract the differential expression of ABA signalling related genes in rice under salt stress. These sequences were BLASTed against the MSU rice genome annotation project database to extract the MSU rice locus number for ease of comparison, as MSU rice locus number have been used previously in chapter 3 and 4. This led to the identification of eight PP2CAs, ten PYR/PYL/RCARs, six SnRK2s (subfamily II and III) and two ABA transport related ABC transporters were identified from Mizuno *et al* (2010). Six of the eight rice PP2CAs, LOC_Os05g46040, LOC_Os01g46760, LOC_Os09g15670 (Os04g0167900 and AK063334), LOC_Os01g62760 and LOC_Os03g16170 were significantly up-regulated (FC ≥ +1.5) under salt stress (FC of 1.68, 2.43, 3.37, 3.76, 6.16 and 2.76, respectively), while LOC_Os05g51510 showed non-significant fold change (FC 1.34) and LOC_Os05g49730 was down-regulated (FC -1.88) (Table 4.14). Three putative rice PYR/PYL/RCARs, LOC_Os06g36670, LOC_Os02g13330 and LOC_Os02g15640 showed significant differential expression (FC of 1.74, 1.54 and -1.61 respectively), while two LOC_Os06g33490 and LOC_Os06g33640 showed down regulated fold change of infinity which indicates that the no reads were observed for these sequence in salt stressed plants. LOC_Os06g33690 showed fold change of positive infinity which indicates that no reads for this locus were observed in control sample. Rest of the four PYR/PYL/RCARs, LOC_Os03g18600, LOC_Os05g39580, LOC_Os10g42280 and LOC_Os05g12260 showed non-significant fold change (Table 4.14). Of the six SnRK2s (subfamily II and III) only LOC_Os07g42940 has shown significant fold change (FC 1.84), whereas rest of the five identified SnRK2s showed non-significant fold change. Of the two ABA-related ABC transporters only LOC_Os11g07600 (ABCG25) has shown significant fold change (FC 1.54), whereas LOC_Os08g43120 (ABCG40) showed non-significant fold change (FC -1.06) (Table 4.14).

Table 4.14 Fold change of ABA-signalling related sequences in rice and barley.

RAP-DB rice locus	NCBI accession no	Fold change	MSU rice locus	Barley IBSC accession no	Fold change	
					mRNA-seq I	mRNA-seq II
PP2CA						
Os05g0592800	AK067627	1.34	LOC_Os05g51510	AK358849	-1.57	-
Os05g0537400	Os05g0537400	1.68	LOC_Os05g46040	MLOC_7586.1	-	3.68
Os01g0656200	AK068272	2.43	LOC_Os01g46760	AK374059, MLOC_61822.2	4.37	-
Os05g0572700	AK108969	-1.88	LOC_Os05g49730	-		
Os04g0167900	Os04g0167900	3.37	LOC_Os09g15670	AK367469	11.83	-
Os01g0846300	AK065949	6.16	LOC_Os01g62760.1	MLOC_8131.1	3.71	2.26
Os03g0268600	AK069274	2.76	LOC_Os03g16170	AK357955, AK367207	1.22 6.27	-1.75
Os09g0325700	AK063334	3.76	LOC_Os09g15670	AK367469	-	-
SnRK2 (subfamily II and III)						
Os03g0764800	AB125309	-1.09	LOC_Os03g55600	AK374298	1.24	1.01
Os12g0586100	AK069697	1.22	LOC_Os12g39630	MLOC_3013	-1.57	1.26
Os03g0610900	AB125311	1.11	LOC_Os03g41460	MLOC_22145, MLOC_36654	-1.55	-1.39
Os10g0564500	AB125304	-1.11	LOC_Os10g41490	-		
Os07g0622000	AK070965	1.84	LOC_Os07g42940	MLOC_69212	1.67	-1.14
Os03g0390200	AK068899	-1.01	LOC_Os03g27280	AK372880	1.15	7.76
PYR/PYL/RCAR						
Os03g0297600	AK108414	-1.31-	LOC_Os03g18600	AK376521, MLOC_71349.1	-2.08 -6.13	-1.07, -1.64
Os05g0473000	Os05g0473000	-1.08-	LOC_Os05g39580	MLOC_60739.1, MLOC_65591.1	-	-
Os01g0827800	-	-		-		
Os10g0573400	AK061525	-1.02	LOC_Os10g42280	AK363238	-1.85	1.21
Os06g0562200	AK107448	1.74	LOC_Os06g36670	MLOC_49654	-	1.03
Os02g0226801	Os02g0226801	1.54	LOC_Os02g13330	MLOC_39291	-1.37	
Os02g0255300	-			-	-	
Os05g0213500	AK065280	1.07	LOC_Os05g12260	MLOC_3912.1	1.55	
Os02g0255500	AK059303	-1.61	LOC_Os02g15640	-	-	
Os06g0528300	Os06g0528300	infinity	LOC_Os06g33690	MLOC_46394	-	-
Os06g0526466	Os06g0526466	-infinity	LOC_Os06g33490	-		
Os06g0527800	AK108261	-infinity	LOC_Os06g33640	-		
Os06g0528150	-	-		-		
ABC transporter						
Os08g0544400	AK107688	-1.06	LOC_Os08g43120 (ABCG40)	MLOC_11283, MLOC_54794, MLOC_68581	2.27	-1.44
Os11g0177400	AK120669	1.54	LOC_Os11g07600 (ABCG25)	MLOC_62985	2.07	-1.59

The differential expression of rice ABA signalling related sequences, were extracted from Mizzuno *et al*. (2010).

4.3 Discussion

4.3.1 Comparison of various bioinformatics tools for data mining

The results of the various types of bioinformatics search tools for potential ABA related gene families using mRNA-seq data demonstrated that these can yield different results. One of the methods relied on the annotations in NCBI (Method A). In this case search for different gene families seemed satisfactory as all the genes identified here were annotated with an indicative product name such as 'protein phosphatase 2C'. However these annotations varied between the loci, thus the output relied heavily on the keyword used. Many additional, nonspecific hits were also identified and the whole procedure was time consuming. Second method was to identify the closest rice orthologue of barley in the NGS data. In mRNA-seq data every barley sequence had a orthologue rice locus listed against it. The above extracted rice locus for ABA-signalling pathway related genes were searched in the barley mRNA-seq dataset. The barley sequences from the mRNA-seq data, which are orthologue to selected rice sequences, were selected for further analysis. Method B proved really helpful in extracting the genes for gene families which were unannotated and also the ones whose annotation was not specific. The reciprocal BLAST (RBH; Method C) method which can also be called as the automatic version of method B, was based on the BLASTing the known rice sequences in NGS database and then confirming the sequence with reciprocal BLAST. This method was useful in finding the highly specific sequences. The reciprocal BLAST is totally dependent on the rice orthologues. RBH was unable to find any sequence in the mRNA-database which does not have any orthologue in rice. Thus RBH, being an automated reciprocal BLAST pipeline tool enabled the rapid identification of specific gene families of interest in related species, streamlining the collection of homologs prior to downstream molecular evolutionary analysis. RBH is not recommended in situations where sequence homology is very low, for instance, identifying orthologues between plants and animals. RBH is not suited for the discovery of non-protein-coding genes. The accuracy of return BLAST classification is limited by the length of sequences in the target sequence database. Thus, contig assembly of transcript reads less than 50 bp is not recommended, especially for species pairs with are rich in in-paralogs (Ziemann *et al*. 2013b). Therefore no single method was found to be of complete satisfaction, each of these methods was used to achieve a

different purpose: to identify the genes in annotated data (method A, B, C); to identify genes in unannotated data (method B, C).

4.3.2 ABA related genes form large families in barley

Initially, the knowledge of the cellular and molecular basis of ABA response was unclear; although number of factors related to ABA responses had been reported (Hirayama and Shinozaki 2007; Hirayama and Shinozaki 2010). The previous model of ABA signalling was very complicated as each ABA receptor and binding proteins were randomly placed in different cellular locations and their relationships to ABA signalling pathway were complicated (McCourt and Creelman 2008). However the situation totally changed with the discovery of PYR/PYL/RCAR, as new soluble ABA receptor (Ma *et al*. 2009; Park *et al*. 2009) and the identification of PP2CA-SnRK2 as downstream components of PYR/PYL/RCAR (Umezawa *et al*. 2009; Vlad *et al*. 2009). On the other hand cell to cell ABA transport, which was unclear, was clarified by the discovery of plasma membrane bound transporters ABCG25 and ABCG40 (Kang *et al*. 2010; Kuromori *et al*. 2010). The roles of these genes were proven to regulate a broad range of plant traits and especially important for adaption to environmental conditions. However, despite this great potential, there was limited research on ABA related gene families of the major crops, wheat and barley. While information existed on these families on other monocot and dicot crops, a genome wide picture of these families were missing in wheat and barley, probably due to their genome size and polyploidy of wheat.

The present study reported a comprehensive analysis of the ABA related gene families through mRNA-seq in conjugation with other in-silico methods. Further many unpublished sequences deposited in Genbank resulted in multiple accession numbers and incorrect or missing annotations; this study has led to what we believe is full complement, including new annotations of ABA related genes (Table 4.11). In total, 13 PP2CAs (9 full length and 4 partial), 9 PP2CDs (6 full length and 3 partial), 10 PYR/PYL/RCARs (7 full length and 3 partial), 5 SnRK2 (subfamily II and III; all full length) and 4 ABA transport related ABC transporters (ABCG25 and ABCG40; 2 full length and 2 partial) were identified in IBSC CDS database. The partial or complete length of the sequence was assessed by aligning against the respective rice and

Arabidopsis orthologues. There was significant advancement on 6 PP2Cs and 9 PYR/PYL/RCARs estimated in barley by Seiler *et al*. (2014). The number of SnRK2 (subfamily II and III) found were consistent with the data obtained by Seiler *et al*. (2014). These findings were consistent with the high multiplicity of isoforms found in other higher plants, e.g., Arabidopsis, rice. Ten and nine PP2CAs, ten and nine PP2CDs, thirteen and fourteen PYR/PYL/RCARs and six and five SnRK2s (subfamily II and III) were identified in rice and Arabidopsis respectively (Xue *et al*, 2008; Umezawa *et al*. 2010). In addition to providing information to the physiology of barley, such data is almost relevant to wheat, a close relative, for gene search and analyses of the latter, e.g., it's largely uncharacterised ABA related genes. The putative annotation of ABA-related genes was on the basis of their closest orthologs in rice. For example, the putative PYR/PYL/RCAR from barley that shares highest amino acid sequence identity with OsPYR1 is denoted as HvPYR1 (Table 4.15).

Table 4.15 Annotations assigned on the basis of characteristic features

Barley	InterProScan/Barley database[a]	New Annotations on the basis of characteristic motifs[b]
PP2CA		
AK251854	PP2C	PP2CA(HvPP2C06.1)
AK374059	PP2C	PP2CA(HvPP2C08.1)
MLOC_61822.2	PP2C	PP2CA(HvPP2C08.2)
MLOC_8131.1	PP2C	PP2CA(HvPP2C09)
AK357955	PP2C	PP2CA(HvPP2C30.1)
AK367207	PP2C	PP2CA(HvPP2C30.2)
MLOC_75421.1	PP2C	PP2CA(HvPP2C49)
AK362128	PP2C	PP2CA(HvPP2C50.1)
MLOC_7586.1	PP2C	PP2CA(HvPP2C50.2)
AK376382	PP2C	PP2CA(HvPP2C51.1)
MLOC_15065.1	PP2C	PP2CA(HvPP2C51.2)
AK358849	PP2C	PP2CA(HvPP2C53.1)
AK367469	PP2C	PP2CA(HvPP2C68)
PYR		
MLOC_39291.1	Polyketide cyclase/dehydrase	PYR(HvPYL4)
MLOC_46394	Polyketide cyclase/dehydrase	PYR(HvPYL7)
AK376521	Polyketide cyclase/dehydrase	PYR(HvPYR1.1)
MLOC_71349.1	Polyketide cyclase/dehydrase	PYR(HvPYR1.2)
MLOC_3912.1	Polyketide cyclase/dehydrase	PYR(HvPYL5)
MLOC_60739.1	Polyketide cyclase/dehydrase	PYR(HvPYL1.1)
MLOC_65591.1	Polyketide cyclase/dehydrase	PYR(HvPYL1.2)
MLOC_49654.1	No description	PYR(HvPYL3.1)
MLOC_72289.1	Polyketide cyclase/dehydrase	PYR(HvPYL3.2)
AK363238	Polyketide cyclase/dehydrase	PYR(HvPYL2)
SnRK2		
MLOC_22145.2	Protein kinase, catalytic domain	SnRK2 subtype III(HvSnRK2.6)
AK374298	Protein kinase, catalytic domain	SnRK2 subtype III(HvSnRK2.2)
MLOC_3013.1	Protein kinase, catalytic domain	SnRK2 subtype III(HvSnRK2.3)
AK372880	Protein kinase, catalytic domain	SnRK2 subtype II(HvSnRK2.11)
MLOC_69212	Protein kinase, catalytic domain	SnRK2 subtype II(HvSnRK2.7)
ABC transporter		
MLOC_11283.1	ABC transporter-like	ABC transporter G40
MLOC_54794.6	ABC transporter-like	ABC transporter G40
MLOC_68581.1	ABC transporter-like	ABC transporter G40
MLOC_62985.1	ABC transporter-like	ABC transporter G25

[a]Annotation derived from the barley database using InterProScan (http://www.ebi.ac.uk/Tools/pfa/iprscan/); [b]New putative annotations assigned on the basis of characteristic features.

4.3.3 Differential expression of ABA related genes in barley (mRNA-seq II)

Gene expression patterns can offer significant signs for gene function. The present study is amongst the first genome-wide studies, targeting all expressed leaf genes in salt and drought stress conditions. Six PP2CAs showed differential expression in abiotic stress conditions, except AK251854 all other sequences were significantly affected (FC of $\geq +1.5$ or ≤ -1.5; Smyth 2004; Zhou *et al.* 2010) by drought stress. In salt stress condition only 3 of these sequences MLOC_8131.1, AK357955 and MLOC_7586.1 showed significant expression, whereas all the six sequences were significantly affected

by exogenous ABA stress condition (Table 4.9). This suggests that all of the PP2CA genes are involved in the ABA dependent stress regulatory pathway, but some of these genes are specific to type of abiotic stress applied. These results are consistent with previous observations on the overlap of drought, salinity and exogenous ABA responsive gene expression in Arabidopsis (Shinozaki and Yamaguchi-Shinozaki 2000; Seki *et al*. 2002) and rice (Rabbani *et al*. 2003). The positive regulation of PP2CA genes under abiotic stress is well-reported, e.g., in Arabidopsis and *B. oleracea* under drought, exogenous ABA application (Ludwikow *et al*. 2013); and are in line with rice responses to salinity, drought and exogenous ABA (Xue *et al*. 2008). All of the PP2CA genes have shown same direction of regulation (up-regulation) in exogenous ABA and drought stressed plants, which is in line with rice (Xue *et al*. 2008) and barley studies (Seiler *et al*. 2014). Whereas the expression of PP2CA sequence AK357955 has shown negative regulation in salt stress condition, which is against the finding of rice orthologue expression under similar conditions (Xue *et al*. 2008). Based on the FC of ≥ +1.5 or ≤ -1.5 being indicative of notable change some of the PP2CAs identified in barley were not significantly differentially expressed in response to any of the above three stress conditions, whereas others have shown significant differential expression (Table 4.9). For example, ABI1 (AK25184) was found to be differentially expressed in exogenous ABA stress conditions, whereas ABI2 (AK358849) was not. This could be linked to the two different roles of ABI1 in ABA dependent pathways; whereas ABI2 has only one role (Yoshida *et al*. 2006) and similar results were shown in rice (Xue *et al*. 2008). The non-expression of other sequences from the same sub-family does not mean that these PP2CAs are not involved in response to ABA, since PP2CA activity can be tissue specific such as rice orthologue LOC_Os05g49730 (MLOC_15065, AK376382) and LOC_Os04g08560 showed tissue specific expression in root but not in leaf (Xue *et al*. 2008) or they might express at specific developmental stages or under special conditions (Xue *et al*. 2008). The PP2CAs showed stronger induction under water stress in a drought-tolerant variety of chickpea as compared to a sensitive variety (Jain and Chattopadhyay 2010).

Five PYR/PYL/RCARs showed differential expression in abiotic stress conditions, but only MLOC_71349.1 was significantly down-regulated (FC of ≥ +1.5 or ≤ -1.5; Smyth 2004; Zhou *et al*. 2010) by salinity, drought and exogenous ABA stress. AK376521

showed significant differential expression (down-regulation) in drought and exogenous ABA stress only, whereas MLOC_3912.1 was significantly induced only in salt stress condition. The barley PYR/PYL/RCAR sequences down-regulated in this study were also down regulated in Arabidopsis (Santiago *et al*. 2009a; Szostkiewicz *et al*. 2010) and barley (Seiler *et al*. 2014) under drought and exogenous ABA stress conditions. The non-significant expression of the barley PYR/PYL/RCAR sequences in this study also go in line with observations from Seiler *et al*. (2014), which demonstrated that six of the barley PYR/PYL/RCAR sequences were found to show no expression in short term drought stress but negative regulation was observed in prolonged stress. The Arabidopsis *pyr/pyl/rcar* (Park *et al*. 2009) mutant exhibited ABA insensitivity in seed germination, root growth and ABA regulated gene expression, which demonstrated the impact of PYR/PYL/RCAR in stress tolerance.

Of the 5 SnRK2 found to show differential expression in abiotic stress condition, only MLOC_3013.1 (SnRK2 subtype III) showed significant down-regulation (FC of ≥ +1.5 or ≤ -1.5; Smyth 2004; Zhou *et al*. 2010) in drought stress condition. In salt and exogenous ABA stress conditions only AK372880 (SnRK2 subtype II) was found to be significantly up-regulated. The negative regulation of SnRK2 subtype III genes in abiotic stress condition are in line with previous reports for SnRK2 in Arabidopsis (Umezawa *et al*. 2004) and rice (Kobayashi *et al*. 2004). AK372880 was also found to be up-regulated in barley and in Arabidopsis orthologue (Seiler *et al*. 2014). The non-significant expression of SnRK2 genes identified in barley (Table 4.9) in any of the above three stress conditions could be explained by tissue specific expression of SnRK2 genes, for example SnRK2C (MLOC_69212) are expressed mainly in the root tip as compared to leaves and siliques (Umezawa *et al*. 2004).

ABC transporter G40, ABA importer (MLOC_68581.1) was found to significantly up-regulated only in drought and exogenous ABA stress conditions, whereas MLOC_62985.1 (ABC transporter G25; ABA exporter) was significantly down-regulated in salt stress condition only. The ABC transporter gene up-regulated in this study is also induced in Arabidopsis (Seki *et al*. 2002), rice (Zhou *et al*. 2007) and wheat (Kawaura *et al*. 2006) under salt and drought stress. The possible reason for the low expression of ABCG25 as compared to ABCG40 could be that ABCG25 is

expressed in the hypocotyl and in the vascular bundles of root and leaves, whereas the ABCG40 is expressed only in guard cells. Moreover these genes are found to be expressed more in roots as compared to shoots, which support the hypothesis that ABA may be produced predominantly in shoots, transported away and then stored in roots (Boursiac *et al*. 2013). The Arabidopsis *abcg40* (Kang *et al*. 2009) mutant exhibited ABA insensitivity in seed germination, root growth and ABA regulated gene expression, which demonstrated the impact of ABCG40 gene in stress tolerance. The fold change difference among the mRNA-seq I and II salt samples could be due to the primary determinants of gene expression changes resulted from biological differences, rather than the artefacts of platform (Yauk *et al*. 2004).

4.3.4 Other additional genes may also be regulated by ABA

Gene expression profiling studies in plants have shown that ABA regulates the expression of a complex network of genes and proteins that also include transcription factors for example, MYB, WRKY etc., proteins involved in transport of phosphorylate anions [slow anion channel associated 1 (SLAC1)] and K+ ions [K+ ions channels (KAT1)] to induce stomatal closure in response to ABA (Umezawa *et al*. 2010). Thus, despite the main focus of the current study being PP2CAs, PYR/PYL/RCAR, SnRK2s (subfamily II and III) and ABA transport related ABC transporters, additional ABA-responsive genes were analysed using mRNA-seq. Closure of stomata is an important trait for tolerance to abiotic stress, thus the SLAC1 and KAT1, shown to be important for stomatal movements (Negi *et al*. 2008; Pilot *et al*. 2001) were also analysed for expression patterns under abiotic stress. SLAC1 (MLOC_67350) was not significantly up-regulated regulation (FC of $\geq +1.5$ or ≤ -1.5) in any of the three stress conditions (FCs of -1.12, 1.01 and 1.20 for salt, drought and exogenous ABA respectively; Table 4.12). SnRK2 phosphorylates the SLAC1 and KAT1 during abiotic stress, which activates the SLAC1 and inhibits the function of KAT1 (Harrison 2012). SLAC1 gene in Arabidopsis, rice and barley were noted to be involved in the closure of stomata and increase abiotic stress tolerance (Kulik *et al*. 2011; Kusumi *et al*. 2012; Roelfsema *et al*. 2012). However, KAT1 was not found in the mRNA-seq II dataset.

Prior literature has shown that transcription factors such as MYB, ABF, NAC and WRKY etc. are involved in the response to abiotic stress through ABA dependent stress signalling pathway. Analysis of the transcription factors regulated by ABA revealed large families of transcription factors, differing in their responses at the time of stress, where some were up-regulated and some were down-regulated (examples shown in Table 4.16). In addition, studies in rice, Arabidopsis, maize and tobacco have shown that the over-expression of certain transcription factors has the ability to improve the tolerance of plants to environmental stresses such as cold, salinity and drought (Zhang *et al*. 2004). Thus, given the above background, the transcription factors up-regulated in response to salinity stress could be targeted as candidates for salinity tolerance.

Table 4.16 Differential expression of other ABA signalling related candidate genes as determined by mRNA-seq II

MLOC NO	Salt		Drought		ABA	
	FC	SS	FC	SS	FC	SS
WRKY2						
MLOC_58019	1.18	34.57	-1.22	15.27	1.11	22.60
WRKY33						
MLOC_67851	-1.41	80.01	-2.78	20.88	1.14	42.14
NAC5						
MLOC_61270	-1.59	21.08	-1.76	9.45	-1.18	13.87
NAC6						
MLOC_63743	1.66	2.26	-3.59	0.65	6.9	5.24
ABF4						
MLOC_53580	3.5	32.16	4.07	24.03	5.43	38.11
MYB						
MLOC_37446	-1.77	167.12	-1.88	103.11	-1.06	163.00
MLOC_14118	3.30	119.34	186.14	326.92	1.26	5.18
SLAC1						
MLOC_67350	-1.12	15.51	1.01	13.28	1.20	18.40
Cyclophilin						
MLOC_52274	-1.16	30.45	-1.06	43.02	1.19	45.22

4.3.5 Relative expressions of ABA related genes in salt condition are conserved between species

Expression pattern of genes can provide important information about the gene function. As PP2C makes vast gene family, so it was decided to target the expression of only PP2CAs in leaf of barley plant. The expression pattern of PP2CA in rice under salinity stress was extracted from (Mizuno *et al*. 2010).

The differential expression observed for PP2CAs in barley from mRNA-seq was also observed for rice (Mizuno *et al*. 2010). Taken together, most of the PP2CAs in barley and rice were found to possess similar expression pattern, as reported previously for

Arabidopsis, *B. oleracea* (Ludwikow *et al*. 2013) and rice (Xue *et al*. 2008). The high expression of MLOC_8131.1 by (+3.71 and +2.26 mRNA-seq I and II respectively) in leaves after salt stress matches the reported rice orthologue LOC_Os01g62760.1; which also showed highest expression among the other members of gene subfamily PP2CA in salt conditions (Mizuno *et al*. 2010). Similarly the response of other barley PP2CA, (MLOC_7586.1, AK367469, AK367207 and AK374059 to salt response has been mirrored in their rice orthologue, which were also up regulated during salt stress. The expression of AK357955 (non-significant in mRNA-seq I and -1.75 in mRNA-seq II) and AK358849 (-1.57 mRNA-seq I) was not in line with its orthologue in rice (Table 4.14). Certain PP2CAs are found to be specific to developmental stages and time of stress (Xue *et al*. 2008). The rice experiment was conducted at 10 days old plants and they were stressed only for 1 hour (Mizzuno *et al*. 2010), whereas barley mRNA-seq was conducted using 14 days old plants and they were stressed for 12 hours. This could have resulted in the difference in fold change among barley and rice PP2CA orthologues. The high differential expression of members PP2CA sequences indicate their important role in plant's response to stress. The fold change difference among the mRNA-seq I and II salt samples could be due to the primary determinants of gene expression changes resulted from biological differences, rather than the artefacts of platform (Yauk *et al*. 2004).

The most of PYR/PYL/RCAR orthologues in barley and rice showed similar expression under the salt stress conditions, as expected from the previous studies in Arabidopsis (Santiago *et al*. 2009a,b; Szostkiewicz *et al*. 2010) and barley (Seiler *et al*. 2014). The differential expression of barley MLOC_71349.1 (-6.13 and -1.64 from mRNA-se I and II respectively) in leaves after salt stress matches the reported rice orthologue LOC_Os03g18600 (non-significant in rice); which also showed down regulation in salt conditions. The other orthologue of rice LOC_Os03g18600 in barley AK376521 also showed down regulation at the time of salt stress (-2.08and -1.07 from mRNA-se I and II respectively). AK363238 also showed similar relative expression to their orthologue in rice LOC_Os10g42280, both having down regulation in salt stress conditions (FC; non-significant in mRNA-seq I and -1.85 in mRNA-seq II). Similarly MLOC_49654.1 (FC: non-significant) and MLOC_3912.1 (FC: 1.55) showed similar direction of fold change in salt stress condition as their rice orthologues LOC_Os06g36670 and

LOC_Os05g12260 respectively. The differential expression of barley PYR/PYL/RCAR sequences MLOC_39291 (FC: -1.55) were not in line with the rice orthologue LOC_Os02g13330 (Table 4.14).

The SnRK2s in barley and rice were found to possess broad expression pattern, as reported in Arabidopsis (Umezawa *et al*. 2004) and rice (Kobayashi *et al*. 2004). The response of barley SnRK2 subtype II MLOC_69212 (FC; 1.67 in mRNA-seq I and non-significant in mRNA-seq II) to salt has been mirrored in rice orthologues LOC_Os07g42940, which also showed significant up-regulation in salt stress. Barley AK374298 and its rice orthologue LOC_Os03g55600, both showed non-significant fold change under salt stress condition. The rice orthologues of barley sequences MLOC_3013 (-1.57), MLOC_22145 (-1.55) and AK372880 (7.76) showed non-significant differential expression under salt stress condition. The possible reason for the low expression of SnRK2s in rice as compared to barley could be related barley being one of the most salt tolerant cereal crops along with sorghum (Munns and Tester 2008). The SnRK2 subtype III which has ABA responsive box showed relatively low differential expression as compared to subtype II, which lacks ABA responsive box. This suggests that all of these kinases are regulated at the time of stress but SnRK2 subtype II through some other ABA independent pathway (Table 4.14).

The ABA-transport related ABC transporter orthologues in barley and rice appeared to be up regulated in salt stress condition for all of the sequences. The response of barley ABCG25 MLOC_62985 (FC; 2.07 in mRNA-seq I and -1.59 in mRNA-seq II) to salt has been mirrored in rice orthologues LOC_Os11g07600, which also showed significant up-regulation in salt stress. The rice orthologues of barley ABCG40 sequence MLOC_68581 (FC; 2.27 in mRNA-seq I and -1.44 in mRNA-seq II) showed non-significant differential expression under salt stress condition. The ABC transporter sequences ABCG25 and 40 were found to be expressed highly in salt stressed barley which supports the role of these sequences in the transportation of ABA across the plasma membrane (Table 4.14).

The up regulation of different PP2CAs, SnRK2s, PYR/PYL/RCAR, ABCG25 and ABCG40 in salt stress condition in both the plants indicates the role of these genes in primary response to salt stress in all the plants. The high relative expression was observed in barley as compared to rice. The high differential expression of ABA dependent sequences in barley as compared to rice in salt stress condition also indicates the high tolerance of barley as compared to Arabidopsis and rice as also suggested by Munns and Tester (2008). The high relative expression of PP2CA and PYR/PYL/RCAR sequences as compared to SnRK2 subfamily III indicates that ABA dependent pathway might also involve some other form of kinases for further downstream action in cereal crops, which needs to be experimentally proven.

The comprehensive analysis of the barley ABA mediated abiotic stress signalling genes demonstrate their functions and responses to abiotic stresses. The potential roles as their responses to salinity and drought stresses suggest that ABA mediated abiotic stress signalling genes could be candidates for tolerance to various stresses. Despite the great potential locked in ABA mediated abiotic stress signalling genes, epigenetic factors regulating their transcription have not been characterised, thus the next chapter sought to investigate the distribution of one such modification H3K4me3 (leads to gene activation).

Chapter 5

H3K4me3 analysis of barley genes involved in ABA mediated abiotic stress signalling using ChIP-seq

5.0 Abstract

Epigenetic factors, such as histone modification and DNA methylation, play a significant role in regulating transcription. To understand the role and influence of epigenetic factors is a central issue in understanding transcriptional regulation under abiotic stress conditions in plants. The present study aimed to investigate the distribution of one such modification H3K4me3 (leads to gene activation), in order to understand the transcriptional regulation under abiotic stress conditions in plants using chromatin immune-precipitation sequencing (ChIP-seq). The methylated or un-methylated H3K4 regions were initially identified using ChIP-seq of barley leaf under salt and drought stresses, which resulted in the identification of 19,015 and 19,005 sequences exhibiting differential H3K4 trimethylation under salt and drought stresses respectively. The ChIP-seq dataset was then searched for genes with *a priori* annotations relating to ABA mediated signalling, which lead to identification of H3K4me3 within the genes or on adjacent regulatory elements of three protein phosphatase 2C subtype A(PP2CA), four protein phosphatase 2C subtype D (PP2CD), two pyrabactin resistance (PYR)/PYR-like (PYL)/regulatory components of ABA receptor (RCAR) (PYR/PYL/RCAR), three SNF1-related kinase 2 (SnRK2) and two ABA transport related ATP binding cassette (ABC) transporter Type G (ABC transporters; ABCG25 and ABCG40). H3K4 tri-methylation was found at the 5'UTR, exon and intron regions of PP2CAs, whereas H3K4me3 was found at exon, intron and 3' regions of PP2CDs and SnRK2s. For PYR/PYL/RCARs, H3K4me3 was found at 5'UTR, exon, intron and 3'UTR regions. On the other hand H3K4me3 was restricted only to exon and intron regions of ABC transporters. Three PP2CAs, three PP2CDs, two PYR/PYL/RCARs, three SnRK2s and one ABCG40 were found to be significantly un-methylated at H3K4 during abiotic stress, whereas one each of PP2CDs and ABCG25 were significantly tri-methylated at H3K4. Studies on H3K4 sites with respect to genes involved in ABA mediated signalling and their methylation under different stress conditions would provide a fundamental step in understanding the transcriptional regulation of these genes, which can be employed for development of tolerant crops.

5.1 Introduction

Transcription is a complicated process that is regulated by both genetic and epigenetic factors. Epigenetic factors, such as histone modification and DNA methylation, play a significant role in regulating transcription. These epigenetic factors are inherently flexible and are redistributed during development to preserve cell fate decisions (Bernstein *et al*. 2007). With the widespread influence of these epigenetic factors on gene expression, it is expected to observe epigenetic modifications at the time of stress or abiotic stress conditions (Young *et al*. 2011). To understand the role and influence of epigenetic factors is a central issue in understanding transcriptional regulation under abiotic stress conditions in plants.

Posttranslational modifications of histone are known since 1960s (Allfrey *et al.* 1964). But the role of histone modification in molecular function of epigenetic control over nuclear process began to be revealed in 1990s (Suganuma and Workman 2011). Since then histones have been found to be acetylated, methylated on lysine and arginine, ubiquitinated, sumolated, phosphorylated and ribosylated and also to undergo proline isomerization (Suganuma and Workman 2011). Histone modification may play a direct role in affecting chromatin structure or they may represent marks to be recognized by protein effectors (detailed in section 1.11). Histone modifications are best characterized for histone 3 (H3) and histone 4 (H4) (Collas 2010). Di and tri-methylation of H3 lysine 9 (H3K9) and tri methylation of H3 lysine 27 (H3K27me3) cause the formation of repressive heterochromatin through the recruitment of hetrochromatin protein 1 (Lachner *et al*. 2001) and polycomb group (PcG) proteins, respectively (Cao *et al*. 2002; Pasini *et al*. 2004). Acetylation of histone tails loosens their interaction with DNA and makes it accessible to targeting of transcriptional activators (Kingston and Narlikar 1999; Pray-Grant *et al*. 2005). Thus acetylation on H3K9 and H4K16 along with di or tri-methylation at H3K4 is found in euchromatin, often in association with transcriptionally active genes (Bernstein *et al*. 2006; Struhl 1998; Zhao *et al*. 2007). Generally histone modifications such as acetylation, certain phosphorylation and ubiquitination enhance transcription (Sridhar *et al*. 2007), whereas biotinylation and sumoylation repress expression of genes (Nathan *et al*. 2006; Camporeale *et al*. 2007). Similarly tri-methylation of H3K4 activates transcription, whereas di-methylation of H3K9 and H3K27 represses transcription (Zhang *et al.* 2007; detailed in section 1.11).

Chromatin immunoprecipitation has become the optimal technique to examine protein DNA interaction inside the cell. ChIP has been used for mapping the localization of post-translationally modified histones and histone variants in the genome and for recording DNA target sites for transcriptional factors and other chromosome associated proteins (Collas 2010). The PCR reaction can be directly performed on the immuno-precipitated chromatin. ChIP produced precipitated with agarose beads can also directly analysed by quantitative real time PCR (qRT-PCR; Kohzaki and Murakami 2007). The advent of oligonucleotides microarray has revolutionized the gene expression analysis. Combination of genomic DNA microarrays with ChIP assays enabled the binding of transcription factor binding sites (Iyer 2001; Ren *et al*. 2001) and histone modifications (Clark and Shen 2006; Loden and Van 2005) on larger areas in the genome and this approach was called as ChIP on chip. The most powerful tool for identifying protein binding sites across the genome comprises directly and quantitatively sequencing ChIP products. With ability to sequence tens or hundreds of millions of short DNA fragments in a single run, next generation sequencing preceded by chromatin immuno-precipitation (ChIP-seq) can be used to record transcription complexes and epigenetic modifications throughout the genome (Park 2009; Robertson *et al*. 2007). ChIP-Seq offers higher resolution, less noise, and greater coverage than its array-based predecessor ChIP on chip. With the decreasing in cost of sequencing, ChIP-Seq has become an indispensable tool for studying gene regulation and epigenetic mechanisms (Park 2009). In plants, there is an increasing evidence of regulation of gene expression by histone modification under various stresses such as Arabidopsisis (Kim *et al*. 2008), rice (Du *et al*. 2013; Zhong *et al*. 2013), maize (He *et al*. 2014; Zhang *et al*. 2014; Li *et al*. 2014), barley (Braszewska-Zalewska *et al*. 2014; Kapazoglou *et al*. 2013) and wheat (Hu *et al*. 2013; Diallo *et al*. 2012). Studies on these epigenetic factors and their role in regulating transcription under different stress conditions would provide a fundamental step in screening for candidate genes for development of tolerant crops. Hence the preliminary aims of this study were to investigate the distribution of H3K4me3, both within the genes and on adjacent regulatory elements specifically for the genes involved in ABA mediated abiotic stress signalling studied in Chapter 4, and explore their methylation or un-methylation under salinity and drought stress by analysing ChIP-seq dataset.

5.2 Results

5.2.1 Analysis of barley ChIP-seq data

This study investigated histone modification of H3K4me3 in *Hordeum vulgare cv.* Hindmarsh using ChIP-seq. Using the barley (cv. Hindmarsh) leaf chromatin, three ChIP-seq libraries were prepared: (i) from chromatin pooled from six separate plants treated with 150 mM NaCl for 12 h, (ii) from chromatin pooled from six separate plants treated with 20 % PEG for 12 h, (iii) from chromatin pooled from six control plants (explained in section 2.8). The reference libraries were loaded on Illumina Hi-Seq 2500 platform (explained in section 2.8). The next-generation ChIP-seq of barley leaf chromatin yielded over 42.3 million 50 nucleotide (nt) sequence reads from control and 53.8 million and 62.2 million for salt and drought treated plants respectively (Table 5.1).

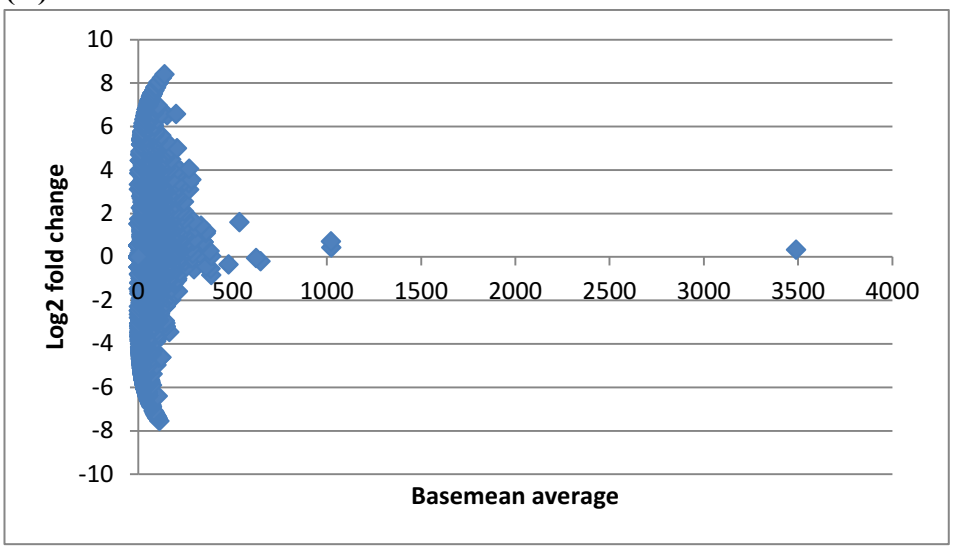
Table 5.1 ChIP-Seq data yield from Illumina Hi-Seq 2500 platform

	Control	Salt stress	Drought stress
Original read length (nt)	50	50	50
Original number of reads	42,371,470	53,825,213	62,166,875
Number of reads after Q30 quality filtering (nt)	37,301,516	37,413,485	41,475,567
Number of reads aligning to IBSC database	28,895,764	28,136,403	27,070,451
% Reads aligned	77.47	75.20	65.27

A computational pipeline was designed, which consisted of published tools and self-developed scripts, for the sequencing data analysis (explained in section 2.7.2). Datasets were filtered for spurious reads using the Fastx Artifacts Filter and poor quality bases were removed from the 3' end with the Fastq Quality Trimmer using a threshold of Q30 (http://hannonlab.cshl.edu/fastx_toolkit/). By aligning 37.3 million, 37.4 million and 41.4 million Q30 quality trimmed ChIP-seq reads from control, salt and drought plants to the International Barley Genome Sequencing Consortium (IBSC) database using Burrows-Wheeler Aligner (BWA) software (Li and Durbin. 2009), it was found that significant proportion of reads (77.47 % control, 75.20 % salt and 65.27 % drought) were mappable to the genome, which suggested that the sequencing data were of high quality (Table 5.1). Model based analysis using MACS (Zhang *et al*. 2008) was used to

obtain the peaks for histone modification. Over twenty thousand of histone modification peaks were obtained from the ChIP-seq data. DESeq, software package (Anders and Hubber 2010) was used to calculate the differential H3K4 trimethylation (Figure 5.1; 6.2). To control the familywise error (the probability of making one or more false discoveries), the false discovery rate (FDR) was the default test for the adjusted p-value in the results from DESEQ using the procedure of Benjamini and Hochberg (1995). The differential H3K4me3 data table can be provided on request.

(A)



(B)

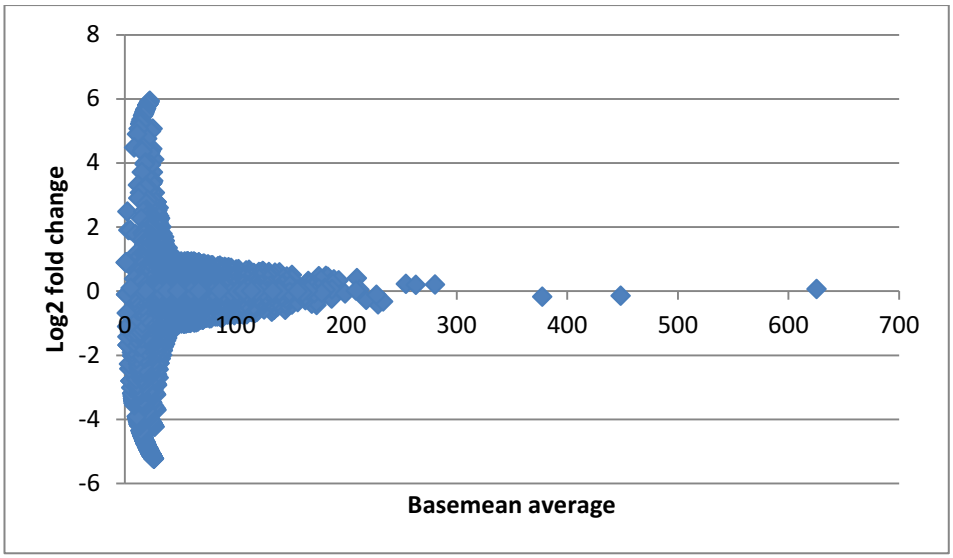


Figure 5.1 Smear plot showing differential H3K4me3 of barley genes under salt and drought stress

(A) Salt stress condition (B) drought stress condition

The graphs were created using Microsoft excel.

5.2.2 Analysis of H3K4me3 in salt stressed plants

Quality trimmed reads from control and salt stressed plants were aligned to barley IBSC genomic database, which led to identification of 19,015 differentially H3K4 trimethylated sequences. DESeq was used to scan for differential histone methylation between control and salinity stress in leaves using a negative binomial model. The sequences with zero reads in control conditions were listed as "infinity" for differential histone methylation and the ones with zero reads in salt stress were listed as "–infinity". Five hundred and eighty four sequences were found to be significantly differentially methylated (FDR adjusted *p* value <0.05). From these, 548 sequences showed increase in H3K4 trimethylation and 36 showed un-methylation at H3K4. The top 20 differentially H3K4 trimethylated genes ranked by methylation change from the barley genome are shown in Table 5.2. The genes with increased H3K4me3 (calculated by DESeq) included a number of genes which have been shown previously to mediate osmotic/ drought/salinity stress tolerance, such as DUF295 domain containing protein (MLOC_57641), DNA binding domain of transcription factor such as leucine zipper (MLOC_53580) and as also several sequences of unknown function (Jiang *et al.* 2013; Kakeda 2009). Genes with decreased H3K4me3 included a number of kinases such as serine/threonine kinase (MLOC_14601) and protein kinase ATP binding site along with calcium binding elongation factor (EF) (MLOC_57566). Serine/threonine kinase is also found to regulate stress responsive gene (Gao *et al*. 2013; Diedhiou *et al*. 2008; Cadenas *et al*. 1999).

Table 5.2 Top 20 genes with increase and decrease in H3K4me3 ranked by methylation change during salt stress

Accession No	Morex Contig No	Differential methylation	P value (FDR adjustment)	Annotation
Genes with increase in H3K4me3				
	Morex_contig_2558999	Inf	.001	
MLOC_57641	Morex_contig_41950	Inf	.001	Protein of unknown function DUF295
MLOC_57640	Morex_contig_41950	Inf	.001	Small GTPase superfamily Small GTPase superfamily Rab type
MLOC_70826	Morex_contig_5862	Inf	.001	Uncharacterised domain XH,XS domain, Zinc finger XS domain
	Morex_contig_1585305	Inf	.001	
MLOC_75821	Morex_contig_69056	Inf	.001	Ribosomal protein S5 S7, Ribosomal protein S7 bacterial organellar type
MLOC_70680	Morex_contig_58347	Inf	.002	Helicase C terminal,Helicase superfamily 1 2 ATP binding domain,SNF2 related 0
MLOC_18725	Morex_contig_1581754	Inf	.002	Glycosyl transferase family 20,HAD like domain,HAD superfamily
MLOC_14761	Morex_contig_1568924	Inf	.003	F box domain cyclin like,Tubby C terminal,Tubby C terminal
MLOC_76709	Morex_contig_72474	Inf	.003	Cys Met metabolism pyridoxal phosphate dependent enzyme,
MLOC_76215	Morex_contig_70567	Inf	.003	NADH ubiquinone oxidoreductase 30kDa subunit 0
	Morex_contig_1975067	Inf	.003	
MLOC_80996	Morex_contig_93981	Inf	.004	DNA repair protein Mre11,Metallophosphoesterase domain,
MLOC_10067	Morex_contig_1558345	Inf	.004	BTB POZ,BTB POZ fold,BTB POZ like,MATH,TRAF like,TRAF
MLOC_21615	Morex_contig_159672	Inf	.004	Mitochondrial transcription termination factor related,
MLOC_71628	Morex_contig_60205	Inf	.005	Apple like,Bulb type lectin domain,Concanavalin A like lectin
MLOC_53580	Morex_contig_38488	Inf	.005	Basic leucine zipper domain,Uncharacterized protein Source

Accession No	Morex Contig No	Differential methylation	P value (FDR adjustment)	Annotation
MLOC_75180	Morex_contig_67303	Inf	.005	ATPase like ATP binding domain,Endoplasmin
	Morex_contig_49512	Inf	.006	
MLOC_18851	Morex_contig_158228	Inf	.006	Ankyrin repeat,Ankyrin repeat containing domain
Genes with decrease in H3K4me3				
	Morex_contig_46733	-Inf	.001	Predicted protein Uncharacterized protein Source UniProtKB TrEMBL Acc F2E3C5 0
MLOC_81111	Morex_contig_94481	-Inf	.01	Plastid lipid associated protein fibrillin conserved domain 0
MLOC_37312	Morex_contig_2548155	-Inf	.011	Major facilitator superfamily,Major facilitator superfamily domain general substrate transporter
MLOC_57566	Morex_contig_41872	-Inf	.020	Calcium binding EF hand,EF Hand 1 calcium binding site,EF HAND 2,EF hand like domain
MLOC_12765	Morex_contig_1564051	-Inf	.028	Aminotransferase class I classII,Aminotransferases class I pyridoxal phosphate binding site
MLOC_13742	Morex_contig_1566402	-Inf	.029	Rpr4901 2 Uncharacterized protein Source UniProtKB TrEMBL Acc A2T582 0
MLOC_20747	Morex_contig_159162	-Inf	.030	
	Morex_contig_12327	-Inf	.031	Protein kinase like domain,Serine threonine tyrosine protein kinase catalytic domain -1909
MLOC_11414	Morex_contig_1561035	-Inf	.031	2 oxoglutarate dehydrogenase E1 component,Dehydrogenase E1 component,Transketolase
MLOC_25246	Morex_contig_168	-Inf	.031	Uncharacterized protein Source UniProtKB TrEMBL Acc M0UV71 0
MLOC_14601	Morex_contig_1568471	-Inf	.036	Protein kinase catalytic domain,Protein kinase like domain,Serine threonine dual specificity
MLOC_66757	Morex_contig_52062	-Inf	.036	Uncharacterized protein Source UniProtKB TrEMBL Acc M0UH32 0
MLOC_10879	Morex_contig_1559973	-Inf	.037	Auxin responsive SAUR protein 0
MLOC_16726	Morex_contig_157429	-Inf	.037	
	Morex_contig_37223	-Inf	.037	Butirosin biosynthesis BtrG like,ChaC like protein,Uncharacterized protein Source UniProtKB

Accession No	Morex Contig No	Differential methylation	P value (FDR adjustment)	Annotation
MLOC_58067	Morex_contig_42316	-Inf	.038	Metridin like ShK toxin,Prolyl 4 hydroxylase alpha subunit,Uncharacterized protein Source
MLOC_59514	Morex_contig_43677	-Inf	.038	
	Morex_contig_6358	-Inf	.038	Glycoside hydrolase catalytic domain,Glycoside hydrolase family 5,Glycoside hydrolase
MLOC_69717	Morex_contig_56862	-Inf	.039	Protein kinase ATP binding site,Protein kinase catalytic domain,Protein kinase like domain
MLOC_14216	Morex_contig_1567582	-Inf	.043	

DESeq was used to scan for differential histone methylation between control and salinity stress in leaves using a negative binomial model. The sequences with zero reads in control conditions were listed as "infinity" for differential histone methylation and the ones with zero reads in salt stress were listed as "–infinity".

5.2.3 Analysis of H3K4me3 in drought stressed plant

Quality trimmed reads from control and drought stressed plants were aligned to barley IBSC genomic database, which led to identification of 19,005 differentially H3K4 trimethylated sequences. DESeq was used to scan for differential histone methylation between control and salinity stress in leaves using a negative binomial model. The sequences with zero reads in control conditions were listed as "infinity" for differential histone methylation and the ones with zero reads in salt stress were listed as "–infinity". Five hundred and three sequences were found to be significantly differentially methylated (FDR adjusted *p* value <0.05). From these, 490 sequences showed increase in H3K4 trimethylation and 13 showed unmethylation at H3K4. The top 20 differentially methylated genes ranked by methylation change from the barley sequences are shown in Table 5.3. Genes with increase in H3K4me3 (calculated by DESeq) included a number of genes which have been shown previously to mediate transcription such as DNA helicase (MLOC_55657) and also leucine zipper transcription factor (MLOC_75510). Leucine zipper protein members were found to belong to the abiotic stress related transcription factors (Komatsuda *et al*. 2007; Xiang *et al*. 2008; Chew *et al.* 2013). The overexpression of DNA helicase members have been shown to confer stress conditions (Sanan-Mishra *et al.* 2005; Vashisht *et al.* 2005; Umate *et al*. 2011). Genes with decrease in H3K4me3 included those in the stress responsive protein zinc finger domain (MLOC_11772), which has been investigated as transcription repressor in Arabidopsis (Sakamoto *et al*. 2004) and some other uncharacterised proteins. A member of serine/threonine protein kinase family (MLOC_72258), which was found to be involved in regulating stress responsive gene (Gao *et al*. 2013; Diedhiou *et al*. 2008; Cadenas *et al*. 1999), was also present in listing of down regulated genes.

Table 5.3 Top 20 genes with increase and decrease in H3K4me3 ranked by methylation change during drought stress

Accession No	Morex Contig No	Differential methylation	P value (FDR adjustment)	Annotation
Genes with increase in H3K4me3				
MLOC_44275	Morex_contig_274444	Inf	0.003	Extracellular ligand binding receptor,
-	Morex_contig_1655373	Inf	0.004	
MLOC_61742	Morex_contig_45823	Inf	0.004	Like Sm LSM domain,Mechanosensitive ion channel MscS,
MLOC_53886	Morex_contig_38719	Inf	0.005	Predicted protein Uncharacterized protein Source UniProtKB TrEMBL Acc F2DFC5 0
-	Morex_contig_1575660	Inf	0.006	
MLOC_55657	Morex_contig_40158	Inf	0.006	DNA helicase,DNA helicase PIF1 ATP dependent,
	Morex_contig_1558108	Inf	0.007	
MLOC_12034	Morex_contig_1562400	Inf	0.007	Glycosyl transferase family 14,
	Morex_contig_57969	Inf	0.007	
MLOC_75510	Morex_contig_6809	Inf	0.007	Basic leucine zipper domain 0
	Morex_contig_275583	Inf	0.007	
	Morex_contig_140717	Inf	0.008338	
MLOC_76383	Morex_contig_71199	Inf	0.008	Uncharacterized protein Source UniProtKB TrEMBL Acc M0Z3F3 0
MLOC_44102	Morex_contig_274185	Inf	0.008	Armadillo type fold,Cell morphogenesis protein C terminal,
MLOC_48405	Morex_contig_322900	Inf	0.008	Protein of unknown function DUF789 0
MLOC_66147	Morex_contig_51154	Inf	0.009	NB ARC 0
MLOC_66532	Morex_contig_51714	Inf	0.009	Mediator complex subunit Med7 0
	Morex_contig_2553190	Inf	0.009	

Accession No	Morex Contig No	Differential methylation	P value (FDR adjustment)	Annotation
	Morex_contig_274629	Inf	0.009	
MLOC_64140	Morex_contig_48411	Inf	0.009	Lactate dehydrogenase glycoside hydrolase family 4 C terminal,
Genes with decrease in H3K4me3				
MLOC_53684	Morex_contig_38555	-Inf	0.006	GH3 auxin responsive promoter 0
	Morex_contig_1570678	-Inf	0.016	
MLOC_72258	Morex_contig_61322	-Inf	0.024	Protein kinase catalytic domain,Protein kinase like domain,Serine threonine protein kinase
MLOC_63693	Morex_contig_47868	-Inf	0.030538	Uncharacterized protein Source UniProtKB TrEMBL Acc M0XWH7 0
MLOC_52202	Morex_contig_37365	-Inf	0.032	Uncharacterized protein Source UniProtKB TrEMBL Acc M0WJW1 0
	Morex_contig_1591518	-Inf	0.041	
MLOC_4785	Morex_contig_135500	-Inf	0.043	Lecithin cholesterol phospholipid diacylglycerol acyltransferase -1878
MLOC_10457	Morex_contig_1559130	-Inf	0.045	Aldolase type TIM barrel,Triosephosphate isomerase,Triosephosphate isomerase active site 0
	Morex_contig_275095	-Inf	0.046	
MLOC_52515	Morex_contig_37581	-Inf	0.047	Domain of unknown function CP12,Glyceraldehyde 3 phosphate dehydrogenase active site,
MLOC_1154	Morex_contig_109819	-Inf	0.048	Uncharacterized protein Source UniProtKB TrEMBL Acc M0UJJ5 0
MLOC_54227	Morex_contig_38979	-Inf	0.048	AP2 ERF domain,DNA binding integrase type 0
MLOC_11772	Morex_contig_1561798	-Inf	0.049	Rubredoxin type fold,Zinc finger CHY type,Zinc finger CTCHY type

DESeq was used to scan for differential histone methylation between control and salinity stress in leaves using a negative binomial model. The sequences with zero reads in control conditions were listed as “infinity” for differential histone methylation and the ones with zero reads in salt stress were listed as “–infinity”.

5.2.4 Identification of genes involved in ABA mediated abiotic stress signalling in ChIP-seq

As detailed in Chapter 4, the genes involved in ABA mediated abiotic stress signalling were extracted from the International Barley Genome Sequencing Consortium (IBSC) CDS database using reciprocal BLAST method (Method C; as mentioned in section 2.22.3). Briefly, this involved a tBLASTn search of each rice protein sequence (Xue *et al*. 2008) against the CDS database of the International Barley Genome Sequencing Consortium (http://mips.helmholtz-muenchen.de/plant/barley/index.jsp/; last accessed May 2014), to find the candidate barley CDSs (start codon to stop codon), which were re-BLASTed (reciprocal BLAST) against the MSU rice genome annotation project database (http://rice.plantbiology.msu.edu/). The barley CDSs which extracted the same rice locus in reciprocal BLAST as the one used in initial BLAST were called 'RBH best match'. The extracted CDS were aligned against their respective rice and Arabidopsis orthologues and analysed for their characteristic features. This led to identification of 13 PP2CAs, 9 PP2CDs, 10 PYR/PYL/RCARs, 5 SnRK2s (subfamily II and III) and 4 ABA transport related ABC transporters (ABCG25 and ABCG40) (Table 4.1, 4.2, 4.3, 4.4). These sequences were BLASTed using the IBSC BLAST tool (http://webblast.ipk-gatersleben.de/barley/viroBlast.php) to extract the Morex and Bowmen contig numbers (Table 4.5). The above extracted Morex_contig numbers were searched in the ChIP-seq dataset, which was prepared by aligning the ChIP-seq, reads with the IBSC barley genome database. This resulted in the identification of three PP2CAs, four PP2CDs, two PYR/PYL/RCARs, three SnRK2 (subfamily II and III) and two ABA-transport related ABC transporters.

5.2.5 Position analysis of tri-methylation at H3K4 with respect to genes involved in ABA mediated abiotic stress signalling

It was found that for the above identified PP2CAs, the histone H3K4 trimethylation was predominately enriched at partial 5' UTR and exon regions. Of the three identified PP2CAs, MLOC_8131 showed H3K4me3 at two different regions (1788-2489 and 2532-2972 regions on Morex_contig_141242; Table 5.4). The first methylation was spanned at the partial 5'UTR and the exon regions, whereas the second one at exon and intron region of the sequence (Figure 5.2). The other two identified PP2CAs, AK374059 and AK357955 showed H3K4 trimethylation only at one region each. The histone modification covered part of exon (4708-5085 region on Morex_contig_50301)

and 5'UTR-exon-intron (3589-4351 region on Morex_contig_40533) regions of AK374059 and AK357955, respectively (Table 5.4; Figure 5.2).

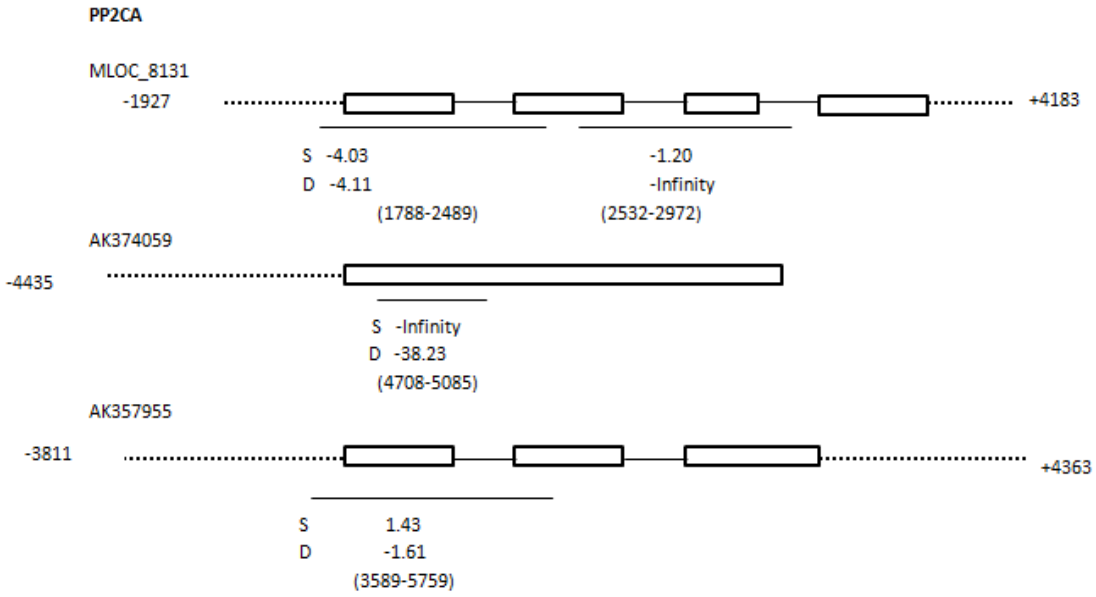


Figure 5.2 Position of H3K4me3 with respect to gene structure of PP2CAs.
The putative exons are shown as rectangles, the putative introns are shown as black lines and un-translated regions are shown as dotted line.; Thin black line mark the region of differential H3K4me3 with respect to gene structure; S: Differential H3K4me3 under salinity; D: Differential H3K4me3 under drought; numbers in the bracket represent the chromosomal positioning of the H3K4me3.

Of the four barley PP2CDs identified in the ChIP-seq, three (AK364849, MLOC_51800 and MLOC_13716.1) showed H3K4me3 only at one region each, whereas MLOC_61942 showed at two different regions. The first H3K4me3 of MLOC_61942 covers the partial exon and 3'UTR (6171-6596 region on Morex_contig_46025), whereas the second covers only exon region (5280-5588) of the same sequence (Table 5.4). The other three PP2CDs AK364849, MLOC_51800 and MLOC_13716.1 demonstrated H3K4 trimethylation at part of 5'UTR-exon (223-609 region on Morex_contig_130636), intron (6240-6688 region on Morex_contig_37090) and 3' UTR (7132-7421 region on Morex_contig_1566336) regions respectively (Table 5.4; Figure 5.3).

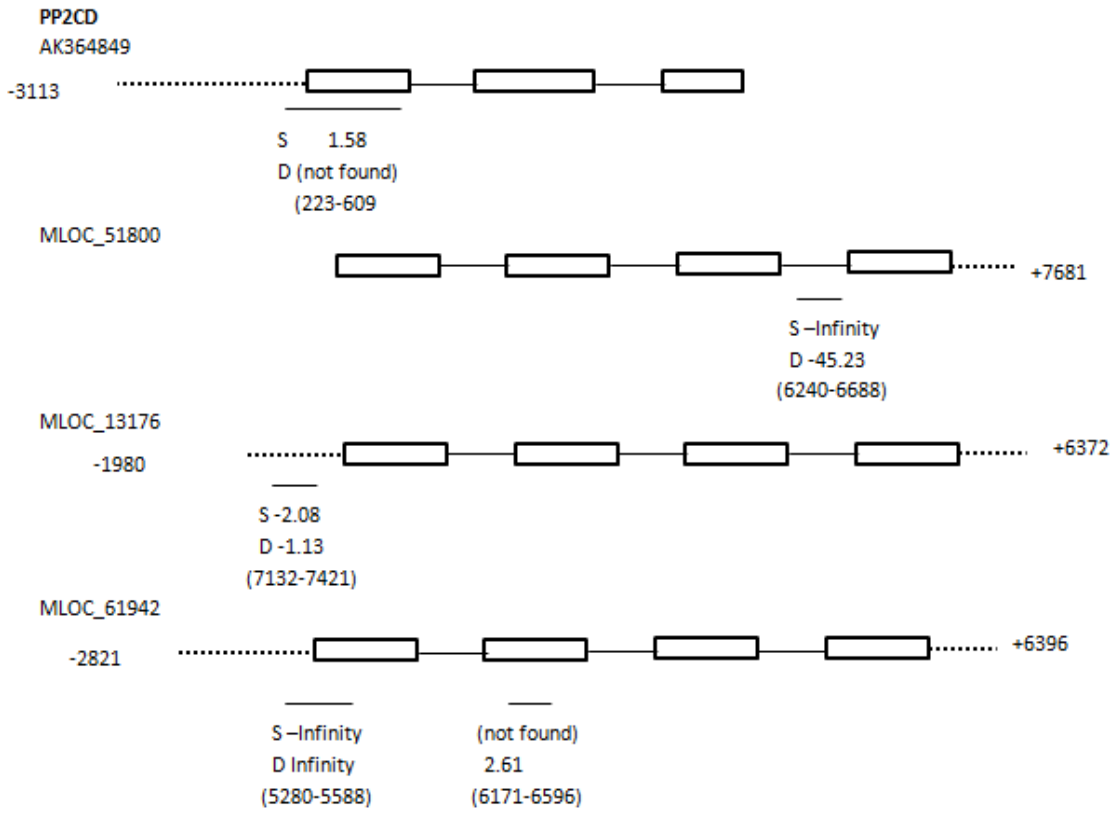


Figure 5.3 Position of H3K4me3 with respect to the gene structure of PP2CDs.
The putative exons are shown as rectangles, the putative introns are shown as black lines and un-translated regions are shown as dotted line.; Thin black line mark the region of differential H3K4me3 with respect to gene structure; S: Differential H3K4me3 under salinity; D: Differential H3K4me3 under drought; numbers in the bracket represent the chromosomal positioning of the H3K4me3.

H3K4 of two PYR/PYL/RCARs identified in the ChIP-seq were trimethylated only at one region each. MLOC_3912 and MLOC_72289.1 demonstrated H3K4me3 at 5' UTR-exon-Intron-exon-3' UTR (1116-2169 region on Morex_contig_134370) and partial 5'UTR-exon (3042-3287 region on Morex_contig_61377) regions respectively (Table 5.4; Figure 5.4).

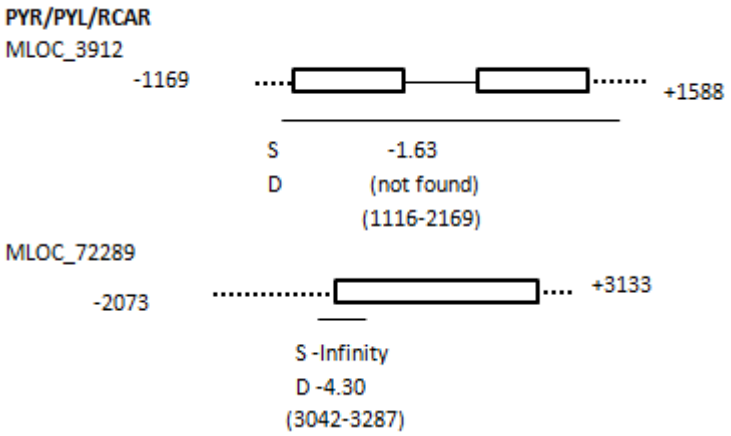


Figure 5.4 Position of H3K4me3 with respect to the gene structure of PYR/PYL/RCARs.
The putative exons are shown as rectangles, the putative introns are shown as black lines and un-translated regions are shown as dotted line.; Thin black line mark the region of differential H3K4me3 with respect to gene structure; S: Differential H3K4me3 under salinity; D: Differential H3K4me3 under drought; numbers in the bracket represent the chromosomal positioning of the H3K4me3.

Three barley SnRK2s (subfamily II and III) were found to have H3K4 trimethylation. H3K4me3 occurred at two different regions in MLOC_3013 (5584-5864, 4964-5317 regions on Morex_contig_127028) and one region at AK251684 (8057-8422 region on Morex_contig_135022), which were located in the exon and intron regions of both the sequences (Table 5.4). On the other hand, MLOC_22145 (2679-3072 region on Morex_contig_160302) was found to have H3K4me3 at 3'-UTR along with exons and intron regions (Table 5.4; Figure 5.5).

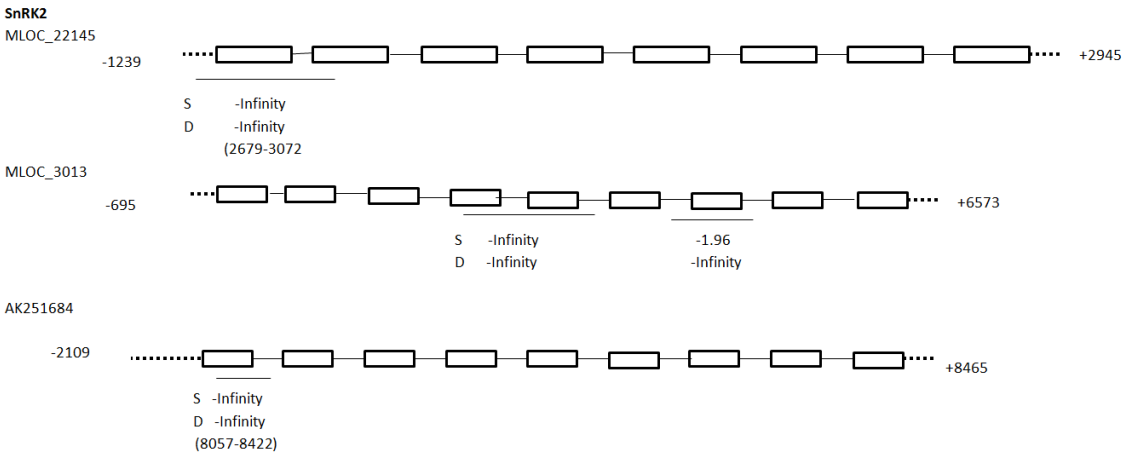


Figure 5.5 Position of H3K4me3 with respect to the gene structure of SnRK2s.
The putative exons are shown as rectangles, the putative introns are shown as black lines and un-translated regions are shown as dotted line.; Thin black line mark the region of differential H3K4me3 with respect to gene structure; S: Differential H3K4me3 under salinity; D: Differential H3K4me3 under drought; numbers in the bracket represent the chromosomal positioning of the H3K4me3.

It was found that for the above-identified ABC transporters, H3K4me3 was predominately enriched in the intron region. Of these, MLOC_62985 showed H3K4 trimethylation at two different regions (4388-4729 and 5724-6136 regions on Morex_contig_47185; Table 5.4). The first H3K4 trimethylation was spanned only at the partial intron region, whereas the second one at partial exon and intron regions of the sequence. The other identified ABC transporter, MLOC_68581 showed H3K4me3 only at one region (18962-19361 region on Morex_contig_54987) covering the intron (Table 5.4; Figure 5.6).

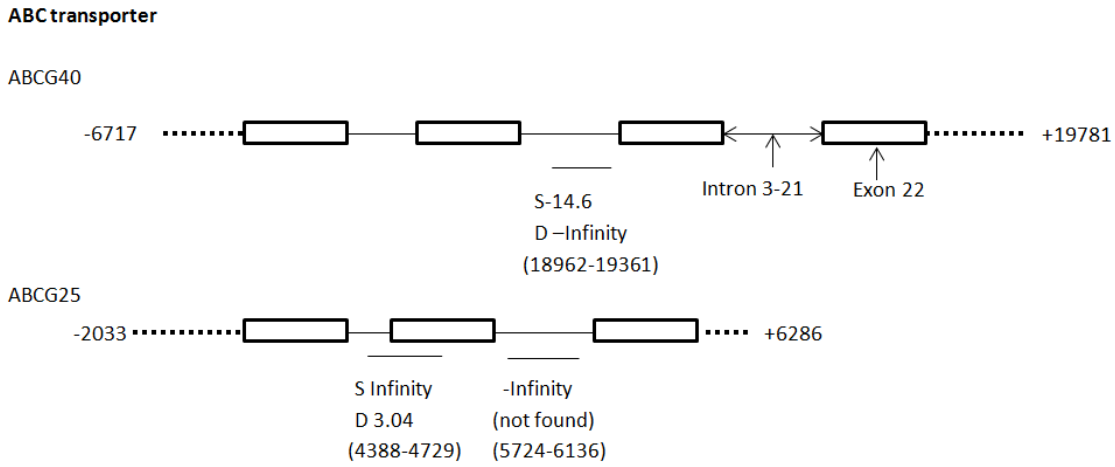


Figure 5.6 Position of H3K4me3 with respect to the gene structure of ABC transporters.
The putative exons are shown as rectangles, the putative introns are shown as black lines and un-translated regions are shown as dotted line.; Thin black line mark the region of differential H3K4me3 with respect to gene structure; S: Differential H3K4me3 under salinity; D: Differential H3K4me3 under drought; numbers in the bracket represent the chromosomal positioning of the H3K4me3.

Table 5.4 Identification and quantitative analysis of differential tri-methylation at H3K4 of candidate barley ABA-signalling related genes under abiotic stress conditions

MLOC No	Morex contig No (+/- strand)	Position of histone modification	Position of gene	Position of modification with respect to contig	Fold Change Salt		Fold Change Drought	
					ChIP-Seq	mRNA-seq	ChIP-Seq	mRNA-seq
PP2CA								
MLOC_8131	Morex_contig_141242(+)	1788-2489	1529-3823	5'-Exon	-4.03	4.60	-4.11	3.03
		2532-2972	1529-3823	Exon-Intron-Exon	-1.20		-infinity	
AK374059	Morex_contig_50301(+)	4708-5085	4256-5957	Exon	-infinity	4.37	-38.23	-
AK357955	Morex_contig_40533(+)	3589-4351	3564-5759	5'-Exon-Intron-Exon	1.43	-1.75	-1.61	2.74
PP2CD								
AK364849	Morex_contig_130636(-)	223-609		Exon	1.58	-1.75	-	-1.40
MLOC_51800	Morex_contig_37090(-)	6240-6688	4953-7681	Intron	-infinity	-1.91	-45.23	2.00
MLOC_13716.1	Morex_contig_1566336(-)	7132-7421	3359_7881	3'	-2.08	-1.23	-1.13	-1.00
MLOC_61942	Morex_contig_46025(-)	5280-5588	413-7103	Exon	-infinity	1.41	Infinity	1.07
		6171-6596	413_7103	Exon-3'	NA		2.61	

MLOC No	Morex contig No (+/- strand)	Position of histone modification	Position of gene	Position of modification with respect to contig	Fold Change Salt		Fold Change Drought	
					ChIP-Seq	mRNA-seq	ChIP-Seq	mRNA-seq
PYR/PYL/RCAR								
MLOC_3912	Morex_contig_134370(+)	1116-2169	911-2562	5'-Exon-Intron-Exon-3'	-1.63	1.55	-	-1.23
MLOC_72289	Morex_contig_61377(-)	3042-3287	2139-3759	5'-Exon	-infinity	-	-4.30	-
SnRK2 (subfamily II and III)								
MLOC_22145	Morex_contig_160302(-)	2679-3072	583-3228	Exon-Intron-Exon-3'	-infinity	-1.55	-infinity	-1.06
MLOC_3013	Morex_contig_127028(-)	5584-5864	4183_6783	Intron-Exon-Intron-Exon	-infinity	-1.57	-infinity	-1.66
		4964-5317		Exon-Intron-Exon-Intron	-1.96		-infinity	
AK251684	Morex_contig_135022(-)	8057-8422		Intron-Exon	-infinity	14.78	-infinity	2.49
ABC transporters								
MLOC_68581	Morex_contig_54987(-)	18962-19361	290-19917	Intron	-14.6	2.27	-infinity	4.76
MLOC_62985	Morex_contig_47158(-)	4388-4729	2404-6610	Intron	Infinity	2.07	3.04	-1.38
		5724-6136	2404-6610	Intron-Exon	-infinity		NA	

DESeq was used to scan for differential histone methylation between control and salinity stress in leaves using a negative binomial model. The sequences with zero reads in control conditions were listed as "infinity" for differential histone methylation and the ones with zero reads in salt stress were listed as "–infinity".

5.2.6 Quantitative analysis of differential tri-methylation at H3K4 of candidate barley ABA-signalling related genes under abiotic stress conditions

Of the 13 putative PP2CAs identified from barley IBSC genome that had the characteristic motifs of PP2CAs (Table 4.1), only three were found in the ChIP-seq dataset (Table 5.4). These genes demonstrated H3K4 trimethylation change under drought (differential methylation up to -1.61) and salt (differential methylation up to 1.43) stresses (Table 5.4). One PP2CA (MLOC_8131) showed significant (differential methylation of $\geq$ +1.5 or $\leq$ -1.5) decrease in H3K4 tri-methylation (differential methylation salt -4.03; differential methylation drought: -4.11) in both stresses and was noted showing highest increase in expression in mRNA-seq (Table 4.9) as compared to other PP2CAs. AK374059 also showed significant decrease under salt (-Infinity) and drought (-38.23) stresses, whereas AK357955 exhibited significant decrease in H3K4me3, only under drought stress (-1.61).

Of the nine putative PP2CDs identified from barley IBSC genome that had the characteristic motifs of PP2CDs (Table 4.1), only four were found in the ChIP-seq dataset (Table 5.4). Similar to PP2CAs, PP2CDs have also shown H3K4 trimethylation changes under drought (differential methylation up to +1.58) and salt (differential methylation up to -2.08) stresses (Table 5.4). One PP2CD, MLOC_61942 exhibited increase in H3K4 trimethylation at two different regions under drought (Infinity and 2.61), whereas only one region was found to show decrease in H3K4 tri-methylation under salinity (-Infinity). MLOC_51800 exhibited relatively high negative H3K4 trimethylation under salt (differential methylation –Infinity) and drought (differential methylation -45.23) stress compared to other PP2CDs. As stated in section 4.2.11, none of the PP2CDs were identified as showing significant differential expression in mRNA-seq, whereas PP2CDs have shown strong H3K4 trimethylation under drought and salt stress.

Of the ten putative PYR/PYL/RCARs identified from barley IBSC genome that had the characteristic residues (Table 4.2), only two were found in the ChIP-seq dataset (Table 5.4). MLOC_72289 showed decrease in H3K4 tri-methylation under salt and drought stresses with differential methylation of –infinity (none of the reads were found under stress) and -4.30 respectively, whereas MLOC_3912 showed significant decrease in

H3K4 trimethylation only under salt stress (differential methylation -1.63) (Table 5.4). On comparing to the transcriptome study, MLOC_72289 was not found to show any expression in the barley leaves under either drought or salt stress, whereas MLOC_3912 was found to show significant negative expression (<=1.5) under salt stress only in mRNA-seq.

Of the six SnRK2s (subfamily II and III), identified from barley IBSC genome, that had characteristic residues (Table 4.3), only three were found in the ChIP-seq dataset (Table 5.4). MLOC_3013 demonstrated decrease in H3K4 trimethylation at two different regions of the gene. Under salt stress, one of H3K4me3 region of MLOC_3013 (4964-5317) showed H3K4me3 of -1.96, whereas in drought stress it exhibited a change of –infinity (Table 5.4). All other SnRK2s have shown –infinity of H3K4 trimethylation under drought and salt stress. MLOC_3013 was found to show highest differential expression among SnRK2 subfamily III under salt and drought stress in mRNA-seq.

Of the six ABA-transport related ABC transporters, identified from barley IBSC genome that had the characteristic residues (Table 4.4), only two were found in the ChIP-seq dataset (Table 5.4). Both of the genes MLOC_68581 and MLOC_62985 showed one region for H3K4me3 each in drought stress, whereas MLOC 62985 showed two H3K4 trimethylation regions in salt stress. MLOC_68581 and MLOC_62985 showed –infinity and 3.04 for H3K4 trimethylation respectively under drought stress, whereas -14.6 and –infinity respectively were observed for above mentioned genes under salt stress (Table 5.4).

5.2.7 Validation of ChIP-seq

Nuclei were isolated and ChIP was performed using antibodies for biological triplicates. PCR of 5 selected sequences from the ChIP-seq data resulted in successful amplification of the bands of expected sizes. Immunoprecipitated DNA was analysed by qRT-PCR, and largely supported the direction of change of H3K4 trimethylation as detected by ChIP-seq for all genes (Table 5.5)

Table 5.5 Validation of DNA quantification from ChIP-Seq using qRT-PCR

BSC accession No	Salt (Fold change)		Drought (Fold change)	
	ChIP-Seq	qRT-PCR	ChIP-Seq	qRT-PCR
MLOC_3912	-1.63	-1.13±0.26	1.03	-1.58±0.07
MLOC_22145	-infinity	-1.54±0.41	-infinity	-1.96±0.23
MLOC_3013	-1.96	-2.21±0.12	-infinity	-1.74±0.36
MLOC_68581	-1.46	-1.48±0.39	-infinity	-1.12±0.18
MLOC_62985	-infinity	-1.54±0.48	3.04	-1.87±0.52

5.3Discussion

5.3.1 Barley H3K4me3 analysis

Gene expression under developmental and stress depends upon nucleosome histone post-translational modifications and also on DNA methylation. Chromatin immunoprecipitation (ChIP) followed by high-throughput DNA sequencing (ChIP-seq) is a valuable and extensively used method for mapping the genomic location of transcription factor binding and histone modifications in living cells. In this study, a genome wide pattern of histone H3 tri-methylation at lysine 4 (H3K4me3) in control, salinity (150mM NaCl for 12h) and drought (20% PEG 12h) treated barley leaves were reported, which generated over 150 million 50 nucleotide reads. This amounted to 7.92 Gbp after quality trimming. Abiotic stress, are major causes of cereal crop yield losses and significantly impact on sustainability, and this study aimed to determine H3K4me3 epigenetic changes marked the chromatin of abiotic stressed plant and if these correlated to changes in gene expression.

When plants are exposed to salinity or drought in laboratory conditions a rapid temporary drop in growth is observed, which is followed by a gradual recovery to a new reduced rate of growth. The temporary effects are due to rapid and transient changes in plant water relations (Munns 2002). Stress specific effects can become visible after few days at high salinity or drought. So ideally, to study the response of plants to acute stress, one should investigate H3K4me3 over a range of times to study the dynamics of H3K4 trimethylation under acute salinity or drought, as done for other studies such as transcriptomics (Seki *et al*. 2002). As explained in section 3.3.1 investigators have undertaken studies in acclimation models such as exposure to cyclic or continuous mild stress (Watkinson *et al*. 2003), which are relevant to study of plants exposed to chronic stress in the field. However, these stresses negatively impact developmental pathways, which then confound the comparison of control and stress samples due to growth retardation of stressed plants. Hence, in order to avoid the potential retardation obscuring the responses of plants to high salinity and drought, researchers often investigate the response to acute salinity and drought stress. While it is preferred to have profiled many time points, these deep sequencing experiments are still too costly. As such, the 12 h time point was selected as the one likely to produce the clearest

effects including differential H3K4 trimethylation of several key transcription factor encoding genes as well as response genes.

This work was aimed to identify barley genes which were differentially trimethylated at H3K4 under acute salt and drought stress in the barley leaf. Of the top 20 genes showing maximum increase (differential methylation of ≥ +1.5) in H3K4me3 under salt stress, only two were found to show significant increase in expression (FC of ≥ +1.5) under salt stress from mRNA-seq experiment (MLOC_10067 and MLOC_53580; Table 3.5; 5.2). MLOC_10067 "BTB POZ, BTB POZ fold, BTB POZ like, MATH, TRAF like, TRAF" and MLOC_53580 "Basic leucine zipper domain, Uncharacterized protein Source" showed 1.74 and 3.50 fold increase in expression under salt stress. The BTB/POZ domain protein found in MYB transcription factor was found to be regulator of hormone (such as gibberellin and abscisic acid) responsive gene expression in barley (Woodger *et al*. 2004). BTB/POZ has also been investigated in Arabidopsis (Guo *et al*. 2013), wherein this gene is not only induced in salinity stress but also in drought conditions. These results were also in line with findings of wheat (Pandey *et al*. 2013), rice (Kimura and Kagawa 2006) and maize (Poland *et al*. 2011). The members of leucine zipper family were found to belong to the abiotic stress related transcription factors of barley (Komatsuda *et al*. 2007), rice (Xiang *et al*. 2008) and wheat (Chew *et al.* 2013). Other important candidates for future functional testing and development of tolerant plants through genetic modification or breeding, which also showed increase in H3K4 trimethylation under salt stress, could include DUF295 (MLOC_57641) domain containing protein and several sequences of unknown function. Domain of unknown function 295 (DUF295) was also found to be involved in abiotic stress response in rice (Jiang *et al.* 2013) and barley (Kakeda 2009). On the other hand, of the top 20 genes showing decrease in H3K4 trimethylation, none exhibited significant differential expression under salt stress in mRNA-seq. Other gene which showed decrease in H3K4me3 under salt stress include a number of kinases such as member of serine/threonine kinase family (MLOC_14301) and protein kinase ATP binding site along with calcium binding elongation factor (EF; MLOC_57566)). This serine/threonine kinase is also found to regulate stress responsive genes in Arabidopsis (Gao *et al*. 2013), rice (Diedhiou *et al*. 2008) and barley (Cadenas *et al*. 1999).

Of the top 20 genes showing increase (differential methylation of ≥ +1.5) in H3K4me3 under drought stress, only two were found to show significantly negative expression (FC of ≤ -1.5) under drought (MLOC_44102 and MLOC_66147; Table 3.6, 5.3). According to the IBSC annotation, MLOC_44102 and MLOC_66147 were annotated as "Armadillo type fold, Cell morphogenesis protein C terminal" and "NB ARC" respectively. Cell morphogenesis genes were recently found to be involved in abiotic and biotic stress response in Arabidopsis and rice (Gachomo *et al.* 2013; Kim *et al*. 2009). Of the 13 genes found to show significant decrease(differential methylation of ≤ -1.5) in H3K4me3 under drought stress, only one MLOC_54227 also exhibited significant negative expression ("AP2 ERF domain, DNA binding integrase type"; Table 3.63, 5.3). According to Dong *et al.* (2012) AP2/ARF domain containing ERF4 transcription factor were down-regulated under abiotic stress in wheat. Other important candidates for future functional testing and development of tolerant plants through genetic modification or breeding, which also showed increase in H3K4 trimethylation under drought stress, could include members of DNA helicase and leucine zipper families. The overexpression of DNA helicase (MLOC_55657) has been shown to confer stress conditions in tobacco (Sanan-Mishra *et al.* 2005). These results were in line with findings in pea (Vashisht *et al.* 2005), rice and Arabidopsis (Umate et al. 2011). The leucine zipper protein (MLOC_75510) was found to belong to the abiotic stress related transcription factors is as explained above. Sequences with un-methylation at H3K4 include those in the stress responsive protein zinc finger domain and some other uncharacterised proteins. A serine/threonine protein kinase (MLOC_72258) was also present in listing of down regulated genes. Zinc finger domain (MLOC_11772) protein has been investigated as transcription repressor in Arabidopsis, wherein this gene has not only differentially expressed under drought but also under salinity and cold conditions (Sakamoto *et al*. 2004). The role of zinc finger proteins under salt and drought tolerance in rice (Huang *et al.* 2009), cold and drought in barley (Mare *et al.* 2004) and salinity in wheat (Li *et al.* 2010b) also support the results. The role of serine /threonine protein kinase in stress tolerance is as explained above. The Fisher exact test value (Fisher 1954) H3K4for trend of H3K4me3 associated with gene activation is 1, which is non-significant (<0.5). As very few genes which have shown significant increase or decrease in H3K4me3 also exhibited significant differential expression

under similar conditions using mRNA-seq, it was hard to conclude that the general trend of H3K4me3 associated with gene activation holds for barley or not.

5.3.2 Tri-methylation at H3K4 of genes involved in ABA mediated stress tolerance pathway

As detailed in Chapter 4, the ABA-signalling related genes were extracted from the International Barley Genome Sequencing Consortium (IBSC) CDS database and mRNA-seq dataset using reciprocal BLAST method. This led to identification of 13 PP2CAs, 9 PP2CDs, 10 PYR/PYL/RCARs, 6 SnRK2s (subfamily II and III) and 4 ABA transport related ABC transporters (ABCG25 and ABCG40) (Table 4.9). Of these 11 PP2CAs, 9 PP2CDs, 7 PYR/PYL/RCARs, 6 SnRK2s (subfamily II and III) and 2 ABA transport related ABC transporters (ABCG25 and ABCG40) showed differential expression under salt and drought stress conditions altogether by mRNA-seq. Epigenetic mechanism such as histone modifications may have a decisive function in regulating expression of plant genes under abiotic stress. The present study is amongst the first genome-wide studies, targeting H3K4 trimethylation of leaf genes in salt and drought stress conditions. Three PP2CAs showed significant decrease in H3K4me3 under salt and drought stress except MLOC_357955, which demonstrated non-significant increase in H3K4me3 under salt stress. The meta profile of H3K4me3 in humans and Arabidopsis suggests that an increase in the H3K4me3 is related to up-regulated expression of the gene and decrease in H3K4me3 is linked to the down-regulated expression of gene (Young *et al*. 2011; Brusslan *et al*. 2012), whereas this study did not find this for identified PP2CAs in barley. The H3K4me3 marks for three PP2CAs were found at or near ATG start codons (Figure 5.2), which are in line with findings in rice and maize (Du *et al*. 2013; Wang *et al*. 2009b), whereas these methylations are found upstream of transcription start site (TSS) in humans (Barski *et al*. 2007; Wang *et al*. 2008).

Four PP2CDs were identified with differential H3K4me3 under salt and drought stress. Of these, only two (MLOC_51800 and MLOC_13716) showed significant decrease in H3K4me3 and gene expression under salt stress. The other two PP2CDs exhibited discordant H3K4me3 and gene expression (Table 5.4). Similarly two PP2CDs, AK364849 and MLOC_61942 showed conservation in direction for H3K4me3 and

gene expression under drought stress. MLOC_13716 showed no differential expression under drought stress. These results are in line with general trend of activation of genes by H3K4me3 as explained above (Young *et al*. 2011; Brusslan *et al*. 2012). The H3K4me3 sites at these PP2CDs are also at or near ATG except, MLCO_51800 which are in accordance to other studies in rice and maize (Du *et al*. 2013; Wang *et al*. 2009b).

Two PYR/PYL/RCARs were extracted from the ChIP-seq. Both of these sequences exhibited significant decrease in H3K4 tri-methylation under salt and drought stress, except MLOC_3912 showing increase in H3K4me3 under drought stress (Table 5.4). MLOC_3912 was found to show different directions in H3K4me3 and gene expression, which is against the general trend of activation of genes by H3K4me3 as explained above (Young *et al*. 2011; Brusslan *et al*. 2012). The other PYR/PYL/RCAR, MLOC_72289 was not found to express in mRNA-seq. Both PYR/PYL/RCARs demonstrated H3K tri-methylation at ATG, which is in line with previous studies in rice and maize (Du *et al*. 2013; Wang *et al*. 2009b).

Three SnRK2s were found to show H3K4 tri-methylation under salt and drought stress. All of these sequences showed significant decrease in H3K4me3 under salt and drought stress (Table 5.4). MLOC_22145 and MLOC_3013 demonstrated significant decrease in H3K4me3 and gene expression under salt and drought stress, which are in line with results from humans and Arabidopsis (Young *et al*. 2011; Brusslan *et al*. 2012). The other SnRK2, AK251684 showed significant decrease in H3K4me3 under salt and drought stress, whereas significant increase in transcription was observed in mRNA-seq under similar conditions. Of the three SnRK2s, two MLOC_22145 and AK251684 demonstrated H3K4me3 at or near ATG, whereas the MLOC_3013 exhibited these sites downstream of ATG (Figure 5.5). These results are also in line with findings from rice and maize (Du *et al*. 2013; Wang *et al*. 2009b).

Both of the ABA signalling related ABC transporters were found to show H3K4me3 under salt and drought stress. MLOC_68581 showed similar direction of H3K4me3 and gene expression under salt stress only. MLOC_62985 demonstrated two sites for H3K4 tri-methylation, one closer to ATG site was considered and it demonstrated similar direction of H3K4me3 and gene expression under salt stress. Both of the ABA

signalling related ABC transporters showed H3K4me3 downstream of ATG, which is different from the findings in rice and maize (Du *et al*. 2013; Wang *et al*. 2009b).

For the genes involved in ABA mediated stress tolerance pathway, a large proportion 60 % (21 of the 35) of these genes were expressed in salinity and drought stresses without significant level of H3K4me3. Two striking examples are AK251854 (PP2CA) and MLOC_71349 (PYR; Section 4.2.11, 4.2.12; Table 4.9), which were significantly expressed during salt and drought stresses, yet completely lacked H3K4me3 modification. It is possible that other epigenetic modifications, such as histone acetylation or cytosine methylation along with modification at enhancers may be involved as the plants are exposed to abiotic stress condition and these modifications need to be evaluated in the future work with respect to expression of ABA signalling related genes.

In this study, the focus was to understand how differences in the distribution of H3K4me3 impact gene expression; however other functional properties such as promoter usage, alternative splicing, antisense transcription and replication timing should also be considered. Many new insights into the biology of gene expression will come as we continue to map more chromatin modifications and uncover new mechanisms that influence transcription. The understanding of relation between the epigenetic modifications and plant's response to environmental stress response is highly desirable, as it is required not only for better understanding molecular mechanism of plant's stress response but also for possible application in the genetic manipulation in plants. Despite the great potential for abiotic stress tolerance in ABA mediated abiotic stress tolerance signalling genes, they have not been employed for determining salt tolerance of different varieties, thus the next chapter sought to identify salt tolerance levels of barley varieties by key physiological and molecular assays.

Chapter 6

Identification of salt tolerant and salt sensitive barley varieties by a consolidated physiological and molecular approach

6.0 Abstract

Soil salinity is a significant international environmental problem and severely affects the yield of cereal crops due to numerous effects on plant-water relations, ion homeostasis and salt toxicity. This work assessed the salt tolerance levels of sixteen barley varieties by key physiological and molecular tests. Seedlings were exposed to acute salinity stress and their physiological responses including relative water content, levels of the stress hormone abscisic acid (ABA) and Na^+/K^+ ratio were compared. A significant variation was noted amongst the varieties, with Calmariout, Hindmarsh and Mundah identified as tolerant varieties and Franklin as the most sensitive variety. Further, the differential expression of key genes in the ABA-regulated stress response pathway, i.e., pyrabactin resistance (PYR)/PYR-Like (PYL)/regulatory components of ABA receptors (RCAR), protein phosphatase 2C (PP2C), sucrose non-fermenting 1-related kinase (SnRK2) and specific ABC transporters, was analysed by quantitative real-time PCR of leaf RNA. The PP2CA and ABC transporters were up-regulated while the PYR/PYL/RCARs and SnRK2s were down-regulated under salt stress. Importantly, tolerance ranking by gene expression closely correlated that by physiological indices. Thus expression analysis of the ABA pathway can be used for rapid identification of potentially salt-tolerant barley varieties before undertaking physiological studies. The outcomes are significant for varietal selection and comparing the alleles for genetics of tolerance.

6.1 Introduction

The phytohormone abscisic acid (ABA) is crucial for plant's adaptive response to salinity, drought and other abiotic stresses. ABA accumulates in plant cells and induces stomatal closure, leaf abscission and expression of many genes, the products of which may protect vegetative tissues (Umezawa *et al*. 2010). ABA also regulates developmental events such as seed germination and dormancy (explained in section 1.2.1). Some aspects of the cellular and molecular basis of ABA dependent stress response pathway were unclear before the discovery of the pyrabactin resistance (PYR)/PYR-Like (PYL)/regulatory components of ABA receptors (RCAR) as the soluble receptors for ABA, followed by that of the protein phosphatases of class 2C subfamily A (PP2CA) and sucrose non-fermenting 1-related protein kinase 2s (SnRK2s) sub family III as downstream components of PYR/PYL/RCAR (Umezawa *et al*. 2010). The double-negative regulatory system of ABA signalling is comprised of PYR/PYL/RCAR (ABA receptors), the PP2CA and SnRK2 (Subfamily III), the transcription factors such as WRKY, MYB, MYC, ABRE binding factor (ABF) and NAC, and their downstream targets such as salt overly sensitive (SOS) and late embryogenesis abundant (LEA) among others (explained in Section 1.6). The ATP binding cassette (ABC) transporter Type G protein ABCG40 plays a role in the uptake of ABA whereas ABCG25 acts as exporter of ABA through the plasma membrane (Umezawa *et al*. 2010).

Among the plant physiological changes under salinity, the relative water content (RWC) is a measure of amount of water contained in a plant tissue, as compared to its maximum carrying capacity. RWC is considered a superior measure for determining tolerance to salinity and drought, as compared to thermodynamic state variables such as water potential, turgor potential and solute potential (Sinclair and Ludlow 1985), and has been used extensively, e.g., abiotic stress studies in barley, wheat and rice (Munns *et al*. 2010; Fukao *et al*. 2011). Plants subjected to salinity stress take up large amounts of Na^+ ions, while uptake of K^+ ions is impaired, hence low Na^+/K^+ selectivity has been used as an important selection criterion for salt tolerance in cotton, barley and wheat (Ashraf 2004). The level of endogenous ABA was found to be up-regulated under salt stress in several plants such as tobacco, rice (Moons *et al*. 1995), maize (Jia *et al*. 2002) and *Medicago sativa* (Palma *et al*. 2014). Increases in endogenous ABA and ABA-

induced proteins are considered a characteristic of tolerant varieties (Moons *et al*. 1995). Hence these physiological characteristics were chosen to monitor salinity tolerance of barley varieties in the present study.

Plants exposed to acute salinity in laboratory conditions exhibit a rapid temporary drop in growth due to changes in plant-water relations, followed by a gradual recovery to a slower growth rate (Munns *et al*. 2010), and specific phenotypes become visible after a few days at high salinity. To study the response to acute stress, gene expression is ideally profiled at a range of time points (Seki *et al*. 2002), while study of plant exposure to cyclic or continuous mild stress represents chronic stress in the field (Watkinson *et al*. 2003). However, chronic stresses negatively impact on developmental pathways, which confounds accurate comparisons of control and stress samples; as such, application of acute stress avoids this complication. In the present study, 16 barley varieties were screened for acute salt tolerance (150 mM NaCl for 12 hours) by assessing key physiological parameters as well as differential expression of selected genes in the ABA-regulated stress response pathway. The results have led to identification of tolerant varieties which would be excellent candidates for cultivation as well as study of salt tolerance related genes/alleles, and development of molecular tests for rapid identification of such varieties.

6.2 Results

6.2.1 Effects of salinity on the relative water content of shoot

The shoot relative water content index **(**RWCI) describes the reduction of RWC after acute salt stress compared to controls (Kausar *et al.* 2012). Exposure to 12 h salt stress led to a significantly reduced RWC (p <0.05, one-sided Student's t test; explained in Section 2.20) in all 16 varieties (Table 6.1; Figure 6.1). The most severe reduction was in Franklin (RWCI: 71.66%±0.4), followed by Gairdner (73.55%±0.1) and Arivat (77.28%±0.7), while more limited changes occurred in Calmariout (97.22%±0.6), Hindmarsh (96.26%±0.5) and Mundah (96.05%±0.5).

Table 6.1 Relative water content based stress tolerance of 16 varieties of barley.

	Calmariout	Hindmarsh	Mundah	ELB14	Golden promise	Buloke	Maythorpe	Vlamingh
Shoot length (cm) control	29.6±0.42	29.4±0.21	28.8±0.32	27.4±0.48	27.7±0.13	27.8±0.44	27.6±0.11	26.4±0.43
Shoot length (cm) stress	29.3±0.07	29.0±0.27	28.4±0.27	27.0±0.78	27.3±0.43	27.4±0.26	27.2±0.48	26.0±0.16
Root length (cm)control	18.2±0.25	17.9±0.52	18.1±0.13	17.6±0.46	17.2±0.11	17.7±0.22	16.4±0.16	16.5±0.24
Root length (cm) stress	18.0±0.26	17.6±0.16	17.9±0.45	17.2±0.29	16.8±0.26	17.4±0.32	16.1±0.23	16.1±0.37
Fresh weight (Control)	0.602±0.029	0.601±0.034	0.594±0.056	0.575±0.048	0.579±0.022	0.572±0.042	0.576±0.031	0.561±0.034
Dry weight (Control)	0.063±0.012	0.064±0.022	0.06±0.047	0.064±0.037	0.066±0.034	0.059±0.044	0.0596±0.026	0.0592±0.017
Turgid Weight (Control)	0.644±0.049	0.657±0.056	0.659±0.036	0.658±0.021	0.657±0.016	0.659±0.038	0.662±0.029	0.658±0.015
RWC Control	92.7±0.022	90.52±0.037	89.14±0.006	86.03±0.018	86.79±0.049	85.49±0.029	85.71±0.018	83.88±0.018
Fresh weight (Stress)	0.582±0.037	0.57±0.044	0.561±0.025	0.521±0.019	0.527±0.004	0.513±0.021	0.512±0.026	0.502±0.033
Dry weight (Stress)	0.062±0.022	0.059±0.035	0.061±0.027	0.062±0.016	0.061±0.047	0.0601±0.018	0.0611±0.024	0.0615±0.039
Turgid Weight (Stress)	0.639±0.014	0.645±026	0.645±0.03	0.642±0.034	0.642±0.034	0.637±0.044	0.636±0.031	0.639±0.012
RWC Stress	90.13±0.049	87.14±0.046	85.62±0.024	79.15±0.005	80.21±0.031	78.52±0.014	78.37±0.025	76.21±0.031
RWC Index	**97.22±0.6**	**96.26±0.48**	**96.05±0.46**	**92±0.29**	**92.41±0.32**	**91.84±0.48**	**91.43±0.16**	**90.85±0.35**
Ranking on the basis of RWC Index	**1**	**2**	**3**	**5**	**4**	**6**	**7**	**9**

RWC= (FW-DW) / (TW-DW); RWC Index (RWCI)= (RWC (stress)/RWC (control))*100; ±= standard error

Table 6.1 Relative water content based stress tolerance of 16 varieties of barley (cont).

	Beecher	Clipper	Skiff	CM72	Numar	Arivat	Gairdner	Franklin
Shoot length (cm) control	28.1±0.21	25.8±0.33	25.1±0.14	24.5±0.61	25.7±0.47	23.8±0.24	23.2±0.11	23.9±0.26
Shoot length (cm) stress	27.7±0.23	25.4±0.27	24.7±0.38	24.1±0.48	25.3±0.44	23.4±0.35	22.8±0.22	23.5±0.11
Root length (cm)control	16.8±0.34	17.1±0.37	15.4±0.49	15.2±0.28	15.1±0.36	14.3±0.57	14.1±0.28	14.5±0.29
Root length (cm) stress	16.4±0.08	16.8±0.12	15.1±0.34	14.8±0.37	14.7±0.28	14.0±0.13	13.8±0.34	14.1±0.37
Fresh weight (Control)	0.563±0.016	0.539±0.029	0.521±0.045	0.511±0.019	0.487±0.024	0.469±0.039	0.448±0.024	0.45±0.016
Dry weight (Control)	0.0582±0.033	0.0578±0.034	0.0576±0.052	0.0582±0.026	0.0582±0.026	0.058±0.048	0.0584±0.031	0.0586±0.019
Turgid Weight (Control)	0.658±0.048	0.662±0.039	0.664±0.032	0.663±0.034	0.669±0.031	0.67±0.017	0.665±0.033	0.661±0.036
RWC Control	84.17±0.026	79.64±0.019	75.41±0.022	74.86±0.029	70.21±0.017	67.17±0.026	64.21±0.047	64.97±0.041
Fresh weight (Stress)	0.505±0.045	0.47±0.047	0.446±0.013	0.437±0.007	0.416±0.044	0.38±0.005	0.354±0.006	0.35±0.041
Dry weight (Stress)	0.0613±0.017	0.0613±0.009	0.0613±0.027	0.0602±0.019	0.0612±0.042	0.0625±0.026	0.0629±0.022	0.0632±0.039
Turgid Weight (Stress)	0.64±0.018	0.649±0.021	0.651±0.014	0.65±0.025	0.66±0.036	0.674±0.028	0.679±0.049	0.679±0.048
RWC Stress	76.68±0.039	69.51±0.025	65.23±0.046	63.89±0.031	58.21±0.005	51.91±0.028	47.23±0.043	46.56±0.303
RWC Index	**91.10±0.47**	**87.28±0.36**	**86.50±0.33**	**85.34±0.27**	**82.90±0.29**	**77.28±0.72**	**73.55±0.14**	**71.66±0.42**
Ranking on the basis of RWC Index	**8**	**10**	**11**	**12**	**13**	**14**	**15**	**16**

RWC= (FW-DW) / (TW-DW); RWC Index (RWCI) = (RWC (stress)/RWC (control))*100; ±= standard error

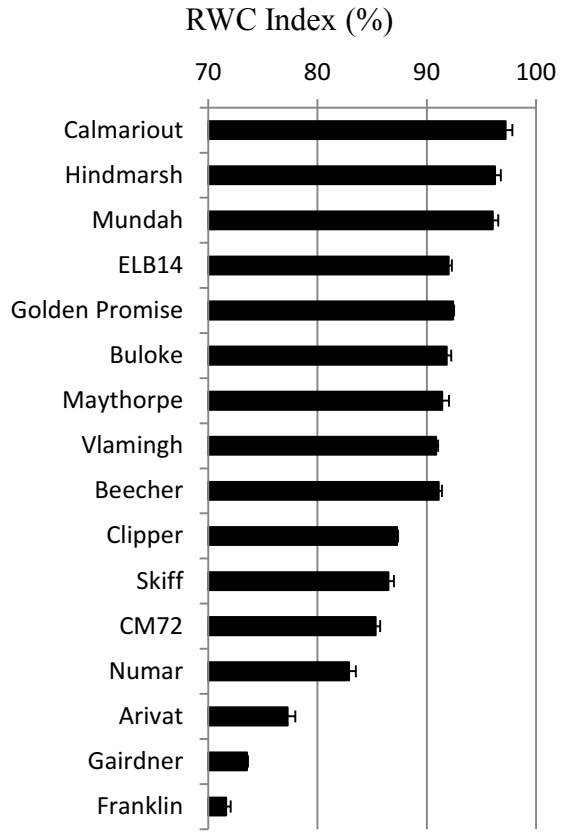


Figure 6.1 Effects of 12 h salt stress on shoot RWC% Index
Relative water content stress tolerance index was calculated as RWCI = (RWC stressed plant / RWC control plant) ×100**.**

6.2.2 Effects of salinity on Na^+/K^+ ratio in plant tissues

Uptake of [Na^+] and loss of [K^+] is another measure of salt sensitivity (Ahmed *et al*. 2013). Salt stress caused a significant (p <0.05) increase in [Na^+] in all varieties and an increase in [Na^+/K^+] ratio in almost all varieties, compared to respective controls. The [Na^+] ranged from 4.91-7.13 mg g^{-1} DW for root and 4.22-6.52 mg g^{-1} DW for shoot in control plants, increasing to 15.89-23.24 mg g^{-1} DW for root and 53.77-65.69 mg g^{-1} DW for shoot under salt stress, the accumulation under stress being more in shoots than in roots (Table 6.2; Figure 6.2). The maximum increase in both shoot and root [Na^+] was observed in Franklin (1274 ± 40 %; 397 ± 20%), followed by Gairdner (1255 ± 10%; 385 ± 15%), while the least increase was in Calmariout (898%±10; 222%±16). All other varieties also showed significant increase in [Na^+] in shoot (987-1235%) and root (243-378 %). In all varieties, the [K^+] decreased significantly (p<0.05) under salt stress, the 48.7-75.87 mg g^{-1} DW for shoot and 8.69-16.56 mg g^{-1} DW for root under control conditions reducing to 36.08-46.44 mg g^{-1} DW for shoot and 4.38-7.82 mg g^{-1}

DW for root (Table 6.2; Figure 6.3). The maximum decrease in shoot [K^+] was observed in Mundah (56.24%±2.3), whereas that in root [K^+] occurred in Numar (32.42%±0.8). The least decreases occurred in Beecher shoot (88.35%±1.2) and Calmariout root (90.03%±0.4). The effect on Na^+/K^+ ratio was more significant in shoots compared to roots, all varieties showing a significant increase (p <0.05) (Fig. 4). The highest ratio was in the controls of Calmariout (0.12±0.13 in shoot, 0.82±0.42 in root) while the lowest ratios were in Numar root (0.37±0.49) and in Franklin and Gairdner shoot (0.08). The highest Na^+/K^+ ratio index was thus recorded for Franklin for both root and shoot, while the lowest index was for Calmariout root and shoot (Table 6.2; Figure 6.4).

Table 6.2 Na^+/K^+ ion ratio based stress tolerance of 16 varieties of barley.

	Calmariout	Hindmarsh	Mundah	ELB14	Golden promise	Buloke	Maythorpe	Vlamingh
Na^+ control (root)	7.13±0.021	7.05±0.34	7.28±	7.09±0.45	6.98±0.35	6.82±0.14	6.95±0.22	6.81±0.23
Na^+ stress (root)	15.89±0.29	15.91±0.27	17.34±	17.23±0.12	17.88±0.33	19.69±0.08	21.50±0.47	21.30±0.12
Na^+ Index (root)	**222.89**	**225.77**	**238.19**	**243.14**	**256.21**	**288.74**	**309.45**	**312.87**
Na^+ control (shoot)	6.52±0.36	6.16±0.03	6.07±0.22	5.89±0.41	5.85±0.44	5.78±0.27	5.72±0.22	5.64±0.24
Na^+ stress (shoot)	58.60±0.39	57.02±0.14	57.61±0.07	58.15±0.78	61.24±0.25	62.33±0.39	63.62±0.36	65.69±0.15
Na^+ Index (shoot)	**898.91**	**925.72**	**949.12**	**987.35**	**1046.98**	**1078.48**	**1112.39**	**1164.77**
Na^+/K^+ control (root)	0.82±0.42	0.8±0.27	0.78±0.22	0.72±0.06	0.66±0.34	0.62±0.36	0.5±0.11	0.51±0.17
Na^+/K^+ stress (root)	2.03±0.36	2.12±0.18	2.33±0.49	2.85±0.47	3.21±0.28	3.66±0.21	3.75±0.18	3.89±0.43
Na^+/K^+ Index (root)	**247.56**	**265**	**298.71**	**395.83**	**486.36**	**590.32**	**750**	**762.74**
Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Na^+/K^+ control (shoot)	0.12±0.13	0.11±0.37	0.11±0.52	0.1±0.12	0.1±0.28	0.11±0.35	0.1±0.07	0.1±0.33
Na^+/K^+ stress (shoot)	1.41±0.33	1.39±0.08	1.46±0.50	1.38±0.17	1.42±0.31	1.48±0.48	1.47±0.38	1.48±0.14
Na^+/K^+ Index (shoot)	**1175.00**	**1263.63**	**1327.27**	**1380.00**	**1420.00**	**1345.45**	**1470.00**	**1480.00**
Rank	**1**	**2**	**3**	**5**	**6**	**4**	**7**	**8**
Average Rank	**1**	**2**	**3**	**4.5**	**5.5**	**5**	**7**	**8**
Rank	**1**	**2**	**3**	**4**	**6**	**5**	**7**	**8**

Na^+ ion Index = (Na^+ (stress)/Na^+ (control))*100; Na^+/K^+ ion Index = (Na^+/K^+ (stress)/Na^+/K^+ (control))*100; ±= standard error

Table 6.2 Na^+/K^+ ion ratio based stress tolerance of 16 varieties of barley (cont).

	Beecher	Clipper	Skiff	CM72	Numar	Arivat	Gairdner	Franklin
Na^+ control (root)	6.77±0.22	6.48±0.25	6.52±0.39	6.42±0.17	6.13±0.5	5.28±0.49	5.1±0.35	4.91±0.31
Na^+ stress (root)	22.42±0.38	22.10±0.07	22.70±0.48	23.24±0.15	23.20±0.48	20.14±0.02	19.68±0.26	19.50±0.15
Na^+ Index (root)	**331.24**	**341.12**	**348.25**	**362.1**	**378.54**	**381.45**	**385.91**	**397.21**
Na^+ control (shoot)	5.59±0.24	5.32±0.36	5.21±0.39	5.09±0.15	4.87±0.27	4.69±0.48	4.51±0.08	4.22±0.12
Na^+ stress (shoot)	65.55±0.38	63.03±0.37	61.87±0.27	61.76±0.26	60.16±0.17	58.59±0.07	56.62±0.43	53.77±0.44
Na^+ Index (shoot)	1172.65	1184.89	1187.67	1213.54	1235.37	1249.45	1255.56	1274.23
$Na^{+/}K^+$ control (root)	0.49±0.82	0.45±0.33	0.41±0.15	0.42±0.36	0.37±0.49	0.4±0.43	0.41±0.82	0.38±0.33
Na^+/K^+ stress (root)	4.17±0.12	4.19±0.21	4.28±0.26	4.33±0.48	4.32±0.08	4.35±0.16	4.39±0.35	4.45±0.42
Na^+/K^+ Index (root)	851.02	931.11	1043.90	1030.952	1167.56	1087.5	1070.73	1171.05
Rank	**9**	**10**	**12**	**11**	**13**	**15**	**14**	**16**
Na^+/K^+ control (shoot)	0.09±0.21	0.09±0.15	0.09±0.34	0.09±0.1	0.09±0.19	0.09±0.15	0.08±0.5	0.08±0.37
Na^+/K^+ stress (shoot)	1.35±0.13	1.36±0.45	1.37±0.09	1.44±0.16	1.46±0.27	1.47±0.49	1.35±0.35	1.49±0.19
Na^+/K^+ Index (shoot)	**1500**	**1511.11**	**1522.22**	**1600**	**1622.22**	**1633.33**	**1687.5**	**1862.5**
Rank	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Average Rank	**9**	**10**	**11.5**	**11.5**	**13**	**14.5**	**14.5**	**16**
Rank	**9**	**10**	**11**	**11**	**13**	**14**	**14**	**16**

Na^+ ion Index = (Na^+ (stress)/Na^+ (control))*100; Na^+/K^+ ion Index = (Na^+/K^+ (stress)/Na^+/K^+ (control))*100; ±= standard error

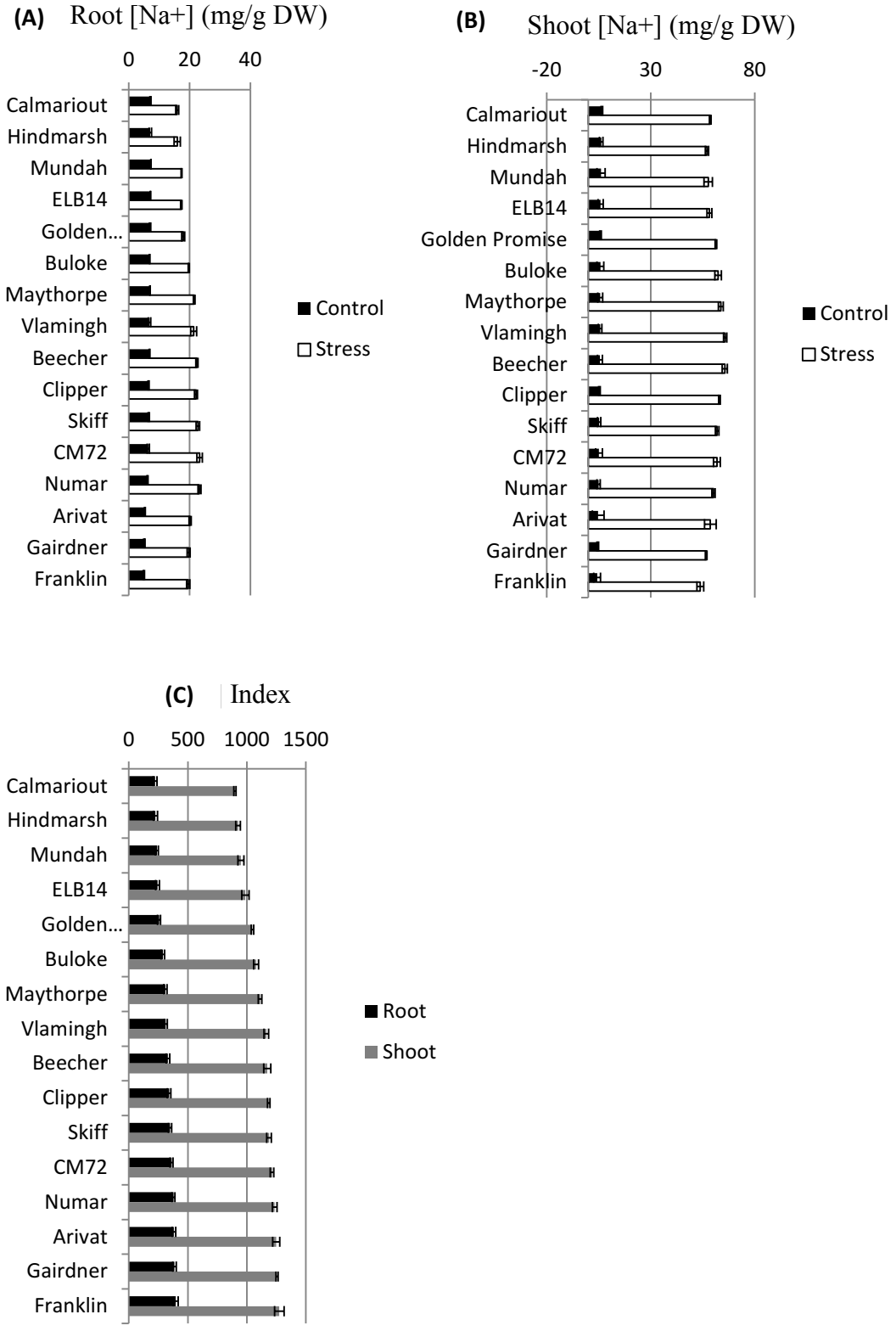


Figure 6.2 Effects of 12h salt stress on [Na^+] and [Na^+] Index

(A)Uptake of [Na^+] in roots under salt stress; (B) Uptake of [Na^+] in shoots under salt stress; (C) [Na^+] Index after salt treatment; Na+ ion stress tolerance index as NaI= (Na+ ion stressed plant / Na+ control plant) ×100.

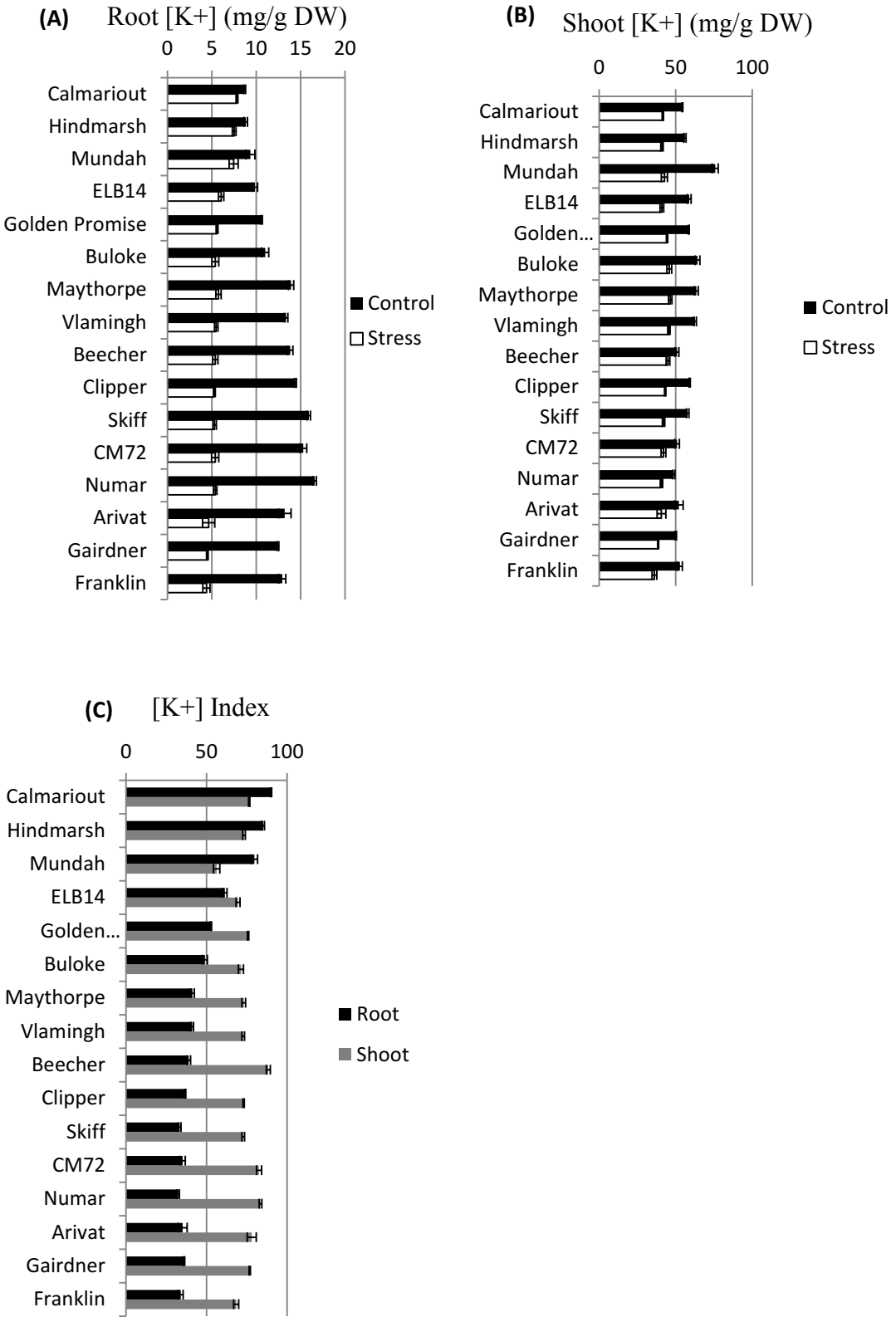


Figure 6.3 Effects of 12 h salt stress on [K^+] and [K+] Index

(A) Loss of [K+] in roots under salt stress (B) Loss of [K+] in shoots under salt stress (C) [K+] Index after salt treatment; K+ ion stress tolerance index as KI = (K+ ion stressed plant / K+ control plant) ×100; K+ ion stress tolerance index as KI = (K+ ion stressed plant / K+ control plant) ×100.

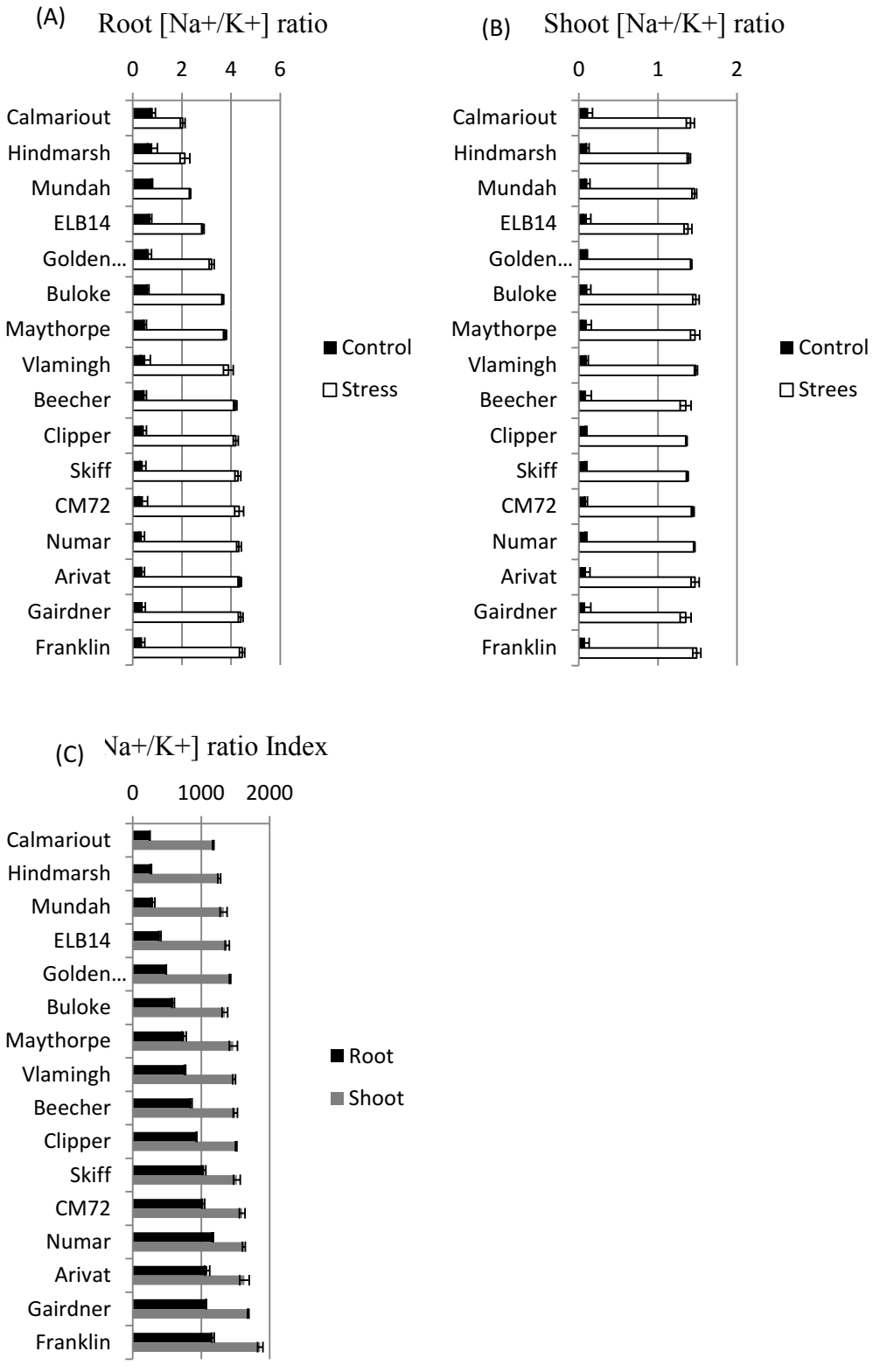


Figure 6.4 Effects of 12 h salt stress on Na^+/K^+ ion ratio Index

(A) N^+/K^+ ratio in roots under salt stress; (B) Na^+/K^+ ratio in shoots under salt stress; (C) Na^+/K^+ ratio Index after salt treatment; Na^+/K^+ ion stress tolerance index as Na/KI = (Na^+/K^+ ion ratio stressed plant / Na^+/K^+ ratio control plant) ×100.

6.2.3 Effects of salinity on induction of endogenous ABA

In order to understand the relationship between salinity stress and induction of endogenous ABA, the [ABA] amounts were measured by ELISA. Salt stress led to 1.5-2 fold (152-202 %) increase of [ABA] in shoot tissues, and much higher increases in roots (8 to 10 fold) (845-1055 %; Figure 6.5). The maximal increase in shoot [ABA] was in Calmariout (210 ± 9%) followed by Hindmarsh (209 ± 10 %), while the least increase was in Franklin and Gairdner (153 ± 20 %). The maximum increase in root [ABA] was also in Calmariout (1110 ± 10%) and the least in Franklin (841 ± 40%) (Table 6.3; Figure 6.5).

Table 6.3 Endogenous ABA concentration based stress tolerance of 16 varieties of barley.

	Calmariout	Hindmarsh	Mundah	ELB14	Golden promise	Buloke	Maythorpe	Vlamingh
ABA control (root)	3.45±0.34	3.39±0.15	3.42±0.49	3.28±0.26	2.21±0.11	3.17±0.35	3.15±0.25	3.01±0.56
ABA stress (root)	38.302±0.25	37.23±0.09	36.74±0.17	34.63±0.14	22.517±0.49	31.61±0.43	30.68±0.27	28.78±0.38
ABA Index (root)	**1110.36±10**	**1098.23±4**	**1074.52±13**	**1055.98±25**	**1018.87±27**	**997.45±33**	**974.22±32**	**956.25±21**
Rank	**1**	**2**	**3**	**4**	**5**	**6**	**8**	**9**
ABA control (shoot)	8.12±0.046	7.89±0.023	7.06±0.024	7.55±0.026	6.98±0.038	6.74±0.049	5.89±0.029	6.21±0.045
ABA stress (shoot)	17.12±0.022	16.55±0.045	14.28±0.038	15.64±0.014	14.08±0.024	13.43±0.055	11.58±0.026	12.26±0.034
ABA Index (shoot)	**210.94±9**	**209.84±10**	**202.35±8**	**207.21±12**	**201.84±32**	**199.29±47**	**196.72±9**	**197.46±16**
Rank	**1**	**2**	**4**	**3**	**5**	**6**	**7**	**8**
Average Rank	**1**	**2**	**3.5**	**3.5**	**5**	**6**	**7.5**	**8.5**
Rank	**1**	**2**	**3**	**3**	**5**	**6**	**7**	**9**

Endogenous ABA index (ABAI) = (ABA (stress) / ABA (control))*100; ±= standard error

Table 6.3 Endogenous ABA concentration based stress tolerance of 16 varieties of barley (cont).

	Beecher	Clipper	Skiff	CM72	Numar	Arivat	Gairdner	Franklin
ABA control (root)	2.89±0.22	2.82±0.37	3.77±0.09	3.73±0.33	3.79±0.45	3.66±0.63	3.58±0.22	3.55±0.13
ABA stress (root)	28.24±0.38	26.67±0.24	35.40±0.17	34.02±0.16	33.28±0.36	31.29±0.52	30.28±0.43	29.86±0.07
ABA Index (root)	**977.39±44**	**945.77±42**	**939.21±24**	**912.24±14**	**878.32±44**	**855.11±7**	**845.84±15**	**841.39±40**
Rank	**7**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
ABA control (shoot)	5.94±0.044	5.78±0.044	5.32±0.036	5.66±0.018	4.87±0.048	4.08±0.011	3.33±0.020	3.59±0.034
ABA stress (shoot)	11.63±0.023	10.19±0.032	9.26±0.015	9.93±0.036	8.37±0.046	6.70±0.026	5.08±0.017	5.51±0.049
ABA Index (shoot)	**195.84±12**	**176.31±25**	**174.24±5**	**175.59±4**	**171.87±3**	**164.22±11**	**152.73±9**	**153.65±20**
Rank	**9**	**10**	**12**	**11**	**13**	**14**	**16**	**15**
Average Rank	**8**	**10**	**11.5**	**11.5**	**13**	**14**	**15.5**	**15.5**
Rank	**8**	**10**	**11**	**11**	**13**	**14**	**15**	**15**

Endogenous ABA index (ABAI) = (ABA (stress) / ABA (control))*100; ±= standard error

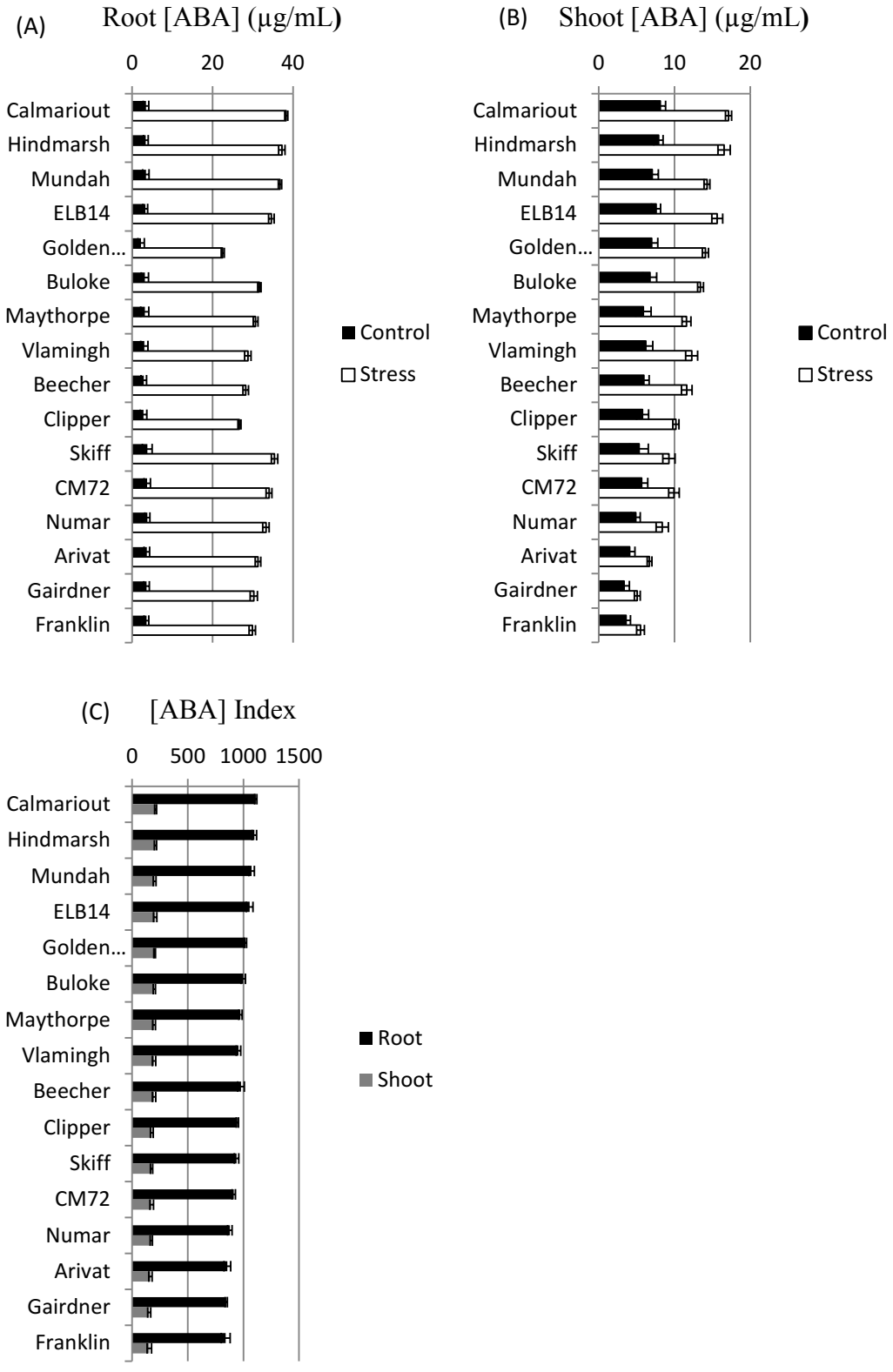


Figure 6.5 Effects of 12 h salt stress on endogenous [ABA] concentration Index
(A) Increase in [ABA] of roots under salt stress; (B) Increase in [ABA] of shoots under salt stress; (C) [ABA] Index after salt treatment; ABA accumulation stress tolerance index (ABAI) was calculated as ABAI = ([ABA] Stressed plant / [ABA] Control plant) ×100

6.2.4 Differential expression of key genes in ABA-mediated abiotic stress response pathway

Two representatives each of the four main gene families in the ABA pathway, i.e., two PYR/PYL/RCARs (putative cytoplasmic soluble receptors of ABA), the PP2CAs, SnRK2s subfamily III (downstream components of PYR/PYL/RCARs), and two ABA-transport related ABC transporters of the G subfamily, ABCG25 and ABCG40 (involved in import and export of ABA across the plasma membrane, respectively) (section 1.6), were analysed in this work, on the basis of their differential expression noted in mRNA-Seq data for the barley variety Hindmarsh (Section 4.2.11-14).

The PP2CAs Hv.19158/AK251854 and Hv.18841/MLOC_8131 showed significant up-regulation (i.e., fold change (FC) of ≥1.5, *p value* <0.05) under salinity in all varieties (Table 6.4), as noted earlier. Strong inductions were noted in Calmariout (Hv.19158/AK251854: FC 4.61±0.4; Hv.18841/MLOC_8131: FC 4.07±0.5), Gairdner (2.85±0.1; 2.12±0.3) and Arivat (3.12±0.5; 2.49±0.3), while the least change was in Franklin (2.71±0.2; 3.58±0.9). Both members (Hv.15651/AK361631 and Hv.9994/AK376521) of the PYR/RYL/RCAR family showed significant down-regulation under salt stress (FC ≤ -1.5, *p value* <0.05) in most varieties. The highest differential expression was in Calmariout (Hv.15651/AK361631: FC -6.13 ± 0.8; Hv.9994/AK376521: FC -2.08 ± 0.2), followed by Hindmarsh (-6.02 ± 0.1; -2.21 ± 0.9), with minimal changes detected in Franklin, Gairdner and Arivat (Table 6.4). The SnRK2s Subfamily III members, Hv.1875/MLOC_3013 and Hv.5014/MLOC_22145, were significantly down-regulated under salt stress in 13 and 8 varieties, respectively. The strongest down-regulations occurred in Calmariout (Hv.1875/MLOC_3013: FC -2.27±0.7; Hv.5014/MLOC_22145: FC -1.81±0.8) and the smallest changes in Gairdner (Table 6.4). Members of ABA- related ABC transporters, Hv.7752/MLOC_62985 (ABCG25) and NTC.9836/MLOC_68581 (ABCG40) showed significant up-regulation under salt stress in 14 and 13 varieties, respectively. The strongest differential expression occurred in Calmariout and Mundah (average FC +2.355) followed by Hindmarsh (Hv.7752/MLOC_62985: FC +2.56±0.3; NTC.9836/MLOC_68581: FC +2.14±0.8) and minimal changes again detected in Franklin, Gairdner and Arivat. The trends of up-regulation of ABC transporter type G members and PP2Cs, and down-

regulations of the PYR/RYL/RCARs and SnRK2s agreed with trends in the previous mRNA-Seq data (Section 4.2.11-15).

Table 6.4 Molecular based stress tolerance ranking of 16 barley varieties.

Gene	Family	FC_NGSI Hindmarsh	FC[a] Calmariout	FC[a] Hindmarsh	FC[a] Mundah	FC[a] ELB14	FC[a] Golden promise	FC[a] Buloke	FC[a] Maythorpe	FC[a] Vlamingh
Hv.19158/AK251854	PP2CA	4.60	4.61±0.35	4.53±0.21	4.32±0.45	3.91±0.19	3.87±0.11	3.81±0.48	3.75±0.17	3.74±0.39
Hv.18841/MLOC_8131	PP2CA	3.71	4.07±0.49	3.82±0.39	3.58±0.23	3.27±0.42	3.31±0.14	3.19±0.15	3.06±0.32	3.01±0.34
Average			**4.34**	**4.175**	**3.95**	**3.59**	**3.59**	**3.5**	**3.405**	**3.375**
Rank			**1**	**2**	**3**	**4**	**4**	**6**	**7**	**8**
Hv.15651/AK361631	PYR	-6.13	-6.02±0.78	-5.89±0.08	-5.91±0.41	-4.15±0.21	-4.13±0.57	-3.98±0.27	-3.87±0.35	-3.89±0.48
Hv.9994/AK376521	PYR	-2.08	-2.21±0.22	-2.17±0.85	-2.03±0.21	-1.88±0.31	-1.87±0.34	-1.85±0.14	-1.83±0.18	-1.88±0.21
Average			**-4.115**	**-4.03**	**-3.97**	**-3.015**	**-3**	**-2.915**	**-2.85**	**-2.885**
Rank			**1**	**2**	**3**	**4**	**5**	**6**	**8**	**7**
Hv.1875/MLOC_3013	SnRK2 subtype III	-1.57	-2.27±0.74	-2.06±0.13	-1.98±0.07	-1.76±0.38	-1.77±0.12	-1.79±0.28	-1.87±0.12	-1.76±0.62
Hv.5014/MLOC_22145	SnRK2 subtype III	-1.55	-1.81±0.79	-1.78±0.27	-1.75±0.19	-1.65±0.25	-1.66±0.09	-1.62±0.29	-1.56±0.33	-1.49±0.12
Average			**-2.04**	**-1.92**	**-1.865**	**-1.705**	**-1.715**	**-1.705**	**-1.715**	**-1.625**
Rank			**1**	**2**	**3**	**6**	**4**	**6**	**4**	**9**
Hv.7752/MLOC_62985	ABC transporter G25	2.07	2.54±0.28	2.56±0.33	2.57±0.28	2.31±0.46	2.17±0.18	2.04±0.11	2.05±0.004	2.02±0.22
NTC.9836/MLOC_68581	ABC transporter G40	2.53	2.17±0.81	2.14±0.49	2.14±0.12	1.94±0.19	1.91±0.23	1.82±0.39	1.78±0.59	1.76±0.07
Average			**2.355**	**2.35**	**2.355**	**2.125**	**2.04**	**1.93**	**1.915**	**1.89**
Rank			**1 +2**	**3**	**1+2**	**4**	**5**	**6**	**8**	**9**
Average of Ranks			**1**	**2**	**2.5**	**4**	**5**	**6**	**7.666667**	**8.25**
Molecular tolerance Rank			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**

Table 6.4 Molecular based stress tolerance ranking of 16 barley varieties (cont).

Gene	Family	FC[a] Beecher	FC[a] Clipper	FC[a] Skiff	FC[a] CM72	FC[a] Numar	FC[a] Arivat	FC[a] Gairdner	FC[a] Franklin
Hv.19158/AK251854	PP2CA	3.75±0.14	3.69±0.47	3.65±0.19	3.63±0.24	3.54±0.09	3.12±0.24	2.85±0.50	2.71±0.1 6
Hv.18841/MLOC_8131	PP2CA	2.98±0.22	2.92±0.38	2.78±0.25	2.86±0.16	2.71±0.46	2.49±0.19	2.12±0.12	2.01±0.94
Average		**3.365**	**3.305**	**3.215**	**3.245**	**3.125**	**2.805**	**2.485**	**2.36**
Rank		**9**	**10**	**12**	**11**	**13**	**14**	**15**	**16**
Hv.15651/AK361631	PYR	-3.86±0.38	-3.75±0.36	-3.71±0.18	-3.69±0.37	-3.58±0.59	-3.31±0.37	-3.13±0.34	-2.88±0.47
Hv.9994/AK376521	PYR	-1.81±0.07	-1.82±0.51	-1.83±0.47	-1.84±0.71	-1.77±0.02	-1.31±0.34	-1.24±0.36	-1.18±0.44
Average		**-2.835**	**-2.785**	**-2.77**	**-2.765**	**-2.675**	**-2.31**	**-2.185**	**-2.03**
Rank		**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Hv.1875/MLOC_3013	SnRK2 subtype III	-1.76±0.42	-1.69±0.09	-1.65±0.41	-1.58±0.23	-1.51±0.25	-1.32±0.48	-1.27±0.27	-1.31±0.25
Hv.5014/MLOC_22145	SnRK2 subtype III	-1.52±0.43	-1.41±0.51	-1.38±0.24	-1.34±0.06	-1.29±0.26	-1.21±0.41	-1.18±0.22	-1.16±0.18
Average		**-1.64**	**-1.55**	**-1.515**	**-1.46**	**-1.4**	**-1.265**	**-1.225**	**-1.235**
Rank		**8**	**10**	**11**	**12**	**13**	**14**	**16**	**15**
Hv.7752/MLOC_62985	ABC transporter G25	2.07±0.37	1.93±0.23	1.88±0.33	1.85±0.11	1.82±0.31	1.53±0.13	1.22±0.17	1.15±0.07
NTC.9836/MLOC_68581	ABC transporter G40	1.77±0.41	1.68±0.29	1.65±0.42	1.67±0.47	1.56±0.18	1.41±0.49	1.19±0.08	1.2±0.19
Average		**1.92**	**1.805**	**1.765**	**1.76**	**1.69**	**1.47**	**1.205**	**1.175**
Rank		**7**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Average of Ranks		**8.25**	**10**	**11.25**	**11.75**	**13**	**14**	**15.25**	**15.75**
Molecular tolerance Rank		**8**	**10**	**11**	**12**	**13**	**14**	**15**	**16**

a: FC (Fold Change) was calculated using ΔΔCt from RT-qPCR; ±= standard error

6.2.5 DNase I hypersensitivity site (DHS) analysis of selected ABA mediated stress signalling genes

Four barley varieties, two being salt tolerant (Calmariout and Hindmarsh) and other being salt sensitive (Gairdner and Franklin) were selected from the varietal analysis (Section 6.2.4) for DNase I hypersensitivity site (DHS) analysis of selected ABA mediated stress signalling genes. The barley plants were grown under the same conditions as those used for varietal analysis (detailed in section 2.3). After 14 days, three independent plants were treated with NaCl as done for varietal analysis (150 mM NaCl for 12 h), while three others remained untreated (controls). The chromatin isolated from the leaf samples of both the control and stressed plants was digested with 4 units of DNase I for 10 minutes (explained in section 2.9.1). The DNA was further extracted and used for qRT-PCR of the eight genes (primers in Table 2.8). Two other sequences, Bowman_contig_222590 and Bowman_contig_68397 were selected as control sequences as their respective orthologue in rice demonstrated non-DH sites (Zhang *et al*. 2012a). The threshold cycle number (Ct) was noted for both the DNase I digested and non-digested DNA and were used to calculate ΔCt using the formulae Ct (digested)-Ct (non-digested). The ΔΔCt was later calculated as ΔCt (stressed)–ΔCt (control) (Zhang *et al*. 2012a).

Two representatives each of the four main gene families in the ABA mediated signalling pathway, i.e., two PYR/PYL/RCARs, the PP2CAs, SnRK2s subfamily III, and two ABA-transport related ABC transporters of the G subfamily, ABCG25 and ABCG40, were analysed for presence of DNase I hypersensitivity sites in the gene (exon and intron region). The PP2CAs AK251854 and MLOC_8131 showed significant increase in DNase I digestion (i.e., ΔΔCt of ≥1.5, *p value* <0.05) under salinity in all four varieties (Table 6.5). Strong digestions were noted in Calmariout (AK251854: ΔΔCt 2.47±0.14; MLOC_8131: ΔΔCt 3.25±0.67) and Hindmarsh (2.26±0.42; 3.09±0.23), while the least change was in Franklin (1.66±0.36; 2.48±0.22). Of the two members of PYR/PYL/RCAR family, only one (MLOC_71349) showed significant decrease in DNase I digestion (i.e., ΔΔCt of ≥-1.5, *p value* <0.05), whereas other PYR/PYL/RCAR, AK376521 demonstrated non-significant ΔΔCt value for all four varieties. The highest decrease in DNase I digestion was in Calmariout (AK361631: ΔΔCt–2.39±0.22), followed by Hindmarsh (-2.18±0.31), with minimal

change detected in Franklin (Table 6.5). The SnRK2s subfamily III member, MLOC_22145 exhibited significant ΔΔCt under salt stress condition for all four varieties, whereas other SnRK2 subtype III exhibited non-significant ΔΔCt. ABA-related ABC transporters, MLOC_62985 (ABCG25) showed significant increase in DNase I digestion for all four varieties, whereas MLOC_685816 (ABCG40) exhibited non-significant ΔΔCt. The strongest differential DNase I digestion occurred in Calmariout (MLOC_62985: ΔΔCt 1.97±0.29) followed by Hindmarsh (1.84±0.17) and minimal changes again detected in Franklin and Gairdner. The control sequences, Bowman_contig_222590 and Bowman_contig_68397 demonstrated non-significant ΔΔCt under salt stress condition for all four varieties.

Table 6.5 DNase I hypersensitivity site analysis

Gene	Family	FC_NGSI Hindmarsh	ΔΔCt Calmariout	ΔΔCt Hindmarsh	ΔΔCt Gairdner	ΔΔCt Franklin
AK251854	PP2CA	4.60	2.47±0.14	2.26±0.42	1.83±0.79	1.66±0.36
MLOC_8131	PP2CA	3.71	3.25±0.67	3.09±0.23	2.56±0.41	2.48±0.22
MLOC_71349	PYR	-6.13	-2.39±0.22	-2.18±0.31	-2.09±0.16	-1.99±0.12
AK376521	PYR	-2.08	-1.03±0.17	-1.10±0.08	-1.07±0.28	-1.00±0.47
MLOC_3013	SnRK2 subtype III	-1.57	1.00±0.38	1.22±0.11	1.01±0.07	-1.14±0.33
MLOC_22145	SnRK2 subtype III	-1.55	2.54±0.49	1.77±0.52	1.56±0.42	1.49±0.53
MLOC_62985	ABC transporter G25	2.07	1.97±0.29	1.84±0.17	1.56±0.37	1.68±0.36
MLOC_68581	ABC transporter G40	2.53	-1.01±0.48	-1.17±0.13	1.12±0.28	-1.03±0.09
Bowman_contig_222590*	polygalacturonase)					
Bowman_contig_68397*	-					

* non-DH sites were selected as control (Zhang *et al*. 2012a); ±= standard error.

6.3 Discussion

6.3.1 Physiological analysis

With the current extent of soil salinity in many countries, and predictions of expansion (Wild 2003), coupled with increased demands on food production due to the growing world population, increasing salinity tolerance of cereal crops is a research priority worldwide. In the present study, 16 barley varieties were selected for screening of tolerance to the acute stress of 150 mM NaCl for 12 h. The application of physiological tests in conjunction with differential expression analysis of key genes of the ABA-mediated stress response pathway to evaluate salt tolerance is described here for the first time.

Changes to physiological characteristics such as RWC, Na^+/K^+ ion ratio and endogenous ABA levels are taken as key indicators of stress response of plants (Sinclair and Ludlow 1985; Ashraf 2004; Moons *et al*. 1995; Chen *et al*. 2007). Decrease in RWC can be related to difference in capability to absorb water from the soil and control its loss through stomata (Khakwani *et al*. 2011). In this study, RWC was found to decrease significantly under salt stress in all varieties. Tolerance rankings were then assigned on the basis of loss in RWC, i.e., RWC index (RWCI) (Table 6.5). Salt tolerant varieties were found to maintain a significantly higher RWC than the salt sensitive varieties of wheat, soybean and barley (Matin *et al*. 1989). Calmariout demonstrated minimum loss of RWC, hence it was classified as the most tolerant variety, whereas Franklin, with maximum loss, was classified as most sensitive. Efflux of K^+ from the root has been found to be inversely proportional to the grain yield of barley under salinity stress (Chen *et al*., 2007), and the ability of the root to retain K^+ has been shown to confer salinity tolerance in barley (Chen *et al*. 2005; 2007) and wheat (Cuin *et al*. 2011). According to Chen *et al*. (2007), barley cultivars, which show high efflux of K^+, yet considerable growth in response to salt stress may have the ability to restrict entry of Na^+ into shoot. The tissue Na^+/K^+ ion ratio is considered to be a critical parameter for tolerance, a lower Na^+/K^+ ratio indicating better salinity tolerance, and the parameter is studied extensively at physiological and genetic levels (Chen *et al*. 2007; Ahmed *et al*. 2013). Calmariout demonstrated the lowest root and shoot Na^+/K^+ ratio, suggesting it to be most tolerant, while Franklin had the highest Na^+/K^+ ratio, suggesting it to be most sensitive. Similar results have been reported for cotton, barley

and wheat, wherein the salt tolerant varieties maintained a lower Na^+/K^+ ratio than sensitive ones (Ashraf 2004). K^+ retention under salt stress is thus an important physiological trait, the K^+ content being suggested to be one of the 'master switches', enabling the plant to transit from normal metabolism to 'hibernated state' during abiotic or biotic stress (Shabala and Pottosin 2014). Based on these reports, Calmariout, which demonstrated the maximum K^+ retention in roots, appears to be most tolerant, while Numar had the minimum K^+ retention, suggesting it to be sensitive. Further, the results of varietal ranking based on the Na^+/K^+ ratio were consistent with those for RWCI (Table 6.2). Increases in endogenous ABA and ABA-induced proteins are considered a characteristic of tolerant varieties (Moons *et al*. 1995), e.g., higher ABA concentrations are noted in the tolerant compared to sensitive varieties of cucumber, *Suaeda maritime,* tobacco and rice (Moons *et al*. 1995). The present observations of the ABA concentration under control conditions being higher in shoot tissue than root in all varieties, and the trend being reversed under salinity stress, are both consistent with those in maize (Jia *et al*. 2002). Higher ABA index in roots as compared to shoot suggest that induced ABA accumulation is differently triggered in root and shoot tissues. Roots may sense the salt in environment sensitively, with or without reduction in cellular volume, and trigger their ABA production by 'osmosensing mechanism', whereas the shoot tissue may produce ABA in response to salt stress but it is triggered by 'dehydration mechanism' (Jia *et al*. 2002). Rankings assigned on the basis of ABA Index (ABAI) suggested Calmariout to be most tolerant and Franklin to be most sensitive (Table 6.5).

6.3.2 Molecular analysis

The Arabidopsis *pyr/pyl/rcar* and *abcg40* (Umezawa *et al*. 2010), mutants exhibit ABA insensitivity in seed germination, root growth and ABA-regulated gene expression, demonstrating the role of the encoded PYR/PYL/RCAR and ABCG40 in stress tolerance. Our previous studies using mRNA-Seq identified transcripts from the barley leaf, with altered expression under salt stress, potentially conferring stress resistance. Two genes from each of the four key gene families involved in the ABA-dependent abiotic stress signalling pathway, i.e., the PYR/PYL/RCAR, PP2CA, SnRK2 subfamily III and ABA-transport related ABC transporter (ABCG25 and ABCG40), were selected. The direction of change of expression for all eight genes under salt stress supports that

of mRNA-Seq data (Section 4.2.11-14). The positive differential expression of barley PP2CA genes under salt stress supports the result in rice under salinity and exogenous ABA application (Xue *et al*. 2008) and in Arabidopsis and *B. oleracea* under drought and exogenous ABA (Ludwikow *et al*. 2013). The ABCG25 and ABCG40 genes found up-regulated here are also induced in Arabidopsis (Seki *et al*. 2002) and rice and wheat (Zhou *et al*. 2007) under salt and/or drought stress. The PP2CAs show stronger induction under water stress in a drought-tolerant variety of chickpea as compared to a sensitive variety (Jain and Chattopadhyay 2010). The SnRK2s subtype III exhibited down-regulation, in line with those in Arabidopsis and rice (Umezawa *et al*. 2010). There are no reports to date on studies of expression of these key genes together with physiological metrics. Thus we propose that the varieties having significant modulations of expression of these key genes (irrespective of the direction of change), could be tentatively considered as most tolerant. An overall tolerance ranking by 'molecular method' was then applied by taking the average of rankings assigned based on each family (explained in materials and methods).

The highest differential expression for all genes was observed for Calmariout, suggesting that it may be most tolerant among the tested varieties. Hindmarsh and Mundah both attained rank two and three, respectively in the 'molecular method', on the basis of their response in terms of differential expression of studied genes. ELB14 and Golden Promise both ranked 4 for PP2CA. For PYR/PYL/RCAR and ABC transporters, ELB14 showed more response than Golden Promise, whereas this reversed for SnRK2; hence ELB14 was assigned rank 4 and Golden Promise ranked 5. By a similar consideration, Buloke and Maythorpe ranked 6 and 7. Vlamingh and Beecher were assigned rank 8, Clipper was ranked 10 for all gene families. Skiff, CM72, Numar and Arivat ranked 11, 12, 13 and 14 respectively, based on overall responses. For PP2CA, PYR/PYL/RCAR and ABA-transport related ABC transporters (ABCG25 and ABCG40), Franklin showed more sensitivity as compared to Gairdner, whereas this reversed for SnRK2 subtype III*;* hence Franklin was assigned rank 16 and Gairdner as 15. The results suggest that the varietal differences in physiological parameters are the result of varied genetic ability to absorb water in the root zone and increase water reserves, and also indicate the potential of Calmariout, Hindmarsh and Mundah as sources of tolerance-related alleles. The results of Gairdner and Franklin being found to

be salt-sensitive and Numar and CM72 being more tolerant confirm previous reports (Chen *et al*. 2007; Zhou *et al*. 2008). The finding of Golden Promise being more tolerant than Maythorpe also confirms the report of Walia *et al*. (2007), and Calmariout, ELB14 and Beecher being more tolerant than Clipper and Arivat confirm the results of Munns *et al*. (1995). However, the result of CM72 being more sensitive compared to Golden Promise does not agree with the finding of Adem *et al*. (2014) and Chen *et al*. (2007). This discrepancy may be due to differences in the stress conditions; in the Adem *et al*. (2014) report, the plants were subjected to chronic stress at a lower salt concentration (100 mM NaCl for 8 days), whereas in the present work the plants were subjected to acute stress (150 mM NaCl for 12 hours). This suggests a need for longer term studies at both physiological and genetic levels for better intervarietal comparisons. Franklin and Gairdner are high yield malt varieties, whereas Calmariout, Hindmarsh and Mundah are animal feed varieties (Smith 2006; Garstang *et al*. 2011). Gairdner and Franklin are late maturing, whereas Calmariout, Hindmarsh and Mundah are considered very early maturing (DEPI 2011; Smith *et al*. 2006; Wiebe and Reid 1961). These characteristics should be considered together with their salinity tolerance for assessing their suitability to particular locales.

Table 6.6 Comparison of ranks assigned to the barley varieties based on different traits.

Barley variety	Ranking basis			
	RWC	Na^+/K^+	ABA	Molecular
Calmariout	1	1	1	1
Hindmarsh	2	2	2	2
Mundah	3	3	3	3
ELB14	5	4	3	4
Golden Promise	4	6	5	5
Buloke	6	5	6	6
Maythorpe	7	7	7	7
Vlamigh	9	8	9	8
Beecher	8	9	8	8
Clipper	10	10	10	10
Skiff	11	11	11	11
CM72	12	11	11	12
Numar	13	13	13	13
Arivat	14	14	14	14
Gairdner	15	14	15	15
Franklin	16	16	15	16

6.3.3 DNase I hypersenstivity site analysis

Nucleosome packaging, which is organization of DNA structure acts as the regulator of transcription by permitting or restricting protein binding and therefore enabling replication and co-ordination of gene activity. The chromatin accessibility, which has been classified traditionally as regions of "open" or "closed" conformation, is subject to changing events at accessible *cis*-regulatory element (Madrigal and Krajewski 2012). "Open" chromatin regions are associated with active chromatin and can be identified by their hypersensitivity to nuclease digestion (Ling and Waxman 2013). The open chromatin or nucleosome loss in eukaryotic genome is an important factor in elucidating potential regulatory activity (Zhang *et al*. 2012a). DNase I hypersensitivity sites can be identified by DNase I hypersentivity assay (DHS), in which isolated nuclei is subjected to regulated digestion with enzyme DNase I (Ling and Waxman 2012). The standard approach for locating the DHS involves southern blotting and high throughput sequencing. An alternative approach, permitting the targeted DHS mapping of specific loci can be carried out using quantitative real time PCR (qRT-PCR). This technique depends on the quantification of relative PCR signal observed when PCR primers amplify across regions of digested DNA compared with amplification of undigested DNA (Follows *et al*. 2007).

Of the eight ABA mediated signalling genes, only five demonstrated DHS in the gene. Both of the PP2CAs (AK251854 and MLOC_8131), and one each of SnRK2 subtype III (MLOC_22145) and ABCG25 (MLOC_62985) showed significant DNase I digestion under salt stress condition. On the other hand PYR/PYL/RCAR, MLOC_71349 exhibited significant decrease in DNase I digestion under salt stress condition. According to Zhang *et al* (2012a), DHS sites are specific to the tissue and developmental stage, which support the absence of DHS in other three genes and these results are in line with findings of rice and Arabidopsis, as only 27 % and 15 % of DHS were found in the genes (exon and intron) of rice and Arabidopsis respectively (Zhang *et al*. 2012a, b). The analysis of gene expression and DNase I digestion of the genes demonstrated that both PP2CAs and ABCG25, which were up-regulated under salt stress also showed increased digestion with DNase I. On the other hand the down regulated PYR/PYL/RCAR (MLOC_71349) exhibited decrease in DNase I digestion under salt stress condition. These findings are explained by literature as DHS marks

almost all regulatory elements such as promoters, enhancers, suppressors, insulators, and locus control regions (Zhang *et al*. 2012a) and are consistent with reports in humans, rice and Arabidopsis (Lister *et al*. 2009; Zhang *et al*. 2012a,b). On the other hand, one of the SnRK2 subtype III gene, MLOC_22145 was found to be down regulated under salt stress, but it has shown increase in DNase I digestion under similar stress condition. This is in line with DHS results in rice, which suggest that many of the down-regulated genes are repressed by regulatory elements (Zhang *et al*. 2012a).

These preliminary DHS results suggest that differentially expressed genes are associated with DHS and is consistent with the hypothesis that gene under complex pattern of regulation require regulatory elements and therefore region of "open" chromatin. For this preliminary DNase I hypersenstivity analysis full length PCR products were used for qRT-PCR. The use of PCR products longer than >200bp is a limitation of qRT-PCR and so these results need to be verified using other techniques. The finding of DNase I hypersentivity sites in the region upto 200bp upstream of transcription start site (TSS) in humans, rice and Arabidopsis suggest that a genome wide mapping of DHS using high throughput sequencing is required for better understanding of regulated gene expression.

The same three varieties, Calmariout, Hindmarsh and Mundah were identified to be salt-tolerant by the physiological and molecular ranking; hence these would be excellent candidates for expansion of cultivation as well as identification of alleles related to tolerance. Further, a simple but effective tolerance ranking system was developed, consolidating the analyses of key physiological parameters and differential expression studies of selected genes. This methodology can serve as a rapid preliminary tool for screening of germplasm collections for potentially salt tolerant lines before undertaking time-consuming physiological studies.

CHAPTER 7

General discussion and conclusion

7.1 Summary of outcomes

Globally, barley is the fourth most-produced cereal crop. In Australia alone, based on quantity produced, area cultivated and revenue generated, barley is one of the most important grain crops. Further, barley also has very important nutritive value for animals and humans. Given its economic and nutritive value, it is critical to keep this crop both abundant and affordable. However, the yields of barley are frequently challenged by environmental stresses such as salinity and drought. The magnitude of losses in crop yields is further exacerbated by climate change. Over and above that, the rapidly expanding human population places additional pressure on commercial farmers to ensure adequate food supply. Therefore, extension of cultivation into areas that would not normally be cultivated would aid in easing the pressure on available land, increase production of food crops and help to keep these grains both plentiful and inexpensive. This could be achievable through the development of more resilient crops. One strong step towards this objective is to enhance the knowledge base regarding the genetics of stress tolerance in plants. Multiple genes have been implicated in these processes, of which the current study focused on abscisic acid (ABA) mediated abiotic stress signalling mediated genes.

The first chapter presented a comprehensive review of literature pertaining to ABA synthesis and stress signalling in plants. ABA is a phytohormone crucial for plant's adaptive responses to abiotic stresses such as drought and salinity. The events leading up the first discovery of ABA (abscisin II) from young cotton fruit were detailed. Subsequently, ABA was also reported from other plants and humans. Since the discovery of the chemical structure of ABA, steady progress has been made in explaining its metabolism, biosynthesis and genes responsive to it. On the other hand, the early stage of the ABA perception by the plant cells was a puzzle until the pyrabactin resistance (PYR)/ PYR like (PYL)/ regulatory component of ABA receptor (RCAR) was identified as soluble ABA receptor (Ma *et al*. 2009; Park *et al*. 2009). This discovery was a major breakthrough in ABA signalling studies as molecular interactions of the PYR/PYL/RACR family of soluble protein in Arabidopsis with already known ABA signalling factors such as protein phosphatase 2C (PP2C) was established. The double-negative regulatory system of ABA signalling is comprised of PYR/PYL/RCAR (ABA receptors), the enzymes of the class PP2C and SNF1-related

kinase 2 (SnRK2), transcription factors such as WRKY, MYB, MYC, NAC and their downstream target genes such as salt overly sensitive (SOS) and late embryogenesis abundant (LEA) among other (explained in section 1.6). Chapter 1 also explored the literature on role and significance of the transport of ABA between cells, organs and tissues. Cell to cell ABA transport was clarified by the discovery of two specific plasma membrane bound ABA transporters, ATP binding cassette (ABC) named ABCGG25 and ABCG40 in Arabidopsis (Kang *et al*. 2010; Kuromori *et al*. 2010). The mechanism of signalling response mediated by ABA during abiotic stress condition was also reviewed (explained in section 1.7).

The review of literature in Chapter 1 showed that in recent years, research into the ABA mediated stress signalling pathway has attracted the attention of a vast number of researchers. The analysis of literature indicated that research on barley and wheat ABA mediated abiotic stress signalling related candidate genes is lagging behind, possibly due to the complexity of their genomes and the lack of sequenced genomes until the release of the barley genome sequences in 2012). Thus, the major aim of the current study was to characterise these genes from barley as a major cereal crop and as a model for wheat. This major aim was addressed through the specific aims outlined below:

1. To analyse the transcriptome of barley (*Hordeum vulgare* cv. Hindmarsh) leaf under control and salt, drought and exogenous ABA stresses using mRNA-seq.
2. To develop bioinformatics based tools to identify specific gene families of interest in the mRNA-seq data or any other database using orthologues from related species and extract the ABA mediated abiotic stress signalling related genes.
3. To investigate the expression modulation of the genes involved in ABA mediated signalling pathway under the above abiotic stresses.
4. To investigate the epigenetic mechanism of histone modifications in regulating ABA mediated abiotic stress signalling genes.
5. To explore salinity tolerance of selected barley varieties employing analyses of key physiological parameters.
6. To test the expression of selected genes involved in ABA mediated abiotic stress signalling in above varieties.

The above aims were addressed using a combination of bioinformatics and experimental techniques as outlined in Chapter 2. The methods used included molecular techniques such as next-generation mRNA-seq, chromatin immunoprecipitation (ChIP)-seq, DNase I hypersentivity site anlaysis, chromatin, DNA and RNA extraction, cDNA synthesis, quantitative real-time PCR, standard PCR, cloning, and DNA sequencing and bioinformatics techniques such as primer design, sequence alignments of putative DNA, protein and sequence consensus generation, use of published tools such as SAMtools, DESeq etc. and self-developed scripts in UNIX such as reciprocal BLAST search for the sequencing data analysis.

The findings addressing the first aim were presented in Chapter 3. Chapter 3 sought to investigate the genes, which may confer resistance to acute salinity, drought and exogenous ABA in barley were carried out using the next generation mRNA-seq of total RNA from control, salt, drought and exogenous ABA stressed plants. The identified genes were analysed for their gene ontologies to determine descriptions of gene products from the available rice database. Semi quantitative reverse transcriptase PCR (sqRT PCR) was also used to validate the differential expression data for the genes generated by mRNA-seq. These analyses revealed that: i) over seven thousand genes demonstrated differential expression in all three stress conditions; ii) around 90% of the genes showing differential expression for salinity or drought were also differentially expressed in exogenous ABA stressed plants; iii) genes such as chlorophyll a/b binding protein, aquaporins and MYB transcription factors were among the most differentially regulated genes during salt stress, iv) genes such as serine threonine kinase, chlorophyll a/b binding protein and MYB transcription factors were among the most differentially regulated genes during drought stress, v) genes such as DUF581 domain containing protein and chlorophyll a/b binding protein were among the most differentially regulated genes during exogenous ABA stress, vi) analysis of gene ontology demonstrated that maximum genes belong to cell and cell part for cellular component in all three abiotic stresses studied, vii) of the various functions maximum genes showed binding as molecular function in all three stress conditions, viii) maximum no of genes showed metabolic process in salt, drought and exogenous stresses. The comparison of mRNA-seq and sqRT-PCR or qRT-PCR differential expression data demonstrated general agreement between the expression changes observed by these techniques. The

mRNA-seq workflow was shared with a then fellow PhD student and was published (Ziemann *et al*. 2013a).

The results to address aims 2 and 3 were reported in Chapter 4. The identification of barley ABA signalling related candidate genes was carried out using data obtained from the sequencing of leaf RNA from control and salt-stressed plants using next-generation mRNA-seq and in other silico methods. Several methods attempted to extract genes of interest from the mRNA-seq dataset, as the International Barley Genome Sequencing Consortium (IBSC) database was not available at the start of the project, and the purpose was to generate a transcriptome for a non-model organism. Reciprocal BLAST search (RBH), an automated reciprocal BLAST pipeline tool was developed enabling the rapid identification of specific gene families of interest in related species, streamlining the collection of homologs prior to downstream molecular evolutionary analysis (Ziemann *et al.* 2013b).

Several reiterative searches of the mRNA-seq data and IBSC CDS database resulted in the identification of twenty three PP2CAs, eleven PP2CDs, twelve PYR/PYL/RCARs, seven SnRK2 (subfamily II and III) and six ABA transport related ABC transporters. The identified sequences were analysed on the basis of certain functional motifs and key residues in the putative proteins, which resulted in 13 PP2CAs, 9 PP2CDs, 10 PYR/PYL/RCARs, 5 SnRK2 (subfamily II and III) and 4 ABA transport related ABC transporters. The amino acid involved in the binding of PYR/PYL/RCAR and PP2C were found to be strictly conserved among all members of PP2CA and PYR/PYL/RCAR families in rice and barley. The total number of genes identified is a significant advancement on previous studies (Section 4.2). The consensus for the characteristic motifs of the ABA signalling related genes was also generated for rice and barley. The thorough appraisal of the barley ABA signalling related genes provided in this chapter is highly significant for further studies in barley and also gives a strong framework for studies in wheat, an evolutionarily-related and among the world's largest cereal. In addition, chapter 4 focussed on the investigation of the expression patterns of the barley ABA signalling related genes in response to different treatments using mRNA-seq. Using mRNA-seq, the current study is likely the first to report the genome-wide expression of leaf barley ABA signalling related genes in response to salt, drought

and exogenous ABA application. The results showed that five PP2CAs, three PYR/PYL/RCARs, two SnRK2s and one each of ABCG25 and ABCG40 were differentially expressed. The expression of other genes such as WRKY, slow anion channel (SLAC1) and MYB transcription factors was also modulated. The in-depth expression analysis of genes involved in ABA mediated abiotic stress signalling make it possible to identify candidates for future functional testing and development of tolerant plants through genetic modification or breeding. A manuscript that focuses on the analysis of ABA signalling related genes from the mRNA-seq is in preparation (Kamboj *et al*. unpublished).

The findings addressing the aim four were presented in Chapter 5. Chapter 5 sought to investigate the distribution of H3K4me3, both within the genes and on adjacent regulatory elements of the barley genes involved in ABA mediated signalling in order to gain a picture of transcriptional regulation under abiotic stress conditions in plants. The methylated or un-methylated H3K4 regions were initially identified using chromatin immuno-precipitation sequencing (ChIP-seq) of barley leaf under salt and drought stresses. qRT PCR was also used to validate the differential expression data for the genes generated by mRNA-seq. ChIP-Seq resulted in the identification of 19,015 and 19,005 sequences exhibiting differential H3K4 trimethylation under salt and drought stresses, respectively. Genes such as serine threonine kinase and leucine zipper protein were found to show differential H3K4 trimethylation under both salt and drought stresses. The comparison of ChIP-seq and qRT-PCR differential H3K4 trimethylation data demonstrated general agreement between the trimethylation observed by these techniques. Searches of the ChIP-seq data resulted in the identification of H3K4me3 within the genes or on adjacent regulatory elements of three PP2CA, four PP2CD, two PYR/PYL/RCAR, three SnRK2 and two ABC transporters (ABCG25 and ABCG40). Generally, PP2CAs, PP2CDs, PYR/PYL/RCARs, SnRK2s and ABCG40 were found to be un-methylated at H3K4 during abiotic stress, whereas ABCG25 was tri-methylated at H3K4. 60 % (21 of the 35) of genes involved in ABA mediated abiotic stress signalling were expressed in salinity and drought stresses without significant level of H3K4 trimethylation, which suggest that other epigenetic modifications, such as histone acetylation or cystosine methylation along with modification at enhancers may be involved as the plants are exposed to abiotic stress condition and these modifications

need to be evaluated in the future work with respect to expression of ABA signalling related genes. The understanding of relation between the epigenetic modifications and plant's response to environmental stress response is highly desirable, as it is required not only for better understanding molecular mechanism of plant's stress response but also for possible application in the genetic manipulation in plants.

Aims 5 and 6 and their findings were detailed in chapter 6. In brief, chapter 6 focussed on the identification of salt tolerant barley varieties using physiological and molecular analysis. Using the physiological responses including relative water content, endogenous ABA accumulation and Na^+/K^+ ratio, the current study analysed 16 barley varieties for their salinity tolerance. The results i) significant variation among varieties was observed in these physiological measures; ii) a salinity-tolerance ranking system was developed, consolidating the analyses of all key physiological parameters; iii) Calmariout was found to be the most tolerant variety among those tested, whereas Franklin was the most sensitive variety. The use of qRT-PCR of leaf cDNA for expression analysis of key genes in the abiotic stress tolerance pathway mediated by ABA, i.e., PYR/PYL/RCAR, PP2C, SnRK2 and ABC transporters lead to results i) highest differential expression for all the studied genes were observed for the variety Calmariout, suggesting that it may be more tolerant for salinity stress than the other varieties studied; ii) overall salinity tolerance ranking by 'molecular method' was calculated by taking the average of rankings assigned as per transcriptional responses of each gene family. In addition, gene expression analysis of genes in the ABA-dependent stress tolerance pathway appeared to be a good predictor of salinity tolerance, indicating it can serve as the first tool in identifying potentially salt-tolerant varieties before undertaking the time-consuming physiological studies. Three separate physiological experiments and differential gene expression analysis of genes in the ABA-dependent stress tolerance pathway led to identification of the same three salt-tolerant barley varieties (Calmariout, Hindmarsh and Mundah), which would be excellent candidates for expansion of cultivation as well as for identification of genes/alleles related to salt tolerance. This work has been submitted and is currently under minor revisions. Four barley varieties, two being salt tolerant (Calmariout and Hindmarsh) and other being salt sensitive (Gairdner and Franklin) were selected from the varietal analysis for DNase I hypersensitivity site (DHS) analysis of selected ABA mediated abiotic stress

signalling genes using qRT-PCR. The results i) of the eight genes studied, only five demonstrated DHS in the gene (start to stop codon); ii) both of the PP2CAs (AK251854 and MLOC_8131), and one each of SnRK2 subtype III (MLOC_22145) and ABCG25 (MLOC_62985) showed significant DNase I digestion under salt stress condition; iii) PYR/PYL/RCAR, MLOC_71349 exhibited significant decrease in DNase I digestion. The analysis of differential expression of these genes and preliminary DHS results suggest that differentially expressed genes are associated with DHS and is consistent with the hypothesis that gene under complex pattern of regulation require regulatory elements and therefore region of "open" chromatin.

In conclusion, the findings from the current study demonstrated that: i) barley PP2CA, PP2CD, PYR/PYL/RCAR, SnRK2 (subfamily II and III) and ABA transfer related ABC transporters form large families currently consisting of thirteen, nine, ten, five and four isoforms respectively; ii) five PP2CAs, three PYR/PYL/RCARs, two SnRK2s and one each of ABCG25 and ABCG40 were differentially expressed in the leaf during abiotic stress conditions; iii) the barley ABA signalling related genes may be regulated by exposure to salinity, drought and other abiotic stresses; iv) mRNA-seq is a robust technology for expression analysis and can also be employed for gene identification in non-model plants which lack sequenced genomes; v) RBH software will satisfy a need for an automated pipeline to rapidly mine gene families of interest from a database; vi) three PP2CAs, four PP2CDs, two PYR/PYL/RCARs, three SnRK2s and one each of ABCG25 and ABCG40 were differentially H3K4 trimethylated in the leaf during abiotic stress conditions; vii) gene expression analysis of genes in the ABA mediated stress tolerance pathway appeared to be a good predictor of salinity tolerance; viii) considering the number of ABA signalling related genes in rice and Arabidopsis, more genes may yet be identified in barley; ix) other important candidate genes for future crop improvement could be transcription factors such as MYB and bZIP found among top 20 up or down regulated genes of mRNA-seq data.

7.2 Future directions

Previous studies seeking to understand the mechanism of ABA mediated stress response have mainly focussed on the identification and interaction of different members of the pathway in Arabidopsis (Section 1.6). The current study resulted in the identification of

candidate genes involved in ABA mediated abiotic stress signalling in cereal crop (barley) and analysis of their expression in the abiotic stresses. This leads to the development of method to employ expression analysis of selected genes in the prediction of salinity tolerance of different barley varieties. These findings provide a further understanding of the molecular mechanisms involved in abiotic stress response. Sequencing of these genes in tolerant and sensitive varieties may provide information on allelic differences, which are potentially related to tolerance.

To gain further insights into the functions, regulation of the barley ABA mediated stress signalling genes and genetic variations; the following future directions are suggested:

- Several transgenic expression studies of downstream genes regulated by ABA mediated abiotic stress signalling such as aquaporins from different plants have yielded desirable traits in response to environmental stresses (Hanba *et al.* 2004; Ayadi *et al.* 2011). Thus, it would be highly informative to carry out transgenic or CSIP gene deletion studies using the barley ABA mediated abiotic stress signalling genes especially PP2CAs and PYR/PYL/RCARs which exhibited the greatest magnitude of change in response to abiotic stresses.
- In this study, we focussed on understanding how differences in the distribution of H3K4me3 impact gene expression, however other functional properties such as other histone modifications, DNA methylations, promoter usage, alternative splicing, antisense transcription and replication timing should also be considered. Many new insights into the biology of gene expression will come as we continue to map more chromatin modifications and uncover new mechanisms that influence transcription using techniques such as ChIP-Seq and DNase I-hypersensitivity assays.
- Numerous *cis*-regulatory elements associated with abiotic stress responsiveness and phytohormone response were identified for barley ABA signalling related genes (Seiler *et al*. 2007). However, these need to be tested functionally under abiotic stress and between varieties using methods such as ChIP-Seq or DNase I-hypersensitivity assays.
- Use of the list of barley ABA signalling related genes for mining of wheat genes from databases such as CerealsDB database (http://www.cerealsdb.uk.net/).

These findings help to provide a strong framework for studies in the evolutionarily-related wheat, which is economically most important cereal crop.

- Single nucleotide polymorphisms (SNPs) account for a large portion of the variations of plants species/varieties and may be related to plant development and stress tolerance. Over the years SNPs have gained momentum as molecular markers for selection in breeding (Mammadov *et al.* 2012). Thus, it will be highly relevant to use high-throughput sequencing technologies to identify the occurrence of SNPs in tolerant and sensitive barley lines.
- Farnesylation, phosphorylation and poly(ADP-ribosyl)ation are the major post translational protein modification, which function to regulate different growth and developmental processes. The role of fernesylated, phosphorylated and poly(ADP-ribosyl)ated proteins in abiotic stress response is known for a long time (Pei *et al*. 1998; Furihata *et al*. 2006). Thus ABA induced post-translational modifications of proteins can be used for genetic engineering of plants.

Since the first discovery of abscisic acid more than fifty years ago, the quest to gain deeper understanding into their structure, functional roles and regulation mechanisms has resulted in significant findings. Yet, it seems that the emergence of new knowledge is unearthing further questions about this intriguing phytohormone and its mediated stress response pathway will be worth investigating in the future.

References

ABARES - Australian bureau of agricultural and resource economics and sciences (2011) Agricultural Commodities Statistics 2011

Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Arabidopsis *AtMYC2* (bHLH) and *AtMYB2* (MYB) function as transcriptional activators in abscisic acid signaling. Plant Cell 15: 63-78

Addicott FT, Carns HR (1983) History and introduction. In Abscisic Acid. Praeger Sci 2: 1-21.

Addicott FT, Lyon JL (1969) Physiology of abscisic acid and related substances. Ann Rev Plant Physiol 20: 139-164.

Adem GD, Roy SJ, Zhou M, Bowman JP, Shabala S (2014) Evaluating contribution of ionic, osmotic and oxidative stress components towards salinity tolerance in barley. BMC Plant Biol 14: 113-125

Ahmed IM, Dai H, Zheng W, Cao F, Zhang G, Sun D, Wu F (2013) Genotypic differences in physiological characteristics in the tolerance to drought and salinity combined stress between Tibetan wild and cultivated barley. Plant Physiol Biochem 63: 49-60

Allfrey VG, Faulkner R, Mirsky AE (1964) Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. P Natl A Sci 51: 786-794

Aman P, Newman CW (1986) Chemical composition of some different types of barley grown in Montana, USA. J Cereal Sci 4: 133-141

Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol 11: 106-111

Anderson BE, Ward JM, Schroeder JI (1994) Evidence for an extracellular reception site for abscisic acid in *commelina* guard cells. Plant Physiol 104: 1177-1183

Ando K, Grumet R (2010) Transcriptional profiling of rapidly growing cucumber fruit by 454-pyrosequencing analysis. J Amer Soc Hort Sci 135: 291-302

Ashraf M (2004) Some important physiological selection criteria for salt tolerance in plants. Flora Rev 199: 361-376

Atkinson NJ, Urwin PE (2012) The interaction of plant biotic and abiotic stresses: from genes to the field. J Expt Bot 63: 3523-3543

Ayadi M, Cavez D, Miled N, Chaumont F, Masmoudi K (2011) Identification and characterization of two plasma membrane aquaporins in durum wheat (*Triticum turgidum* L. subsp. durum) and their role in abiotic stress tolerance. Plant Physiol Biochem 49: 1029-1039

Bai Y, Du F, Bai Y, Liu H (2010) Determination strategies of phytohormones: recent advances. Anal Methods 2: 1867-1873

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

Bairoch A (1992) Prosite: a dictionary of sites and patterns in proteins. Nucleic Acids Res 20: 2013-2018

Bannister AJ, Kouzarides T (2011) Regulation of chromatin by histone modifications. Cell Res 21: 381-395

Barley Australia: Barley (2014) http://www.barleyaustralia.com.au/industry-information. Accessed May 2014

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. Cell 129: 823-837

Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T (2010) Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell 143: 470-484

Bäurle I, Dean C (2006) The timing of developmental transitions in plants. Cell 125: 655-664

Belin C, De Franco PO, Bourbousse C, Chaignepain S, Schmitter JM, Vavasseur A, Giraudat J, Brygoo HB, Thomine S (2006) Identification of features regulating OST1 kinase activity and OST1 function in guard cells. Plant Physiol 141: 1316-1327

Benjamini Y, Hochberg Y (1995) Controlling the false rate discovery rate: A practical and powerful approach to multiple testing. J Royal Stat Society 57: 289-300

Bensmihen S, Rippa S, Lambert G, Jublot D, Pautot V (2002) The Homologous ABI5 and EEL transcription factors function antagonistically to fine-tune gene expression during late embryogenesis. Plant Cell 14: 1391-403

Bernstein BE, Meissner A, Lander ES (2007) The mammalian epigenome. Cell 128: 669-681

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125: 315-326

Blumwald E (2000) Sodium transport and salt tolerance in plants. Curr Opin Cell Biol 12: 431-434

Bothmer R von, Sato K, Komatsuda T, Yasuda S, Fischbeck G (2003) The domestication of cultivated barley. In: von Bothmer R, van Hintum TH, Knüpffer H, Sato K (eds) Diversity in Barley (*Hordeum vulgare*). Elsevier Science BV, Netherlands, pp 9-27

Boursiac Y, Leran S, Corratge-Faillie C, Gojon A, Krouk G, Lacombe B (2013) ABA transport and transporters. Trends Plant Sci 18: 325-333

Braszewska-Zalewska A, Tylikowska M, Kwasniewska J, Szymanowska-Pulka J (2014) Epigenetic chromatin modifications in barley after mutagenic treatment. J Appl Genet 10: 722-733

Bräutigam A, Gowik U (2010) What can next generation sequencing do for you? Next generation sequencing as a valuable tool in plant research. Plant Biol 12: 831-841

Bray EA, Bailey-Serres J, Weretilnyk E (2000) Responses to abiotic stresses. In: Buchanan B, Gruissem W, Jones R (eds) Biochemistry and molecular biology of plants. American Society of Plant Physiologists, Rockville, pp 1158-1203

Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, Allis CD (1996) Tetrahymena histone acetyltransferase A: a homolog to yeast *Gcn5p* linking histone acetylation to gene activation. Cell 84: 843-51

Brusslan JA, Alvarez-Canterbury AMR , Nair NU, Rice JC, Hitchler MJ, Pellegrini M (2012) Genome-wide evaluation of histone methylation changes associated with leaf senescence in Arabidopsis. P Natl A Sci 10: 65-78

Bruzzone S, Bodrato N, Usai C, Guida L, Moreschi I, Nano R, Antonioli B, Fruscione F, Magnone M, Scarfì S, De Flora A, Zocchi E (2008). Abscisic acid is an endogenous stimulator of insulin release from human pancreatic islets with cyclic ADP ribose as second messenger. J Biol Chem 283: 32188-32197.

Bruzzone S, Moreschi I, Usai C, Guida L, Damonte G, Salis A, Scarfì S, Millo E, De Flora A, Zocchi E (2007) Abscisic acid is an endogenous cytokine in human granulocytes with cyclic ADP-ribose as second messenger. P Natl A Sci 104: 5759-5764

Burbidge A, Grieve TM, Jackson A, Thompson A, McCarty DR, Taylor IB (1999) Characterization of the ABA-deficient tomato mutant notabilis and its relationship with maize *Vp14*. Plant J 17: 427-431.

Cabrera-Vera TM, Vanhauwe J, Thomas TO, Medkova M, Preininger A, Mazzoni MR, Hamm HE (2003) Insights into G protein structure, function, and regulation. Endocr Rev 24: 765-781

Cadenas AG, Verhey SD, Holappa LD, Shen Q, Ho THD, Simmons W (1999) An abscisic acid-induced protein kinase, PKABA1, mediates abscisic acid supressed gene expression in barley aleurone layers. Plant Biol 96: 1767-1772

Camporealea G, Oommena AM, Griffina JB, SarathbJB, Zemplenia J (2007) K12-biotinylated histone H4 marks heterochromatin in human lymphoblastoma cells. J Nut Biochem 18: 760-768

Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RSR, Zhang Y (2002) Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 298: 1039-1043

Carrow RN, Duncan RR (1998) Salt-affected turfgrass sites: Assessment and management. Sleeping Bear Press, Farmington Hills, MI, pp 68-95

Chaves MM, Flexas J, Pinheiro C (2009) Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. Ann Bot 103: 551-560

Chen FSC, Mactaggart JM, Lawrence C, Wang H, Westley JC (1988) Analysis of abscisic acid in the brains of rodents and ruminanats. Agr Biol Chem 52: 1273-1274

Chen Z, Newman I, Zhou M, Mendham N, Zhang G, Shabala S (2005) Screening plants for salt tolerance by measuring K^+ flux: a case study for barley. Plant Cell Environ 28: 1230-1246

Chen Z, Zhou M, Newman I, Mendham NJ, Zhang G, Shabala S (2007) Potassium and sodium relations in salinised barley tissues as a basis of differential salt tolerance. Funct Plant Biol 34: 150-162

Cheong YH, Pandey GK, Grant JJ, Batistic O, Li L, Kim BG, Lee SC, Kudla J, Luan S (2007) Two calcineurin B-like calcium sensors, interacting with protein kinase CIPK23, regulate leaf transpiration and root potassium uptake in Arabidopsis. Plant J 52: 223-239

Cherel I, Michard E, Platet N, Mouline K, Alcon C, Sentenac H, Thibaud JB (2002) Physical and functional interaction of the Arabidopsis K+ channel AKT2 and phosphatase *AtPP2CA*. Plant Cell 14: 1133-1146

Chernys JT, Zeevaart JAD (2000) Characterization of the 9-cis-epoxycarotenoid dioxygenase gene family and the regulation of abscisic acid biosynthesis in avocado. Plant Physiol 124: 343-354

Chew W, Hrmova M, Lopato S (2013) Role of homeodomain leucine zipper (HD-Zip) IV transcription factors in plant development and plant protection from deleterious environmental factors. J Mol Sci 14: 8122-8147

Choi HI, Park HJ, Park JH, Kim S, Im MY, Seo HH, Kim YW, Hwang I, Kim SY (2005) Arabidopsis calcium-dependent protein kinase AtCPK32 interacts with ABF4, a transcriptional regulator of abscisic acid-responsive gene expression and modulates its activity. Plant Physiol 139:1 750-61

Clark DJ, Shen CH (2006) Mapping histone modifications by nucleosome immunoprecipitation. Method in Enzymol 410: 416-430.

Collas P (2010) The current state of chromatin immunoprecipitation. Mol Biotechnol 45: 87-100

Cramer GR, Urano K, Delrot S, Pezzotti M, Shinozaki K (2011) Effects of abiotic stress on plants: a systems biology perspective. BMC Plant Biol 11: 163-177

Cuin TA, Zhou M, Parsons D, Shabala S (2011) Genetic behaviour of physiological traits conferring cytosolic K+/Na+ homeostasis in wheat. Plant Biol 14: 438-446

Cunningham FX, Gantt E (1998) Genes and enzymes of carotenoid biosynthesis in plants. Annu Rev Plant Phys 49: 557-583

Cutler AJ, Krochko JE (1999) Formation and breakdown of ABA. Trends Plant Sci 4: 472-478

Cutler SR, Rodriguez PL, Finkelstein RR, Abrams SR (2010) Abscisic acid: emergence of a core signaling network. Annu Rev Plant Biol 61: 651-679

Danquah A, de Zelicourt A, Colcombet J, Hirt H (2013) The role of ABA and MAPK signaling pathways in plant abiotic stress responses. Biotechnol Adv 32: 40-52

Dawson MA, Bannister AJ, Gottgens B, Foster SD, Bartke T, Green AR, Kouzarides T (2009) JAK2 phosphorylates histone H3Y41 and excludes HP1alpha from chromatin. Nature 461: 819-822

Department of Environment and Primary Industries (DEPI): Victorian Winter Crop Summary (2011) www.grdc.com.au/~/.../B92E5B829E21494F83D966AF541747CC.pdf. Accessed May 2014

Devos KM, Gale MD (1997) Comparative genetics in the grasses. Plant Mol Biol 35: 3-15

Diallo AO, Ali-Benali MA, Badawi M, Houde M, Sarhan F (2012) Expression of vernalization responsive genes in wheat is associated with histone H3 trimethylation. Mol Genet Genom 287: 575-590

Diedhiou C, Popova OV, Dietz KJ, Golldack D (2008) The SNF type serine-threonine protein kinase SAPK4 regulates stress responsive gene expression in rice. BMC Plant Biol 8: 49-61

Dietz KJ, Sauter A, Wichert K, Messdaghi D, Hartung W (2000) Extracellular beta-glucosidase activity in barley involved in the hydrolysis of ABA glucose conjugate in leaves. J Expt Bot 51: 937-944

Dolferus R, Ji X, Richards RA (2011) Abiotic stress and control of grain number in cereals. Plant Sci 181: 331-341

Dong X, Greven MC, Kundaje A, Djebali S, Brown JB, Cheng C, Gingeras TR, Gerstein M, Guig R, BirneyE, Weng Z (2012) Modeling gene expression using chromatin features in various cellular contexts. Gene Biol 13: 53-63

Du Z, Lib H, Weia Q, Zhaob X, Wanga C, Zhub Q, Yia X, Xua W, Liuc XS, Jinb W, Su Z (2013) Genome-wide analysis of histone modifications: H3K4me2, H3K4me3, H3K9ac, and H3K27ac in *Oryza sativa* L. Japonica. Mol Plant 18: 32-40

Du Z, Zhou X, Ling Y, Zhang Z, Su Z (2010) agriGO: a GO analysis toolkit for the agricultural community. Nucleic Acids Res 38: 64-70

Dugas DV, Monaco MK, Olson A, Klein RR, Kumari S, Ware D, Klein PE (2011) Functional annotation of the transcriptome of *Sorghum bicolor* in response to osmotic stress and abscisic acid. BMC Genom 12: 514-518

Dutnall RN, Ramakrishnan V (1997) Twists and turns of the nucleosome: tails without ends. Structure 5: 1255-59

Egan AN, Schlueter J, Spooner DM (2012) Applications of next generation sequencing in plant biology. Am J Bot 99: 175-185

Eversole K, Graner A, Stein N (2009) Wheat and barley genome sequencing. In: Feuillet C, Muehlbauer G (eds) Genetics and Genomics of the Triticeae. Springer, New York, pp 713-742

Fan CM, Wang YW, Hu RB, Zhang XM, Chen JX, Fu YF (2013) Genome-wide expression analysis of soybean MADS genes showing potential function in the seed development. P Natl A Sci 10: 215-237

Finkelstein R, Gampala SSL, Lynch TJ, Thomas TL, Rock CD (2005) Redundant and distinct functions of the ABA response loci ABA-INSENSITIVE(ABI)5 and ABRE-BINDING FACTOR (ABF)3. Plant Mol Biol 59: 253-67

Finkelstein RR (1994) Mutations at two new Arabidopsis ABA response loci are similar to the *abi3* mutations. Plant J 5: 765-771

Finkelstein RR, Gampala SSL, Rock CD (2002) Abscisic acid signaling in seeds and seedlings. Plant Cell 14: 15-45

Fisher RA (1954) Statistical Methods for Research Workers. Oliver and Boyd

Fletcher TM, Hansen JC (1996) The nucleosomal array: structure/function relationships. Crit Rev Eukaryot Gene Expr 6: 149-88

Follows MJ, Dutkiewicz S, Grant S, Chisholm SW (2007) Emergent biogeography of microbial communities in a model ocean. Science 315: 1843-1846

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

Food and Agriculture Organisation of the United Nations (2005) Management of irrigation-induced salt-affected soils ftp://ftp.fao.org/agl/agll/docs/salinity_brochure_eng.pdf

Frey A, Audran C, Marin E, Sotta B, Marion-Poll A (1999) Engineering seed dormancy by the modification of *zeaxanthin epoxidase* gene expression. Plant Mol Biol 39: 1267-1274

Fujii H, Chinnusamy V, Rodrigues A, Rubio S, Antoni R, Park SY, Cutler SR, Sheen J, Rodriguez PL, Zhu JK. (2009) *In vitro* reconstitution of an abscisic acid signalling pathway. Nature 462: 660-664

Fujii H, Verslues PE, Zhu JK (2007) Identification of two protein kinases required for abscisic acid regulation of seed germination, root growth, and gene expression in Arabidopsis. Plant Cell 19: 485-494

Fujita Y, Fujita M, Satoh R, Maruyama K, Parvez MM (2005) AREB1 is a transcription activator of novel ABRE-dependent ABA signaling that enhances drought stress tolerance in Arabidopsis. Plant Cell 17: 3470-88

Fujita Y, Fujita M, Satoh R, Maruyama K, Parvez MM, Seki M, Hiratsu K, Ohme-Takagi M, Shinozaki K, Yamaguchi-Shinozaki K (2005) AREB1 is a transcription activator of novel ABRE-dependent ABA signaling that enhances drought stress tolerance in Arabidopsis. Plant Cell 17: 3470-3488

Fujita Y, Nakashima K, Yoshida T, Katagiri T, Kidokoro S, Kanamori N, Umezawa T, Fujita M, Maruyama K, Ishiyama K, Kobayashi M, Nakasone S, Yamada K, Ito T, Shinozaki K, Yamaguchi-Shinozaki K (2009) Three SnRK2 protein kinases are the main positive regulators of abscisic acid signaling in response to water stress in Arabidopsis. Plant Cell Physiol 50: 2123-2132

Fukao T, Yeung E, Serres JB (2011) The submergence tolerance regulator SUBA mediates crosstalk between submergence and drought tolerance in rice. Plant Cell 23: 412-427

Furihata T, Maruyama K, Fujita Y, Umezawa T, Yoshida R (2006) Abscisic acid-dependent multisite phosphorylation regulates the activity of a transcription activator AREB1. P Natl A Sci 103: 1988-93

Furihata T, Maruyama K, Fujita Y, Umezawa T, Yoshida R,Shinozaki K, Yamaguchi-Shinozaki K (2006) Abscisic acid-dependent multisite phosphorylation regulates the activity of a transcription activator AREB1. P Natl A Sci 103: 1988-1993

Gachomo EW, Jimenez-Lopez JC, Smith SR, Cooksey AB, Oghoghomeh OM, Johnson N, Baba-Moussa L, Kotchoni SO (2013) The cell morphogenesis *ANGUSTIFOLIA* (*AN*) gene, a plant homolog of CtBP/BARS, is

involved in abiotic and biotic stress response in higher plants. BMC Plant Genet 13: 79-89

Gao R, Duan K, Guo G, Du Z, Chen Z, Li L, He T, Lu R, Huang J (2013) Comparative transcriptional profiling of two Contrasting barley genotypes under salinity stress during the seedling stage. International J Genom 11: 1-19

Garbers C, DeLong A, Deruere J, Bernasconi P, Soll D (1996) A mutation in protein phosphatase 2A regulatory subunit A affects auxin transport in Arabidopsis. Embo J 15: 2115-2124

Garg R, Patel RK, Tyagi AK, Jain M (2011) *De novo* assembly of chickpea transcriptome using short reads for gene discovery and marker identification. DNA Res 18: 53-63

Garstang JR, Spink HH, Suleimenov M, Schillinger WF, Mckenzie RH, Tanake DL, Ceccarelli S, Grando S, Paynter BH, Fettell NA (2011) Cultural Practises: Focus on major barley producing regions. In: SE Ullrich (ed) Barley: Production, improvement and uses, 10th Edition, Wiley, New York, pp 221-281

Gaut BS (2002) Evolutionary dynamics of grass genomes. New Phytol 154: 15-28

Gill SS, Tuteja N (2010) Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant Physiol Biochem 48: 909-930

Gillard DF, Walton DC (1976) Abscisic acid metabolism by a cell-free preparation from *Echinocystis lobata* liquid endoserum. Plant Physiol 58: 790-795

Gillies SA, Futardo A, Henry RJ (2012) Gene expression in the developing aleurone and starchy endosperm of wheat. Plant Biotechnol J 10: 668-679

Gomes-Filho E, Lima CR, Costa JH, da Silva AC, da Guia Silva, Lima M, de Lacerda CF, Prisco JT (2008) Cowpea ribonuclease: Properties and effect of NaCl-salinity on its activation during seed germination and seedling establishment. Plant Cell Rep 27: 147-157

Gómez-Cadenas A, Verhey SD, Holappa LD, Shen Q, Ho THD, Walker-Simmons MK (1999) An abscisic acid-induced protein kinase, PKABA1, mediates abscisic acid-suppressed gene expression in barley aleurone layers. P Natl A Sci 96: 1767-1772

Gosti F, Beaudoin N, Serizet C, Webb AAR, Vartanian N, Giraudat J (1999) ABI1 protein phosphatase 2C is a negative regulator of abscisic acid signaling. Plant Cell 11: 1897-1909

Grains Research & Development Corporation (2008) GRDC impact assessment report series: an economic analysis of GRDC's investment in barley breeding http://www.grdc.com.au/uploads/documents/GRDC_ImpAss_BarleyBreeding1.pdf

Guerriero G, Martin N, Golovko A, Sundstrom JF, Rask L, Ezcurra I (2009) The RY/Sph element mediates transcriptional repression of maturation genes from late maturation to early seedling growth. New Phytol 184: 552-565

Guo G, Sun X, Chen C, Wu S, Huang P, Li Z, Dean M, Huang Y, Jia W, Zhou Q, Tang A, Yang Z, Li X, Song P, Zhao X, Ye R, Zhang S, Lin Z, Qi M, Wan S, Xie L, Fan F, Nickerson ML, Zou X, Hu X, Xing L, Lv Z, Mei H, Gao S, Liang C, Gao Z, Lu J,Yu Y, Liu C, Li L, Fang X, Jiang Z, Yang J, Li C, Zhao X, Chen

J, Zhang F, Lai Y, Lin Z, Zhou F, Chen H, Chan HC, Tsang S, Theodorescu D, Li Y, Zhang X, Wang J, Yang H, Gui Y, Wang J, Cai Z (2013) Whole-genome and whole-exome sequencing of bladder cancer identifies frequent alterations in genes involved in sister chromatid cohesion and segregation. Nat Genet 45: 1459-1463

Guo S, Zheng Y, Joung JG, Liu S, Zhang Z, Crasta OR, Sobral BW, Xu Y, Huang S, Fei Z (2010) Transcriptome sequencing and comparative analysis of cucumber flowers with different sex types. BMC Genom 11: 384-392

Guo Y, Xiong LM, Song CP, Gong DM, Halfter U, Zhu JK (2002) A calcium sensor and its interacting protein kinase are global regulators of abscisic acid signaling in Arabidopsis. Developmental Cell 3: 233-244

Guzman GM, Pizzio GA, Antoni R, Vera-Sirera F, Merilo E, Bassel GW, Fernandez MA, Holdsworth MJ, Perez-Amador MA, Kollist H, Rodriguez PL (2012) Arabidopsis PYL/PYR/RCAR receptors play a major role in quantitative regulation of stomatal aperture and transcriptional response to ABA. Plant Cell Pre 10: 1-14

Hanba YT, Shibasaka M, Hayashi Y, Hayakawa T, Kasamo K, Terashima I, Katsuhara M (2004) Overexpression of the barley aquaporin HvPIP2;1 increases internal CO2 assimilation in the leaves of transgenic rice plants. Plant Cell Physiol 45: 521-529

Haring M, Offermann S, Danker T, Horst I, Peterhansel C, Stam M (2007) Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization. Plant Methods 3: 11-27

Harrison MA (2012) Cross talk between phytohormone signalling pathways under both optimal and stressful environmental conditions. In Khan NA (ed.), Nazar R (ed.), Iqbal N (ed.), Anjum NA (ed.) Phytohormones and abiotic stress tolerance in plants. Springer, New York, pp 49-76

Hartung W, Slovik S (1991) Physicochemical properties of plant growth regulators and plant tissue determine their distribution and redistribution stomatal regulation by abscisic acid in leaves. New Phytol 119: 361-382

Hauser F, Horie T (2010) A conserved primary salt tolerance mechanism mediated by HKT transporters: a mechanism for sodium exclusion and maintenance of high $K(^{+})/Na(^{+})$ ratio in leaves during salinity stress. Plant Cell Environ 33: 552-565

He S, Yan S, Wang P, Zhu W, Wang X, Shen Y, Shao K, Xin H, Li S, Li L (2014) Comparative analysis of genome-wide chromosomal histone modification patterns in maize cultivars and their wild relatives. P Natl A Sci 15: 125-136

Hetherington AM (2001) Guard Cell Signaling. Cell 107: 711-714

Higgins CF (1992) ABC transporters from microorganism to man. Annu Rev Cell Biol 8: 67-113

Himmelbach A, Hoffmann T, Leube M, Hohener B, Grill E (2002) Homeodomain protein ATHB6 is a target of the protein phosphatase ABI1 and regulates hormone responses in Arabidopsis. EMBO J 21: 3029-3038

Hirayama T, Shinozaki K (2007) Perception and transduction of abscisic acid signals: keys to the function of the versatile plant hormone ABA. Trends Plant Sci 12: 343-351

Hirayama T, Shinozaki K (2010) Research on plant abiotic stress responses in the post-genome era: past, present and future. Plant J 61: 1041-1052

Hiremath PJ, Farmer A, Cannon SB, Woodward J, Kudapa H, Tuteja R, Kumar A, BhanuPrakash A, Mulaosmanovic B, Gujaria N, Krishnamurthy L, Gaur PM, KaviKishor PB, Shah T, Srinivasan R, Lohse M, Xiao Y, Town CD, Cook RD, May GD, Varshney RK (2011) Large-scale transcriptome analysis in chickpea (*Cicer arietinum* L.), an orphan legume crop of the semi-arid tropics of Asia and Africa. Plant Biotechnol J 9: 922-931

Hoagland DR, Arnon DI (1950) The water culture method for growing plants without soil. University of California Agric. Expt station, Berkley Circular 347

Hobo T, Kowyama Y, Hattori T (1999) A bZIP factor, TRAB1, interacts with VP1 and mediates abscisic acid-induced transcription. P Natl A Sci 96: 15348-15353

Hong MJ, Kim DY, Lee TG, Jeon WB (2010) Functional characterization of pectin methylesterase inhibitor (PMEI) in wheat. Genes Genet Syst 85: 97-106

Hoth S, Morgante M, Sanchez JP, Hanafey MK, Tingey SV, Chua NH (2002) Genome-wide gene expression profiling in *Arabidopsis thaliana* reveals new targets of abscisic acid and largely impaired gene regulation in the *abi1-1* mutant. J Cell Sci 115: 4891-4900

Hou X, Tong H, Selby J, Dewitt J, Peng X, He ZH (2005) Involvement of a cell wall-associated kinase, WAKL4, in Arabidopsis mineral responses. Plant Physiol 139: 1704-1716

Houde M, Belcaid M, Ouellet F, Danyluk J, Monroy AF, Dryanova A, Gulick P, Bergeron A, Laroche A, Links MG, MacCarthy L, Crosby WL, Sarhan F (2006) Wheat EST resources for functional genomics of abiotic stress. BMC Genom 7: 149-170

Hrabak EM, Chan CWM, Gribskov M, Harper JF, Choi JH, Halford N, Kudla J, Luan S, Nimmo HG, Sussman MR, Thomas M, Walker-Simmons K, Zhu JK, Harmon AC (2003) The Arabidopsis CDPK-SnRK superfamily of protein kinases. Plant Physiol 132: 666-671

Hu Z, Han Z, Song N, Chai L, Yao Y, Peng H, Ni N, Sun Q (2013) Epigenetic modification contributes to the expression divergence of three *TaEXPA1* homoeologs in hexaploid wheat (*Triticum aestivum*). New Phytol 197: 145-156

Huai J, Wang M, He J, Zheng J, Dong Z, Lv H, Zhao J, Wang G (2008) Cloning and characterization of the SnRK2 gene family from *Zea mays*. Plant Cell Rep 27: 1861-1868

Huang J, Zhou W, Dong W, Watson AM, Hong Y (2009) Directed, efficient, and versatile modifications of the Drosophila genome by genomic engineering. P Natl A Sci 106: 8284-8289

Huang Y, Fang J, Bedford MT, Zhang Y, Xu RM (2006) Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. Science 312: 748-751

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

International Barley Genome Sequencing Consortium, Mayer KF, Waugh R, Brown JW, Schulman A, Langridge P, Platzer M, Fincher GB, Muehlbauer GJ, Sato K, Close TJ, Wise RP, Stein N (2012) A physical, genetic and functional sequence assembly of the barley genome. Nature 29: 711-716

International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. Nature 436: 793-800

Ishitani M, Nakamura T, Han SY, Takabe T (1995) Expression of the betaine aldehyde dehydrogenasegene in barley in response to osmotic stress and abscisic acid. Plant Mol Biol 27: 307-315

Islam AKMR, Shepherd KW, Sparrow DHB (1981) Isolation and characterization of euplasmic wheat-barley chromosome addition lines. Heredity 46: 16l-174

Iuchi S, Kobayashi M, Yamaguchi-Shinozaki K, Shinozaki K (2000) A stress-inducible gene for 9-cis-epoxycarotenoid dioxygenase involved in abscisic acid biosynthesis under water stress in drought-tolerant cowpea. Plant Physiol 123: 553-562

Iyer LM (2001) Adaptations of the helix-grip fold for ligand binding and catalysis in the START domain superfamily. Proteins 43: 134-139

Jain D, Chattopadhyay D (2010) Analysis of gene expression in response to water deficit of chickpea (*Cicer arietinum* L.) varieties differing in drought tolerance. BMC Plant Biol 10: 467-474.

Jakoby M, Weisshaar B, Dröge-Laser W, Vicente-Carbajosa J, Tiedemann J, Kroj T, Parcy F (2002) bZIP transcription factors in Arabidopsis. Trends Plant Sci 7: 106-111

James RA, Davenport RJ, Munns R (2006) Physiological characterisation of two genes for Na^+ exclusion in durum wheat: *Nax1* and *Nax2*. Plant Physiol 142: 1537-1547

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

Jeong J, Park Y, Jung H, Park SH, Kim JK (2009) Rice NAC proteins act as homodimers and heterodimers. Plant Biotechnol Rep 3: 127-134

Jia W, Wang Y, Shuquiu Z, Zhang J (2002) Salt stress induced ABA accumulation is more sensitively triggered in roots than in shoots. J Expt Bot 53: 2201-2206

Jia W, Zhang J, Zhang DP (1996) Metabolism of xylem-delivered ABA in relation to ABA flux and concentration in leaves of maize and *Commelina communis*. J Exp Bot 47: 1085-1091

Jiang F, Hartung W (2008) Long-distance signalling of abscisic acid (ABA): the factors regulating the intensity of the ABA signal. J Expt Bot 59: 37-43

Jiang SY, González JM, Ramachandran S (2013) Comparative genomic and transcriptomic analysis of tandemly and segmentally duplicated genes in rice. P Natl A Sci 104: 5759-5764

Kakeda K (2009) S locus linked F box genes expressed in anthers of *Hordeum bulbosum*. Plant Cell Rep 28: 1453-1460

Kakumanu A, Ambavaram MM, Klumas C, Krishnan A, Batlang U, Myers E, Grene R, Pereira A (2012) Effects of drought on gene expression in maize reproductive and leaf meristem tissue revealed by RNA-Seq. Plant Physiol 160: 846-867

Kane MD, Jatkoe TA, Stumpf CR, Lu J, Thomas JD, Madore S (2000) Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. Nucleic Acid Res 28: 4552-4557

Kang J, Hwang JU, Lee M, Kim YY, Assmann SM, Martinoia E, LeeY (2010) PDR-type ABC transporter mediates cellular uptake of the phytohormone abscisic acid. P Natl A Sci 107: 2355-2360

Kapazoglou A, Drosou V, Argiriou A, Tsaftaris AS (2013) The study of a barley epigenetic regulator, *HvDME,* in seed development and under drought. BMC Plant Biol 13: 172-178

Kasuga M, Liu Q, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1999) Improving plant drought, salt and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Nat Biotechnol 17: 287-291

Kausar A, Ashraf YM, Ali I, Niaz M, Abbass Q (2012) Evaluation of Sorghum varieties/lines for salt tolerance using physiological indices as screening tool. Pak J Bot 44: 47-52

Kawaura K, Mochida K, Ogihara Y (2008) Genome-wide analysis for identification of salt-responsive genes in common wheat. Func Integr Genom 8: 277-286

Kawaura K, Mochida K, Yamazaki Y, Ogihara Y (2006) Transcriptome analysis of salinity stress responses in common wheat using 22k oligo-DNA microarray. Funct Integr Genom 6: 132-142

Kerckhoffs DA, Brouns F, Hornstra G, Mensink RP (2002) Effect on the human lipoprotein profile of β-glucan, soy protein and isoflavones, plant sterols, garlic and tocotrienols. J Nutr 132: 2494-2505

Khakwani AA, Dennett MD, Munir M (2011) Drought tolerance screening of wheat varieties by inducing water stress conditions. Songklanakarin J Sci Technol 33: 135-142

Kiefer CM, Hou C, Little JA, Dean A (2008) Epigenetics of beta-globin gene regulation. Mutat Res 647: 68-76

Kim H, Hwang H, Hong JW, Lee YN, Ahn IP, Yoon IS, Yoo SD, Lee S, Lee SC, Kim BG (2012) A rice orthologue of the ABA receptor, *OsPYL/RCAR5*, is a positive regulator of the ABA signal transduction pathway in seed germination and early seedling growth. J Expt Bot 63: 1013-1024

Kim S (2007) Recent advances in ABA signaling. J Plant Biol 50: 117-121

Kim S, Park SY, Kim S, Rho HS, Chi MH, Choi J, Park J, Kong S, Park J, Goh J, Lee YH (2009) Homeobox transcription factors are required for conidiation and

appresorium development in the rice blast fungus *Maganthorpe oryzae*. P Natl A Sci 5: 41-52

Kim TH, Böhmer M, Hu H, Nishimura N, Schroeder JI (2010) Guard cell signal transduction network: advances in understanding abscisic acid, CO_2, and Ca^{2+} signalling. Annu Rev Plant Biol 61: 561-591

Kim JM, To TK, Ishida J, Morosawa T, Kawashima M, Matsui A, Toyoda T, Kimura H , Shinozaki K, Seki M (2008) Alterations of lysine modifications on the histone H3 N tail under drought stress conditions in *Arabidopsis thaliana*. Plant Cell Physiol 49: 1580-1588

Kimura M, Kagawa T (2006) Phototropin and light-signaling in phototropism. Curr Opin Plant Biol 9: 503-508

Kingston RE, Narlikar GJ (1999) ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. Genes Dev 13: 2339-2352

Kline KG, Sussman MR, Jones AM (2010). Abscisic acid receptors. Plant Physiol 154: 479-482.

Klingler JP, Batelli G, Zhu JK (2010) ABA receptors: the START of a new paradigm in phytohormone signalling. J Exp Bot 61: 3199-3120

Kobayashi Y, Yamamoto S, Minami H, Kagaya Y, Hattori T (2004) Differential activation of the rice sucrose nonfermenting1-related protein kinase2 family by hyperosmotic stress and abscisic acid. Plant Cell 16: 1163-1177

Kohzaki H, Murakami Y (2007) Faster and easier chromatin immunoprecipitation assay with high sensitivity. Proteomics 7: 10-14

Komatsu K, Nishikawa Y, Ohtsuka T, Taji T, Quatrano R, Tanaka S, Sakata Y (2009) Functional analyses of the ABI1-related protein phosphatase type 2C reveal evolutionarily conserved regulation of abscisic acid signaling between Arabidopsis and the moss. Plant Mol Biol 70: 327-340

Komatsuda T, Pourkheirandish M, He C, Azhaguvel P, Kanamori H, Perovic D, Stein N, Graner A, Wicker T, Tagiri A, Lundqvist U, Fujimura T, Matsuoka M, Matsumoto T, Yano M (2007) Six rowed barley originated from a mutation in a homeodomain leucine zipper I class homeobox gene. P Natl A Sci 104: 1424-1429

Kouzarides T (2007) Chromatin modifications and their function. Cell 128: 693-705

Kulik A, Wawer I, Krzywinska E, Bucholc M, Dobrowolska G (2011) SnRK2 protein kinases-key regulators of plant response to abiotic stresses. J Integ Biol 15: 859-871

Kuromori T, Miyaji T, Yabuuchi H, Shimizu H, Sugimoto E, Kamiya A, Moriyama Y, Shinozaki K (2010) ABC transporter *AtABCG25* is involved in abscisic acid transport and responses. P Natl A Sci 107: 2361-2366

KushiroT, Okamoto M, Nakabayashi K, Yamagishi K, Kitamura S, Asami T, Hirar N, Koshiba T, Kamiya Y, Nambara E (2004) The Arabidopsis cytochrome P450 CYP707A encodes ABA 8′-hydroxylases: key enzymes in ABA catabolism. EMBO J 23: 1647-1656

Kusumi K, Hirotsuka S, Kumamaru T, Iba K (2012) Increased leaf photosynthesis caused by elevated stomatal conductance in rice mutant deficient in SLAC1, a guard cell anion channel protein. J Expt Bot 63: 5635-5644

Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature 410: 116-120

Lan F, Shi Y (2009) Epigenetic regulation: methylation of histone and non-histone proteins. Sci China C Life Sci 52: 311-322

Langridge P, Paltridge N, Fincher G (2006) Functional genomics of abiotic stress tolerance in cereals. Briefings Funct Genom Proteom 4: 343-354

Lee HS (1997) Endogenous biosynthetic precursors of (+)-abscisic acid. V. inhibition by tungstate and its removal by cinchonine shows that xanthoxal is oxidised by a molybdo-aldehyde oxidase. Aust J Plant Physiol 24: 727-735

Lee KH, Piao HL, Kim HY, Choi SM, Jiang F, Hartung W, Kwak JM, Lee IJ, Hwang I (2006) Activation of glucosidase via stress-induced polymerization rapidly increases active pools of abscisic acid. Cell 126: 1109-1120

Lee SC, Lan W, Buchanan BB, Luan S (2009) A protein kinase-phosphatase pair interacts with an ion channel to regulate ABA signaling in plant guard cells. P Natl A Sci 24: 727-735

Leung J, Bouvier-Durand M, Morris P, Guerrier D, Chefdor F, Giraudat J (1994) Arabidopsis ABA response gene ABI1: features of a calcium-modulated protein phosphatase. Science 264: 1448-1452

Leung J, Giraudat J (1998) Abscisic acid signal transduction. Annu Rev Plant Physiol Plant Mol Biol 49: 199-222

Li P, Ponnala L, Gandotra N, Wang L, Si Y, Tausta SL, Kebrom TH, Provart N, Patel R, Myers CR, Reidel EJ, Turgeon R, Liu P, Sun Q, Nelson T, Brutnell TP (2010) The developmental dynamics of the maize leaf transcriptome. Nat Genet 42: 1060-1067

Li C, Lv J, Zhao X, Ai X, Zhu X, Wang M, Zhao S, Xia G (2010b) *TaCHP*: A wheat zinc finger protein gene down-regulated by abscisic acid and salinity stress plays a positive role in stress tolerance. Plant Physiol 154: 211-221

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinforma 25: 1754-1760

Li M, Hong Y, Wang X (2009a) Phospholipase D- and phosphatidic acid-mediated signaling in plants. Biochimica et Biophysica Acta 1791: 927-935

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009b) The sequence alignment/map format and SAMtools. Bioinforma 25: 2078–2079

Li H, Yan S, Zhao L, Tan J, Zhang Q, Gao F, Wang P, Hou H, Li L (2014) Histone acetylation associated up-regulation of the cell wall related genes is involved in salt stress induced maize root swelling. BMC Plant Biol 14: 156-172

Li J, Wang XQ, Watson MB, Assmann SM (2000) Regulation of abscisic acid-induced stomatal closure and anion channels by guard cell AAPK kinase. Science 287: 300-303

Lin BL, Wang HJ, Wang JS, Zaharia LI, Abrams SR (2005). Abscisic acid regulation of heterophylly in *Marsilea quadrifolia* L.: effects of R-(−) and S-(+) isomers. J Exp Bot 56: 2935-2948.

Lindgren LO, Stalberg KG, Hoglund AS (2003) Seed-specific overexpression of an endogenous Arabidopsis phytoene synthase gene results in delayed germination and increased levels of carotenoids, chlorophyll and abscisic acid. Plant Physiol 132: 779-785

Lindsay MP, Lagudah ES, Hare RA, Munns R (2004) A locus for sodium exclusion (*Nax1*), a trait for salt tolerance, mapped in durum wheat. Funct Plant Biol 31: 1105-1114

Ling G, Waxman DJ (2013) DNase I digestion of isolated nulcei for genome-wide mapping of DNase hypersensitivity sites in chromatin. Methods Mol Bio 977: 21-33

Lister L, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini G, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462: 315-322

Liu YH, Li HY, Shi YS, Song YC, Wang TY, Li Y (2009a) A maize early responsive to dehydration gene, *ZmERD4*, provides enhanced drought and salt tolerance in Arabidopsis. Plant Mol Biol Rep 27: 542-548

Liu ZQ, Gao J, Dong AW, Shen WH (2009b) A truncated Arabidopsis nucleosome assembly protein I, AtNAP1;3T, alters plant growth responses to abscisic acid and salt in the *Atnap1;3-2* Mutant Mol Plant 2: 688-699

Loden M, van SB (2005) Whole-genome views of chromatin structure. Chromosome Res 13: 289-298

Lopez-Molina L, Chua N-H (2000) A null mutation in a bZIP factor confers ABA-insensitivity in *Arabidopsis thaliana*. Plant Cell Physiol 41: 541-7

Luan S (2003) Protein phosphatases in plants. Annu Rev Plant Biol 54: 63-92

Ludwikow A, Skowronska DB, Szczepaniak M, Belter N, Dominiak E, Sadowski J (2013) Expression profiles and genomic organisation of group A protein phosphatase 2C genes in *Brassica oleracea*. Ann Appl Biol 163: 124-134

Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389: 251-260

Luger K, Richmond TJ (1998) The histone tails of the nucleosome. Curr Opin Genet Dev 8: 140-46

Lumba S, Cutler S, McCourt P (2010) Plant nuclear hormone receptors: a role for small molecules in protein-protein interactions. Annu Rev Cell Dev Biol 26: 445-469

Lv DW, Zhang PGM, Cheng ZW, Li XH, Yan YM (2014) Integrative network analysis of the signaling cascades in seedling leaves of bread wheat by large-scale phosphoproteomic profiling. J Proteo Res 13: 2381-2395

Ma Y, Szostkiewicz I, Korte A, Moes D, Yang Y, Christmann A, Grill E (2009) Regulators of PP2C phosphatase activity function as abscisic acid sensors. Science 324: 1064-1068.

Macknight R, Bancroft I, Page T, Lister C, Schmidt R, Love K, Westphal L, Murphy G, Sherson S, Cobbett C, Dean C (1997) FCA, a gene controlling flowering time in Arabidopsis, encodes a protein containing RNA-binding domains. Cell 89: 737-745

Macknight R, Duroux M, Laurie R, Dijkwel P, Simpson G, Dean C (2002) Functional significance of the alternative transcript processing of the Arabidopsis floral promoter FCA. Plant Cell 14: 877-888

Madrigal P, Krajewski P (2012) Current bioinformatic approaches to identify DNase I hypersensitive sites and genomic footprints from DNase-seq data. Front Genet 3: 230-236

Mammadov J, Aggarwal R, Buyyarapu R, Kumpatla S (2012) SNP markers and their impact on plant breeding. Int J Plant Genom 2012: 1-11

Manfre AJ, Lanni LM, Marcotte WR (2006) The Arabidopsis group 1 late embryogenesis abundant protein ATEM6 is required for normal seed development. Plant Physiol 140: 140-149

Mare C, Mazzucotelli E, Crosatti C, Francia E, Stanca AM, Cattivelli L (2004) Hv-WRKY38: a new transcription factor involved in cold- and drought-response in barley. Plant Mol Biol 55: 399-416

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen Y-J, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Goodwin BC (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437: 376-380

Marin E, MarionPoll A (1997) Tomato flacca mutant is impaired in ABA aldehyde oxidase and xanthine dehydrogenase activities. Plant Physiol Biochem 35: 369-372.

Marin, E. Nussaume L, Quesada A, Gonneau M, Sotta B, Hugueney P, Frey A, Marion-Poll A (1996) Molecular identification of *zeaxanthin epoxidase* of *Nicotiana plumbaginifolia*, a gene involved in abscisic acid biosynthesis and corresponding to the ABA locus of *Arabidopsis thaliana*. Embo J 15: 2331-2342

Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y (2008) RNAseq: An assessment of technical reproducibility and comparisonwith gene expression arrays. Genom Res 18: 1509-1517

Martínez-Gómez P, Crisosoto C, Bonghi C, Rubio M (2011) New approaches to Prunus transcriptome analysis. Genetica 139: 755-769

Matin MA, Brown JH, Ferguson H (1989) Leaf water potential, relative water contentand diffusive resistance as screening techniques for drought tolerance in barley. Agronomy J 81: 100-105

Matsui A, Ishida J, Morosawa Y, Kaminuma E, Endo TA, Okamoto M, Nambara E, Nakajima M, Kawashima M, Satou M, Kim JM, Kobayashi N, Toyoda T, Shinozaki K, Seki M (2008) Arabidopsis transcriptome analysis under drought,

cold, high-salinity and ABA treatment conditions using tiling array. Plant Cell Physiol 49: 1135-1149

Mayer KFX, Martis M, Hedley PE, Simkova H, Liu H, Morris JA, Steuernagel B, Taudien S, Roessner S, Gundlach H, Kubalakova M, Suchankova P, Murat F (2011) Unlocking the barley genome by chromosomal and comparative genomics. Plant Cell 23: 1249-1263

Mayer KFX, Taudien S, Martis M, Ŝimková H, Suchánková P, Gundlach H, Wicker T, Petzold A, Felder M, Steuernagel B, Scholz U, Graner A, Platzer M, Doležel J, Stein N (2009) Gene content and virtual gene order of barley chromosome 1H. Plant Physiol 151: 496-505

McCallum CM, Comai L, Greene EA, Henikoff S (2000) Targeting induced local lesions in genomes (TILLING) for plant functional genomics. Plant Physiol 123: 439-442.

McCourt P, Creelman R (2008) The ABA receptors - we report you decide. Curr Opin Plant Biol 11: 474-478

Melcher K, Ng LM, Zhou XE, Soon FF, Xu Y, Suino-Powell KM, Park SY, Weiner JJ, Fujii H, Chinnusamy V, Kovach A, Li J, Wang Y, Li J, Peterson FC, Jensen DR, Yong EL, Volkman BF, Cutler SR, Zhu JK, Xu HE (2009) A gate-latch-lock mechanism for hormone signalling by abscisic acid receptors. Nature 462: 602-608

Michaels SD, Amasino RM (1999) Flowering locus C encodes a novel MADS domain Protein that acts as a repressor of flowering. Plant Cell 11: 949-956

Milborrow BV (2001) The pathway of biosynthesis of abscisic acid in vascular plants: a review of the present state of knowledge of ABA biosynthesis. J Expt Bot 52: 1145-1164

Miyazono K, Miyakawa T, Sawano Y, Kubota K, Kang H, Asano A, Miyauchi Y, Takahashi M, Zhi Y, Fujita Y, Yoshida T, Kodaira KS, Yamaguchi-Shinozaki K, Tanokura M (2009) Structural basis of abscisic acid signalling. Nature 462: 609-614

Mizuno H, Kawahara Y, Sakai H, Kanamori H, Wakimoto H, Yamagata H, Oono Y, Wu J, Ikawa H, Itoh T, Matsumoto T (2010) Massive parallel sequencing of mRNA in identification of unannotated salinity stress-inducible transcripts in rice (*Oryza* sativa L). BMC Genom 11: 683-695

Mochizuki N, Brusslan JA, Larkin R, Nagatani A, Chory J (2001) Arabidopsis genomes uncoupled 5 (*GUN5*) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. P Natl A Sci 98: 2053-2058

Modra Seeds Fact Sheet, February 2011 http://www.modraseeds.com.au/pdf/Hindmarsh_Factsheet_Feb_2011.pdf

Moes D, Himmelbach A, Korte A, Haberer G, Grill E (2008) Nuclear localization of the mutant protein phosphatase *abi1* is required for insensitivity towards ABA responses in Arabidopsis. Plant J 54: 806-819

Moffat AS (2002) Finding new ways to protect drought-stricken plants. Sci 296: 1226-1229

Mohammadi M, Kav NNV, Deyholos MK (2007) Transcriptional profiling of hexaploid wheat (*Triticum aestivum* L.) roots identifies novel, dehydration-responsive genes. Plant Cell Environ 30: 630-645

Moons A, Bauw G, Prinsen E, Montagu MV, Straeten VD (1995) Molecular and physiological response to abscisic acid and salts in roots of salt sensitive and salt tolerant indica rice varieties. Plant Physiol 107: 177-186

Mori IC, Murata Y, Yang YZ, Munemasa S, Wang YF, Andreoli S, Tiriac H, Alonso JM, Harper JF, Ecker JR, Kwak JM, Schroeder JI (2006) CDPKs, CPK6 and CPK3 function in ABA regulation of guard cell S-type anion- and Ca^{2+}-permeable channels and stomatal closure. P Natl A Sci Biology 4: 1749-1762

Mullis KB, Faloona FA (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods Enzymol 155: 335-350

Munns R (1993) Physiological processes limiting plant growth in saline soil: some dogmas and hypotheses. Plant Cell Environ 16: 15-24

Munns R (2002) Comparative physiology of salt and water stress. Plant Cell Environ 25: 239-250

Munns R (2005) Genes and salt tolerance: bringing them together. New Phytol 167: 645-663

Munns R, James RA, Sirault XRR, Furbank RT, Jones HG (2010) New phenotyping methods for screening wheat and barley for beneficial responses to water deficit. J Expt Bot 61: 3499-3507

Munns R, James RA, Xu B, Athman A, Conn SJ, Jordans C, Byrt CS, Hare RA, Tyerman SD, Tester M, Plett D, Gilliham M (2012) Wheat grain yield on saline soils is improved by an ancestral Na^+ transporter gene. Nature Biotech 30: 360-364

Munns R, Rebetzke GJ, Husain S, James RA, Hare RA (2003) Genetic control of sodium exclusion in durum wheat. Aust J Agric Res 54: 627-635

Munns R, Schachtman DP, Condon AG (1995) The significance of a two-phase growth response to salinity in wheat and barley. Aust J Plant Physiol 22: 561-569

Munns R, Tester M (2008) Mechanisms of salinity tolerance. Annu Rev Plant Biol 59: 651-681

Mustilli AC, Merlot S, Vavasseur A, Fenzi F, Giraudat J (2002) Arabidopsis OST1 protein kinase mediates the regulation of stomatal aperture by abscisic acid and acts upstream of reactive oxygen species production. Plant Cell 14: 3089-3099

Nagashima A, Uehara Y, Sakai T (2008) The ABC subfamily B auxin transporter AtABCB19 is involved in the inhibitory effects of N-1-naphthyphthalamic acid on the phototropic and gravitropic responses of Arabidopsis hypocotyls. Plant Cell Physiol 49: 1250-1255

Nakashima K, Ito Y, Yamaguchi-Shinozaki K (2009a) Transcriptional regulatory networks in response to abiotic stresses in Arabidopsis and grasses. Plant Physiol 149: 88-95

Nakashima K, Fujita Y, Kanamori N, Katagiri T, Umezawa T, Kidokoro S, Maruyama K, Yoshida T, Ishiyama K, Kobayashi M, Shinozaki K, Yamaguchi-Shinozaki K

(2009b) Three Arabidopsis SnRK2 protein kinases, SRK2D/SnRK2.2, SRK2E/SnRK2.6/OST1 and SRK2I/SnRK2.3, involved in ABA signaling are essential for the control of seed development and dormancy. Plant Cell Physiol 50: 1345-1363

Nakashima K, Tran LS, Van Nguyen D, Fujita M, Maruyama K, Todaka D, Ito Y, Hayashi N, Shinozaki K, Yamaguchi-Shinozaki K (2007) Functional analysis of a NAC-type transcription factor OsNAC6 involved in abiotic and biotic stress-responsive gene expression in rice. Plant J 51: 617-630

Nambara E (1992) A mutant of Arabidopsis which is defective in seed development and storage protein accumulation is a new *abi3* allele. Plant J 2: 435-439

Nambara E, Marion-Poll A (2005) Abscisic acid biosynthesis and catabolism. Annu Rev Plant Biol 56: 165-185

Nambara E, Suzuki M, Abrams S, McCarty DR, Kamiya Y, McCourt P (2002) A screen for genes that function in abscisic acid signaling in *Arabidopsis thaliana*. Genetics 161: 1247-55

Nambara E, Suzuki M, Abrams S, McCarty DR, Kamiya Y, McCourt P (2002) A screen for genes that function in abscisic acid aignaling in *Arabidopsis thaliana*. Genetics 161: 1247-1255

Narusaka Y, Nakashima K, Shinwari ZK, Sakuma Y, Furihata T, Abe H, Narusaka M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Interaction between two *cis*-acting elements, ABRE and DRE, in ABA-dependent expression of Arabidopsis *rd29A* gene in response to dehydration and high-salinity stresses. Plant J 34: 137-148

Narusaka Y, Narusaka M, Seki M, Umezawa T, Ishida J, Nakajima M, Enju A, Shinozaki K (2004) Crosstalk in the responses to abiotic and abiotic stresses in Arabidopsis: Analysis of gene expression in cytochrome P450 gene superfamily by cDNA microarray. Plant Mol Biol 55: 327-34

Nathan D, Ingvarsdottir K, Sterner DE, Bylebyl GR, Dokmanovic M, Dorsey JA, Whelan KA, Krsmanovic M, Lane WS, Meluh PB, Johnson ES, Berger SL (2006) Histone sumoylation is a negative regulator in *Saccharomyces cerevisiae* and shows dynamic interplay with positive-acting histone modifications. Genes Dev 20: 966-76

National Land and Water Resources Audit (2001) http://lwa.gov.au/products/pr010158; Last accessed August 2014

Negi J, Matsuda O, Nagasawa T, Oba Y, Takahashi H, Kawai-Yamada M, Uchimiya H, Hashimoto M, Iba K (2008) CO2 regulator SLAC1 and its homologues are essential for anion homeostasis in plant cells. Nature 452: 483-486

Neill SJ, Burnett EC, Desikan R, Hancock JT (1998) Cloning of a wilt-responsive cDNA from an *Arabidopsis thaliana* suspension culture cDNA library that encodes a putative 9-cis-epoxy-carotenoid dioxygenase. J Expt Bot 49: 1893-1894.

Nemhauser JL, Hong F, Chory J (2006) Different plant hormones regulate similar processes through largely non overlapping transcriptional responses. Cell 126: 467-475

Neuman H, Galpaz N, Cunningham FX Jr, Zamir D, Hirschberg J (2014) The tomato mutation *nxd1* reveals a gene necessary for neoxanthin biosynthesis and demonstrates that violaxanthin is a sufficient precursor for abscisic acid biosynthesis. Plant J 78: 80-93

Newman RK, Newman CW (1991) Barley as a food grain. Cereal Food World 36: 800-805

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

Ng SS, Yue WW, Oppermann U, Klose RJ (2009) Dynamic protein methylation in chromatin biology. Cell Mol Life Sci 66: 407-422

Nietzsche M, Schiebi I, Bornke F (2014) The complex becomes more complex: protein-protein interactions of SnRK1 with DUF581 family proteins provide a framework for cell- and stimulus type-specific SnRK1 signaling in plants. Fro Plant Sci 5: 13-19

Nieves M, Nieves-Cordones M, Poorter H, Simon MD (2011) Leaf nitrogen productivity is the major factor behind the growth reduction induced by long-term salt stress. Tree Physiol 31: 92-101

Nishimura N, Hitomi K, Arvai AS, Rambo RP, Hitomi C, Cutler SR, Schroeder JI, Getzoff ED (2009) Structural mechanism of abscisic acid binding and signaling by dimeric PYR1. Science 326: 1373-1379

Nishimura N, Sarkeshik A, Nito K, Park SY, Wang A, Carvalho PC, Lee S, Caddell DF, Cutler SR, Chory J, Yates JR, Schroeder JI (2010) PYR/PYL/RCAR family members are major in-vivo ABI1 protein phosphatase 2C-interacting proteins in Arabidopsis. Plant J 61: 290-299

Nishimura N, Yoshida T, Kitahata N, Asami T, Shinozaki K, Hirayama T (2007) ABA-Hypersensitive Germination1 encodes a protein phosphatase 2C, an essential component of abscisic acid signaling in Arabidopsis seed. Plant J 50: 935-949

North HM, Almeida AD, Boutin JP, Frey A, To A, Botran L, Sotta B, Marion-Poll A (2007) The Arabidopsis ABA-deficient mutant *aba4* demonstrates that the major route for stress-induced ABA accumulation is via neoxanthin isomers. Plant J 50: 810-824

Offermanns S (2003) G-proteins as transducers in transmembrane signalling. Progress Biophy Mol Biol 83: 101-130

Ohkuma K, Addicott FT, Smith OE, Thiessen WE (1965) The structure of abscisin II. Tetrehadron letters 29: 2529-2535

Ohkuma K, Lyon JL, Addicott FT, Smith OE (1963) Abscisin II, an abscission accelerating substance from young cotton fruit.Science 142: 1592-1593

Ohta M, Guo Y, Halfter U, Zhu JK (2003) A novel domain in the protein kinase SOS2 mediates interaction with the protein phosphatase 2C ABI2. P Natl A Sci 100: 11771-11776

Oki M, Aihara H, Ito T (2007) Role of histone phosphorylation in chromatin dynamics and its implications in diseases. Subcell Biochem 41: 319-336

Olvera-Carrillo Y, Luis Reyes J, Covarrubias AA (2011) Late embryogenesis abundant proteins: versatile players in the plant adaptation to water limiting environments. Plant Signal Behav 6: 586-589

Osakabe Y, Maruyama K, Seki M, Satou M, Shinozaki K, Yamaguchi-Shinozaki K (2005) Leucine-rich repeat receptor-like kinase1 is a key membrane-bound regulator of abscisic acid early signaling in Arabidopsis. Plant Cell 17: 1105-1119

Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee Y., Zheng L, Orvis J, Haas B, Wortman J, Buell CR (2007) The TIGR rice genome annotation resource: improvements and new features. Nucleic acid research 35 Database Issue: D846-851

Ozturk ZN, Talame V, Deyholos M, Michalowski CB, Galbraith DW, Gozukirmizi N, Tuberosa R, Bohnert HJ (2002) Monitoring large scale changes in transcript abundance in drought and salt stressed barley. Plant Mol Biol 48: 551-573

Palma F, Lopez-Gomez M, Tejera NA, Lluch C (2014) Involvement of abscisic acid in the response of *Medicago sativa* plants in symbiosis with *Sinorhizobium meliloti* to salinity. Plant Sci 223: 16-24

Pandey BP, Basnet KB, Bhatta MR, Sah SR, RB Thapa, Kandel TP (2013) Effect of row spacing and direction of sowing on yield and yield attributing characters of wheat cultivated in Western Chitwan, Nepal. Agri Sci 4: 309-316

Pandey S, Nelson DC, Assmann SM (2009) Two novel GPCR-type G proteins are abscisic acid receptors in Arabidopsis. Cell 136: 136-148

Park PJ (2009) ChIP-seq: advantages and challenges of a maturing technology. Nat Rev Genet 10: 669-680

Park SY, Fung P, Nishimura N, Jensen DR, Fujii H, Zhao Y, Lumba S, Santiago J, Rodrigues A, Chow TF, Alfred SE, Bonetta D, Finkelstein R, Provart NJ, Desveaux D, Rodriguez PL, McCourt P, Zhu JK, Schroeder JI, Volkman BF, Cutler S (2009) Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL Family of START proteins. Science 324: 1068-1071

Pasini D, Bracken AP, Jensen MR, Lazzerini DE, Helin K (2004) Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J 23: 4061-4071

Pei ZM, Ghassemian M, Kwak CM, McCourt P, Schroeder JI (1998) Role of farnesyltransferase in ABA regulation of guard cell anion channels and plant water loss. Science 282: 287-290

Peleg Z, Apse MP, Blumwald E (2011) Engineering salinity and water-stress tolerance in crop plants: Getting closer to the field. Adv Bot Res 57: 405-443

Pernas M, García-Casado G, Rojo E, Solano R, Sánchez-Serrano JJ (2007) A protein phosphatase 2A catalytic subunit is a negative regulator of abscisic acid signalling1. Plant J 51: 763-778

Pilot G, Lacombe BT, Gaymard F, Chérel I, Boucherez J, Thibaud JB, Sentenac H (2001) Guard cell inward K+ channel activity in Arabidopsis involves expression of the twin channel subunits *KAT1* and *KAT2*. J Biol Chem 276: 3215-3221

Pizzio GA, Rodriguez L, Antoni R, Gonzalez-Guzman M, Yunta C, Merilo E, Kollist H, Albert A, Rodriguez PL (2013) The *PYL4* A194T mutant uncovers a key role of PYR1-LIKE4/PROTEIN PHOSPHATASE 2CA interaction for abscisic acid signaling and plant drought resistance**.** Plant Biol 163: 441-455

Polanda JA, Bradbury PJ, Bucklera ES, Nelsona RJ (2011) Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. P Natl A Sci 108: 6893-6898

Popova LP, Tsonko TD, Lazova GN, Stoinova ZG (1996) Drought and ABA induced changes in photosynthesis of barley plants. Physiol Plant 96: 623-629

Pray-Grant MG, Daniel JA, Schieltz D, Yates JR,. Grant PR (2005) Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. Nature 433: 434-438

Qin X, Zeevaart JAD (1999) The 9-cis-epoxycarotenoid cleavage reaction is the key regulatory step of abscisic acid biosynthesis in water-stressed bean. P Natl A Sci 96: 15354-15361

Rabbani MA, Maruyama K, Abe H,Khan MA, Katsura K, Ito Y, Yoshiwara K, Seki M, Shinozaki K, Yamaguchi Shinozaki K (2003) Monitoring expression profiles of rice (*Oryza sativa* L.) genes under cold, drought and high-salinity stresses and ABA application using both cDNA microarray and RNA gel blot analyses. Plant Physiol 133: 1755-1767

Radauer C, Lackner P, Breiteneder H (2008) The Bet v 1 fold: an ancient, versatile scaffold for binding of large, hydrophobic ligands. BMC Evol Biol 8: 286-292

Ramakrishna W, Dubcovsky J, Park YJ, Busso C, Emberton J, San-Miguel P, Bennetzen JL (2002) Different types and rates of genome evolution detected by comparative sequence analysis of orthologous segments from four cereal genomes. Genet 162: 1389-1400

Razem FA, El-Kereamy A, Abrams SR, Hill RD (2006) The RNA-binding protein FCA is an abscisic acid receptor. Nature 439: 290-294

Razem FA, Luo M, Liu JH, Abrams SR, Hill RD (2004) Purification and characterization of a barley aleurone abscisic acid-binding protein. J Biol Chem 279: 9922-9929

Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD, Jenuwein T (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature 406: 593-599

Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings E, Simon I, Zeitlinger J, Schreiber J, Hannett N, Kanin E, Volkert TL, Wilson CJ, Bell SP, Young RA (2000) Genome-wide location and function of DNA binding proteins. Science 290: 2306-2309

Rengasamy (2010) Soil processes affecting production in salt-affected soils. Funct Plant Biol 37: 613-620

Rizhsky L, Liang H, Shuman J, Shulaev V, Davletova S, Mittler R (2004) When defense pathways collide: The response of Arabidopsis to a combination of drought and heat stress. Plant Physiol 134:1683-1696

Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, Zeng T, Euskirchen G, Bernier B, Varhol R, Delaney A, Thiessen N, Griffith OL, He A, Marra M, Snyder M, Jones S (2007) Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. Nat Methods 4: 651-657

Rock CD, Zeevaart JA (1991). The *aba* mutant of *Arabidopsis thaliana* is impaired in epoxy-carotenoid biosynthesis. P Natl A Sci 88: 7496-7499.

Roelfsema MRG, Hedrich R, Geiger D (2012) Anion channels: master switches of stress responses. Trends Plant Sci 4: 221-229

Rook F, Corke F, Card R, Munz G, Smith C, Bevan MW (2001) Impaired sucrose-induction mutants reveal the modulation of sugar-induced starch biosynthetic gene expression by abscisic acid signalling. Plant J 26: 421-433

Ross EM (2008) Coordinating Speed and Amplitude in G-Protein Signaling. Current Biology 18: 777-783

Rubio S, Rodrigues A, Saez A, Dizon MB, Galle A, Kim TH, Santiago J, Flexas J, Schroeder JI, Rodriguez PL (2009) Triple loss of function of protein phosphatases type 2C leads to partial constitutive response to endogenous abscisi acid. Plant Physiol 150: 1345-1355

Rybicki E (2005) PCR primer design and reaction optimisation. In VE Coyne, MD James, SJ Reid and EP Rybicki (eds) Molecular Biology Techniques, Academic press, New York, pp 154-175

Saha J, Chatterjee C, Sengupta A, Gupta K, Gupta B (2013) Genome –wide analysis and evolutionary study of sucrose non-fermenting 1-related protein kinase 2 (SnRK2), gene family members in Arabidopsis and *Oryza*. Comput Biol Chem 49: 59-70

Saito S, Hirai N, Matsumoto C, Ohigashi H, Ohta D, Sakata K, Mizutani M (2004) Arabidopsis *CYP707A*s encode (+)-abscisic acid 8′-hydroxylase, a key enzyme in the oxidative catabolism of abscisic acid. Plant Physiol 134: 1439-1449

Sakamoto T, Miura K, Itoh H, Tatsumi T, Ueguchi-Tanaka M, Ishiyama K, Kobayashi M, Agrawal GK, Takeda S, Abe K, Miyao A, Hirochika H, Kitano H, Ashikari M, Matsuoka M. (2004) An overview of gibberellin metabolism enzyme genes and their related mutants in rice. Plant Physiol 134: 1642-1653

Salamini F, Özkan H, Brandolini A, Schäfer-Pregl R, Martin W (2002) Genetics and geography of wild cereal domestication in the near east. Nat Rev Genet 3: 429-441

Salse J, Feuillet C (2007) Comparative genomics of cereals. In: Varshney R, Tuberosa R (eds) Genomics-Assisted Crop Improvement, The Netherlands: Springer-Verlag, pp 177-205

Sambrook J and Russell DW (2001) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press: New York

Sanan-Mishra N, Pham XH, Sopory SK, Tuteja N (2005) Pea DNA helicase 45 overexpression in tobacco confers high salinity tolerance without affecting yield. P Natl A Sci 102: 509-514

Sansen S, Ranter CJD, Gebruers K, Brijs K, Courtin CM, Delcour JA, Rabijns A (2004) Structural basis for inhibition of *Aspergillus niger* Xylanase by *Triticum aestivum* xylanase inhibitor-I. J Biol Chem 279: 6022-3602

Santiago J, Rodrigues A, Saez A, Rubio S, Antoni R, Dupeux F, Park SY, Márquez JA, Cutler SR, Rodriguez PL (2009a) Modulation of drought resistance by the abscisic acid receptor PYL5 through inhibition of clade A PP2Cs. Plant J 60: 575-588

Santiago J, Dupeux F, Round A, Antoni R, Park SY, Jamin M, Cutler SR, Rodriguez PL, Marquez JA (2009b). The abscisic acid receptor PYR1 in complex with abscisic acid. Nature 462: 665-668

Schachtman DP, Goodge, JQD (2008) Chemical root to shoot signaling under drought. Trends Plant Sci 13: 281-287

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD (2009) The B73 maize genome: complexity, diversity and dynamics. Science 326: 1112-1115

Schneider TD, Stephens RM (1990) Sequence logos: A new way to display consensus sequences. Nucleic Acids Res 18: 6097-6100

Schreiber AW, Sutton T, Caldo RA, Kalashyan E, Lovell B, Mayo G, Muehlbauer GJ, Druka A, Waugh R, Wise RP, Langridge P, Baumann U (2009) Comparative transcriptomics in the *Triticeae*. BMC Genom 10: 285-302

Schulte D, Close TJ, Graner A, Langridge P, Matsumoto T, Muehlbauer G, Sato K, Schulman AH, Waugh R, Wise RP, Stein N (2009) The international barley sequencing consortium - At the threshold of efficient access to the barley genome. Plant Physiol 149: 142-147

Schwartz SH, Leon-Kloosterziel KL, Koornneef M, Zeevaart JA (1997a) Biochemical characterization of the *aba2* and *aba3* Mutants in *Arabidopsis thaliana*. Plant Physiol 114: 161-166.

Schwartz SH, Tan BC, Gage DA, Zeevaart JAD, McCarty DR (1997b) Specific oxidative cleavage of carotenoids by VP14 of Maize. Science 276: 1872-1874

Seiler C, Harshavardhan VT, Rajesh K, Reddy PS, Strickert M, Rolletschek H, Scholz U, Wobus U, Sreenivasulu N (2011) ABA biosynthesis and degradation contributing to ABA homeostasis during barley seed development under control and terminal drought-stress conditions. J Expt Bot 446: 1-18

Seiler C, Harshavardhan VT, Reddy PS, Hensel G , Kumlehn J, Eschen-Lippold L, Rajesh K, Korzun K, Wobus U, Lee J, Selvaraj G, Sreenivasulu N (2014) Abscisic acid flux alterations result in differential abscisic acid signaling responses and impact assimilation efficiency in barley under terminal drought stress. Plant Physiol 164: 1677-1696

Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M, Oono Y, Kamiya A, Nakajima M, Enju A, Sakurai T, Satou M, Akiyama K, Taji T, Yamaguchi-Shinozaki K, Carninci P, Kawai J, Hayashizaki Y, Shinozaki K (2002) Monitoring the

expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J 31: 279-292

Sekimoto H, Seo M, Kawakami N, Komano T, Desloire S, Liotenberg S, Marion-Poll A, Caboche M, Kamiya Y, Koshiba T (1998) Molecular cloning and characterization of aldehyde oxidases in *Arabidopsis thaliana*. Plant Cell Physiol 39: 433-442

Seo M, Koiwai H, Akaba S, Komano T, Oritani T, Kamiya Y, Koshiba T (2000a) Abscisic aldehyde oxidase in leaves of *Arabidopsis thaliana*. Plant J 2: 481-488

Seo M, Peters AJM, Koiwai H, Oritani T, Marion-Poll A, Zeevaart JAD, Koornneef M, Kamiya Y, Koshiba T (2000b) The Arabidopsis aldehyde oxidase 3 (AAO3) gene product catalyzes the final step in abscisic acid biosynthesis in leaves. P Natl A Sci 32: 12908-12913

Seo M, Koshiba T (2002) Complex regulation of ABA biosynthesis in plants. Trends Plant Sci 7: 41-48.

Severin AJ, Woody JL, Bolon YT, Joseph B, Diers BW, Farmer AD, Muehlbauer GJ, Nelson RT, Grant D, Specht JE, Graham MA, Cannon SB, May GD, Vance CP, Shoemaker RC (2010) RNASeq atlas of Glycine max: a guide to the soybean transcriptome. BMC Plant Biol 10: 160-175

Shabala S, Pottosin I (2014) Regulation of potassium transport in plants under hostile conditions: implications for abiotic and biotic stress tolerance. Physiologia Plantarum 151: 257-279

Sheen J (1998) Mutational analysis of protein phosphatase 2C involved in abscisic acid signal transduction in higher plants. P Natl A Sci 95: 975-980

Sheldon CC, Rouse DT, Finnegan EJ, Peacock WJ, Dennis ES (2000) The molecular basis of vernalization: The central role of flowering locus C (FLC). P Natl A Sci 97: 3753-3758

Shen YY, Wang XF, Wu FQ, Du SY, Cao Z, Shang Y, Wang XL, Peng CC, Yu XC, Zhu SY, Fan RC, Xu YH, Zhang DP (2006) The Mg-chelatase H subunit is an abscisic acid receptor. Nature 443: 823-826

Shinozaki K, Yamaguchi-Shinozaki K (2007) Gene networks involved in drought stress response and tolerance. J Expt Bot 58: 221-227

Shinozaki K, Yamaguchi-Shinozaki K (2000) Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. Curr Opin Plant Biol 3: 217-223

Shinozaki K, Yamaguchi-Shinozaki K, Seki M (2003) Regulatory network of gene expression in the drought and cold stress responses. Curr Opin Plant Biol 6: 410-417

Simpson GG, Dijkwel PP, Quesada V, Henderson I, Dean C (2003) FY is an RNA 32 end-processing factor that interacts with FCA to control the Arabidopsis floral transition. Cell 113: 777-787

Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I (2009) ABySS: a parallel assembler for short read sequence data. Genome Res 19: 1117-1123

Sinclair T, Ludlo M (1985) Who taught plants thermodynamics? The unfulfilled

potential of plant water potential. Aust J Plant Physiol 12: 213-217

Sinha AK, Jaggi M, Raghuram B, Tuteja N (2011) Mitogen-activated protein kinase signaling in plants under abiotic stress. Plant Signal Behav 6: 196-203

Skribanek A, Tomcsanyi A (2008) Predicting water stress tolerance of malting barley varieties with seedling PEG-reactions. Acta Biologica Szegediensis 52: 187-189

Skriver K, Mundy J (1990) Gene expression in response to abscisic acid and osmotic stress. Plant Cell 2: 503-512

Smith A (2006) Western region barley variety guide 2006 http://donotpublish.sponsored.uwa.edu.au/lfgn/__data/page/2984/2006_barley_variety_guide.pdf. Accessed May 2014

Smyth GK (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol 3: Article 3

Sokol A, Kwiatkowska A, Jerzmanowski A, Prymakowska-Bosak M (2007) Up-regulation of stress-inducible genes in tobacco and Arabidopsis; cells in response to abiotic stresses and ABA treatment correlates with dynamic changes in histone H3 and H4 modifications. Planta 227: 245-254

Sreenivasulu N, Sopory SK, Kavi Kishor PB (2007) Deciphering the regulatory mechanisms of abiotic stress tolerance in plants by genomic approaches. Gene 388: 1-13

Sridhar VV, Kapoor A, Zhang K, Zhu J, Zhou T, Hasegawa PM, Bressan RA, Zh JK (2007) Control of DNA methylation and heterochromatic silencing by histone H2B deubiquitination. Nature 447: 735-738

Struhl K (1998) Histone acetylation and transcriptional regulatory mechanisms. Genes Dev 12: 599-606

Suganuma T, Workman JL (2011) Signal and combinatorial functions of histone modifications. Ann Rev Biochem 80: 473-499

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

Szostkiewicz I, Richter K, Kepka M, Demmel S, Ma Y, Korte A, Assaad FF, Christmann A, Grill E (2010) Closely related receptor complexes differ in their ABA selectivity and sensitivity. Plant J 61: 25-35

Talame V, Ozturk NZ, Bohnert HJ, Tuberosa R (2007) Barley transcript profiles under dehydration shock and drought stress treatments: a comparative analysis. J Expt Biol 58: 229-240

Tamura K, Yonemaru J (2010) Next-generation sequencing for comparative transcriptomics of perennial ryegrass (*Lolium perenne L.*) and meadow fescue (*Festuca pratensis* Huds.) during cold acclimation. Grassland Science 56: 230-239

Tan BC, Schwartz SH, Zeevaart JAD, McCarty DR (1997) Genetic control of abscisic acid biosynthesis in maize. P Natl A Sci 94: 12235-12240

Tanake H, Osakabe Y, Katsura S, Mizuno S, Maruyama K, Kusakabe K, Mizoi J, Shinozaki K, Shinozaki KY (2012) Abiotic stress-inducible receptor like kinases negatively control ABA signaling in Arabidopsis. Plant J 4: 599-613

Taunton J, Hassig CA, Schreiber SL (1996) A mammalian histone deacetylase related to the yeast transcriptional regulator *Rpd3p*. Science 272: 408-11

Taylor HF, Burden RS (1972) Xanthoxin, a recently discovered plant growth inhibitor. Proceedings of the Royal Society of London. Series B. Biological Sciences 180: 317-346.

Taylor IB, Linforth RST, Al-Naieb RJ, Bowman WR, Marples BA (1988) The wilty tomato mutants flacca and sitiens are impaired in the oxidation of ABA-aldehyde to ABA. Plant Cell Environ 11: 739-745.

Tewolde H, Fernandez CJ, Erickson CA (2006) Wheat cultivars adapted to post-heading high temperature stress. J Agron Crop Sci 192: 111-120

The International Wheat Genome Sequencing Consortium (IWGSC), Mayen KX (2014) A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Sci 345: 6194-6199

Tian S, Mao X, Zhang H, Chen S, Zhai C, Yang S, Jing R (2013) Cloning and characterization of *TaSnRK2.3*, a novel SnRK2 gene in common wheat. J Exp Bot 64: 2063-2080

Tilman D, Balzer C, Hill J, Belfort BL (2011) Global food demand and the sustainable intensification of agriculture. P Natl A Sci 108: 20260-20264

Travis J (2008) Uncorking the grape genome. Science 320: 475-477

Turner NC (1981) Techniques and experimental approaches for the measurement of plant water status. Plant Soil 58: 339-366

Tyerman S, Bohnert H, Maurel C, Steudle E, Smith J (1999) Plant aquaporins: their molecular biology, biophysics and significance for plant water relations. J Expt Bot 50: 1055-1071

Ueda A, Kathiresan A, Inada M, Narita Y, Nakamura T, Shi W, Takabe T, Bennett J (2004) Osmotic stress in barley regulates expression of a different set of genes than salt stress does. J Expt Bot 55: 2213-2218

Ullrich S (2011) In: Ullrich SE (ed) Barley: Production, improvement and uses. John Wiley and sons, West Sussex, UK, pp 3-13

Umate P, Tuteja N, Tuteja R (2011) Genome-wide comprehensive analysis of human helicases Comm Inter Biol 4: 118-137

Umezawa T, Nakashima K, Miyakawa T, Kuromori T, Tanokura M, Shinozaki K, Yamaguchi-Shinozaki K (2010) Molecular basis of the core regulatory network in ABA responses: sensing, signaling and transport. Plant Cell Physiol 51: 1821-1839

Umezawa T, Sugiyama N, Mizoguchi M, Hayashi S, Myouga F, Yamaguchi-Shinozaki K, Ishihama Y, Hirayama T, Shinozaki K (2009) Type 2C protein phosphatases directly regulate abscisic acid-activated protein kinases in Arabidopsis. P Natl A Sci 106: 17588-17593

Umezawa T, Yoshida R, Maruyama K, Yamaguchi-Shinozaki K, Shinozaki K (2004) SRK2C, a SNF1-related protein kinase 2, improves drought tolerance by controlling stress-responsive gene expression in *Arabidopsis thaliana*. P Natl A Sci 101: 17306-17311

Uno Y, Furihata T, Abe H, Yoshida R, Shinozaki K, Yamaguchi-Shinozaki K (2000) Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. P Natl A Sci 97: 11632-11637

Urano K, Maruyama K, Ogata Y, Morishita Y, Takeda M, Sakurai N, Suzuki H, Saito K, Shibata D, Kobayashi M, Yamaguchi-Shinozaki K, Shinozaki K (2009) Characterization of the ABA-regulated global responses to dehydration in Arabidopsis by metabolomics. Plant J 57: 1065-1078

van den Brûle S, Smart C (2002) The plant PDR family of ABC transporters. Planta 216: 95-106

van Gool D, Vernon L (2006) Potential impacts of climate change on agricultural land use suitability: Lupins. Resource Management Technical Report 305. Department of Agriculture, Western Australia, South Perth

Vashisht AA, Pradhan A, Tuteja R, Tuteja N (2005) Cold and salinity stress-induced bipolar pea DNA helicase 47 is involved in protein synthesis and stimulated by phosphorylation with protein kinase C. Plant J 44: 76-87

Vasil IK (2007) Molecular genetic improvement of cereals: transgenic wheat (*Triticum aestivum L.*). Plant Cell Rep 26: 1133-1154

Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG, Mann M. (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. Cell 142: 967-980

Vlad F, Rubio S, Rodrigues A, Sirichandra C, Belin C, Robert N, Leung J, Rodriguez PL, Lauriere C, Merlot S (2009) Protein phosphatases 2C regulate the activation of the Snf1-related kinase OST1 by abscisic acid in Arabidopsis. Plant Cell 21: 3170-3184

Wahid A, Farooq M, Basra SMA, Rasul E, Siddique KHM (2010) Germination of seeds and propagules under salt stress. In: Pessarakli M (ed) Hand book on plant and crop stress, CRC Press, Boca Raton, USA, pp 321-337

Walia H, Wilson C, Condamine P, Ismail A, Cui X, Close T (2007) Array-based genotyping and expression analysis of barley cv Maythorpe and Golden Promise. BMC Genom 8: 87-101

Walia H, Wilson C, Condamine P, Liu X, Ismail AM, Zeng L, Wanamaker SI, Mandal J, Xu J, Cui X, Close TJ (2005) Comparative transcriptional profiling of two contrasting rice genotypes under salinity stress during the vegetative growth stage. Plant Physiol 139: 822-835

Walia H, Wilson C, Wahid A, Condamine P, Cui X, Close TJ (2006) Expression analysis of barley (*Hordeum vulgare* L.) during salinity stress. Funct Integr Genom 6: 143-156

Walker CJ, Willows RD (1997) Mechanism and regulation of Mg-chelatase. Biochem J 327: 321-333

Walker-Simmons MK, Anderberg RJ, Rose PA, Abrams SR (1992) Optically pure abscisic acid analogs-tools for relating germination inhibition and gene expression in wheat embryos. Plant Physiol 99: 501-507

Wang W, Vinocur B, Altman A (2003) Plant responses to drought, salinity and extreme temperature: towards genetic engineering for stress tolerance. Planta 218: 1-14

Wang W, Vinocur B, Shoseyov O, Altman A (2004) Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends Plant Sci 9: 244-252

Wang Z, Gerstein M, Snyder M (2009a) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10: 57-63

Wang X, Ellingc AA, Lib X, Lie N, Penga Z, Heb G, Sunc H, Qib Y, Liud XS, Deng XW (2009b) Genome-wide and organ-specific landscapes of epigenetic modifications and their relationships to mRNA and Small RNA transcriptomes in maize. A Soc Plant Biol 21: 1053-1069

Wang XF, Zhang DP (2008) Abscisic acid receptors: Multiple signal-perception sites. Annals of Botany 101: 311-317

Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh TY, Peng W, Zhang MQ, Zhao K (2008) Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet 40: 897-903

Wasilewska A1, Vlad F, Sirichandra C, Redko Y, Jammes F, Valon C, Frey FN, Leung J (2008) An update on abscisic acid signaling in plants and more. Mol Plant 1: 198-217

Watkinson JI, Sioson AA, Vasquez-Robinet C, Shukla M, Kumar D, Ellis M, Heath LS, Ramakrishnan N, Chevone B, Watson LT, Van ZL, Egertsdotter U, Sederoff RR, Grene R (2003) Photosynthetic acclimation is reflected in specific patterns of gene expression in drought-stressed Loblolly Pine. Plant Physiol 133: 1702-1716

Weber APM, Weber KL, Carr K, Wilkerson C, Ohlrogge JB (2007) Sampling the Arabidopsis transcriptome with massively parallel pyrosequencing. Plant Biol 144: 32-42

Wei Y, Mizzen CA, Cook RG, Gorovsky MA, Allis CD (1998) Phosphorylation of histone H3 at serine 10 is correlated with chromosome condensation during mitosis and meiosis in *Tetrahymena*. P Natl A Sci 95: 7480-7484

Wicker T, Schlagenhauf E, Graner A, Close TJ, Keller B, Stein N (2006) 454 sequencing put to the test using the complex genome of barley. BMC Genom 7: 275-286

Wiebe GA, Reid DA (1961) Classification of barley varieties. In: GA Weibe and DA Reid (eds) Classification of barley varieties grown in United States and Canada in 1958. United States Department of Agriculture Press, Washington DC, pp 69-217

Wild A (2003) Soils, land and food: managing the land during the twenty-first century. Ann Bot London 93: 785-786

Winter W, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ (2007) An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. P Natl A Sci 2: 718-724

Woodger FJ, Jacobsen JV, Gubler F (2004) GMPOZ, a BTB/POZ domain nuclear protein, is a regulator of hormone responsive gene expression in barley aleurone. Plant Cell Physiol 45: 945-950

Xhemalce B, Dawson MA, Bannister AJ. (2011) Histone modifications In: Meyers R (ed). Encyclopedia of Molecular Cell Biology and Molecular Medicine. John Wiley and Sons, New York, pp 354-365

Xiang Y, Tang N, Du H, Ye H, Xiong L (2008) Characterization of OsbZIP23 as a key player of the basic leucine zipper transcription factor family for conferring abscisic acid sensitivity and salinity and drought tolerance in rice. Soc Plant Biol 148: 1938-1952

Xiong L, Zhu JK (2003) Regulation of abscisic acid biosynthesis. Plant Physiol 133: 29-36

Xiong LM, Lee BH, Ishitani M, Lee H, Zhang CQ, Zhu JK (2001) FIERY1 encoding an inositol polyphosphate 1-phosphatase is a negative regulator of abscisic acid and stress signaling in Arabidopsis. Genes Dev 15: 1971-1984

Xu D, Duan X, Wang B, Hong B, Ho T, Wu R (1996) Expression of a late embryogenesis abundant protein gene, HVA1, from barley confers tolerance to water deficit and salt stress in transgenic rice. Plant Physiol 110: 249-257

Xu ZY, Lee KH, Dong T, Jeong JC,Jin JB, Kanno Y, Kim DH, Kim SY, Seo M, Bressan RA, Yun DJ, Hwang I (2012a) A vacuolar beta-glucosidase homolog that possesses glucose-conjugated abscisic acid hydrolyzing activity plays an important role in osmotic stress responses in Arabidopsis. Plant Cell 24: 2184-2199

Xu YH, Liu R, Yan L, Liu ZQ, Jiang SC, Shen YY, Wang XF, Zhang DP (2012b) Light-harvesting chlorophyll a/b binding proteins are required for stomatal response to abscisic acid in Arabidopsis. J Expt Bot 63: 1095-1106

Xue T, Wang D, Zhang S, Ehlting J, Ni F, Jakab S, Zheng C, Zhong Y (2008) Genome wide and expression analysis of protein phosphatase 2C in rice and Arabidopsis. BMC Genom 9: 550-570

Yamaguchi-Shinozaki K, Shinozaki K (1994) A novel *cis*-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress. Plant Cell 6: 251-264

Yang J, Zhang J, Wang Z, Zhu Q, Wang W (2001) Hormonal changes in the grains of rice subjected to water stress during grain filling. Plant Physiol 127: 315-323

Yauk CL, Berndt ML, Williams A, Douglas GR (2004) Comprehensive comparison of six microarray technologies. Nucleic Acids Res 32: 124-130

Yin P, Fan H, Hao Q, Yuan X, Wu D, Pang Y, Yan C, Li W, Wang J, Yan N (2009) Structural insights into the mechanism of abscisic acid signaling by PYL proteins. Nat Struct Mol Biol 16: 1230-1236

Yoshida R, Hobo T, Ichimura K, Mizoguchi T, Takahashi F, Aronso J, Ecker JR, Shinozaki K (2002) ABA-activated SnRK2 protein kinase is required for dehydration stress signaling in Arabidopsis. Plant Cell Physiol 43: 1473-1483

Yoshida R, Umezawa T, Mizoguchi T, Takahashi S, Takahashi F, Shinozaki K (2006) The regulatory domain of SRK2E/OST1/SnRK2.6 interacts with ABI1 and integrates abscisic acid (ABA) and osmotic stress signals controlling stomatal closure in Arabidopsis. J Biol Chem 281: 5310-5318

Young MD, Willson TA, Wakefield MJ, Trounson E, Hilton DJ, Blewitt ME, Oshlack A, Majewski IJ (2011) ChIP-seq analysis reveals distinct H3K27me3 profiles that correlate with transcriptional activity. Nucleic Acids Res 39: 7415-27

Zeevaart JAD, Creelman RA (1988) Metabolism and physiology of abscisic acid. Annu Rev Plant Phys 39: 439-473

Zenoni S, Ferrarini A, Giacomelli E, Xumerle L, Fasoli M, Malerba G, Bellin D, Pezzotti M, Delledonne M (2010) Characterization of transcriptional complexity during berry development in *Vitis vinifera* using RNA-Seq. Plant Physiol 152: 1787-1795

Zhang DP, Wu ZY, Li XY, Zhao ZX (2002) Purification and identification of a 42-Kilodalton abscisic acid-specific-binding protein from epidermis of broad bean leaves. Plant Physiol 128: 714-725

Zhang JZ, Creelman RA, Zhu JK (2004) From laboratory to field. Using information from Arabidopsis to engineer salt, cold, and drought tolerance in crops. Plant Physiol 135: 615-621

Zhang M, Lu S, Li G, Mao Z, Yu X, Sun W, Tang Z, Long M, Su W (2010) Identification of a residue in helix 2 of rice plasma membrane intrinsic proteins that influences water permeability. J Biol Chem 285: 41982-41992

Zhang M, Xie S, Dong X, Zhao X, Zeng B, Chen J, Li H, Yang W, Zhao H, Wang G, Chen Z, Sun S, Hauck A, Jin W, Lai J (2014) Genome-wide high resolution parental-specific DNA and histone methylation maps uncover patterns of imprinting regulation in maize. Genet Res 24: 167-176

Zhang W, Wu Y, Schnable JC, Zeng Z, Freeling M, Crawford GE, Jiang J (2012a) High-resolution mapping of open chromatin in the rice genome. Genome Res 22: 151-162

Zhang W, Ruan J, Ho THD, You Y, Yu T, Quatrano RS (2005) *Cis*-regulatory element based targeted gene finding: genome-wide identification of abscisic acid- and abiotic stress-responsive genes in *Arabidopsis thaliana*. Bioinformatics 21: 3074-3081

Zhang W, Zhang T, Wu Y, Jiang J (2012b) Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in Arabidopsis. Plant Cell 24: 2719-2731

Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, Jacobsen SE (2007) Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis. P Natl A Sci Bio 10: 785-792

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS (2008) Model-based Analysis of ChIP-Seq (MACS). Genome Biol 9: 137-175

Zhao J, Ren W, Zhi D, Wang L, Xia G (2007) Arabidopsis *DREB1A*/*CBF3* bestowed transgenic tall fescue increased tolerance to drought stress. Plant Cell Rep 26: 1521-1528

Zheng J, Fu J, Gou M, Huai J, Liu Y, Jian M, Huang Q, Guo X, Dong Z, Wang H, Wang G (2010) Genome-wide transcriptome analysis of two maize inbred lines under drought stress. Plant Mol Biol 72: 407-421

Zhong X, Zhang H, Zhao Y, Sun Q, Hu Y , Peng H, Zhou DX (2013) The rice NAD+-dependent histone deacetylase OsSRT1 targets preferentially to stress and metabolism-related genes and transposable elements. P Natl A Sci 10: 371-379

Zhou J, Wang X, Jiao Y, Qin Y, Liu X, He K, Chen C, Ma L, Wang J, Xiong L, Zhang Q, Fan L, Deng WX (2007) Global genome expression analysis of rice in response to drought and high salinity stresses in shoot, flag leaf and panicle. Plant Mol Biol 63: 591-608

Zhou L, Liu Y, Liu Z, Kong D, Duan M, Luo L (2010) Genome-wide identification and analysis of drought-responsive microRNAs in *Oryza sativa*. J Expt Bot 61: 4157-468

Zhou MX, Li HB, Chen ZH and Mendham NJ (2008) Combining ability of barley flour pasting properties. J Cereal Sci 48: 789-793

Zhou R, Cutler AJ, Ambrose SJ, Galka MM, Nelson KM, Squires TM, Loewen K, Jadhav AS, Ross ARD, Taylor DC, Abrams SR (2004) A new abscisic acid catabolic pathway. Plant Physiol 134: 361-369

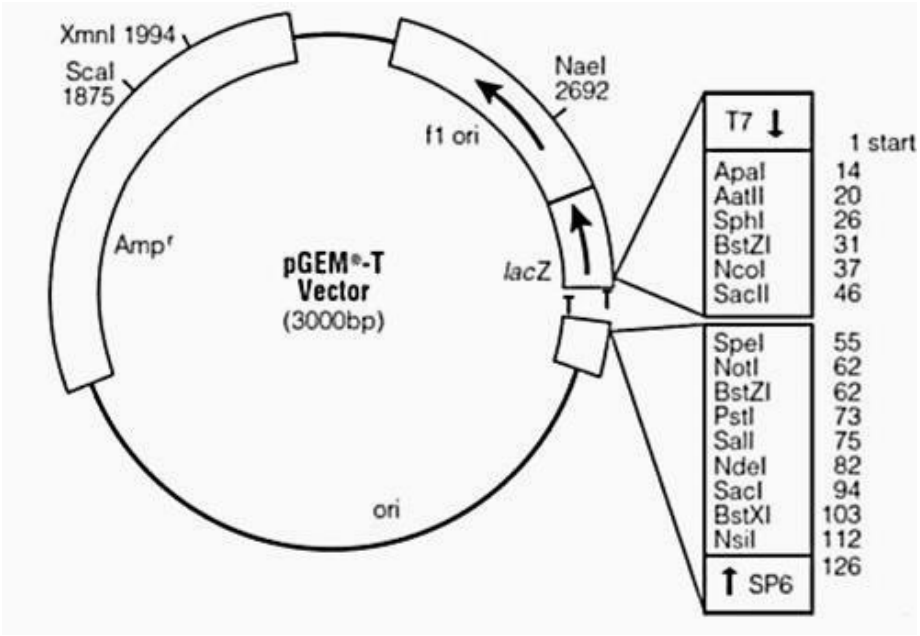
Ziemann M, Kamboj A, Bhave M (2013b) OrthoRBH: A streamlined pipeline for mining large gene family sequences in related species. Bioinfo 9: 267-269

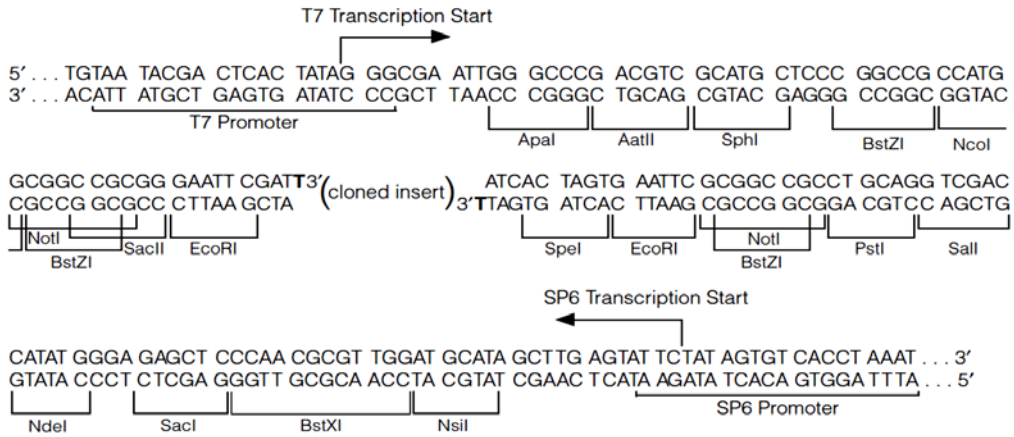
Ziemann M, Kamboj A, Hove RM, Loveridge S, El-Osta A, Bhave M (2013a) Analysis of the barley leaf transcriptome under salinity stress using mRNA-Seq. Acta Physiol Plant 35: 1915-24

Zohary D, Hopf M (2000) Domestication of plants in the old world: the origin and spread of cultivated plants in West Asia, Europe, and the Nile Valley, 3rd edn. New York, Oxford University Press. pp 15-57

Appendices

Appendix I pGEM®-T Easy vector map





The T7 and SP6 promoter sequences were used as primer annealing sites for amplification and insert sequencing.

Source:

http://www.promega.com/~/media/files/resources/protocols/technical%20manuals/0/pgem-t%20and%20pgem-t%20easy%20vector%20systems%20protocol.pdf?la=en; last accessed August 2014

Appendix II Details of mRNA-seq procedure

A2.1 Library preparation

A2.1.1 Purification of mRNA from total RNA

This step allowed the separation of poly-A tailed mRNA from total RNA, and was conducted using poly-T magnetic beads (Ilumina).

1. The DNA-free total RNA (isolated as in section 2.5.2) from two separate salt-stressed plants was pooled in equal amounts (0.5 µg each). The pooled RNA (1 µg) was diluted to 50 µL with nuclease-free water in a 1.5 mL tube, incubated at 65 °C for 5 minutes to disrupt the secondary structures, then chilled on ice.
2. Step 1 was also carried out for RNA from two control (untreated) plants.
3. 15 µL of the Sera-Mag oligo (dT) beads were washed twice with 100 µL of Bead Binding Buffer*, supernatant removed after the beads have fully pelleted against the magnetic stand and the beads resuspended in 50 µL of Bead Binding Buffer*.
4. The 50 µL pooled RNA sample was added to this suspension, the tube rotated at RT for 5 minutes and the 100 µL supernatant removed. The beads were then washed twice with 200 µL of Washing Buffer* and the supernatant removed after the beads have fully pelleted to the magnetic stand.
5. 50 µL of 10 mM Tris-HCl* was then added to the beads and heated on a heating block at 80 ºC for 2 minutes to elute the mRNA from the beads. The tube was then immediately placed on the magnetic stand and the 50 µL supernatant containing the mRNA population transferred to a 1.5 mL tube containing 50 µL of Bead Binding Buffer making a total volume of 100 µL.
6. The tube containing the 100 µL mRNA mixed with Bead Binding Buffer* eluted in step 5 was heated on a heat block at 65 ºC for 5 minutes in order to disrupt the secondary structures and then chilled on ice.
7. The used beads from step 5 were washed twice with 200 µL of Washing Buffer* and the supernatant removed.
8. The chilled 100 µL mRNA mixed with Bead Binding Buffer* from step 6 was added back to the washed beads in step 7, rotated at RT for 5 minutes and the supernatant removed. The beads were washed again twice with 200 µL of Washing Buffer and the supernatant removed.
9. 17 µL of 10 mM Tris-HCl* was added to the washed beads and heated on a heating block at 80 ºC for 2 minutes to elute the mRNA. After heating, the tube was placed on a

magnet stand and the supernatant containing the mRNA (approx. 16 µL) transferred to a fresh 200 µL tube.

A2.1.2 mRNA fragmentation

This procedure allowed the fragmentation of mRNA into small pieces using divalent cations and increased temperatures.

10. 4 µL of 5× Fragmentation Buffer* was added to the 16 µL purified mRNA from step 9 and the tube incubated in a PCR thermal cycler at 94 ºC for 5 minutes.

11. 2 µL of Fragmentation Stop Solution* was added, the mixture chilled on ice and transferred to a 1.5 ml tube.

12. 3M sodium acetate, pH 5.2 (2 µL), 20 mg/mL glycogen* (2 µL) and 100% ethanol (60 µL) were added to the tube and chilled at -80 ºC for 30 minutes.

13. The mixture was centrifuged at 14,000 rpm for 25 minutes at 4 ºC and the ethanol pipetted carefully to avoid dislodging the RNA pellet.

14. The pellet was washed with 300 µL 70% ethanol, centrifuged and the ethanol removed carefully. It was then air-dried for 10 minutes and resuspended in 11.1 µL of RNase-free water in a 200 µL tube.

A2.1.3 Synthesis of first strand cDNA

This step allowed the synthesis of first strand cDNA using mRNA as template, with the help of reverse transcriptase and random primers.

15. The fragmented mRNA (11.1 µL) was mixed with random primers* (1 µL) and incubated at 65 ºC for 5 minutes, then chilled on ice.

16. 5× First Strand Buffer (4 µL; Invitrogen), 100 mM DTT (2 µL; Invitrogen), 25 mM dNTP mix* (0.4 µL) and RNase inhibitor (0.5 µL) were added to the above mixture, mixed well and incubated on a heating block at 25 ºC for 2 minutes.

17. 1 µL of SuperScript™ II reverse transcriptase (200U/µL; Invitrogen) was added to the sample (total volume 20 µL) and the tube incubated as follows on a PCR thermal cycler: 25 ºC for 10 minutes, 42 ºC for 50 minutes, 70 ºC for 15 minutes, then hold at 4 ºC. The tube was then chilled on ice.

A2.1.4 Synthesis of the second strand cDNA and purification of double-stranded cDNA

This procedure allowed the synthesis of the second strand cDNA by removal of the RNA template and then the synthesis of DNA strand using the first strand cDNA as template.

18. 62.8 µL of ultra-pure water*, 10 µL of GEX Second Strand Buffer* and 1.2 µL of 25 mM dNTP Mix* were added to the 20 µL first strand cDNA preparation (section A2.1.3) and the mixture held on ice for 5 minutes.

19. 1 µL of RNaseH* (5U/µL) and 5 µL of DNA Polymerase I (10U/µL) were then added (total volume 100 µL), mixed well and incubated at 16 ºC on a thermocycler for 2.5 hours. The double-stranded cDNA was then purified from the above mixture using the QIAquick PCR purification kit (Qiagen). 5 volumes (500 µL) of Buffer PB were added to 1 volume (100 µL) of the second strand cDNA mix.

20. The mixture was applied to a QIAquick column assembly and centrifuged at 13,000 rpm for 60 seconds. The flow-through was discarded, the column washed with 750 µL of Buffer PE and centrifuged at 13,000 rpm for 60 seconds. The flow-through was discarded and the column centrifuged for another 60 seconds. The dry column was placed in a clean 1.5 mL tube and the bound double-stranded cDNA was eluted in 50 µL Buffer EB by centrifugation.

A2.1.5 End repair and adenylation of the 3´ end

End repair involves conversion of any overhangs created on the cDNA during the above steps to blunt ends using T4 DNA polymerase and Klenow DNA polymerase. The 3´adenylation step then adds an ‘A’ nucleotide to both 3´ ends of the cDNA strands in preparation for ligation to the adapters, which have one ‘T’ overhang.

21. End-repair was carried out in a 100 µL reaction volume, consisting of the 50 µL double-stranded cDNA (above), 27.4 µL ultra-pure water, 10 µL 10× End Repair Buffer*, 1.6 µL 25 mM dNTP Mix*, 5 µL T4 DNA Polymerase*, 1 µL Klenow DNA Polymerase* and 5 µL T4 Polynucleotide Kinase*. The mixture was incubated in a heat block at 20 ºC for 30 minutes.

22. The blunt-ended cDNA was purified from the mixture using the QIAquick PCR purification kit (Qiagen) as above, except that it was eluted in 32 µL of Buffer EB.

23. The 3´adenylation was carried out in 50 µL volume, consisting of 32 µL of above cDNA, 5 µL of A-Tailing Buffer*, 10 µL of 1 mM dATP* and 3 µL of Klenow exo* (3´ to 5´ exo minus). The mixture was incubated on a heating block at 37 ºC for 30 minutes.
24. The 3´A-cDNA was purified using the MinElute PCR Purification kit (Qiagen). Five volumes (250 µL) of Buffer PB were added to 1 volume (50 µL) of the 3´ adenylation mix. The mixture was applied to a MinElute column assembly and centrifuged at 13,000 rpm for 60 seconds. The flow-through was discarded and the column washed with 750 µL of Buffer PE and centrifuged. The flow-through was discarded and the column centrifuged for an 1 minute.
25. The dry MinElute column was placed in a clean 1.5 ml tube and the double-stranded cDNA was eluted in 23 µL of Buffer EB.

A2.1.6 Ligation of the adapters

This step prepares the cDNA fragments for hybridization (binding) to the flow-cells.
26. The ligation of adapters was carried out in a 50 µL volume consisting of 23 µL of the eluted 3´A cDNA, 25 µL of 2× Rapid T4 DNA Ligase Buffer*, 1 µL PE Adapter Oligo Mix* and 1 µL T4 DNA Ligase*. The mixture was incubated at RT for 15 minutes. The adapter-ligated cDNA molecules were purified using the MinElute PCR Purification kit (Qiagen) as above, except that they were eluted in 10 µL of Buffer EB.

A2.1.7 Gel purification of the cDNA templates

This step allows selecting a size range of the cDNAs for later enrichment.
27. The above cDNA (10 µL) was mixed with 2 µL of 6× Loading Dye and electrophoresed on a 2% agarose gel at 120 V for 60 minutes, alongside a 100 bp DNA Ladder. The DNA population in the 200 bp size range was excised from the gel and purified using the QIAquick Gel Extraction kit (Qiagen).
28. 3 volumes of Buffer QG was added to 1 unit weight of the excised gel slice (e.g., 0.3 mL Buffer QG for a 0.1g gel slice) and incubated at 50 ºC for 10 minutes to melt the gel slice.
29. 1 gel volume (e.g., 100 µL of isopropanol for a 0.1g gel slice) of isopropanol was added to precipitate the cDNA.

30. The mixture was applied to the QIAquick column assembly and centrifuged at 13,000 rpm for 60 seconds. The flow-through was discarded, 500 µL of Buffer QG was added to the column and centrifuged as above.
31. The flow-through was discarded, the column washed with 750 µL of Buffer PE and centrifuged as above. The flow-through was discarded and the column centrifuged for 60 seconds.
32. The dry column was placed in a clean 1.5 mL tube and the bound cDNA eluted in 30 µL of Buffer EB.

A2.1.8 PCR enrichment of the purified cDNA templates

The procedure involves the use of the polymerase chain reaction (PCR) to amplify the cDNA templates, in order to create the final cDNA library.
33. The reactions were carried out in 50 µL volumes consisting of 30 µL of purified cDNA (above), 10 µL of 5× Phusion Buffer (Finnzymes), 1 µL of PCR Primer PE 1.0 (5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT 3′), 1 µL of PCR Primer PE 2.0 (5' CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT 3′), 0.5 µL of 25 mM dNTP Mix*, 0.5 µL of Phusion DNA Polymerase (2U/µL, Finnzymes) and 7 µL of ultra-pure water*.

34. Amplification was conducted using the following cycling conditions: initial denaturation (98 ºC, 30 seconds), then 15 cycles of denaturation (98 ºC, 10 seconds), annealing (65 ºC, 30 seconds) and extension (72 ºC, 30 seconds), then a final extension (72 ºC, 5 minutes).
35. The amplified cDNA was then purified using the QIAquick PCR purification kit (Qiagen) as detailed above, except that it was eluted in 30 µL of Buffer EB. Before cluster amplification, the size, quality and concentration of this DNA were tested on the MCE®-202 MultiNA, Microchip Electrophoresis System (Shimadzu, Japan) as described in section 2.5.5.

A2.2 Cluster generation

A2.2.1 Sample preparation for cluster generation

This step involves denaturation and dilution of the cDNA template. Firstly, the concentration of the eluted DNA (step 35 above) was adjusted to 10 nM (~1.9 ng/μL) using 10 mM Tris-Cl, pH 8.5. 2 μL of the 10 nM template DNA prepared above was mixed with 10mM Tris-Cl, pH8.5 (17 μL) and HP3 (2 N NaOH; see below) (1 μL) in a 20 μL volume, to give a final template cDNA concentration of 1 nM. The mixture was briefly vortexed, pulse-centrifuged and incubated at RT for 5 minutes in order to denature the double-stranded cDNA to single strands. The denatured DNA was diluted with chilled Hybridization Buffer (HT1; see below) to a total volume of 1 000 μL and a final DNA concentration of 15 pM. The reagents were mixed by inversion and pulse-centrifuged. 120 μL of the denatured cDNA was added to an eight-tube strip, labelled 'B'.

A2.2.2 Preparation of reagents for cluster amplification

Amplification Mix (AMX1): The AMX1 was mixed by inversion and centrifuged at 1,000 rpm for 60 seconds. The tube labelled 'Reagent 1', the weight recorded, and chilled on ice.

Wash Buffer (HT2): HT2 was mixed and centrifuged as above, labelled 'Reagent 3' and the weight recorded. 100 μL of HT2 was pipetted into an eight-tube strip labelled 'C'. Another 100 μL of HT2 was pipetted into an eight-tube strip labelled 'G'. Both were stored at RT.

100% Formamide (AT1): AT1 was mixed and centrifuged as above. The weight was recorded; the reagent labelled 'Reagent 9' and stored RT.

AMX1 Premix (APM1): APM1 was mixed by inversion, labelled 'Reagent 11' and the weight recorded. 100 μL was pipetted into an eight-tube strip labelled 'D' and stored at RT.

Hybridisation Buffer (HT1): HT1 was mixed and centrifuged as above, labelled 'Reagent 12' and the weight recorded. 140 μL was pipetted into an eight-tube strip labelled 'A'.

1× Phusion Master Mix, Finnzymes Oy (HFE): HFE was mixed by inversion, pulse centrifuged and 120 μL pipetted into an eight-tube strip labelled 'E'.

2 N NaOH (HP3): HP3 was mixed by inversion, pulse centrifuged and diluted to 0.1 N NaOH in PW1 (Wash Solution). The diluted HP3 was pipetted into an eight-tube strip labelled 'F'.

A2.2.3 Cluster amplification

Before amplification, the Cluster Station was washed with deionized water and the air gaps removed. The reagents were loaded in positions 1 (AMX1), 3 (HT2), 9 (AT1), 11 (APM1) and 12 (HT1) (Figure A2.1). 'OK' was then clicked to resume the amplification recipe v7 (SR_Amplification_only_v<#>.xml). When prompted, the flow cell v4, hybridisation manifold and the tube strip holder were loaded onto the Station. The tube strip labelled 'A' (HT1) was loaded into the holder and OK was clicked. When prompted by the amplification recipe, tubes B (template mix), C (HT2), D (APM1), E (HFE), F (diluted HP3) and G (HT2) were loaded onto the tube strip holder in that order. When prompted, the manifold inlets were removed, 'OK' clicked to proceed and then the amplification manifold was loaded instead of the hybridisation manifold. Amplification was allowed to resume by clicking 'OK'. After amplification, the cluster station was washed with deionized water.

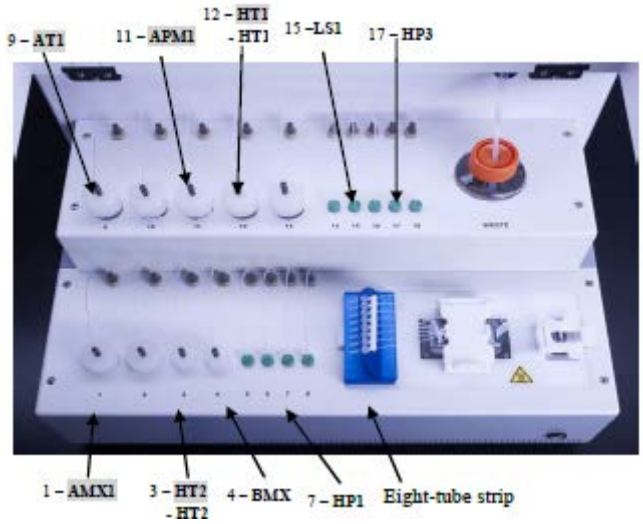


Figure A2.1: Cluster station showing the reagent positions
Shown in bold are the reagents loaded onto the particular positions. Highlighted in grey are the reagents used for cluster amplification, while the ones not highlighted are the ones used for linearisation, blocking and primer hybridisation.

A2.2.4 Preparation of reagents for linearisation, blocking and primer hybridisation

Wash Buffer (HT2): HT2 ('Reagent 3') from the amplification step was used.

Blocking Mix (BMX): BMX was mixed by inversion, pulse centrifuged, weighed, and the container labelled 'Reagent 4'.

Sequencing Primer Mix (HP1): The Sequencing Primer Buffer (HP4) was mixed by inversion and 6.6 µL of Genomic DNA Sequencing Primer (5' ACACTCTTTCCCTACACGACGCTCTTCCGATCT 3′) was added. The mixture was vortexed, pulse centrifuged, weight recorded, and the reagent labelled 'Reagent 7'.

Hybridisation Buffer (HT1): The Hybridisation Buffer (HT1) from the amplification step was used. The weight of the reagent was recorded and labelled 'Reagent 12'.

Linearisation Solution (LS1): 1.4 µL of 3-amino-1-propanol (3-APL) was added to LS1, mixed by inversion, pulse centrifuged and the weight recorded. The tube was labelled 'Reagent 15'and chilled on ice.

2 N NaOH (HP3): HP3 was mixed by inversion, pulse centrifuged and diluted to 0.1 N NaOH in PW1 (Wash Solution). The diluted HP3 was labelled 'Reagent 17' and its weight was recorded.

A2.2.5 Linearisation, blocking and primer hybridisation

Before these procedures were carried out, the Cluster Station was washed with deionized water and the air gaps removed. The reagents were loaded in positions 3 (HT2), 4 (BMX), 7 (HP1), 12 [(HT1), 15 (LS1) and 17 (HP3) (Figure A2.1). After loading the reagents, 'OK' was clicked to resume the linearisation, blocking and hybridization recipe v7 (SR_Linearisation_Blocking_PrimerHyb_v<#>.xml). When prompted, the amplification manifold was loaded onto the Station and OK clicked to proceed. After completion of the recipe, the flow cell v4 was removed from the Station for sequencing. The Cluster station was washed with deionized water.

A2.3 Sequencing

A2.3.1 Preparation of reagents for sequencing

Incorporation Mix (IMX36): 3.52 ml of Long-Read FFN (LFN36) was added to the Incorporation Mix (IMX36). The Rapid DNA Polymerase (RDP36) was pulse-centrifuged prior to adding 220 µL of it to the IMX36+LFN36 mix. The IMX36+LFN36+RDP36 reagent was mixed by inversion, centrifuged at 3,900 rpm for 60 seconds at RT and the weight recorded.

High Salt Buffer (PR1), Incorporation Buffer (PR2) and Cleavage Buffer (PR3): These buffers were mixed by inversion and their weights recorded.

Scan Mix (SMX36): SMX36 was mixed by inversion and its weight recorded.

Cleavage Mix (CLM36): CLM36 was mixed by inversion, centrifuged at 3,900 rpm for 60 seconds at RT, and the weight of the reagent recorded.

A2.3.2 Pre-run wash and priming reagents

The Genome Analyzer IIx (GAIIx) was loaded with a used flow cell, 10 mL of PW1 at positions 1, 3 and 6 and 40 mL of PW1 in Nalgene bottles at positions 2, 4, 5 and 7 as shown in Figure A2.2. The Genome Analyzer bundled tubes were placed into a 50 mL tube and the pre-wash run recipe v7 (GA2_PreWash_v<#>.xml) was run. After the pre-run wash, the reagents were loaded at positions 1 (IMX36), 2 (PW1), 3 (SMX36), 4 (PR1), 5 (PR2), 6 (CLM36) and 7 (PR3) (Figure A2.2) on the GA. Next, the positions 1-7 were primed using the priming recipe v7 (GA2_Prime_v<#>.xml). After priming, the used flow cell v4 was removed and the prism also removed for cleaning. The prism and a new flow cell v4 (from section A2.2.5) were loaded onto the GAIIx, a check for leaks and proper reagent delivery done and then immersion oil was applied between the prism and flow cell v4.

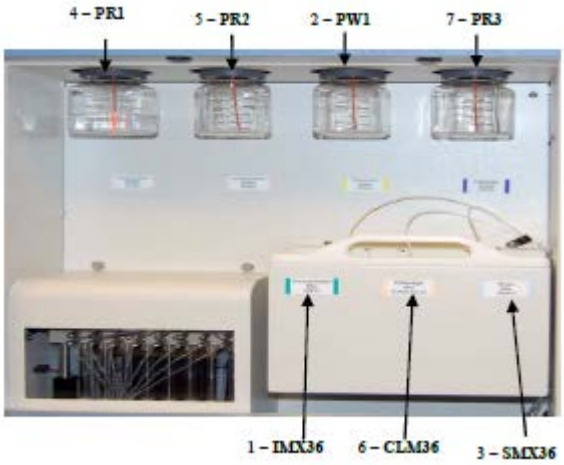


Figure A2.2: Genome analyzer IIx showing the reagent positions

A2.3.3 First-base incorporation and sequencing

This step was performed according to the first-base incorporation recipe v7 (GA2_FirstBase_v<#>.xml). The results were evaluated by checking the goodness of fit (≥ 0.990) and sensitivity (350 - 400) parameters in the calibration results tab. After evaluation, the 76-cycle sequencing recipe (GA2_76Cycle_SR_v<#>.xml) was followed.

A2.3.4 Post-run wash

The GAIIx was loaded with 10 mL of PW1 at positions 1, 3 and 6 and 40 mL of PW1 at positions 2, 4, 5 and 7 as mentioned above (Figure A2.2). The Genome Analyzer bundled tubes were placed into a 50 mL tube and the post-run recipe v7 (GA2_PostWash_v<#>.xml) was run.

Appendix III Validation of differential expression from mRNA-Seq I using sqRT-PCR.

Unigene	Differential expression (Fold change)	
	mRNA-seq I	sqRT-PCT
Hv.10251	-69.03	-3.63±0.65
Hv.20929	2.13	4.21±0.59
Hv.22598	1.17	1.13±0.43
Hv.469	-2.08	-1.11±0.51
Hv.8276	-1.04	1.15±0.02
Hv.8888	-17.88	-2.68±0.77

Appendix IV Significant GO terms for up-regulated transcripts in salt stress condition (mRNA-Seq II)

GO term	Ontology	Description	Number in input list	Number in BG/Ref	p-value	FDR
GO:0006468	P	protein amino acid phosphorylation	19	1584	0.00022	0.0098
GO:0043687	P	post-translational protein modification	22	1818	7.3e-05	0.0098
GO:0006464	P	protein modification process	23	1997	0.00011	0.0098
GO:0043412	P	macromolecule modification	23	2046	0.00015	0.0098
GO:0006796	P	phosphate metabolic process	21	1879	0.00028	0.0098
GO:0006793	P	phosphorus metabolic process	21	1879	0.00028	0.0098
GO:0044267	P	cellular protein metabolic process	28	2983	0.00057	0.017
GO:0016310	P	phosphorylation	19	1761	0.00075	0.02
GO:0005529	F	sugar binding	10	272	7e-07	0.00014
GO:0030246	F	carbohydrate binding	10	327	3.6e-06	0.00036
GO:0004713	F	protein tyrosine kinase activity	18	1426	0.00017	0.01
GO:0016301	F	kinase activity	20	1756	0.0003	0.01
GO:0016773	F	phosphotransferase activity, alcohol group as acceptor	20	1738	0.00026	0.01
GO:0004672	F	protein kinase activity	19	1587	0.00023	0.01
GO:0004674	F	protein serine/threonine kinase activity	17	1424	0.00046	0.013

The GO was performed using agriGO (http://bioinfo.cau.edu.cn/agriGO/analysis.php)

Appendix V Significant GO terms for down-regulated transcripts in drought stress condition (mRNA-Seq II)

GO term	Ontology	Description	Number in input list	Number in BG/Ref	p-value	FDR
GO:0006468	P	protein amino acid phosphorylation	22	1584	6.6e-05	0.0055
GO:0043687	P	post-translational protein modification	24	1818	7.2e-05	0.0055
GO:0030001	P	metal ion transport	8	225	2.8e-05	0.0055
GO:0016310	P	phosphorylation	22	1761	0.00027	0.007
GO:0044267	P	cellular protein metabolic process	33	2983	0.00013	0.007
GO:0006464	P	protein modification process	24	1997	0.00027	0.007
GO:0043412	P	macromolecule modification	25	2046	0.00017	0.007
GO:0006796	P	phosphate metabolic process	23	1879	0.00027	0.007
GO:0006793	P	phosphorus metabolic process	23	1879	0.00027	0.007
GO:0006811	P	ion transport	9	450	0.00068	0.016
GO:0006812	P	cation transport	8	390	0.0011	0.023
GO:0004713	F	protein tyrosine kinase activity	22	1426	1.5e-05	0.0017
GO:0004674	F	protein serine/threonine kinase activity	22	1424	1.5e-05	0.0017
GO:0005524	F	ATP binding	35	3055	4.8e-05	0.0023
GO:0016773	F	phosphotransferase activity, alcohol group as acceptor	24	1738	3.7e-05	0.0023
GO:0032559	F	adenyl ribonucleotide binding	35	3058	4.9e-05	0.0023
GO:0016301	F	kinase activity	23	1756	0.00011	0.0026
GO:0032555	F	purine ribonucleotide binding	35	3238	0.00014	0.0026
GO:0032553	F	ribonucleotide binding	35	3238	0.00014	0.0026

GO:0030554	F	adenyl nucleotide binding	35	3210	0.00012	0.0026
GO:0001883	F	purine nucleoside binding	35	3210	0.00012	0.0026
GO:0001882	F	nucleoside binding	35	3220	0.00012	0.0026
GO:0004672	F	protein kinase activity	22	1587	6.8e-05	0.0026
GO:0046873	F	metal ion transmembrane transporter activity	5	104	0.00022	0.0038
GO:0000166	F	nucleotide binding	38	3794	0.00032	0.0049
GO:0017076	F	purine nucleotide binding	35	3394	0.0003	0.0049

The GO was performed using agriGO (http://bioinfo.cau.edu.cn/agriGO/analysis.php)

Appendix VI Significant GO terms for up-regulated transcripts in drought stress condition (mRNA-Seq II)

GO term	Ontology	Description	Number in input list	Number in BG/Ref	p-value	FDR
GO:0005506	F	iron ion binding	16	733	1.1e-06	0.0002
GO:0020037	F	heme binding	13	622	1.6e-05	0.00081
GO:0046906	F	tetrapyrrole binding	13	629	1.7e-05	0.00081
GO:0009055	F	electron carrier activity	16	919	1.8e-05	0.00081
GO:0016491	F	oxidoreductase activity	25	1995	3.5e-05	0.0013
GO:0004497	F	monooxygenase activity	9	423	0.00024	0.0074

The GO was performed using agriGO (http://bioinfo.cau.edu.cn/agriGO/analysis.php)

Appendix VII Significant GO terms for down-regulated transcripts in exogenous ABA stress condition (mRNA-Seq II)

GO term	Ontology	Description	Number in input list	Number in BG/Ref	p-value	FDR
GO:0006091	P	generation of precursor metabolites and energy	10	308	2.5e-06	0.001
GO:0006119	P	oxidative phosphorylation	7	145	5.8e-06	0.0012
GO:0000160	P	two-component signal transduction system (phosphorelay)	5	60	8.9e-06	0.0012
GO:0015985	P	energy coupled proton transport, down electrochemical gradient	5	102	0.00011	0.0081
GO:0015986	P	ATP synthesis coupled proton transport	5	102	0.00011	0.0081
GO:0034220	P	ion transmembrane transport	5	102	0.00011	0.0081
GO:0015992	P	proton transport	5	113	0.00019	0.0098
GO:0006818	P	hydrogen transport	5	113	0.00019	0.0098
GO:0006811	P	ion transport	9	450	0.0003	0.014
GO:0009206	P	purine ribonucleoside triphosphate biosynthetic process	5	163	0.00098	0.02
GO:0009141	P	nucleoside triphosphate metabolic process	5	165	0.001	0.02
GO:0009142	P	nucleoside triphosphate biosynthetic process	5	163	0.00098	0.02
GO:0009144	P	purine nucleoside triphosphate metabolic process	5	165	0.001	0.02
GO:0009145	P	purine nucleoside triphosphate biosynthetic process	5	163	0.00098	0.02
GO:0009199	P	ribonucleoside triphosphate metabolic process	5	165	0.001	0.02
GO:0009205	P	purine ribonucleoside triphosphate metabolic process	5	165	0.001	0.02
GO:0006812	P	cation transport	8	390	0.00053	0.02
GO:0009201	P	ribonucleoside triphosphate biosynthetic process	5	163	0.00098	0.02
GO:0046034	P	ATP metabolic process	5	160	0.0009	0.02

GO:0055085	P	transmembrane transport	6	227	0.00069	0.02
GO:0023052	P	signaling	9	495	0.00059	0.02
GO:0006754	P	ATP biosynthetic process	5	158	0.00085	0.02
GO:0046483	P	heterocycle metabolic process	7	349	0.0013	0.024
GO:0009259	P	ribonucleotide metabolic process	5	184	0.0017	0.025
GO:0015672	P	monovalent inorganic cation transport	5	186	0.0018	0.025
GO:0009150	P	purine ribonucleotide metabolic process	5	181	0.0016	0.025
GO:0009152	P	purine ribonucleotide biosynthetic process	5	179	0.0015	0.025
GO:0009260	P	ribonucleotide biosynthetic process	5	182	0.0016	0.025
GO:0006164	P	purine nucleotide biosynthetic process	5	184	0.0017	0.025
GO:0055086	P	nucleobase, nucleoside and nucleotide metabolic process	6	275	0.0018	0.025
GO:0006163	P	purine nucleotide metabolic process	5	188	0.0018	0.025
GO:0009165	P	nucleotide biosynthetic process	5	207	0.0028	0.037
GO:0006350	P	transcription	17	1698	0.0033	0.041
GO:0023046	P	signaling process	6	318	0.0037	0.045
GO:0023060	P	signal transmission	6	318	0.0037	0.045
GO:0016070	P	RNA metabolic process	13	1180	0.0041	0.048
GO:0000156	F	two-component response regulator activity	5	41	1.3e-06	0.00028
GO:0022890	F	inorganic cation transmembrane transporter activity	7	229	0.00011	0.0076
GO:0015078	F	hydrogen ion transmembrane transporter activity	6	150	7.6e-05	0.0076

GO:0015077	F	monovalent inorganic cation transmembrane transporter activity	6	180	0.0002	0.0087
GO:0022857	F	transmembrane transporter activity	12	722	0.00019	0.0087
GO:0015075	F	ion transmembrane transporter activity	9	449	0.00029	0.01
GO:0060089	F	molecular transducer activity	6	205	0.00041	0.011
GO:0004871	F	signal transducer activity	6	205	0.00041	0.011
GO:0008324	F	cation transmembrane transporter activity	7	342	0.0012	0.027
GO:0022891	F	substrate-specific transmembrane transporter activity	9	573	0.0016	0.034
GO:0005215	F	transporter activity	13	1110	0.0025	0.049
GO:0045263	C	proton-transporting ATP synthase complex, coupling factor F(o)	5	30	2.6e-07	1.9e-05
GO:0033177	C	proton-transporting two-sector ATPase complex, proton-transporting domain	5	41	1.3e-06	4.7e-05
GO:0045259	C	proton-transporting ATP synthase complex	5	55	5.8e-06	0.00014
GO:0016020	C	membrane	25	2185	8.3e-05	0.0015
GO:0016469	C	proton-transporting two-sector ATPase complex	5	106	0.00014	0.002
GO:0044425	C	membrane part	12	1006	0.0031	0.037

The GO was performed using agriGO (http://bioinfo.cau.edu.cn/agriGO/analysis.php)

Appendix VIII Significant GO terms for up-regulated transcripts in exogenous ABA stress condition (mRNA-Seq II)

GO term	Ontology	Description	Number in input list	Number in BG/Ref	p-value	FDR
GO:0004553	F	hydrolase activity, hydrolyzing O-glycosyl compounds	11	454	6e-06	0.0011
GO:0016798	F	hydrolase activity, acting on glycosyl bonds	11	503	1.5e-05	0.0014
GO:0004866	F	endopeptidase inhibitor activity	5	88	3.8e-05	0.002
GO:0030414	F	peptidase inhibitor activity	5	90	4.3e-05	0.002
GO:0004857	F	enzyme inhibitor activity	5	170	0.00082	0.031
GO:0005576	C	extracellular region	7	269	0.00017	0.0057

The GO was performed using agriGO (http://bioinfo.cau.edu.cn/agriGO/analysis.php)

Appendix IX Validation of differential expression from mRNA-Seq II using qRT-PCR.

Barley	Salt (FC)		Drought (FC)		ABA(FC)	
	mRNA-seq	qRT-PCR	mRNA-seq	qRT-PCR	mRNA-seq	qRT-PCR
AK251854	-1.00	4.53	1.19	4.12	3.24	4.78
MLOC_8131.1	2.26	3.82	3.03	3.27	7.26	5.11
AK376521	-1.07	-5.89	-1.84	-5.62	-2.51	-2.18
MLOC_71349.1	-1.64	-2.17	-3.12	-2.87	-13.16	-10.22
MLOC_22145.2	-1.39	-1.78	-1.06	-1.58	-1.47	-1.09
MLOC_3013.1	1.26	-2.06	-1.66	-1.95	1.18	-1.48
MLOC_68581.1	-1.44	2.14	4.76	3.46	1.61	1.42
MLOC_62985.1	-1.59	2.56	-1.38	2.17	1.36	2.33

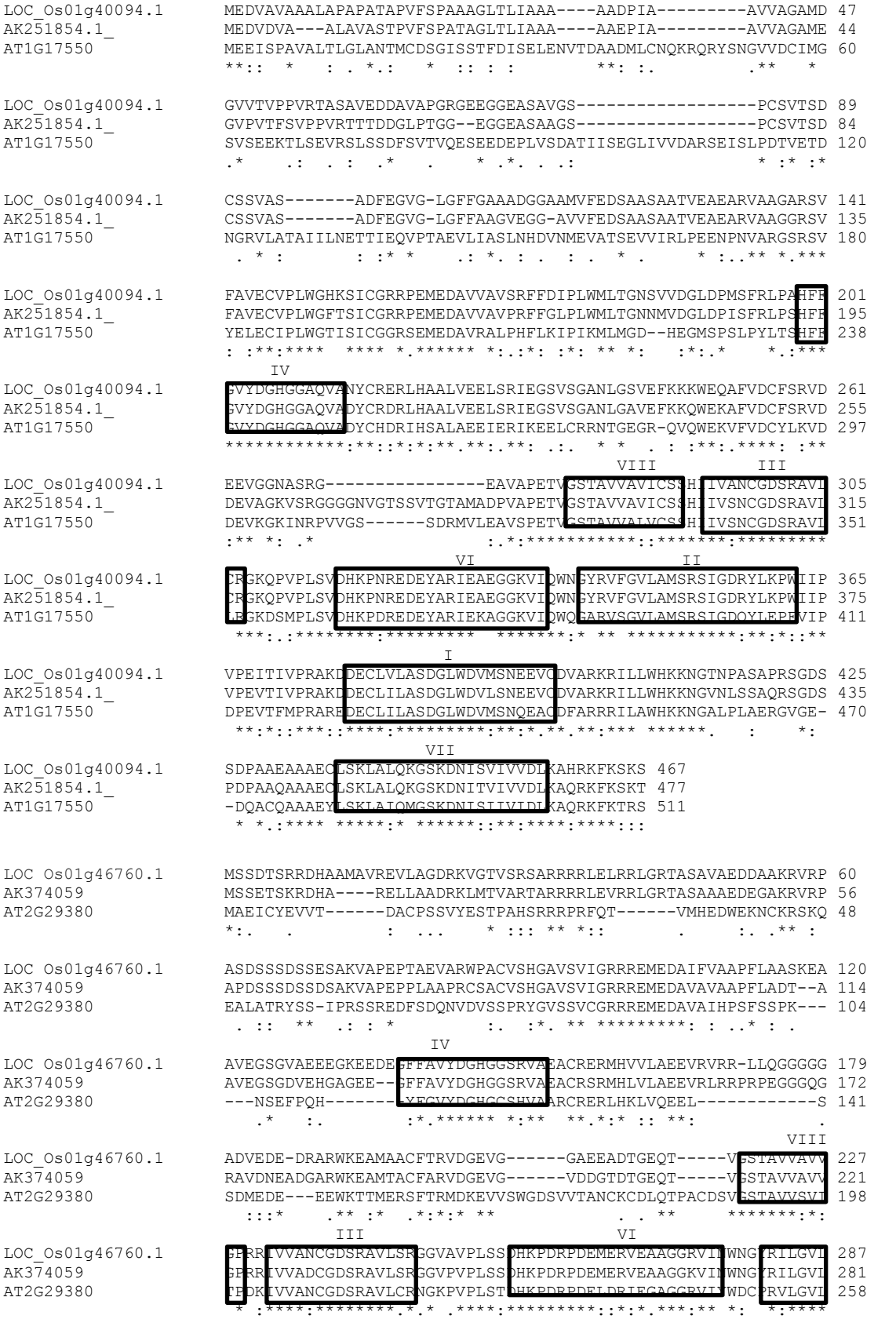
Appendix X ABA-related gene and protein sequences in rice.

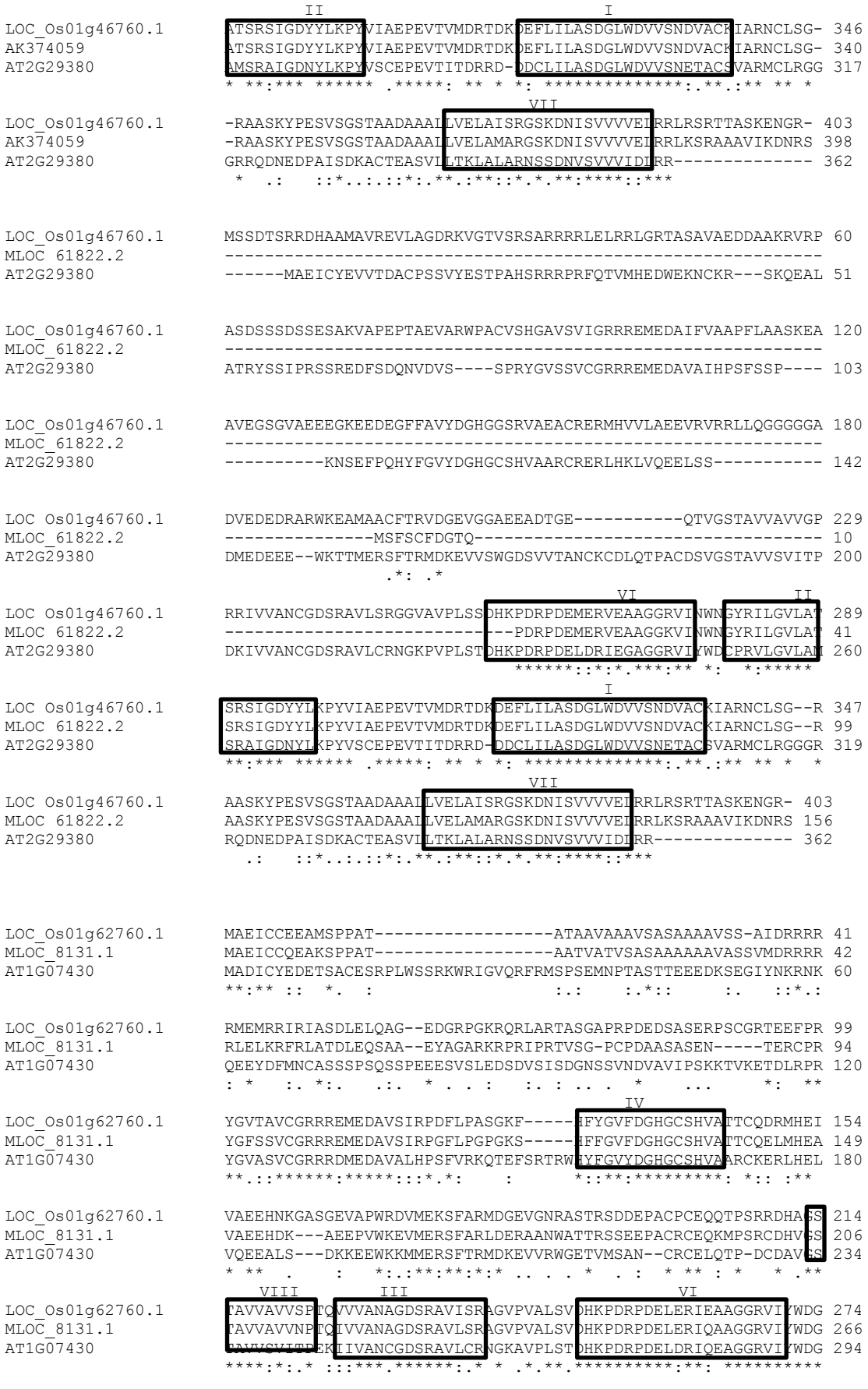
Gene name	Subfamily	Rice locus[a]	Arabidopsis Locus	Putative full-length protein (aa)[b]	Mw (kDa)[c]	PI[d]
PP2C						
OsPP2C01	-	LOC_Os01g07090	AT1G03590	332	35.43	4.69
OsPP2C02	PP2CG	LOC_Os01g19130	AT1G07160	381	41.82	4.59
OsPP2C03	PP2CH	LOC_Os01g32964	AT1G07430	335	35.9	7.91
OsPP2C04	-	LOC_Os01g36080	AT1G07630	658	73.19	6.26
OsPP2C05	-	LOC_Os01g37130	AT1G09160	390	41.74	5.51
OsPP2C06	PP2CA	LOC_Os01g40094	AT1G16220	468	48.63	4.77
OsPP2C07	PP2CF2	LOC_Os01g43100	AT1G17550	392	41.81	8.41
OsPP2C08	PP2CA	LOC_Os01g46760	AT1G18030	404	43	5.44
OsPP2C09	PP2CA	LOC_Os01g62760	AT1G22280	415	43.99	5.25
OsPP2C10	PP2CF2	LOC_Os02g05630	AT1G34750	349	37.57	5.02
OsPP2C11	PP2CI	LOC_Os02g08364	AT1G43900	363	39.71	5.04
OsPP2C12	PP2CE	LOC_Os02g13100	AT1G47380	390	41	5.88
OsPP2C13	PP2CF2	LOC_Os02g15594	AT1G48040	364	38.96	5.53
OsPP2C14	PP2CE	LOC_Os02g27220	AT1G67820	519	55.29	8.71
OsPP2C15	PP2CH	LOC_Os02g35910	AT1G68410	443	48.01	5.72
OsPP2C16	PP2CI	LOC_Os02g38580	AT1G72770	522	56.26	4.66
OsPP2C17	PP2CI	LOC_Os02g38690	AT1G78200	764	82.81	4.81
OsPP2C18	PP2CI	LOC_Os02g38710	AT1G79630	805	84.98	4.21
OsPP2C19	PP2CI	LOC_Os02g38780	AT2G20050	653	71.41	4.65
OsPP2C20	PP2CI	LOC_Os02g38804	AT2G20630	518	56.55	4.51
OsPP2C21	PP2CI	LOC_Os02g39410	AT2G25070	341	53.76	10.8
OsPP2C22	PP2CI	LOC_Os02g39480	AT2G25620	582	63.3	4.88
OsPP2C23	PP2CK	LOC_Os02g42250	AT2G28890	320	34.73	4.99
OsPP2C24	PP2CK	LOC_Os02g42270	AT2G29380	316	33.45	5.28
OsPP2C25	PP2CD	LOC_Os02g46080	AT2G30020	388	43.01	8.55
OsPP2C26	PP2CC	LOC_Os02g46490	AT2G30170	597	64.83	5.75
OsPP2C27	PP2CG	LOC_Os02g55560	AT2G33700	185	39.27	4.9
OsPP2C28	PP2CD	LOC_Os03g04430	AT2G34740	400	44.29	8.77
OsPP2C29	PP2CD	LOC_Os03g10950	AT2G35350	393	42.51	8.22
OsPP2C30	PP2CA	LOC_Os03g16170	AT2G40180	405	43.28	5.71
OsPP2C31	PP2CC	LOC_Os03g16760	AT2G40860	632	68.74	6.39
OsPP2C32	PP2CB	LOC_Os03g18150	AT2G46920	392	41.31	6.37
OsPP2C33	PP2CH	LOC_Os03g18970	AT3G02750	433	45.83	6.69
OsPP2C34	PP2CD	LOC_Os03g55320	AT3G05640	381	41.75	9.79
OsPP2C35	PP2CC	LOC_Os03g60650	AT3G06270	640	69.18	5.31
OsPP2C36	PP2CD	LOC_Os03g61690	AT3G09400	387	42.02	7.25
OsPP2C37	PP2CA	LOC_Os04g08560	AT3G11410	435	44.81	7.24
OsPP2C38	PP2CE	LOC_Os04g25570	AT3G12620	461	49.69	6.78
OsPP2C39	PP2CC	LOC_Os04g33080	AT3G15260	521	57.32	4.57
OsPP2C40	PP2CH	LOC_Os04g37660	AT3G16560	462	48.71	6.32
OsPP2C41	PP2CF1	LOC_Os04g37904	AT3G16800	285	31.01	6.25
OsPP2C42	PP2CI	LOC_Os04g42260	AT3G17090	353	38.17	6.83
OsPP2C43	PP2CD	LOC_Os04g49490	AT3G17250	389	43.19	6.51
OsPP2C44	PP2CF1	LOC_Os04g52000	AT3G23360	322	34.77	8.66
OsPP2C45	PP2CF2	LOC_Os04g56450	AT3G27140	283	30.58	4.75
OsPP2C46	PP2CC	LOC_Os05g02110	AT3G51370	594	63.48	5.28
OsPP2C47	PP2CG	LOC_Os05g04360	AT3G51470	390	42.38	4.55
OsPP2C48	PP2CE	LOC_Os05g29030	AT3G55050	354	38.1	7.38
OsPP2C49	PP2CA	LOC_Os05g38290	AT3G62260	417	43.78	8.33
OsPP2C50	PP2CA	LOC_Os05g46040	AT3G63320	388	41.55	5.97
OsPP2C51	PP2CA	LOC_Os05g49730	AT3G63340	382	40.35	7.83
OsPP2C52	PP2CF2	LOC_Os05g50970	AT4G03415	492	52.92	4.49
OsPP2C53	PP2CA	LOC_Os05g51510	AT4G08260	446	46.68	4.67
OsPP2C54	PP2CG	LOC_Os06g08140	AT4G11040	361	39.35	4.98
OsPP2C55	PP2CF2	LOC_Os06g33530	AT4G16580	354	38.29	10.5
OsPP2C56	PP2CF2	LOC_Os06g33549	AT4G26080	353	37.81	8.48
OsPP2C57	PP2CG	LOC_Os06g39600	AT4G27800	368	39.44	5.17
OsPP2C58	PP2CI	LOC_Os06g44210	AT4G28400	369	40.59	4.71

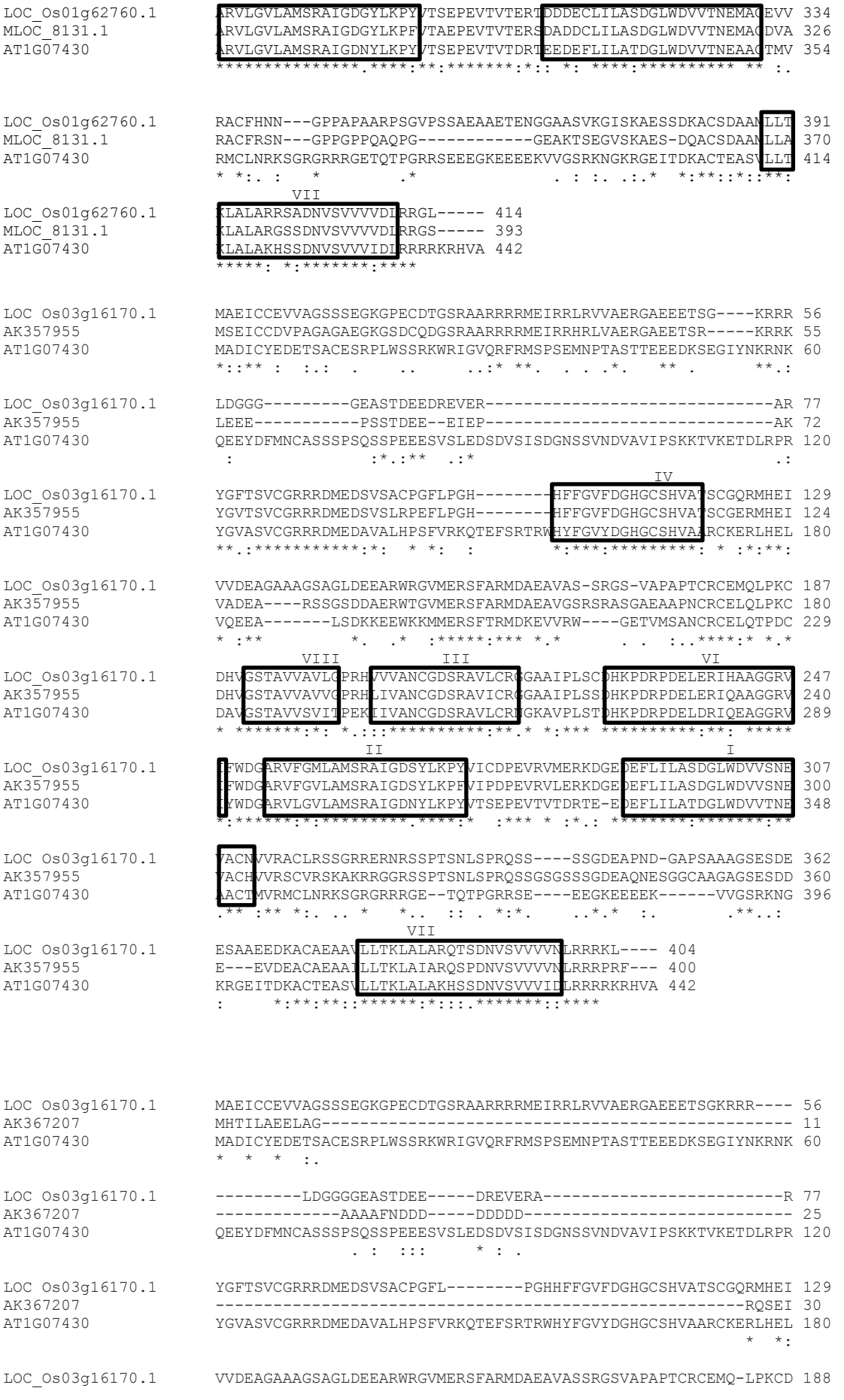
OsPP2C59	PP2CF2	LOC_Os06g48300	AT4G31750	328	34.94	4.39
OsPP2C60	PP2CD	LOC_Os06g50380	AT4G31860	393	43.58	8.65
OsPP2C61	PP2CD	LOC_Os07g02330	AT4G32950	378	40.62	8.24
OsPP2C62	PP2CF1	LOC_Os07g32380	AT4G33500	291	31.74	6.97
OsPP2C63	PP2CF1	LOC_Os07g33230	AT4G33920	224	24.68	7.05
OsPP2C64	PP2CE	LOC_Os07g37890	AT4G38520	428	46.3	5.55
OsPP2C65	-	LOC_Os07g45170	AT5G01700	446	47.22	5.62
OsPP2C66	PP2CE	LOC_Os08g39100	AT5G02400	532	57.67	5
OsPP2C67	-	LOC_Os09g14540	AT5G02760	368	39.39	6.83
OsPP2C68	PP2CA	LOC_Os09g15670	AT5G06750	359	37.7	6.5
OsPP2C69	PP2CH	LOC_Os09g28560	AT5G10740	423	45.81	5.34
OsPP2C70	PP2CI	LOC_Os09g38550	AT5G19280	352	38.47	5.08
OsPP2C71	-	LOC_Os10g22460	AT5G24940	466	48.49	4.08
OsPP2C72	PP2CD	LOC_Os10g39780	AT5G26010	335	36.76	6.84
OsPP2C73	PP2CE	LOC_Os11g01790	AT5G27930	421	46.26	5.85
OsPP2C74	PP2CB	LOC_Os11g13820	AT5G36250	398	41.35	6.64
OsPP2C75	PP2CE	LOC_Os11g22404	AT5G51760	433	47.11	6.12
OsPP2C76	-	LOC_Os11g37540	AT5G53140	1116	27.19	6.9
OsPP2C77	PP2CB	LOC_Os12g09640	AT5G57050	422	**435.57**	7.32
OsPP2C78	PP2CD	LOC_Os12g39120	AT5G59220	393	43.53	8.61
PYR/PYL/RCAR						
OsPYR1	NA	LOC_Os03g18600	AT4G17870	229	23.8	7.2
OsPYL1		LOC_Os05g39580	AT5G46790	217	22.7	8.3
OsPYL2		LOC_Os10g42280	AT2G26040	212	23.0	5.45
OsPYL3		LOC_Os06g36670	AT1G73000	207	22.3	6.07
OSPYL4		LOC_Os02g13330	AT2G38310	207	22.5	6.45
OsPYL5		LOC_Os05g12260	AT5G05440	209	22.2	5.69
OsPYL6		LOC_Os02g15640	AT2G40330	204	23.0	6.46
OsPYL7		LOC_Os06g33690	AT4G01026	206	23.3	6.25
OsPYL8		LOC_Os06g33490	AT5G53160	158	17.7	5.37
OsPYL9		LOC_Os06g33640	AT1G01360	206	23.2	6.65
SnRK2						
OsSnRK2.2	III	LOC_Os03g55600	AT3G50500	371	41.7	4.85
OsSnRK2.3	III	LOC_Os12g39630	AT5G66880	361	40.6	4.81
OsSnRK2.6	III	LOC_Os03g41460	AT4G33950	362	40.7	4.80
OsSnRK2.7	II	LOC_Os07g42940	AT4G40010	339	38.5	5.31
OsSnRK2.8	II	LOC_Os10g41490	AT1G78290	334	37.9	5.67
OsSnRK2.11	II	LOC_Os03g27280	-	342	38.7	5.43
ABA-transport related ABC transporters						
OsABCG25	G25	LOC_Os11g07600	AT1G71960	612	64.9	9.4
OsABCG40	G40	LOC_Os08g43120	AT1G15520	1324	149.9	8.8

a. ESTs retrieved from MSU rice database; b. Size of proteins in amino acids; c,d. Molecular weight (Mw) and isoelectric point (pI) of the putative protein for PP2C were extracted from Xue et al. 2008. whereas for other families' Molecular weight (Mw) and isoelectric point (pI) were predicted in the ExPASy Proteomics Server (http://web.expasy.org/compute_pi/)

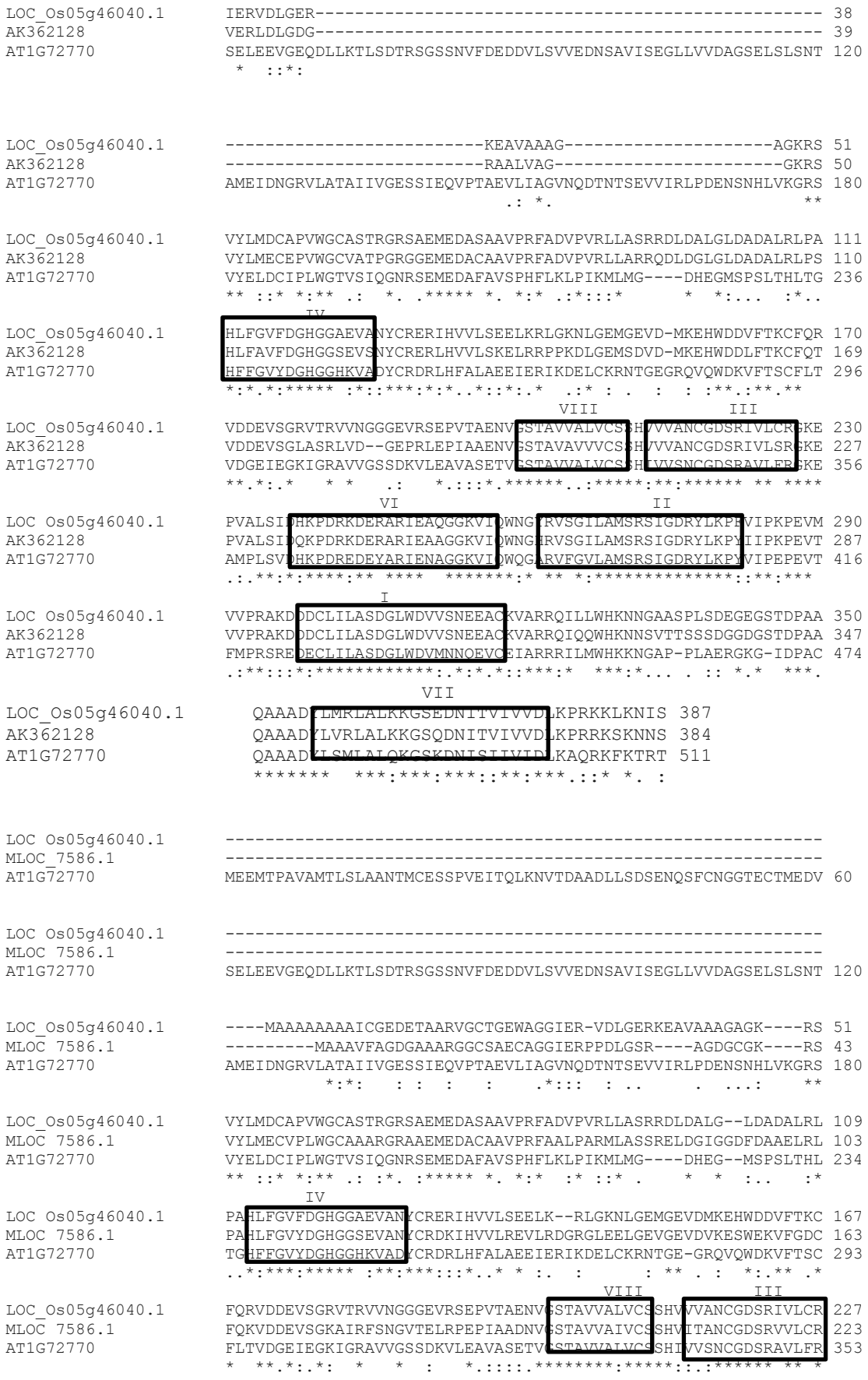
Appendix XI Alignment of putative amino acid sequences of barley PP2CAs with rice and Arabidopsis

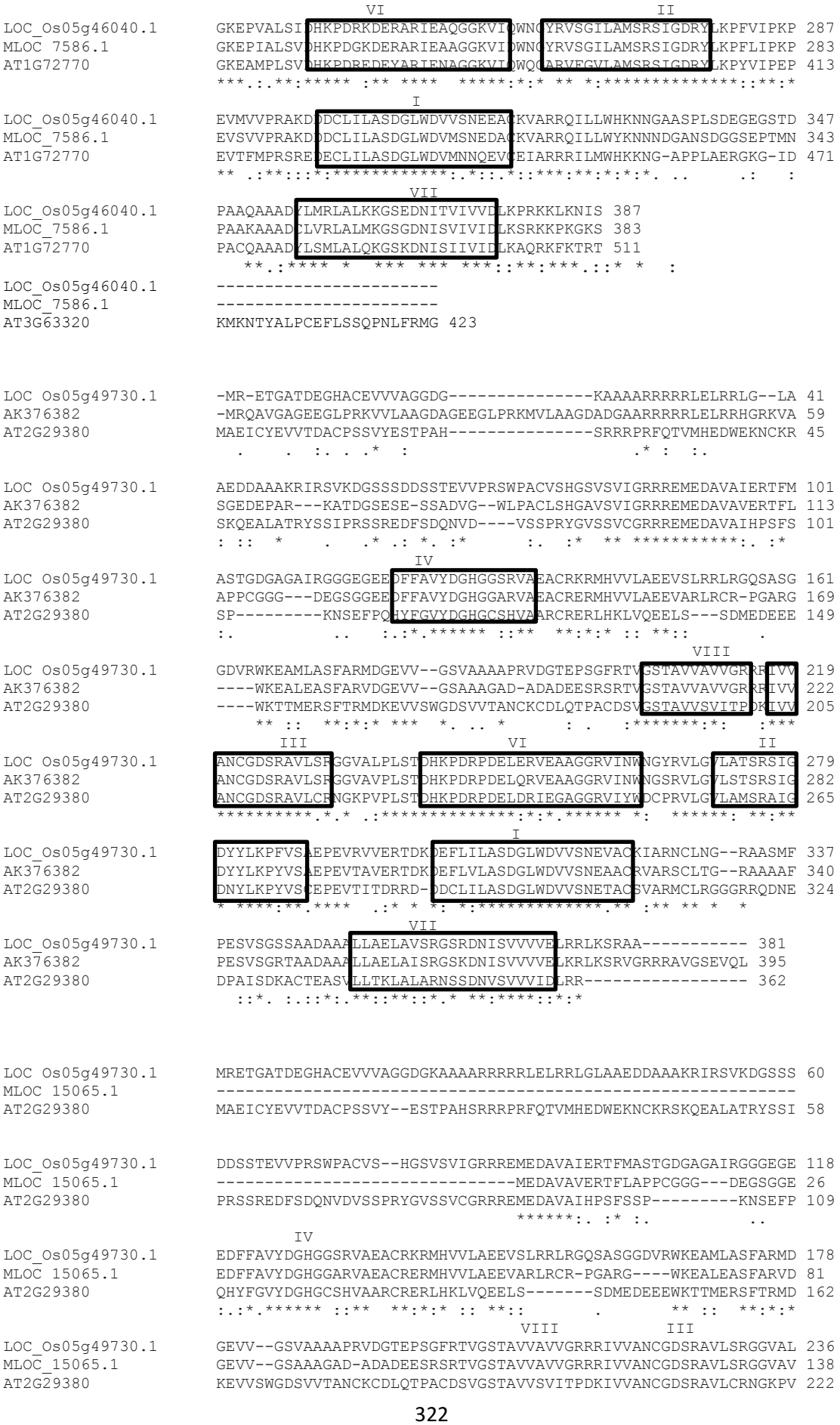


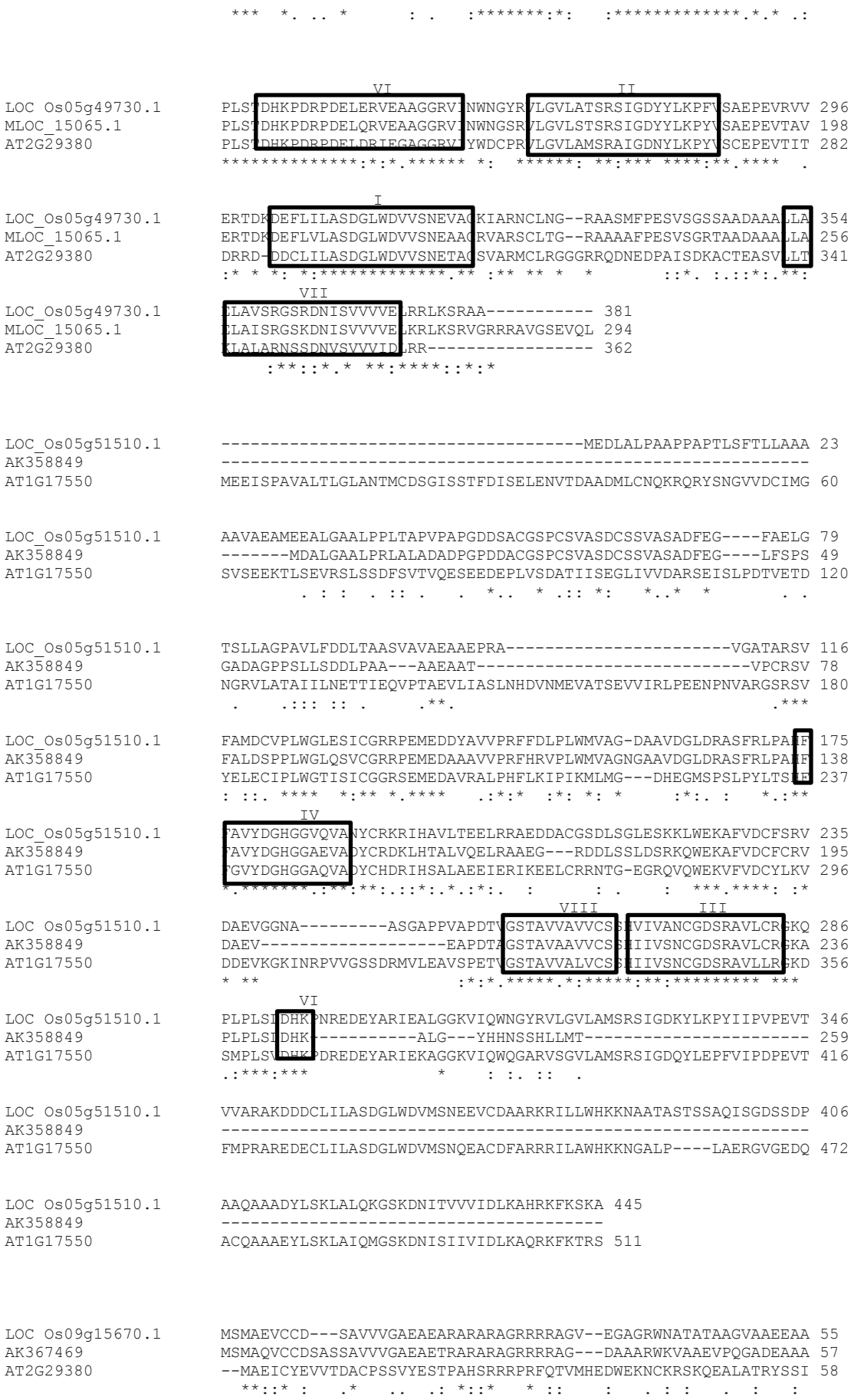

II
I
LOC_Os01g46760.1 ATSRSIGDYYLKPYVIAEPEVTVMDRTDKDEFLILASDGLWDVVSNDVACKIARNCLSG- 346
AK374059 ATSRSIGDYYLKPYVIAEPEVTVMDRTDKDEFLILASDGLWDVVSNDVACKIARNCLSG- 340
AT2G29380 AMSRAIGDNYLKPYVSCEPEVTITDRRD-DDCLILASDGLWDVVSNETACSVARMCLRGG 317
VII
LOC_Os01g46760.1 -RAASKYPESVSGSTAADAAALLVELAISRGSKDNISVVVVELRRLRSRTTASKENGR- 403
AK374059 -RAASKYPESVSGSTAADAAALLVELAMARGSKDNISVVVVELRRLKSRAAAVIKDNRS 398
AT2G29380 GRRQDNEDPAISDKACTEASVLLTKLALARNSSDNVSVVVIDLRR-------------- 362
LOC_Os01g46760.1 MSSDTSRRDHAAMAVREVLAGDRKVGTVSRSARRRRLELRRLGRTASAVAEDDAAKRVRP 60
MLOC_61822.2 --
AT2G29380 ------MAEICYEVVTDACPSSVYESTPAHSRRRPRFQTVMHEDWEKNCKR---SKQEAL 51
LOC_Os01g46760.1 ASDSSSDSSESAKVAPEPTAEVARWPACVSHGAVSVIGRRREMEDAIFVAAPFLAASKEA 120
MLOC_61822.2 --
AT2G29380 ATRYSSIPRSSREDFSDQNVDVS----SPRYGVSSVCGRRREMEDAVAIHPSFSSP---- 103
LOC_Os01g46760.1 AVEGSGVAEEEGKEEDEGFFAVYDGHGGSRVAEACRERMHVVLAEEVRVRRLLQGGGGGA 180
MLOC_61822.2 --
AT2G29380 ----------KNSEFPQHYFGVYDGHGCSHVAARCRERLHKLVQEELSS----------- 142
LOC_Os01g46760.1 DVEDEDRARWKEAMAACFTRVDGEVGGAEEADTGE----------QTVGSTAVVAVVGP 229
MLOC_61822.2 ---------------MSFSCFDGTQ---------------------------------- 10
AT2G29380 DMEDEEE--WKTTMERSFTRMDKEVVSWGDSVVTANCKCDLQTPACDSVGSTAVVSVITP 200
VI
II
LOC_Os01g46760.1 RRIVVANCGDSRAVLSRGGVAVPLSSDHKPDRPDEMERVEAAGGRVINWNGYRILGVLAT 289
MLOC_61822.2 ---------------------------PDRPDEMERVEAAGGKVINWNGYRILGVLAT 41
AT2G29380 DKIVVANCGDSRAVLCRNGKPVPLSTDHKPDRPDELDRIEGAGGRVIYWDCPRVLGVLAM 260
I
LOC_Os01g46760.1 SRSIGDYYLKPYVIAEPEVTVMDRTDKDEFLILASDGLWDVVSNDVACKIARNCLSG--R 347
MLOC_61822.2 SRSIGDYYLKPYVIAEPEVTVMDRTDKDEFLILASDGLWDVVSNDVACKIARNCLSG--R 99
AT2G29380 SRAIGDNYLKPYVSCEPEVTITDRRD-DDCLILASDGLWDVVSNETACSVARMCLRGGGR 319
VII
LOC_Os01g46760.1 AASKYPESVSGSTAADAAALLVELAISRGSKDNISVVVVELRRLRSRTTASKENGR- 403
MLOC_61822.2 AASKYPESVSGSTAADAAALLVELAMARGSKDNISVVVVELRRLKSRAAAVIKDNRS 156
AT2G29380 RQDNEDPAISDKACTEASVLLTKLALARNSSDNVSVVVIDLRR-------------- 362
LOC_Os01g62760.1 MAEICCEEAMSPPAT----------ATAAVAAAVSASAAAAVSS-AIDRRRR 41
MLOC_8131.1 MAEICCQEAKSPPAT----------AATVATVSASAAAAAAVASSVMDRRRR 42
AT1G07430 MADICYEDETSACESRPLWSSRKWRIGVQRFRMSPSEMNPTASTTEEEDKSEGIYNKRNK 60
LOC_Os01g62760.1 RMEMRRIRIASDLELQAG--EDGRPGKRQRLARTASGAPRPDEDSASERPSCGRTEEFPR 99
MLOC_8131.1 RLELKRFRLATDLEQSAA--EYAGARKRPRIPRTVSG-PCPDAASASEN-----TERCPR 94
AT1G07430 QEEYDFMNCASSSPSQSSPEEESVSLEDSDVSISDGNSSVNDVAVIPSKKTVKETDLRPR 120
IV
LOC_Os01g62760.1 YGVTAVCGRRREMEDAVSIRPDFLPASGKF-----HFYGVFDGHGCSHVATTCQDRMHEI 154
MLOC_8131.1 YGFSSVCGRRREMEDAVSIRPGFLPGPGKS-----HFFGVFDGHGCSHVATTCQELMHEA 149
AT1G07430 YGVASVCGRRRDMEDAVALHPSFVRKQTEFSRTRWLYFGVYDGHGCSHVARCKERLHEL 180
LOC_Os01g62760.1 VAEEHNKGASGEVAPWRDVMEKSFARMDGEVGNRASTRSDDEPACPCEQQTPSRRDHAGS 214
MLOC_8131.1 VAEEHDK---AEEPVWKEVMERSFARLDERAANWATTRSSEEPACRCEQKMPSRCDHVGS 206
AT1G07430 VQEEALS---DKKEEWKKMMERSFTRMDKEVVRWGETVMSAN--CRCELQTP-DCDAVLS 234
VIII
III
VI
LOC_Os01g62760.1 TAVVAVVSPTQVVVANAGDSRAVISRAGVPVALSVDHKPDRPDELERIEAAGGRVIYWDG 274
MLOC_8131.1 TAVVAVVNPTQIVVANAGDSRAVLSRAGVPVALSVDHKPDRPDELERIQAAGGRVIYWDG 266
AT1G07430 TAVVSVITPEKLIVANCGDSRAVLCRNGKAVPLSTDHKPDRPDELDRIQEAGGRVIYWDG 294
II
I

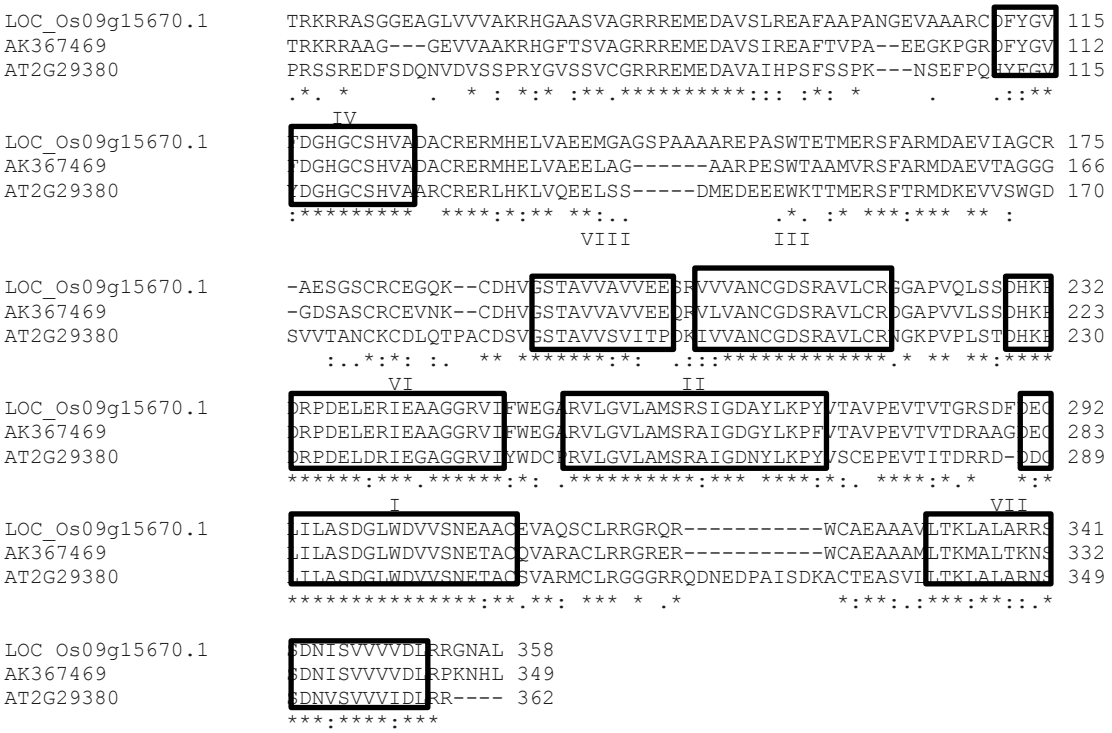

LOC_Os01g62760.1 ARVLGVLAMSRAIGDGYLKPYVTSEPEVTVTERDDDECLILASDGLWDVVTNEMACEVV 334
MLOC_8131.1 ARVLGVLAMSRAIGDGYLKPFVTAEPEVTVTERSDADDCLILASDGLWDVVTNEMACDVA 326
AT1G07430 ARVLGVLAMSRAIGDNYLKPYVTSEPEVTVTDRTEEDEFLILATDGLWDVVTNEAACTMV 354
LOC_Os01g62760.1 RACFHNN---GPPAPAARPSGVPSSAEAAETENGGAASVKGISKAESSDKACSDAANLLT 391
MLOC_8131.1 RACFRSN---GPPGPPQAQPG----------GEAKTSEGVSKAES-DQACSDAANLLA 370
AT1G07430 RMCLNRKSGRGRRRGETQTPGRRSEEEGKEEEEKVVGSRKNGKRGEITDKACTEASVLLT 414
VII
LOC_Os01g62760.1 KLALARRSADNVSVVVVDLRRGL----- 414
MLOC_8131.1 KLALARGSSDNVSVVVVDLRRGS----- 393
AT1G07430 KLALAKHSSDNVSVVVIDLRRRRKRHVA 442
LOC_Os03g16170.1 MAEICCEVVAGSSSEGKGPECDTGSRAARRRRMEIRRLRVVAERGAEEETSG----KRRR 56
AK357955 MSEICCDVPAGAGAEGKGSDCQDGSRAARRRRMEIRRHRLVAERGAEETSR-----KRRK 55
AT1G07430 MADICYEDETSACESRPLWSSRKWRIGVQRFRMSPSEMNPTASTTEEEDKSEGIYNKRNK 60
LOC_Os03g16170.1 LDGGG---------GEASTDEEDREVER----------AR 77
AK357955 LEEE-----------PSSTDEE--EIEP----------AK 72
AT1G07430 QEEYDFMNCASSSPSQSSPEEESVSLEDSDVSISDGNSSVNDVAVIPSKKTVKETDLRPR 120
IV
LOC_Os03g16170.1 YGFTSVCGRRRDMEDSVSACPGFLPGH--------HFFGVFDGHGCSHVATSCGQRMHEI 129
AK357955 YGVTSVCGRRRDMEDSVSLRPEFLPGH--------HFFGVFDGHGCSHVATSCGERMHEI 124
AT1G07430 YGVASVCGRRRDMEDAVALHPSFVRKQTEFSRTRWHYFGVYDGHGCSHVAARCKERLHEL 180
LOC_Os03g16170.1 VVDEAGAAAGSAGLDEEARWRGVMERSFARMDAEAVAS-SRGS-VAPAPTCRCEMQLPKC 187
AK357955 VADEA----RSSGSDDAERWTGVMERSFARMDAEAVGSRSRASGAEAAPNCRCELQLPKC 180
AT1G07430 VQEEA-------LSDKKEEWKKMMERSFTRMDKEVVRW----GETVMSANCRCELQTPDC 229
VIII III VI
LOC_Os03g16170.1 DHVGSTAVVAVLGPRHVVVANCGDSRAVLCRGGAAIPLSCDHKPDRPDELERIHAAGGRV 247
AK357955 DHVGSTAVVAVVGPRHLIVANCGDSRAVICRGGAAIPLSSDHKPDRPDELERIQAAGGRV 240
AT1G07430 DAVGSTAVVSVITPEKLIVANCGDSRAVLCRNGKAVPLSTDHKPDRPDELDRIQEAGGRV 289
II I
LOC_Os03g16170.1 LFWDGARVFGMLAMSRAIGDSYLKPYVICDPEVRVMERKDGEDEFLILASDGLWDVVSNE 307
AK357955 LFWDGARVFGVLAMSRAIGDSYLKPFVIPDPEVRVLERKDGEDEFLILASDGLWDVVSNE 300
AT1G07430 LFWDGARVLGVLAMSRAIGDNYLKPYVTSEPEVTVTDRTE-EDEFLILATDGLWDVVTNE 348
LOC_Os03g16170.1 VACNVVRACLRSSGRRERNRSSPTSNLSPRQSS----SSGDEAPND-GAPSAAAGSESDE 362
AK357955 VACHVVRSCVRSKAKRRGGRSSPTSNLSPRQSSGSGSSSGDEAQNESGGCAAGAGSESDD 360
AT1G07430 VACTMVRMCLNRKSGRGRRRGE--TQTPGRRSE----EEGKEEEEK------VVGSRKNG 396
VII
LOC_Os03g16170.1 ESAAEEDKACAEAAVLLTKLALARQTSDNVSVVVVNLRRRKL---- 404
AK357955 E---EVDEACAEAAILLTKLAIARQSPDNVSVVVVNLRRRPRF--- 400
AT1G07430 KRGEITDKACTEASVLLTKLALAKHSSDNVSVVVIDLRRRRKRHVA 442
LOC_Os03g16170.1 MAEICCEVVAGSSSEGKGPECDTGSRAARRRRMEIRRLRVVAERGAEEETSGKRRR---- 56
AK367207 MHTILAEELAG---------- 11
AT1G07430 MADICYEDETSACESRPLWSSRKWRIGVQRFRMSPSEMNPTASTTEEEDKSEGIYNKRNK 60
LOC_Os03g16170.1 ---------LDGGGGEASTDEE-----DREVERA----------R 77
AK367207 -------------AAAAFNDDD-----DDDDD---------- 25
AT1G07430 QEEYDFMNCASSSPSQSSPEEESVSLEDSDVSISDGNSSVNDVAVIPSKKTVKETDLRPR 120
LOC_Os03g16170.1 YGFTSVCGRRRDMEDSVSACPGFL--------PGHHFFGVFDGHGCSHVATSCGQRMHEI 129
AK367207 ----------RQSEI 30
AT1G07430 YGVASVCGRRRDMEDAVALHPSFVRKQTEFSRTRWHYFGVYDGHGCSHVAARCKERLHEL 180
LOC_Os03g16170.1 VVDEAGAAAGSAGLDEEARWRGVMERSFARMDAEAVASSRGSVAPAPTCRCEMQ-LPKCD 188

```
AK367207            VA-----------------WRGALWRSFARADGLAGLACACGRGSVPSCSCPLSGAHKGA 73
AT1G07430           VQEEALS-------DKKEEWKKMMERSFTRMDKEVVR--WGETVMSANCRCELQ-TPDCD 230
                    *                 *:  : ***:* *  .          ..* * :.   .

                                VIII              III                    VI
LOC_Os03g16170.1    HVGSTAVVAVLGPRHVVVANCGDSRAVLCRGGAAIPLSCDHKPDRPDELERIHAAGGRVI 248
AK367207            IVGSTAVVALLVRDRLIVANCGDSRAVLSRAGVAVPLSHDHKPDRPDEMARIKAAGGKVI 133
AT1G07430           AVGSTAVVSVITPGKIIVANCGDSRAVLCRNGKAVPLSTDHKPDRPDELDRIQEAGGRVI 290
                     *******:::   :::************.* * *:*** *********: **: ***:**
                                    II                                  I
LOC_Os03g16170.1    FWDGARVFGMLAMSRAIGDSYLKPYVICDPEVRVMERKDGEDEFLILASDGLWDVVSNEV 308
AK367207            FINGARVRGILAMSRALGHKVLKPEVICEPEISITERSE-DDDCLILASDGLWDVISNKV 192
AT1G07430           YWDGARVLGVLAMSRAIGDNYLKPYVTSEPEVTVTDRTE-EDEFLILATDGLWDVVTNEA 349
                    : :*****.*******.**.****** .:**: : :*.: :*: ****:******::*:.

LOC_Os03g16170.1    ACNVVRACLRSSGRRERNRSSPTSNLSPRQSSSSGDEAPNDGAPSAAAGSESDEESAAEE 368
AK367207            ACNVARQCLEDG--------SPTR-------------APAAGAPTAPAISQ-------EE 224
AT1G07430           ACTMVRMCLNRKSGRGRRRGETQTPGRRSEEEGKEEEEKVVGSRKNGKRGE-------IT 402
                    ** :.* **.          ..                    *: .    .:
                                      VII
LOC_Os03g16170.1    DKACAEAAVLLTKLALARQTSDNVSVVVVNLRRRKL---- 404
AK367207            EPRCFRAAALLARLALGRESSDNISVVVVDLKARG----- 259
AT1G07430           DKACTEASVLLTKLALAKHSSDNVSVVVIDLRRRRKRHVA 442
                    :  * .*:.**::***.:.:***:****::*: *


LOC_Os05g38290.2    MAAEICREEAAKSMPAAAAGATAIARRRRRVEGFRFAAGSLEPPQ---EDADAGVARCGK 57
AT1G07430           -MADICYEDETSACESRPLWSS--RKWRIGVQRFRMSPSEMNPTASTTEEEDKSEGIYNK 57
MLOC_75421.1        ------------------------------MKMFSCFGIRTRPAQPALAKDQSVVGINNV 30
                                                 :: *       .*.     . :   .  .

LOC_Os05g38290.2    RQRVAG----ARAGAGAATAGPCRPSAGAEFG-------------------------SRW 88
AT1G07430           RNKQEEYDFMNCASSSPSQSSPEEESVSLEDSDVSISDGNSSVNDVAVIPSKKTVKETDL 117
MLOC_75421.1        PAAGDG-----CHLRSPVGKEAEGSGALVR--------------------------AAW 58
                                   ..    .   ..  .                          :
                                                                      IV
LOC_Os05g38290.2    WPRYGVTSVFGRR-REMEDAVSIRPDFLRGSTSSG--HHHFFGVFDGHGCSHVARMCQDR 145
AT1G07430           RPRYGVASVCGRR-RDMEDAVALHPSFVRKQTEFSRTHWHYFGVYDGHGCSHVARCKER 176
MLOC_75421.1        RGAYGAAGKSREHGVAMQDTVSMRPSFCTWVDGSR---MHFFAVFDGHGGPVVSALLRDH 115
                       **.:.   .:   *:*:*:::*.*           *:*.*:**** . *:   :::

LOC Os05g38290.2    MHELVVDAYKKAVSGKEAAAAAPAWKDVMEKGFARMDDEATIWA-KSRTGGEPACRCELQ 204
AT1G07430           LHELVQ---EEALSDK-----KEEWKKMMERSFTRMDKEVVRWG-ETVMSAN--CRCELQ 225
MLOC_75421.1        MHAILAEELTGAAAAHR------------KKQHQDEGAEVCAWKGALRQSFARADELAAS 163
                    :* ::      * : :            :: .   . *.  *      .     .   .
                                VIII          III
LOC_Os05g38290.2    TPARCDHVGSTAVVAVVGDNRVVVANSGDSRAVLCRAGVPVPLSVDHK-PDRPDELERIK 263
AT1G07430           TPD-CDAVGSTAVVSVITPEKIIVANCGDSRAVLCRNGKAVPLSTDHK-PDRPDELDRIQ 283
MLOC_75421.1        AVPRGTIMGSTAVVALVVGGRLIVANCGDSRAVLCRAGRAVPLSQDHKQPARPDEMA--- 220
                    :      :******:::   :::***.********* * .**** *** * ****:

LOC_Os05g38290.2    AAGGRVIYWDGARVLGVLAMSRAIGDGYLKPYVTSEPEVTVTERADDDECLILASDGLWD 323
AT1G07430           EAGGRVIYWDGARVLGVLAMSRAIGDNYLKPYVTSEPEVTVTDRTEEDEFLILATDGLWD 343
MLOC_75421.1        -AGGVMHHNDGPTGSRALGMDNDK----------------------------------- 243
                     *** : : **.    .*.*..

LOC_Os05g38290.2    VVTNEMACEVVRACFRSNGPPSPPGCSRPKAVLPPPAGASGGGGGDAVVKGVDKAE---- 379
AT1G07430           VVTNEAACTMVRMCLNRK---SGRGRRRGETQTPGRRSEEEGKEEEEKVVGSRKNGKRGE 400
MLOC_75421.1        ------------------------------------------------------------

LOC_Os05g38290.2    -SDKACADAALLLAKLAIARRSADNVSVVVVDLRRPVP---- 416
AT1G07430           ITDKACTEASVLLTKLALAKHSSDNVSVVVIDLRRRRKRHVA 442
MLOC_75421.1        -------------------------------------------

LOC Os05g46040.1    ---------MAAAAAAAAICG---------EDETAAR----VGCTGEWAGG--------- 29
AK362128            ------------MAAAAAICG---------EDEPAPRDPAAAAECAAAGGG--------- 30
AT1G72770           MEEMTPAVAMTLSLAANTMCESSPVEITQLKNVTDAADLLSDSENQSFCNGGTECTMEDV 60
                              ** ::*          :: . .      .      .*
```


LOC_Os05g46040.1 IERVDLGER---------- 38
AK362128 VERLDLGDG---------- 39
AT1G72770 SELEEVGEQDLLKTLSDTRSGSSNVFDEDDVLSVVEDNSAVISEGLLVVDAGSELSLSNT 120
LOC_Os05g46040.1 ----------KEAVAAAG----------AGKRS 51
AK362128 ----------RAALVAG----------GKRS 50
AT1G72770 AMEIDNGRVLATAIIVGESSIEQVPTAEVLIAGVNQDTNTSEVVIRLPDENSNHLVKGRS 180
LOC_Os05g46040.1 VYLMDCAPVWGCASTRGRSAEMEDASAAVPRFADVPVRLLASRRDLDALGLDADALRLPA 111
AK362128 VYLMECEPVWGCVATPGRGGEMEDACAAVPRFADVPVRLLARRQDLDGLGLDADALRLPS 110
AT1G72770 VYELDCIPLWGTVSIQGNRSEMEDAFAVSPHFLKLPIKMLMG----DHEGMSPSLTHLTG 236
IV
LOC_Os05g46040.1 HLFGVFDGHGGAEVANYCRERIHVVLSEELKRLGKNLGEMGEVD-MKEHWDDVFTKCFQR 170
AK362128 HLFAVFDGHGGSEVSNYCRERLHVVLSKELRRPPKDLGEMSDVD-MKEHWDDLFTKCFQT 169
AT1G72770 HFFGVYDGHGGHKVADYCRDRLHFALAEEIERIKDELCKRNTGEGRQVQWDKVFTSCFLT 296
VIII III
LOC_Os05g46040.1 VDDEVSGRVTRVVNGGGEVRSEPVTAENVGSTAVVALVCSSHVVVANCGDSRIVLCRGKE 230
AK362128 VDDEVSGLASRLVD--GEPRLEPIAAENVGSTAVAVVVCSSHVVVANCGDSRIVLSRGKE 227
AT1G72770 VDGEIEGKIGRAVVGSSDKVLEAVASETVGSTAVVALVCSSHVVVSNCGDSRAVLFRGKE 356
VI II
LOC Os05g46040.1 PVALSIDHKPDRKDERARIEAQGGKVIQWNGYRVSGILAMSRSIGDRYLKPEVIPKPEVM 290
AK362128 PVALSIDQKPDRKDERARIEAAGGKVIQWNGHRVSGILAMSRSIGDRYLKPYIIPKPEVT 287
AT1G72770 AMPLSVDHKPDREDEYARIENAGGKVIQWQGARVFGVLAMSRSIGDRYLKPYVIPEPEVT 416
I
LOC Os05g46040.1 VVPRAKDDDCLILASDGLWDVVSNEEACKVARRQILLWHKNNGAASPLSDEGEGSTDPAA 350
AK362128 VVPRAKDDDCLILASDGLWDVVSNEEACKVARRQIQQWHKNNSVTTSSSDGGDGSTDPAA 347
AT1G72770 FMPRSREDECLILASDGLWDVMNNQEVCEIARRRILMWHKKNGAP-PLAERGKG-IDPAC 474
VII
LOC_Os05g46040.1 QAAADYLMRLALKKGSEDNITVIVVDLKPRKKLKNIS 387
AK362128 QAAADYLVRLALKKGSQDNITVIVVDLKPRRKSKNNS 384
AT1G72770 QAAADYLSMLALQKGSKDNISIIVIDLKAQRKFKTRT 511
LOC Os05g46040.1 ---------- 0
MLOC_7586.1 ---------- 0
AT1G72770 MEEMTPAVAMTLSLAANTMCESSPVEITQLKNVTDAADLLSDSENQSFCNGGTECTMEDV 60
LOC Os05g46040.1 ---------- 0
MLOC 7586.1 ---------- 0
AT1G72770 SELEEVGEQDLLKTLSDTRSGSSNVFDEDDVLSVVEDNSAVISEGLLVVDAGSELSLSNT 120
LOC_Os05g46040.1 ----MAAAAAAAAICGEDETAARVGCTGEWAGGIER-VDLGERKEAVAAAGAGK----RS 51
MLOC 7586.1 ---------MAAAVFAGDGAAARGGCSAECAGGIERPPDLGSR----AGDGCGK----RS 43
AT1G72770 AMEIDNGRVLATAIIVGESSIEQVPTAEVLIAGVNQDTNTSEVVIRLPDENSNHLVKGRS 180
LOC_Os05g46040.1 VYLMDCAPVWGCASTRGRSAEMEDASAAVPRFADVPVRLLASRRDLDALG--LDADALRL 109
MLOC 7586.1 VYLMECVPLWGCAAARGRAAEMEDACAAVPRFAALPARMLASSRELDGIGGDFDAAELRL 103
AT1G72770 VYELDCIPLWGTVSIQGNRSEMEDAFAVSPHFLKLPIKMLMG----DHEG--MSPSLTHL 234
IV
LOC Os05g46040.1 PAHLFGVFDGHGGAEVANYCRERIHVVLSEELK--RLGKNLGEMGEVDMKEHWDDVFTKC 167
MLOC 7586.1 PAHLFGVYDGHGGSEVANYCRDKIHVVLREVLRDGRGLEELGEVGEVDVKESWEKVFGDC 163
AT1G72770 TGHFFGVYDGHGGHKVADYCRDRLHFALAEEIERIKDELCKRNTGE-GRQVQWDKVFTSC 293
VIII III
LOC_Os05g46040.1 FQRVDDEVSGRVTRVVNGGGEVRSEPVTAENVGSTAVVALVCSSHVVVANCGDSRIVLCRG 227
MLOC 7586.1 FQKVDDEVSGKAIRFSNGVTELRPEPIAADNVGSTAVVAIVCSSHVITANCGDSRVVLCRG 223
AT1G72770 FLTVDGEIEGKIGRAVVGSSDKVLEAVASETVGSTAVVALVCSSHIVVSNCGDSRAVLFRG 353

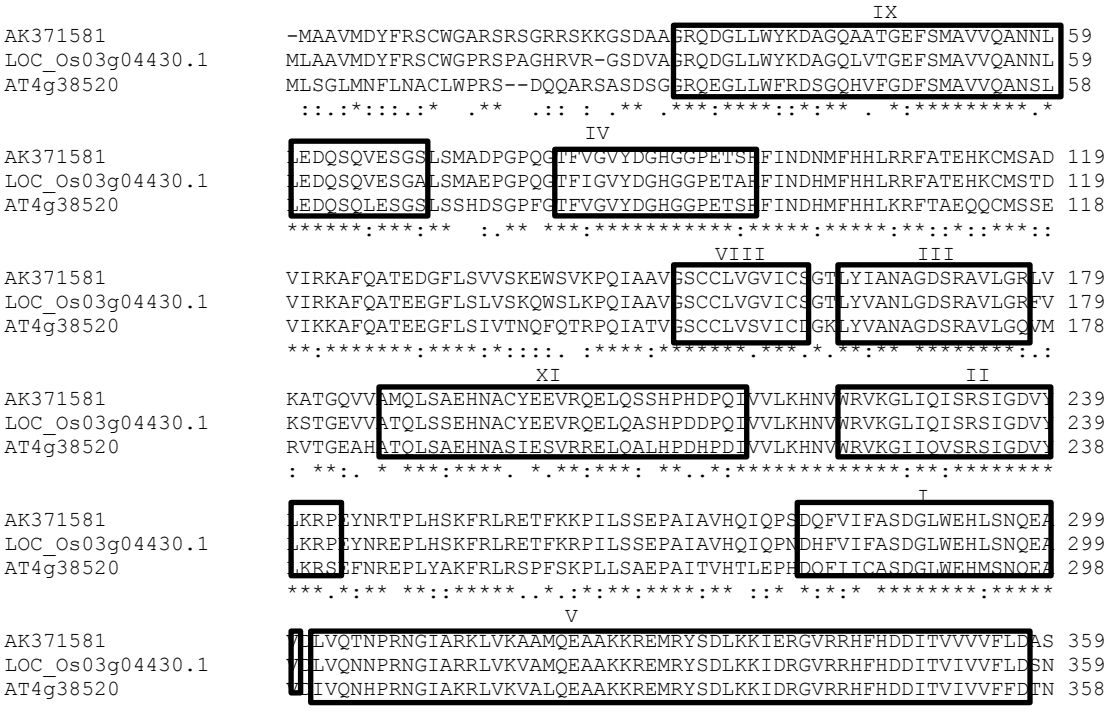

VI
II
LOC_Os05g46040.1 GKEPVALSIDHKPDRKDERARIEAQGGKVIQWNGYRVSGILAMSRSIGDRYLKPFVIPKP 287
MLOC_7586.1 GKEPIALSVDHKPDGKDERARIEAAGGKVIQWNGYRVSGILAMSRSIGDRYLKPFLIPKP 283
AT1G72770 GKEAMPLSVDHKPDREDEYARIENAGGKVIQWQGARVFGVLAMSRSIGDRYLKPYVIPEP 413
I
LOC_Os05g46040.1 EVMVVPRAKDQDCLILASDGLWDVVSNEEACKVARRQILLWHKNNGAASPLSDEGEGSTD 347
MLOC_7586.1 EVSVVPRAKDQDCLILASDGLWDVMSNEDACKVARRQILLWYKNNNDGANSDGGSEPTMN 343
AT1G72770 EVTFMPRSREQECLILASDGLWDVMNNQEVCEIARRRILMWHKKNG-APPLAERGKG-ID 471
VII
LOC_Os05g46040.1 PAAQAAADQLMRLALKKGSEDNITVIVVDLKPRKKLKNIS 387
MLOC_7586.1 PAAKAAADQLVRLALMKGSGDNISVIVIDLKSRKKPKGKS 383
AT1G72770 PACQAAADQLSMLALQKGSKDNISIIVIDLKAQRKFKTRT 511
LOC_Os05g46040.1 ----------
MLOC_7586.1 ----------
AT3G63320 KMKNTYALPCEFLSSQPNLFRMG 423
LOC_Os05g49730.1 -MR-ETGATDEGHACEVVVAGGDG---------------KAAAARRRRRLELRRLG--LA 41
AK376382 -MRQAVGAGEEGLPRKVVLAAGDAGEEGLPRKMVLAAGDADGAARRRRRLELRRHGRKVA 59
AT2G29380 MAEICYEVVTDACPSSVYESTPAH---------------SRRRPRFQTVMHEDWEKNCKR 45
LOC_Os05g49730.1 AEDDAAAKRIRSVKDGSSSDDSSTEVVPRSWPACVSHGSVSVIGRRREMEDAVAIERTFM 101
AK376382 SGEDEPAR---KATDGSESE-SSADVG--WLPACLSHGAVSVIGRRREMEDAVAVERTFL 113
AT2G29380 SKQEALATRYSSIPRSSREDFSDQNVD----VSSPRYGVSSVCGRRREMEDAVAIHPSFS 101
IV
LOC_Os05g49730.1 ASTGDGAGAIRGGGEGEEDFFAVYDGHGGSRVAEACRKRMHVVLAEEVSLRRLRGQSASG 161
AK376382 APPCGGG---DEGSGGEEDFFAVYDGHGGARVAEACRERMHVVLAEEVARLRCR-PGARG 169
AT2G29380 SP---------KNSEFPQHYFGVYDGHGCSHVAARCRERLHKLVQEELS---SDMEDEEE 149
VIII
LOC_Os05g49730.1 GDVRWKEAMLASFARMDGEVV--GSVAAAAPRVDGTEPSGFRTVGSTAVVAVVGRRRIVV 219
AK376382 ----WKEALEASFARVDGEVV--GSAAAGAD-ADADEESRSRTVGSTAVVAVVGRRRIVV 222
AT2G29380 ----WKTTMERSFTRMDKEVVSWGDSVVTANCKCDLQTPACDSVGSTAVVSVITPDKLVV 205
III
VI
II
LOC_Os05g49730.1 ANCGDSRAVLSRGGVALPLSTDHKPDRPDELERVEAAGGRVINWNGYRVLGVLATSRSIG 279
AK376382 ANCGDSRAVLSRGGVAVPLSTDHKPDRPDELQRVEAAGGRVINWNGSRVLGVLSTSRSIG 282
AT2G29380 ANCGDSRAVLCRNGKPVPLSTDHKPDRPDELDRIEGAGGRVIYWDCPRVLGVLAMSRAIG 265
I
LOC_Os05g49730.1 DYYLKPFVSAEPEVRVVERTDKDEFLILASDGLWDVVSNEVACKIARNCLNG--RAASMF 337
AK376382 DYYLKPYVSAEPEVTAVERTDKDEFLVLASDGLWDVVSNEAACRVARSCLTG--RAAAAF 340
AT2G29380 DNYLKPYVSCEPEVTITDRRD-DDCLILASDGLWDVVSNETACSVARMCLRGGGRRQDNE 324
VII
LOC_Os05g49730.1 PESVSGSSAADAAALLAELAVSRGSRDNISVVVVELRRLKSRAA---------- 381
AK376382 PESVSGRTAADAAALLAELAISRGSKDNISVVVVELKRLKSRVGRRRAVGSEVQL 395
AT2G29380 DPAISDKACTEASVLLTKLALARNSSDNVSVVVIDLRR---------------- 362
LOC_Os05g49730.1 MRETGATDEGHACEVVVAGGDGKAAAARRRRRLELRRLGLAAEDDAAAKRIRSVKDGSSS 60
MLOC_15065.1 --
AT2G29380 MAEICYEVVTDACPSSVY--ESTPAHSRRRPRFQTVMHEDWEKNCKRSKQEALATRYSSI 58
LOC_Os05g49730.1 DDSSTEVVPRSWPACVS--HGSVSVIGRRREMEDAVAIERTFMASTGDGAGAIRGGGEGE 118
MLOC_15065.1 ------------------------------MEDAVAVERTFLAPPCGGG---DEGSGGE 26
AT2G29380 PRSSREDFSDQNVDVSSPRYGVSSVCGRRREMEDAVAIHPSFSSP---------KNSEFP 109
IV
LOC_Os05g49730.1 EDFFAVYDGHGGSRVAEACRKRMHVVLAEEVSLRRLRGQSASGGDVRWKEAMLASFARMD 178
MLOC_15065.1 EDFFAVYDGHGGARVAEACRERMHVVLAEEVARLRCR-PGARG----WKEALEASFARVD 81
AT2G29380 QHYFGVYDGHGCSHVAARCRERLHKLVQEELS-------SDMEDEEEWKTTMERSFTRMD 162
VIII
III
LOC_Os05g49730.1 GEVV--GSVAAAAPRVDGTEPSGFRTVGSTAVVAVVGRRRIVVANCGDSRAVLSRGGVAL 236
MLOC_15065.1 GEVV--GSAAAGAD-ADADEESRSRTVGSTAVVAVVGRRRIVVANCGDSRAVLSRGGVAV 138
AT2G29380 KEVVSWGDSVVTANCKCDLQTPACDSVGSTAVVSVITPDKIVVANCGDSRAVLCRNGKPV 222

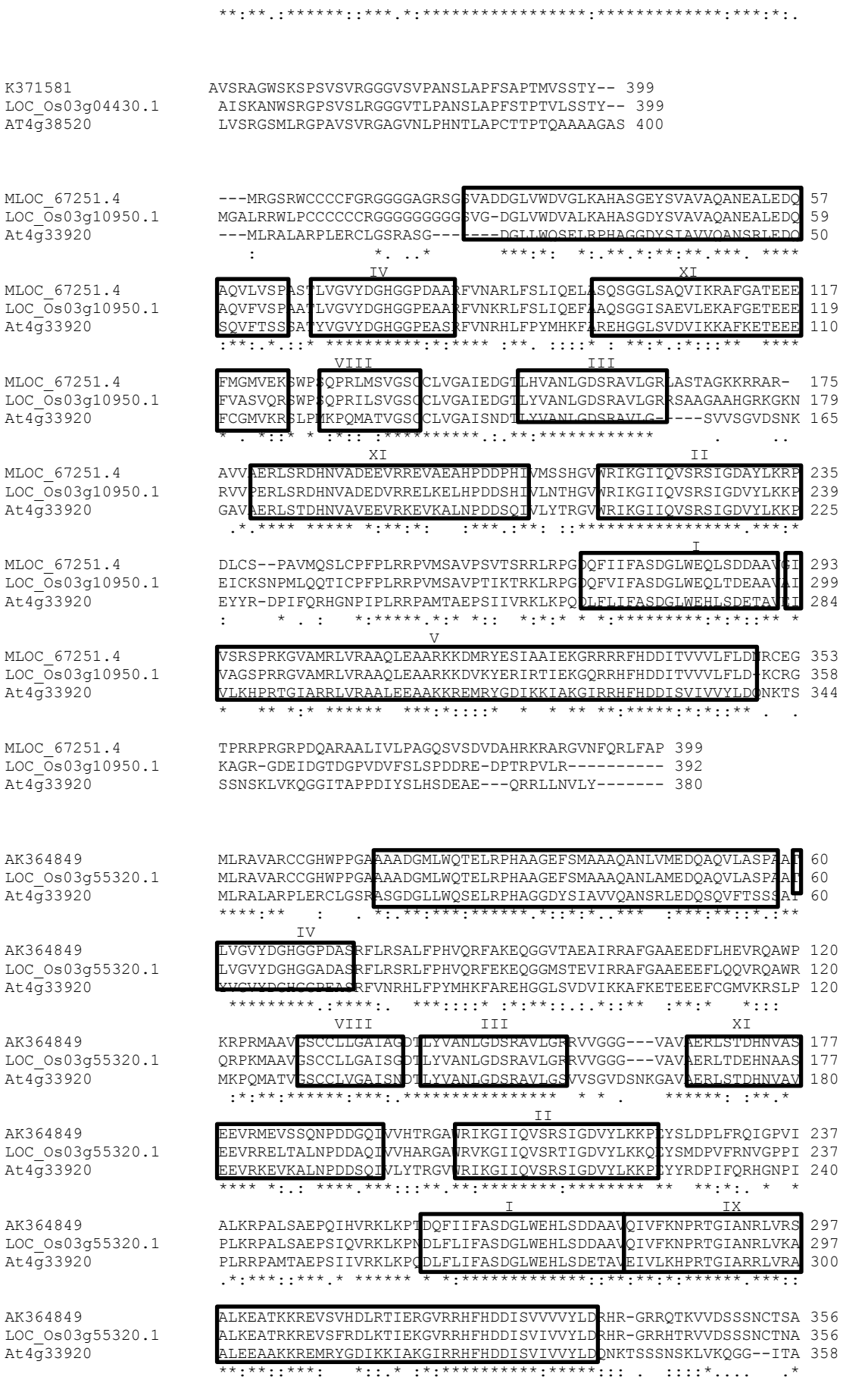


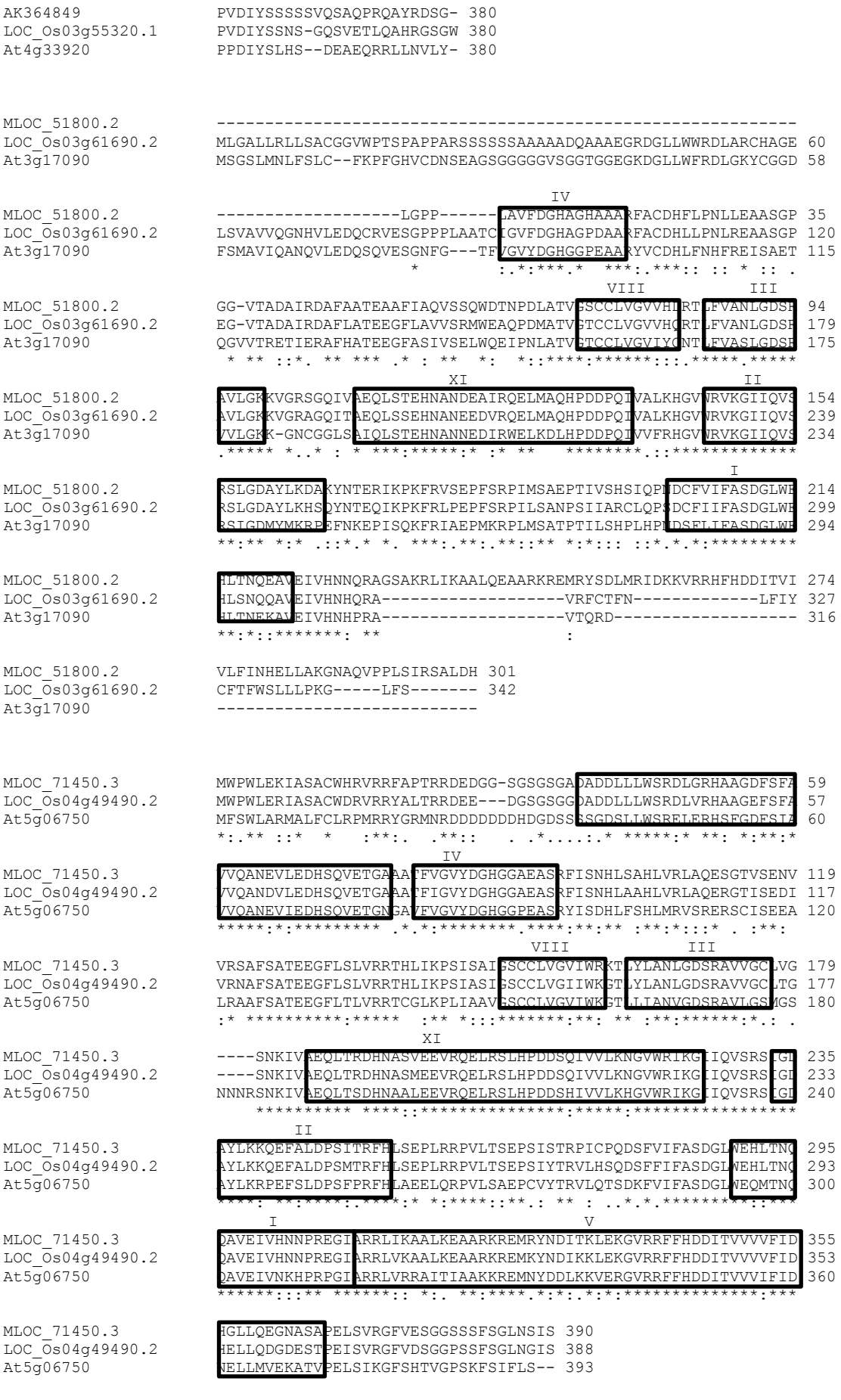


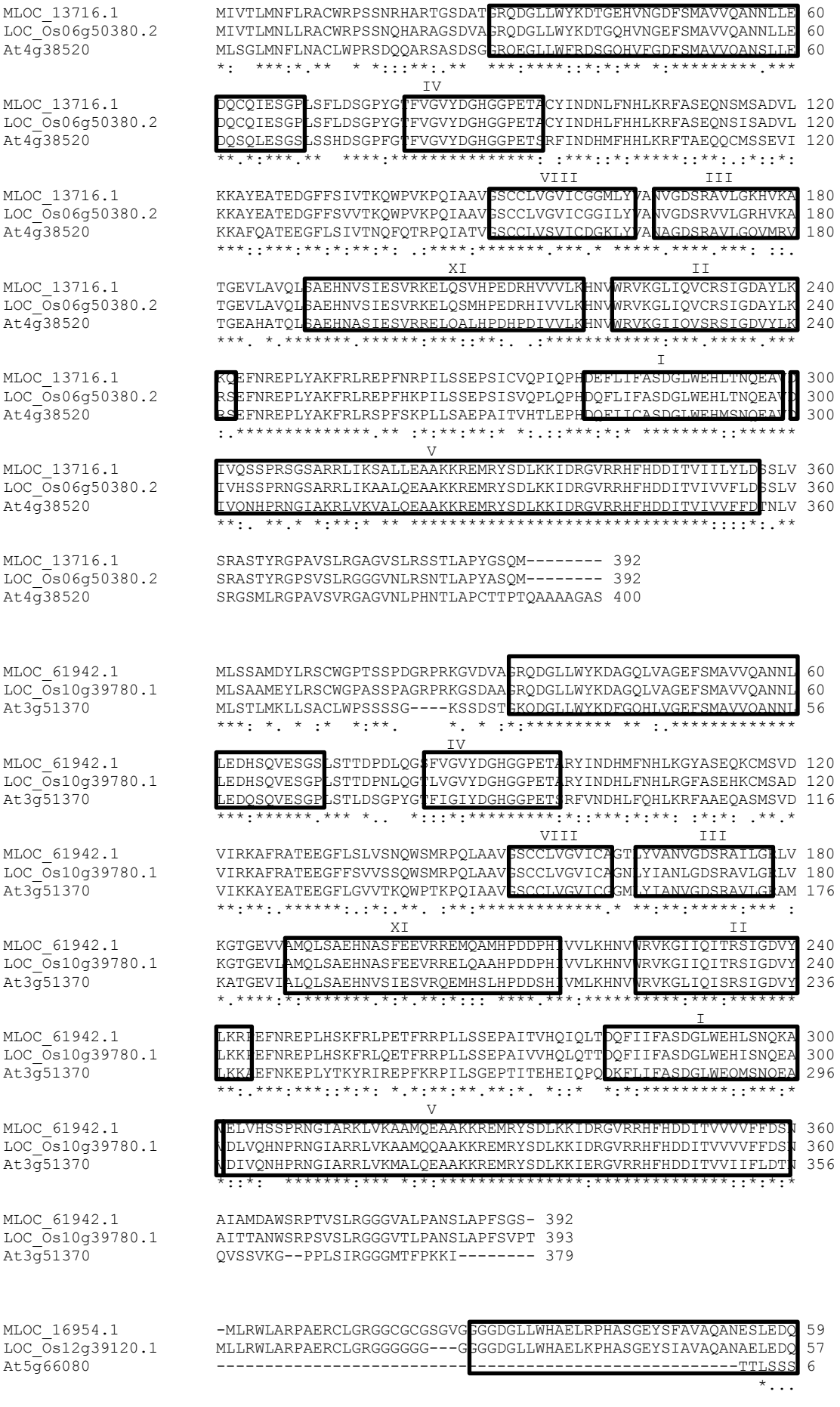
Asterisk indicate the conservation of residue in all the three species and dashes indicate the gaps or absence of sequence data; alignments were made using CLUSTALW2 (http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=clustalw2; last accessed May 2014) and then bioedit was used to align the conserved functional motifs; box: conserved functional motifs; roman numbers: indicate the specific motifs; bold indicate amino acid involved in binding of PP2CA to PYR/PYL/RCAR.; rice sequences were derived from the MSU rice genomic annotation project (http://rice.plantbiology.msu.edu/; last accessed May 2014) and Arabidopsis sequences were derived from TAIR (http://www.arabidopsis.org/index.jsp); motif sequences and the positions were derived from Xue *et al*. (2008).

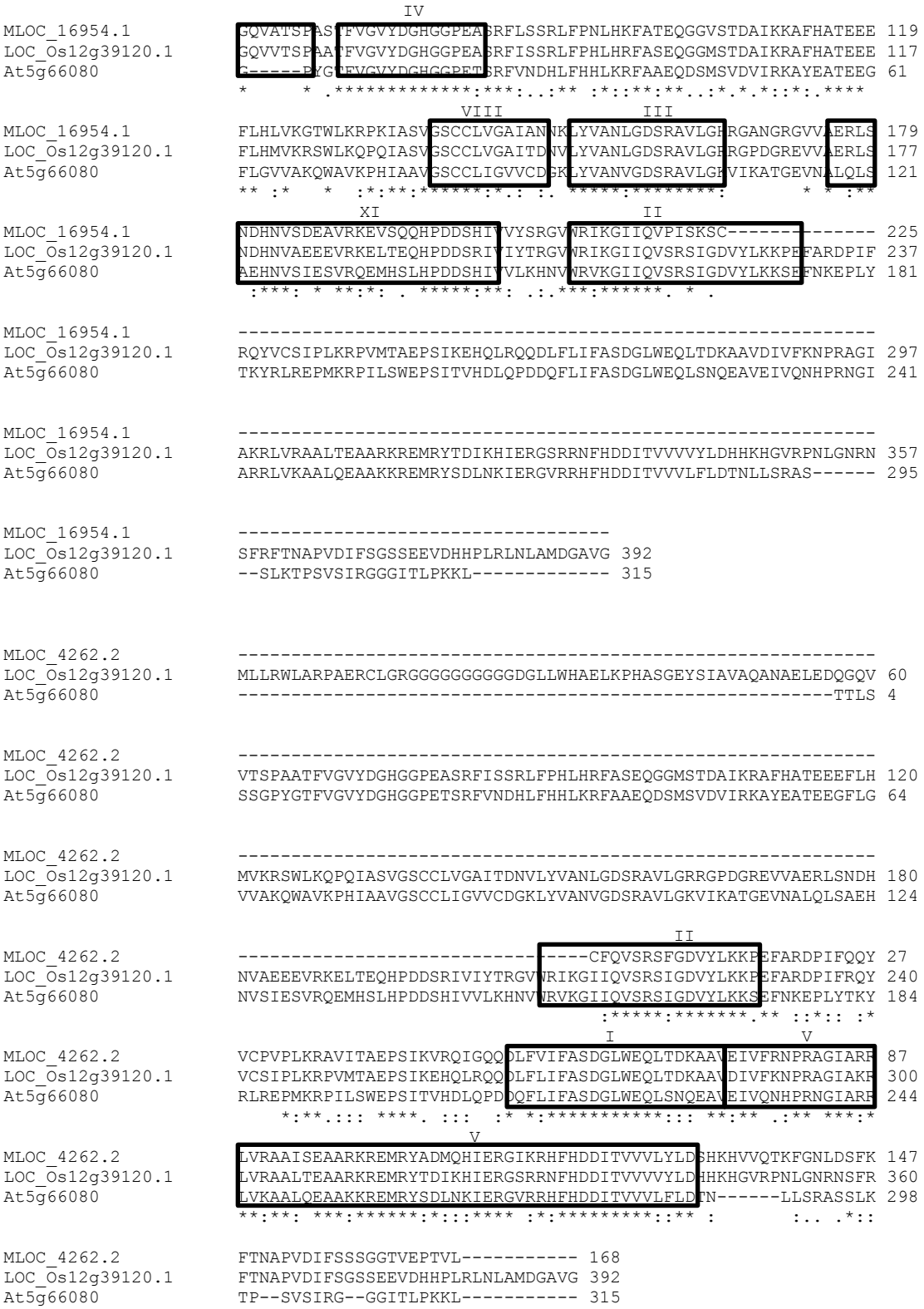
Appendix XII Alignment of putative amino acid sequences of barley PP2CDs with rice and Arabidopsis




K371581 AVSRAGWSKSPSVSVRGGGVSVPANSLAPFSAPTMVSSTY-- 399
LOC_Os03g04430.1 AISKANWSRGPSVSLRGGGVTLPANSLAPFSTPTVLSSTY-- 399
AT4g38520 LVSRGSMLRGPAVSVRGAGVNLPHNTLAPCTTPTQAAAAGAS 400
MLOC_67251.4 ---MRGSRWCCCCFGRGGGGAGRSGSVADDGLVWDVGLKAHASGEYSVAVAQANEALEDQ 57
LOC_Os03g10950.1 MGALRRWLPCCCCCCRGGGGGGGGGSVG-DGLVWDVALKAHASGDYSVAVAQANEALEDQ 59
At4g33920 ---MLRALARPLERCLGSRASG---LGGDGLLWQSELRPHAGGDYSIAVVQANSRLEDQ 50
MLOC_67251.4 AQVLVSPASTLVGVYDGHGGPDAARFVNARLFSLIQELASQSGGLSAQVIKRAFGATEEE 117
LOC_Os03g10950.1 AQVFVSPAATLVGVYDGHGGPEAARFVNKRLFSLIQEFAAQSGGISAEVLEKAFGETEEE 119
At4g33920 SQVFTSSSATYVGVYDGHGGPEASRFVNRHLFPYMHKFAREHGGLSVDVIKKAFKETEEE 110
MLOC_67251.4 FMGMVEKSWPSQPRLMSVGSCCLVGAIEDGTLHVANLGDSRAVLGRLASTAGKKRRAR- 175
LOC_Os03g10950.1 FVASVQRSWPSQPRILSVGSCCLVGAIEDGTLYVANLGDSRAVLGRRSAAGAAHGRKGKN 179
At4g33920 FCGMVKRSLPMKPQMATVGSCCLVGAISNDTLYVANLGDSRAVLG----SVVSGVDSNK 165
MLOC_67251.4 AVVAERLSRDHNVADEEVRREVAEAHPDDPHIVMSSHGVWRIKGIIQVSRSIGDAYLKRP 235
LOC_Os03g10950.1 RVVPERLSRDHNVADEDVRRELKELHPDDSHIVLNTHGVWRIKGIIQVSRSIGDVYLKKP 239
At4g33920 GAVAERLSTDHNVAVEEVRKEVKALNPDDSQIVLYTRGVWRIKGIIQVSRSIGDVYLKKP 225
MLOC_67251.4 DLCS--PAVMQSLCPFFPLRRPVMSAVPSVTSRRLRPGDQFIIFASDGLWEQLSDDAAVSI 293
LOC_Os03g10950.1 EICKSNPMLQQTICPFFPLRRPVMSAVPTIKTRKLRPGDQFVIFASDGLWEQLTDEAAVAI 299
At4g33920 EYYR-DPIFQRHGNPIPLRRPAMTAEPSIIVRKLKPQDLFLIFASDGLWEHLSDETAVAI 284
MLOC_67251.4 VSRSPRKGVAMRLVRAAQLEAARKKDMRYESIAAIEKGRRRRFHDDITVVVLFLDNRCEG 353
LOC_Os03g10950.1 VAGSPRRGVAMRLVRAAQLEAARKKDVKYERIRTIEKGQRRHFHDDITVVVLFLDHKCRG 358
At4g33920 VLKHPRTGIARRLVRAALEEAAKKREMRYGDIKKIAKGIRRHFHDDISVIVVYLDQNKTS 344
MLOC_67251.4 TPRRPRGRPDQARAALIVLPAGQSVSDVDAHRKRARGVNFQRLFAP 399
LOC_Os03g10950.1 KAGR-GDEIDGTDGPVDVFSLSPDDRE-DPTRPVLR---------- 392
At4g33920 SSNSKLVKQGGITAPPDIYSLHSDEAE---QRRLLNVLY------- 380
AK364849 MLRAVARCCGHWPPGAAAADGMLWQTELRPHAAGEFSMAAAQANLVMEDQAQVLASPAAT 60
LOC_Os03g55320.1 MLRAVARCCGHWPPGAAAADGMLWQTELRPHAAGEFSMAAAQANLAMEDQAQVLASPAAT 60
At4g33920 MLRALARPLERCLGSRASGDGLLWQSELRPHAGGDYSIAVVQANSRLEDQSQVFTSSSAT 60
AK364849 LVGVYDGHGGPDASRFLRSALFPHVQRFAKEQGGVTAEAIRRAFGAAEEDFLHEVRQAWP 120
LOC_Os03g55320.1 LVGVYDGHGGADASRFLRSRLFPHVQRFEKEQGGMSTEVIRRAFGAAEEEFLQQVRQAWR 120
At4g33920 VVGVYDGHGGPEASRFVNRHLFPYMHKFAREHGGLSVDVIKKAFKETEEEFCGMVKRSLP 120
AK364849 KRPRMAAVGSCCLLGAIAGDTLYVANLGDSRAVLGRVVGGG---VAVAERLSTDHNVAS 177
LOC_Os03g55320.1 QRPKMAAVGSCCLLGAISGDTLYVANLGDSRAVLGRVVGGG---VAVAERLTDEHNAAS 177
At4g33920 MKPQMATVGSCCLVGAISNDTLYVANLGDSRAVLGSVVSGVDSNKGAVAERLSTDHNVAV 180
AK364849 EEVRMEVSSQNPDDGQIVVHTRGAVRIKGIIQVSRSIGDVYLKKPLYSLDPLFRQIGPVI 237
LOC_Os03g55320.1 EEVRRELTALNPDDAQIVVHARGAVRVKGIIQVSRTIGDVYLKKQLYSMDPVFRNVGPPI 237
At4g33920 EEVRKEVKALNPDDSQIVLYTRGVVRIKGIIQVSRSIGDVYLKKPLYYRDPIFQRHGNPI 240
AK364849 ALKRPALSAEPQIHVRKLKPTDQFIIFASDGLWEHLSDDAAVQIVFKNPRTGIANRLVRS 297
LOC_Os03g55320.1 PLKRPALSAEPSIQVRKLKPNDLFLIFASDGLWEHLSDDAAVQIVFKNPRTGIANRLVKA 297
At4g33920 PLRRPAMTAEPSIIVRKLKPQDLFLIFASDGLWEHLSDETAVEIVLKHPRTGIARRLVRA 300
AK364849 ALKEATKKREVSVHDLRTIERGVRRHFHDDISVVVVYLDRHR-GRRQTKVVDSSSNCTSA 356
LOC_Os03g55320.1 ALKEATRKREVSFRDLKTIEKGVRRHFHDDISVIVVYLDRHR-GRRHTRVVDSSSNCTNA 356
At4g33920 ALEEAAKKREMRYGDIKKIAKGIRRHFHDDISVIVVYLDQNKTSSSNSKLVKQGG--ITA 358


AK364849 PVDIYSSSSSVQSAQPRQAYRDSG- 380
LOC_Os03g55320.1 PVDIYSSNS-GQSVETLQAHRGSGW 380
At4g33920 PPDIYSLHS--DEAEQRRLLNVLY- 380
MLOC_51800.2 ----------
LOC_Os03g61690.2 MLGALLRLLSACGGVWPTSPAPPARSSSSSSAAAAADQAAAEGRDGLLWWRDLARCHAGE 60
At3g17090 MSGSLMNLFSLC--FKPFGHVCDNSEAGSGGGGGVSGGTGGEGKDGLLWFRDLGKYCGGD 58
IV
MLOC_51800.2 -------------------LGPP------LAVFDGHAGHAAARFACDHFLPNLLEAASGP 35
LOC_Os03g61690.2 LSVAVVQGNHVLEDQCRVESGPPPLAATCIGVFDGHAGPDAARFACDHLLPNLREAASGP 120
At3g17090 FSMAVIQANQVLEDQSQVESGNFG---TFVGVYDGHGGPEAARYVCDHLFNHFREISAET 115
VIII III
MLOC_51800.2 GG-VTADAIRDAFAATEAAFIAQVSSQWDTNPDLATVGSCCLVGVVHIRTLFVANLGDSR 94
LOC_Os03g61690.2 EG-VTADAIRDAFLATEEGFLAVVSRMWEAQPDMATVGTCCLVGVVHCRTLFVANLGDSR 179
At3g17090 QGVVTRETIERAFHATEEGFASIVSELWQEIPNLATVGTCCLVGVIYQNTLFVASLGDSR 175
XI II
MLOC_51800.2 AVLGKKVGRSGQIVAEQLSTEHNANDEAIRQELMAQHPDDPQIVALKHGVVRVKGIIQVS 154
LOC_Os03g61690.2 AVLGKKVGRAGQITAEQLSSEHNANEEDVRQELMAQHPDDPQIVALKHGVVRVKGIIQVS 239
At3g17090 VVLGKK-GNCGGLSAIQLSTEHNANNEDIRWELKDLHPDDPQIVVFRHGVVRVKGIIQVS 234
I
MLOC_51800.2 RSLGDAYLKDAKYNTERIKPKFRVSEPFSRPIMSAEPTIVSHSIQPNDCFVIFASDGLWE 214
LOC_Os03g61690.2 RSLGDAYLKHSQYNTEQIKPKFRLPEPFSRPILSANPSIIARCLQPSDCFIIFASDGLWE 299
At3g17090 RSIGDMYMKRPEFNKEPISQKFRIAEPMKRPLMSATPTILSHPLHPLDSFLIFASDGLWE 294
MLOC_51800.2 HLTNQEAVEIVHNNQRAGSAKRLIKAALQEAARKREMRYSDLMRIDKKVRRHFHDDITVI 274
LOC_Os03g61690.2 HLSNQQAVEIVHNHQRA-------------------VRFCTFN-------------LFIY 327
At3g17090 HLTNEKAEIVHNHPRA-------------------VTQRD-------------------- 316
MLOC_51800.2 VLFINHELLAKGNAQVPPLSIRSALDH 301
LOC_Os03g61690.2 CFTFWSLLLPKG-----LFS------- 342
At3g17090 ----------
MLOC_71450.3 MWPWLEKIASACWHRVRRFAPTRRDEDGG-SGSGSGADADDLLLWSRDLGRHAAGDFSFA 59
LOC_Os04g49490.2 MWPWLERIASACWDRVRRYALTRRDEE---DGSGSGGDADDLLLWSRDLVRHAAGEFSFA 57
At5g06750 MFSWLARMALFCLRPMRRYGRMNRDDDDDDDHDGDSSSGDSLLWSRELERHSFGDFSIA 60
IV
MLOC_71450.3 VVQANEVLEDHSQVETGAAAIFVGVYDGHGGAEASRFISNHLSAHLVRLAQESGTVSENV 119
LOC_Os04g49490.2 VVQANDVLEDHSQVETGAAAIFIGVYDGHGGAEASRFISNHLAAHLVRLAQERGTISEDI 117
At5g06750 VVQANEVIEDHSQVETGNGAVFVGVYDGHGGPEASRYISDHLFSHLMRVSRERSCISEEA 120
VIII III
MLOC_71450.3 VRSAFSATEEGFLSLVRRTHLIKPSISAISSCCLVGVIWRKTLYLANLGDSRAVVGCLVG 179
LOC_Os04g49490.2 VRNAFSATEEGFLSLVRRTHLIKPSIASISSCCLVGIIWKGTLYLANLGDSRAVVGCLTG 177
At5g06750 LRAAFSATEEGFLTLVRRTCGLKPLIAAVESCCLVGVIWKGTLLIANVGDSRAVLGSMGS 180
XI
MLOC_71450.3 ----SNKIVAEQLTRDHNASVEEVRQELRSLHPDDSQIVVLKNGVWRIKGIIQVSRSIGL 235
LOC_Os04g49490.2 ----SNKIVAEQLTRDHNASMEEVRQELRSLHPDDSQIVVLKNGVWRIKGIIQVSRSIGL 233
At5g06750 NNNRSNKIVAEQLTSDHNAALEEVRQELRSLHPDDSHIVVLKHGVWRIKGIIQVSRSIGL 240
II
MLOC_71450.3 AYLKKQEFALDPSITRFHLSEPLRRPVLTSEPSISTRPICPQDSFVIFASDGLWEHLTNQ 295
LOC_Os04g49490.2 AYLKKQEFALDPSMTRFHLSEPLRRPVLTSEPSIYTRVLHSQDSFFIFASDGLWEHLTNQ 293
At5g06750 AYLKRPEFSLDPSFPRFHLAEELQRPVLSAEPCVYTRVLQTSDKFVIFASDGLWEQMTNQ 300
I V
MLOC_71450.3 QAVEIVHNNPREGIARRLIKAALKEAARKREMRYNDITKLEKGVRRFFHDDITVVVVFID 355
LOC_Os04g49490.2 QAVEIVHNNPREGIARRLVKAALKEAARKREMKYNDIKKLEKGVRRFFHDDITVVVVFID 353
At5g06750 QAVEIVNKHPRPGIARRLVRRAITIAAKKREMNYDDLKKVERGVRRFFHDDITVVVIFID 360
MLOC_71450.3 HGLLQEGNASAPELSVRGFVESGGSSSFSGLNSIS 390
LOC_Os04g49490.2 HELLQDGDESTPEISVRGFVDSGGPSSFSGLNGIS 388
At5g06750 NELLMVEKATVPELSIKGFSHTVGPSKFSIFLS-- 393


MLOC_13716.1
LOC_Os06g50380.2
At4g38520
IV
VIII
III
XI
II
I
V
MLOC_61942.1
LOC_Os10g39780.1
At3g51370
IV
VIII
III
XI
II
I
V
MLOC_16954.1
LOC_Os12g39120.1
At5g66080



Asterisk indicate the conservation of residue in all the three species and dashes indicate the gaps or absence of sequence data; alignments were made using CLUSTALW2 (http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=clustalw2; last accessed May 2014) and then bioedit was used to align the conserved functional motifs; box: conserved functional motifs; roman numbers: indicate the specific motifs; bold indicate amino acid involved in binding of PP2CA to PYR/PYL/RCAR.; rice sequences were derived from the MSU rice genomic annotation project (http://rice.plantbiology.msu.edu/; last accessed May 2014) and Arabidopsis sequences were derived from TAIR (http://www.arabidopsis.org/index.jsp); motif sequences and the positions were derived from Xue *et al*. (2008).

Appendix XIII Alignment of putative amino acid sequences of barley PYRs with rice and Arabidopsis

```
LOC_Os02g13330.1      MEPHMERALREAVASEAERRELEGVVRAHHTFPAAERAAGPGRRPTCTSLVAQRVDAPLA 60
MLOC_39291.1          MEHHMESALRQGLT-EPERRELEGVVEEHHTFPG--RASG-----TCTSLVTQRVQAPLA 52
AT2G26040             ---MSSSPAVKGLT-DEEQKTLEPVIKTYHQFEP---DPT-----TCTSLITQRIHAPAS 48
                           . .  :.:: : *:: ** *:. :* *      .      *****::**:.** :

LOC_Os02g13330.1      AVWPIVRGFANPQRYKHFIKSCELAAGDGATVGSVREVAVVSGLPASTSTERLEILDDDR 120
MLOC_39291.1          AVWDIVRGFANPQRYKHFIKSCALAAGDGATVGSVREVTVVSGLPASTSTERLEILDDDR 112
AT2G26040             VVWPLIRRFDNPERYKHFVKRCRLISGDGD-VGSVREVTVISGLPASTSTERLEFVDDDH 107
                      .** ::* * **:*****:* * * :***  *******:*:**************::***:

LOC_Os02g13330.1      HVLSFRVVGGDHRLRNYRSVTSVTEFSSPSSPPRPYCVVVESYVVDVPEGNTEEDTRMFT 180
MLOC_39291.1          HILSFCVVGGEHRLRNYRSVTSVTEFTDQPSGP-SYCVVVESYVVDVPEGNTEEDTRMFT 171
AT2G26040             RVLSFRVVGGEHRLKNYKSVTSVNEFLNQDSGK-VYTVVLESYTVDIPEGNTEEDTKMFV 166
                      ::*** ****:***:**:*****.** .  *    * **:***.**:**********:**.

LOC_Os02g13330.1      DTVVKLNLQKLAAVATS-SSPPAAGNHH 207
MLOC_39291.1          DTVVKLNLQKLAAIATTTSSPPPLDGQS 199
AT2G26040             DTVVKLNLQKLGVAATS---APMHDDE- 190
                      ***********.. **:   .*  ...



LOC_Os06g33690.1      MNGVGGAGGAAAGKLPMVSHRRVQWRLADERCELREEEMEYIRRFHRHEPSSNQCTSFAA 60
MLOC_46394.1          MD--GGSSGVGADGI---------WRPWDEHTMLCLKEMEYVRRFHQHELGANQCTSFIA 49
AT5G53160             ME----ANGI-------------------ENLTNPNQEREFIRRHHKHELVDNQCSSTLV 37
                      *:    :.*                    *.     :* *::**.*:**   ***:*  .

LOC_Os06g33690.1      KHIKAPLHTVWSLVRRFDQPQLFKPFVRNCVMRENIIATGCIREVNVQSGLPATRSTERL 120
MLOC_46394.1          KHIKAPLQTVWSVVRRFDKPQVFKPFVEKCVMQGN-IEPGCVREVTVKSGLPAKWSIERL 108
AT5G53160             KHINAPVHIVWSLVRRFDQPQKYKPFISRCVVKGN-MEIGTVREVDVKSGLPATRSTERL 96
                      ***:**:: ***:*****:** :***: .**:: * :  * :*** *:*****. * ***

LOC_Os06g33690.1      ELLDDNEHILKVNFIGGDHMLKNYSSILTVHSEVIDGQLGTLVVESFIVDVPEGNTKDDI 180
MLOC_46394.1          ELLDDNEHILRVKFIDGNHPLKNYSSILTVHHEVIDGHPGALVIESFVVDVPEENTENEI 168
AT5G53160             ELLDDNEHILSIRIVGGDHRLKNYSSIISLHPETIEGRIGTLVIESFVVDVPEGNTKDET 156
                      ********** :.::.*:* *******:::* *.*:*: *:**:***:***** **:::

LOC_Os06g33690.1      SYFIENVLRCNLRTLADVSEERLANP------ 206
MLOC_46394.1          FYLVGNFLKVNHKLLADVSEGRIDGRALN--- 197
AT5G53160             CYFVEALIKCNLKSLADISERLAVQDTTESRV 188
                       *::  .:: * : ***:**

LOC_Os03g18600.1      MPCIPASSPGIPHQHQHQHHRALAGVGMAVGCAAEAAVAAAGVAGTRCGAHDGEVPMEVA 60
AK376521              MP-TPYSAAAL-----QQHHRLVSSSG-GLGSAAGAGAGAR-----RCGEHDGTVPPEVA 48
AT5G05440             MR----SPVQL-----QHGSDATNGFHTLQPHDQTDGPIKR-----VCLTRGMHVPEHVA 46
                      *     *.  :     ::      .           .          *  :.  ** .**

LOC_Os03g18600.1      RHHEHAEPGSGRCCSAVVQHVAAPAAAVWSVVRRFDQPQAYKRFVRSCALLAGDG-GVGT 119
AK376521              RHHEHAAPGGRCCCSAVVQRVAAPAADVWAVVRRFDQPQAYKSFVRSCALLDGDG-GVGT 107
AT5G05440             MHHTHD-VGPDQCCSSVVQMIHAPPESVWALVRRFDNPKVYKNFIRQCRIVQGDGLHVGD 105
                       ** *   *   ***:*** : **.  **::*****:*:.** *:*.* :: ***  **

LOC_Os03g18600.1      LREVRVVSGLPAASSRERLEILDDESHVLSFRVVGGEHRLKNYLSVTTVHPSP---SAPT 176
AK376521              LREVRVVSGLPAASSRERLEILDDERHVLSFSVVGGEHRLRNYRSVTTVHPAPGEGASPS 167
AT5G05440             LREVMVVSGLPAVSSTERLEILDEERHVISFSVVGGDHRLKNYRSVTTLHASD------D 159
                      **** *******.** *******:* **:** ****:***:** ****:*.:

LOC_Os03g18600.1      AATVVVESYVVDVPPGNTPEDTRVFVDTIVKCNLQSLAKTAEKLAAGARAAGS-- 229
AK376521              PSTLVVESYVVDVPPGNTPEDTRVFVDTIVKCNLQSLARTAEKLAGRGAAYGAPP 222
AT5G05440             EGTVVVESYIVDVPPGNTEEETLSFVDTIVRCNLQSLARSTNRQ----------- 203
                       .*:*****:******** *:*  ******:*******:::::
```

```
LOC_Os03g18600.1      MPCIPASSPGIPHQHQHQHHRALAGVGMAVGCAAEAAVAAAGVAGTRCGAHDGEVPMEVA 60
MLOC_71349.1          MPCIPASSPSIQHHNHHHRVLAGVGVGVGMGCGAEAVVAAAGTAGMRCGEHDCEVPAEVA 60
AT5G05440             -----MRSP-VQLQHG-----SDATNGFHT---LQPHDQTDGPIKRVCLTRGMHVPEHVA 46
                             ** :  ::      : .  *.      :.   : *     *  :. .** .**


LOC_Os03g18600.1      RHHEHAEPGSGRCCSAVVQHVAAPAAAVWSVVRRFDQPQAYKRFVRSCALLAGDG-GVGT 119
MLOC_71349.1          RHHEHAEPGSGQCCSAVVQHVAAPAAAVWSVVRRFDQPQAYKRFVRSCALVAGDG-GVGT 119
AT5G05440             MHHTHD-VGPDQCCSSVVQMIHAPPESVWALVRRFDNPKVYKNFIRQCRIVQGDGLHVGD 105
                       ** *   *..:***:*** : **. :**::*****:*:.**.*:*.* :: ***  **

LOC_Os03g18600.1      LREVRVVSGLPAASSRERLEILDDESHVLSFRVVGGEHRLKNYLSVTTVHPSPSAPTAAT 179
MLOC_71349.1          LREVHVVSGLPAASSRERLEILDDESHVLSFRVVGGEHRLKNYLSVTTVHPSPAAPSSAT 179
AT5G05440             LREVMVVSGLPAVSSTERLEILDEERHVISFSVVGGDHRLKNYRSVTTLHAS---DDEGT 162
                      **** *******.** *******:* **:** ****:****** ****:*.*      .*

LOC_Os03g18600.1      VVVESYVVDVPPGNTPEDTRVFVDTIVKCNLQSLAKTAEKLAAGARAAGS 229
MLOC_71349.1          VVVESYVVDVPAGNTIDDTRVFIDTIVKCNLQSLAKTAEKLAAVS----- 224
AT5G05440             VVVESYIVDVPPGNTEEETLSFVDTIVRCNLQSLARSTNRQ--------- 203
                      ******:****.*** ::*  *:****:*******:::::


LOC_Os05g12260.1      MVGLVGGGG--WRVGDDAAGGGGGGAVAAGAAAAAEAEHMRRLHSHAPGEHQCSSALVKH 58
MLOC_3912.1           MVGLVGGGARVWRLSHEPASGAGGG------GAATEADYMRRLHGHAPGENQCTSALVKH 54
AT5G53160             ----------------MEANGIENLT-----NPNQEREFIRRHHKHELVDNQCSSTLVKH 39
                                     *.*  .        .  * :.:** * *   ::**:*:****

LOC_Os05g12260.1      IKAPVHLVWSLVRSFDQPQRYKPFVSRCVVRGGDLEIGSVREVNVKTGLPATTSTERLEL 118
MLOC_3912.1           IKAPVHLVWSLVRSFDQPQRYKPFVSRCVVRGGDLEIGSVREVNVKTGLPATTSTERLEQ 114
AT5G53160             INAPVHIVWSLVRRFDQPQKYKPFISRCVVKG-NMEIGTVREVDVKSGLPATRSTERLEL 98
                      *:****:****** *****:****:*****:* ::***:****:**:***** ******

LOC_Os05g12260.1      LDDDEHILSVKFVGGDHRLRNYSSIVTVHPESIDGRPGTLVIESFVVDVPDGNTKDETCY 178
MLOC_3912.1           LDDDEHILSVKFVGGDHRLR----------------------VRAFTSHF---------LL 144
AT5G53160             LDDNEHILSIRIVGGDHRLKNYSSIISLHPETIEGRIGTLVIESFVVDVPEGNTKDETCY 158
                      ***:*****:::*******:                      :.:*. ..

LOC_Os05g12260.1      FVEAVIKCNLTSLAEVSERLAVQSPTSPLEQ 209
MLOC_3912.1           YFLLVPKCSITSLV----------PNAPLQV 165
AT5G53160             FVEALIKCNLKSLADISER-LAVQDTTESRV 188
                      :.  : **.:.**.          .:  .


LOC_Os05g39580.1      MMPYTAPRPSPPQHSRIGGCGGGGVLKAAGAAGHAASCVAVPAEVARHHEHAAGVGQCCS 60
MLOC_60739.1          ----------------------------------------------ARHHEHAAGAGQCCS 15
AT5G05440             MRSPVQLQHGSDATNGFHTLQPHDQTDGPIKRVCLTRGMHVPEHVAMHHTHDVGPDQCCS 60
                                                                  * ** * .* .****

LOC_Os05g39580.1      AVVQAIAAPVDAVWSVVRRFDRPQAYKHFIRSCRLLDGDGDGGAVAVGSVREVRVVSGLP 120
MLOC_60739.1          AVVQAIEAPVGAVWAVVRRFDRPQAYKHFIRSCRLVDGDGG----AVGSVREVRVVSGLP 71
AT5G05440             SVVQMIHAPPESVWALVRRFDNPKVYKNFIRQCRIVQGDGL----HVGDLREVMVVSGLP 116
                      :*** * **  :**::*****.*:.**:***.**:::***      **.:*** ******

LOC_Os05g39580.1      ATSSRERLEILDDERRVLSFRVVGGEHRLSNYRSVTTVHETAAGAAAAVVVESYVVDVPH 180
MLOC_60739.1          ATTSRERLEILDDERRVLSFRVVGGEHRLSNYRSVTTVHETAS-AGGAVVVESYVVDVPP 130
AT5G05440             AVSSTERLEILDEERHVISFSVVGGDHRLKNYRSVTTLHASDD--EGTVVVESYIVDVPP 174
                      *.:* *******:**:*:** ****:***.*******:* :     .:******:****

LOC_Os05g39580.1      GNTADETRMFVDTIVRCNLQSLARTAEQLALAAPRAA 217
MLOC_60739.1          GNTDDETRTFVDTIVRCNLQSLARTAQQLALAA---- 163
AT5G05440             GNTEEETLSFVDTIVRCNLQSLARSTNRQ-------- 203
                      *** :**  ***************::::
```

```
LOC_Os05g39580.1      MMPYTAPRPSPPQHSRIGGCGGGGVLKAAGAAGHAASCVAVPAEVARHHEHAAGVGQCCS 60
MLOC_65591.1          -MPYAAARPSPQQHSRIS----------AACKALVAQGAAVPGEVARHHEHAAGAGQCCS 49
AT5G05440             MRSPVQLQHGSDATNGFHTLQPHDQTDGPIKRVCLTRGMHVPEHVAMHHTHDVGPDQCCS 60
                        . .  : ..   . :           .      :    ** .** ** * .* .****

LOC_Os05g39580.1      AVVQAIAAPVDAVWSVVRRFDRPQAYKHFIRSCRLLDGDGDGGAVAVGSVREVRVVSGLP 120
MLOC_65591.1          AVVQAIAAPVEAVWSVVRRFDRPQAYKRFIKSCRLVDGDGG----AVGSVREVRVVSGLP 105
AT5G05440             SVVQMIHAPPESVWALVRRFDNPKVYKNFIRQCRIVQGDGL----HVGDLREVMVVSGLP 116
                      :*** * ** ::**::*****.*:.**.**:.**:::***      **.:*** ******

LOC_Os05g39580.1      ATSSRERLEILDDERRVLSFRVVGGEHRLSNYRSVTTVHETAAGAAAA----VVVESYVV 176
MLOC_65591.1          GTSSRERLEILDDERRVLSFRIVGGEHRLANYRSVTTVNEVASTVAGAPRVTLVVESYVV 165
AT5G05440             AVSSTERLEILDEERHVISFSVVGGDHRLKNYRSVTTLHASDD--EGT----VVVESYIV 170
                      ..** *******:**:*:** :***:*** *******::       .:    :*****:*

LOC_Os05g39580.1      DVPHGNTADETRMFVDTIVRCNLQSLARTAEQLALAAPRAA 217
MLOC_65591.1          DVPPGNTGDETRMFVDTIVRCNLQSLARTAEQLALAAPRVN 206
AT5G05440             DVPPGNTEEETLSFVDTIVRCNLQSLARSTNRQ-------- 203
                      *** *** :**  ***************::::



LOC_Os06g36670.1      MEAHVERALREGLTEEERAALEPAVMAHHTFPPSTTTATTAAATCTSLVTQRVAAPVRAV 60
MLOC_49654.1          ----------------------------------------MSLVAQRMPASGCAA 15
AT2G26040             ---MSSSPAVKGLTDEEQKTLEPVIKTYHQFEP-------DPTTCTSLITQRIHAPASVV 50
                                                               **::**: *.  ..

LOC_Os06g36670.1      WPIVRSFGNPQRYKHFVRTCALAAGDGASVGSVREVTVVSGLPASTSTERLEMLDDDRHI 120
MLOC_49654.1          WPIVRNFGNPQRYKHFVRTSALVAGDEASIGKAWEDTVVSGLPTSTSSERLEILDDGRHI 75
AT2G26040             WPLIRRFDNPERYKHFVKRCRLISGDGD-VGSVREVTVISGLPASTSTERLEFVDDDHRV 109
                      **::* *.**:**.***: . * :**   :*.. * **:****:***:****::**.:::

LOC_Os06g36670.1      ISFRVVGGQHRLRNYRSVTSVTEFQPPAAGPGPAPPYCVVVESYVVDVPDGNTAEDTRMF 180
MLOC_49654.1          LSFSIVDGEHCLCNYRSVSSVTEIL---LG-----PYRAVVESYVV--PDG--------- 116
AT2G26040             LSFRVVGGEHRLKNYKSVTSVNEFLNQDSGK----VYTVVLESYTVDIPEGNTEEDTKMF 165
                      :** :*.*:* * **:**:**.*:     *      * .*:***.*  *:*

LOC_Os06g36670.1      TDTVVKLNLQMLAAVAEDSSSASRRRD 207
MLOC_49654.1          ---------------------------
AT2G26040             VDTVVKLNLQKLGVAATSAPMHDDE-- 190



LOC_Os06g36670.1      MEAHVERALREGLTEEERAALEPAVMAHHTFPPSTTTATTAAATCTSLVTQRVAAPVRAV 60
MLOC_72289.1          MEAHMERALREGVTEAERAALEGTVRAHHTFP-----GRAPGATCTSLVAQRVAAPVRAV 55
AT2G26040             MSS---SPAVKGLTDEEQKTLEPVIKTYHQFE-------PDPTTCTSLITQRIHAPASVV 50
                      *.:    .  :*:*: *: :** .: ::* *        .  :*****::**: **. .*

LOC_Os06g36670.1      WPIVRSFGNPQRYKHFVRTCALAAGDGASVGSVREVTVVSGLPASTSTERLEMLDDDRHI 120
MLOC_72289.1          WPIVRSFGNPQRYKHFVRTCALAAGDGASVGSVREVTVVSGLPASTSTERLEILDDDRHI 115
AT2G26040             WPLIRRFDNPERYKHFVKRCRLISGDGD-VGSVREVTVISGLPASTSTERLEFVDDDHRV 109
                      **::* *.**:******: * * :***  *********:**************::***:::

LOC_Os06g36670.1      ISFRVVGGQHRLRNYRSVTSVTEFQPPAAGPGPAPPYCVVVESYVVDVPDGNTAEDTRMF 180
MLOC_72289.1          LSFSVVGGEHRLRNYRSVTSVTEFQP-----G---PYCVVVESYVVDVPDGNTEEDTRMF 167
AT2G26040             LSFRVVGGEHRLKNYKSVTSVNEFLNQ----DSGKVYTVVLESYTVDIPEGNTEEDTKMF 165
                      :** ****:***:**:*****.**       .    * **:***.**:*:*** ***:**

LOC_Os06g36670.1      TDTVVKLNLQMLAAVAEDSSSA--SRRRD 207
MLOC_72289.1          TDTVVKLNLQKLASVAEESGAAPGSRRRD 196
AT2G26040             VDTVVKLNLQKLG-VAATSAPMHDDE--- 190
                      .********* *. **  *..   ..



LOC_Os10g42280.1      MEQQ------EEVPPPPAGLGLTAEEYAQVRATVEAHHRYAVGPGQCSSLLAQRIHAPPA 54
AK363238              MEQQPVAAAPAAEPEVPAGLGLTAAEYAQLLPTVEAYHRYAVGPGQCSSLVAQRIEAPPA 60
AT4G17870             ---------------MPS--ELTPEERSELKNSIAEFHTYQLDPGSCSSLHAQRIHAPPE 43
                                     *:   **. * :::  ::  .* * :.**.**** ****.***
```

```
LOC_Os10g42280.1      AVWAVVRRFDCPQVYKHFIRSCVLRPDPHHDDNGNDLRPGRLREVSVISGLPASTSTERL 114
AK363238              AVWAIVRRFDCPQVYKHFIRSCALRPDP---EAGDDLRPGRLREVSVISGLPASTSTERL 117
AT4G17870             LVWSIVRRFDKPQTYKHFIKSCSVEQNF-------EMRVGCTRDVIVISGLPANTSTERL 96
                       **::***** **.*****:** :. :        ::* *  *:* *******.******

LOC_Os10g42280.1      DLLDDAHRVFGFTITGGEHRLRNYRSVTTVSQLD-----EICTLVLESYIVDVPDGNTED 169
AK363238              DLLDDARRAFGFTITGGEHRLRNYRSVTTVSELSPAAPAEICTVVLESYVVDVPDGNSEE 177
AT4G17870             DILDDERRVTGFSIIGGEHRLTNYKSVTTVHRFEKE--NRIWTVVLESYVVDMPEGNSED 154
                      *:*** :*. **:* ****** **:***** .:.     .* *:*****:**:*:**:*:

LOC_Os10g42280.1      DTRLFADTVIRLNLQKLKSVSEANANAAAAAAAPPPPPPAAAE 212
AK363238              DTRLFADTVVRLNLQKLKSVAET--NAAAAAAAPP------AE 212
AT4G17870             DTRMFADTVVKLNLQKLATVAEAMARNSGDGSGSQ------VT 191
                      ***:*****::****** :*:*:  . :. .:..       .
```

Asterisk indicate the conservation of residue in all the three species and dashes indicate the gaps or absence of sequence data; alignments were made using CLUSTALW2 (http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=clustalw2; last accessed May 2014) and then bioedit was used to align the conserved functional motifs; bold indicate residues that are directly involved in ABA binding and grey shading indicate residues that form hydrogen bond with ABA through water; rice sequences were derived from the MSU rice genomic annotation project (http://rice.plantbiology.msu.edu/; last accessed May 2014) and Arabidopsis sequences were derived from TAIR (http://www.arabidopsis.org/index.jsp); positioning of amino acid residues were derived from Yin *et al*. (2009).

Appendix XIV Alignment of putative amino acid sequences of barley SnRK2s with rice and Arabidopsis

```
LOC_Os03g41460.1      MDRAALTVGPGMDMPIMHDGDRYELVRDIGSGNFGVARLMRSRADGQLVAVKYIERGDKI 60
MLOC_22145.2          MDRAALTVGPGMDMPIMHDGDRYELVKDIGSGNFGVARLMRNRADGQLVAVKYIERGEKI 60
AT5G66880             MDRAPVTTGP-LDMPIMHDSDRYDFVKDIGSGNFGVARLMRDKLTKELVAVKYIERGDKI 59
                      ****.:*.** :*******.***::*:**************.:   :**********:**

LOC_Os03g41460.1      DENVQREIINHRSLRHPNIIRFKEVILTPTHLAIVMEYASGGELFERICNAGRFSEDEAR 120
MLOC_22145.2          DENVQREIINHRSLRHPNIIRFKEVILTPTHLAIVMEYASGGELFERICNAGRFSEDEAR 120
AT5G66880             DENVQREIINHRSLRHPNIVRFKEVILTPTHLAIIMEYASGGELYERICNAGRFSEDEAR 119
                      ******************:**************:*********:***************

LOC_Os03g41460.1      FFFQQLISGVSYCHSMQVCHRDLKLENTLLDGSTAPRLKICDFGYSKSSVLHSQPKSTVG 180
MLOC_22145.2          FFFQQLLSGVSYCHSMQVCHRDLKLENTLLDGSTAPRLKICDFGYSKSSVLHSQPKSTVG 180
AT5G66880             FFFQQLLSGVSYCHSMQICHRDLKLENTLLDGSPAPRLKICDFGYSKSSVLHSQPKSTVG 179
                      ******:**********:***************.*************************

LOC_Os03g41460.1      TPAYIAPEVLLKKEYDGKIADVWSCGVTLYVMLVGAYPFEDPDEPKNFRKTIQRILGVQY 240
MLOC_22145.2          TPAYIAPEVLLKKEYDGKIADVWSCGVTLYVMLVGAYPFEDPDEPKNFRKTIQRILSVQY 240
AT5G66880             TPAYIAPEVLLRQEYDGKIADVWSCGVTLYVMLVGAYPFEDPEEPRDYRKTIQRILSVKY 239
                      **********::*****************************:**:::********.*:*

LOC_Os03g41460.1      SIPDYVHISPECRDLIARIFVANPATRISIPEIRNHPWFLKNLPADLMDDSKMSSQYEEP 300
MLOC_22145.2          SIPDYVHISSECRDLIAKIFVGNPATRITIPEIRNHPWFLKNLPADLVDDSTMSSQYEEP 300
AT5G66880             SIPDDIRISPECCHLISRIFVADPATRISIPEIKTHSWFLKNLPADLMNESNTGSQFQEP 299
                      **** ::**.** .**::***.:*****:****:.*.**********:::*. .**::**

LOC_Os03g41460.1      EQPMQSMDEIMQILAEATIPAAGSGGINQFLNDGLDLDDDMEDLDSDPDLDVESSGEIVY 360
MLOC_22145.2          EQPMQSMDEIMQILAEATIPAAGSR-INQFLNDGLDLDDDMDDLDSDADLDVESSGEIVY 359
AT5G66880             EQPMQSLDTIMQILSEATIPAVRNRCLDDFMTDNLDLDDDMDDEDSESEIDIDSSGEIVY 359
                      ******:* ****::******. .  :::*:.*.*******:*:**:.::*::*******

LOC_Os03g41460.1      AM 362
MLOC_22145.2          AM 361
AT5G66880             AI 361
                      *:

LOC_Os03g55600.1      MAAAGAGAGAPDRAALTVGPGMDMPIMHDSDRYELVRDIGSGNFGVARLMRDRRTMELVA 60
AK374298              MAGA-----APDRAALTVGPGMDMPIMHDSDRYELVRDIGSGNFGVARLMRDRRTMELVA 55
AT3G50500             ----------MDPATNSPIMPIDLPIMHDSDRYDFVKDIGSGNFGVARLMTDRVTKELVA 50
                                * *: :    :*:**********::*:************* ** * ****
```

```
LOC_Os03g55600.1      VKYIERGEKIDDNVQREIINHRSLKHPNIIRFKEVILTPTHLAIVMEYASGGELFERICK 120
AK374298              VKYIERGEKIDENVQREIINHRSLKHPNIIRFKEVILTPTHLAIVMEYASGGELFERICK 115
AT3G50500             VKYIERGEKIDENVQREIINHRSLRHPNIVRFKEVILTPSHLAIVMEYAAGGELYERICN 110
                      ***********:************:****:*********:*********:****:****:

LOC_Os03g55600.1      NVRFSEDEARYFFQQLISGVSYCHSMQVCHRDLKLENTLLDGSPAPRLKICDFGYSKSSV 180
AK374298              NIRFSEDEARYFFQQLISGVSYCHSMQVCHRDLKLENTLLDGSPAPRLKICDFGYSKSSV 175
AT3G50500             AGRFSEDEARFFFQQLISGVSYCHAMQICHRDLKLENTLLDGSPAPRLKICDFGYSKSSV 170
                        ********:************:**:*******************************

LOC_Os03g55600.1      LHSQPKSTVGTPAYIAPEVLLKKEYDGKTADVWSCGVTLYVMVVGAYPFEDPEEPKNFRK 240
AK374298              LHSQPKSTVGTPAYIAPEVLLKKEYDGKIADVWSCGVTLYVMVVGAYPFEDPEEPKNFRK 235
AT3G50500             LHSQPKSTVGTPAYIAPEILLRQEYDGKLADVWSCGVTLYVMLVGAYPFEDPQEPRDYRK 230
                      ******************:**::***** *************:*********:**:::**

LOC_Os03g55600.1      TIQRILNVQYSIPENVDISPECRHLISRIFVGDPSLRITIPEIRSHGWFLKNLPADLMDD 300
AK374298              TIQRILSVQYSIPDNVDISPECRHLISRIFVGDPALRITIPEIRSHNWFLKNLPADLMDD 295
AT3G50500             TIQRILSVTYSIPEDLHLSPECRHLISRIFVADPATRITIPEITSDKWFLKNLPGDLMDE 290
                      ******.* ****:::.:*************.**: ******* *. *******.****:

LOC_Os03g55600.1      DSMSSQYEEPDQPMQTMDQIMQILTEATIEPACSR-INHILTDGLDLDDDMDDLDSDSDI 359
AK374298              DSMSSQYEEPEQPMQTMDQIMQILTEATIEPACSR-INHILTDGFDMDDDMDDLESDSDL 354
AT3G50500             NRMGSQFQEPEQPMQSLDTIMQIISEATIEPTVRNRCLDDFMADNLDLDDDMDDFDSESEI 350
                      : *.**::**:****::* ****::*****.. .* ::.:::*.:*:*******::*:*::

LOC_Os03g55600.1      DVDSSGEIVYAM 371
AK374298              DIDSSGEIVYAM 366
AT3G50500             DVDSSGEIVYAL 362
                      *:*********:


LOC_Os12g39630.1      MERAAAGPLGMEMPIMHDGDRYELVKEIGSGNFGVARLMRNRASGDLVAVKYIDRGEKID 60
MLOC_3013.1           MERGTMG----DVPVMLDGDRYELVRSIGSGNFGVARLMRNRASGELVAVKYIDRGEKID 56
AT4G33950             MDRPAVSG-PMDLPIMHDSDRYELVKDIGSGNFGVARLMRDKQSNELVAVKYIERGEKID 59
                      *:* : .    ::*:* *.******:.************** *.:*******:******

LOC_Os12g39630.1      ENVQREIINHRSLRHPNIIRFKEVILTPTHLAIVMEYASGGELFERICSAGRFSEDEARF 120
MLOC_3013.1           ENVQREIINHRSLRHPNIIRFKEVILTPTHLAIVMEYASGGELFDRICTAGRFSVDEARF 116
AT4G33950             ENVKREIINHRSLRHPNIVRFKEVILTPTHLAIVMEYASGGELFERICNAGRFSEDEARF 119
                      ***:**************:*************************:***.***** *****

LOC_Os12g39630.1      FFQQLISGVSYCHSMQVCHRDLKLENTLLDGSTAPRLKICDFGYSKSSVLHSQPKSTVGT 180
MLOC_3013.1           FFQQLISGVSYCHSMQVCHRDLKLENTLLDGSTTPRLKICDFGYSKSSVLHSQPKSTVGT 176
AT4G33950             FFQQLISGVSYCHAMQVCHRDLKLENTLLDGSPAPRLKICDFGYSKSSVLHSQPKSTVGT 179
                      *************:******************.:*************************

LOC_Os12g39630.1      PAYIAPEVLLKKEYDGKIADVWSCGVTLYVMLVGAYPFEDPEDPKNFRKTIQKILGVQYS 240
MLOC_3013.1           PAYIAPEVLIKKEYDGKIADVWSCGVTLYVMLVGAYPFEDPENPKNFKMTIQKILGVQYS 236
AT4G33950             PAYIAPEVLLKKEYDGKVADVWSCGVTLYVMLVGAYPFEDPEEPKNFRKTIHRILNVQYA 239
                      *********:*******:************************:****: **::**.***:

LOC_Os12g39630.1      IPDYVHISPECRDLITRIFVGNPASRITMPEIKNHPWFMKNIPADLMDDGMVSNQYEEPD 300
MLOC_3013.1           IPDYIHIPMDCRNLLSRIFVANPATRITIPEIKNHPWFLKNLPADLMDGPTVSNQYEEPD 296
AT4G33950             IPDYVHISPECRHLISRIFVADPAKRISIPEIRNHEWFLKNLPADLMNDNTMTTQFDESD 299
                      ****:**. :**.*::****.:**.**::***:** **:**:*****:.  ::.*::*.*

LOC_Os12g39630.1      QPMQNMNEIMQILAEATIPAAGTSGINQFLTDSLDLDDDMED-MDSDLD-LDIESSGEIV 358
MLOC_3013.1           QPMQNMNEIMQIMAEATIPAASALGINKFLPDGLDLDDDMDD-LDSDLD-IDMDSSGEIV 354
AT4G33950             QPGQSIEEIMQIIAEATVPPAGTQNLNHYLTGSLDIDDDMEEDLESDLDDLDIDSSGEIV 359
                      ** *.::******:****:*.*.:: .:*::*...**:****:: ::**** :*::******

LOC_Os12g39630.1      YAM 361
MLOC_3013.1           YAM 357
AT4G33950             YAM 362
                      ***


LOC_Os03g27280.1      MERYEVMRDIGSGNFGVAKLVRDVATNHLFAVKFIERGLKIDEHVQREIMNHRSLKHPNI 60
AK372880              MDRYEVVRDIGSGNFGVAKLVRDVRTREHFAVKFIERGHKIDEHVQREIMNHRSLKHPNI 60
AT4G40010             MERYDILRDLGSGNFGVAKLVREKANGEFYAVKYIERGLKIDEHVQREIINHRDLKHPNI 60
                      *:**:::**:************:  . . :***:**** **********:***.******
```

```
LOC_Os03g27280.1      IRFKEVVLTPTHLAIVMEYAAGGELFERICNAGRFSEDEARFFFQQLISGVSYCHSMQVC 120
AK372880              IRFKEVVLTPTHLAIVMEYASGGELFQRICNAGRFSEDEGRFFFQQLISGVSYCHSMQVC 120
AT4G40010             IRFKEVFVTPTHLAIVMEYAAGGELFERICNAGRFSEDEGRYYFKQLISGVSYCHAMQIC 120
                      ******.:************:*****:************.*:::*:**********:**:*

LOC_Os03g27280.1      HRDLKLENTLLDGSVTPRLKICDFGYSKSSVLHSQPKSTVGTPAYIAPEVLSRKEYDGKV 180
AK372880              HRDLKLENTLLDGSVAPRLKICDFGYSKSSVLHSQPKSTVGTPAYIAPEVLSRREYDGKV 180
AT4G40010             HRDLKLENTLLDGSPSSHLKICDFGYSKSSVLHSQPKSTVGTPAYVAPEVLSRKEYNGKI 180
                      ************** :.:*****************************:*******:**:**:

LOC_Os03g27280.1      ADVWSCGVTLYVMLVGAYPFEDPDDPRNFRKTITRILSVQYSIPDYVRVSADCRHLLSRI 240
AK372880              ADVWSCGVTLYVMLVGAYPFEDPDEPRNFRKTIARILSVQYSVPDYVRVSMDCIHLLSRI 240
AT4G40010             ADVWSCGVTLYVMLVGAYPFEDPEDPRNIRNTIQRILSVHYTIPDYVRISSECKHLLSRI 240
                      ***********************::***:*:** *****:*::*****:* :* ******

LOC_Os03g27280.1      FVGNPEQRITIPEIKNHPWFLKNLPIEMTDEYQRSMQLADMNTPS----QSLEEVMAIIQ 296
AK372880              FVGNPQQRITIPEIKNHPWFLKRLPVEMTDEYQRGMQLADMNTPS----QSLEEAMAIIQ 296
AT4G40010             FVADPDKRITVPEIEKHPWFLKGPLVVPPEEEKCDNGVEEEEEEEEKCKQSVEEIVKIIE 300
                      **.:*::***:***::******   :  .:* : .  : : :  .    **:** : **:

LOC_Os03g27280.1      EARKPGDAMKLAGAGQVACLGSMDLDDID-----DIDDIDIENSGDFVCAL 342
AK372880              EAQKPG-HNALGLAGQVACLGSMDLDDIDF----DVDDIDIENSGDFVCPL 342
AT4G40010             EARKGVNGTDNN-GGLGLIDGSIDLDDIDDADIYDDVDDDEERNGDFVCAL 350
                      **:*         .*     **:******     *  * * *..*****.*


LOC_Os07g42940.1      MERYEVIKDIGSGNFGVAKLVRDVRTKELFAVKFIERGQKIDENVQREIMNHRSLRHPNI 60
MLOC_69212.1          MERYEVIKDIGSGNFGVAKLVRDVRTKELFAVKFIERGHKIDENVQREIMNHRSLRHPNI 60
AT1G78290             MERYEIVKDIGSGNFGVAKLVRDKFSKELFAVKFIERGQKIDEHVQREIMNHRSLIHPNI 60
                      *****::****************  :************:****:*********** ****

LOC_Os07g42940.1      VRFKEVVLTPTHLAIVMEYAAGGELFERICSAGRFSEDEARFFFQQLISGVSYCHSMQIC 120
MLOC_69212.1          VRFKEVVLTPTHLAIVMEYAAGGELFERICGSGKFSENEARFFFQQLLSGVSYCHSMQIC 120
AT1G78290             IRFKEVLLTATHLALVMEYAAGGELFGRICSAGRFSEDEARFFFQQLISGVNYCHSLQIC 120
                      :*****:**.****:********** ***.:*:***:*********:***.****:***


LOC_Os07g42940.1      HRDLKLENTLLDGSIAPRLKICDFGYSKSSLLHSQPKSTVGTPAYIAPEVLARKEYDGKV 180
MLOC_69212.1          HRDLKLENTLLDGSEAPRLKICDFGYSKSSVLHSQPKSTVGTPAYIAPEVLSRREYDGKV 180
AT1G78290             HRDLKLENTLLDGSEAPRVKICDFGYSKSGVLHSQPKTTVGTPAYIAPEVLSTKEYDGKI 180
                      ************** ***:**********.:******:*************: :*****:

LOC_Os07g42940.1      ADVWSCGVTLYVMLVGAYPFEDPDEPRNFRKTITRILSVQYMVPDYVRVSMECRHLLSRI 240
MLOC_69212.1          ADVWSCGVTLYVMLVGAYPFEDPDEPKNFRKTITRILSVQYSVPDYVRISMECRHLLSRI 240
AT1G78290             ADVWSCGVTLYVMLVGAYPFEDPSDPKDFRKTIGRILKAQYAIPDYVRVSDECRHLLSRI 240
                      **********************.:*::***** ***..** :*****:* *********

LOC_Os07g42940.1      FVANPEQRITIPEIKNHPWFLKNLPIEMTDEYQMSVQMNDINTPSQGLEEIMAIIQEARK 300
MLOC_69212.1          FVANPEQRITIQEIKNHPWFLKNLPIEMTDEYQMSLHMVGVNAPPQSLEEIMAIIQEARI 300
AT1G78290             FVANPEKRITIEEIKNHSWFLKNLPVEM---YEGSLMMNGPST--QTVEEIVWIIEEARK 295
                      ******:**** *****.*******:**   *: *: * . .:  * :***: **:***

LOC_Os07g42940.1      P--------GDGSKFSGQI--PGLGSMELDDVDTDDIDVEDSGDFVCAL 339
MLOC_69212.1          P--------GDGSKFAGQLSVPGLGSMELDDIDDVDADVEDSGDFVCAL 341
AT1G78290             PITVATGLAGAGGSGGSSNGAIGSSSMDLDDLDT-DFDDIDTADLLSPL 343
                      *        * *.. ...    * .**:***:*  * *  *:.*::..*
```

Asterisk indicate the conservation of residue in all the three species and dashes indicate the gaps or absence of sequence data; alignments were made using CLUSTALW2 (http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=clustalw2; last accessed May 2014) and then bioedit was used to align the conserved functional motifs; box: conserved functional motifs; bold: indicate residues important in binding to PP2C; rice sequences were derived from the MSU rice genomic annotation project (http://rice.plantbiology.msu.edu/; last accessed May 2014) and Arabidopsis sequences were derived from TAIR (http://www.arabidopsis.org/index.jsp); motif sequences and the positions were derived from Belin *et al*. (2006).

Appendix XV Alignment of putative amino acid sequences of barley ABC transporters with rice and Arabidopsis

```
MLOC_11283.1          ------------------------------------------------------------
LOC_Os08g43120.1      ------------------------------------------------------MAAA 4
ABCG40                MEGTSFHQASNSMRRNSSVWKKDSGREIFSRSSREEDDEEALRWAALEKLPTFDRLRKGI 60

MLOC_11283.1          ------------------------------------------------------------
LOC_Os08g43120.1      VELTGDGGTTAETRWLSP-------------PLTHDDNRGFLQMLREKKERLGVGAAKVE 51
ABCG40                LTASHAGGPINEIDIQKLGFQDTKKLLERLIKVGDDEHEKLLWKLKKRIDRVGIDLPTIE 120

MLOC_11283.1          ------------------------------------------------------------
LOC_Os08g43120.1      VRLEKLTVEADVRVGRRAVPTLLNCAINAAQELAACAHMCTTRKKPMKIINEATGTIRPS 111
ABCG40                VRFDHLKVEAEVHVGGRALPTFVNFISNFADKFLNTLHLVPNRKKKFTILNDVSGIVKPG 180

MLOC_11283.1          ------------------------------------------------------------
LOC_Os08g43120.1      RMTLLLGAPGSGKTTLLKALAGKLDSSLKMKGKVTYNGEEVNSSTPQYLHAYVSQYDLHH 171
ABCG40                RMALLLGPPSSGKTTLLLALAGKLDQELKQTGRVTYNGHGMNEFVPQRTAAYIGQNDVHI 240

MLOC_11283.1          ------------------------------------------------------------
LOC_Os08g43120.1      AEMTLSS-----------------LKKLICTGHNLWRRKQPYNKLYYQAIKIE------- 207
ABCG40                GEMTVRETFAYAARFQGVGSRYDMLTELARREKEANIKPDPDIDIFMKAMSTAGEKTNVM 300

MLOC_11283.1          ------------------------------------------------------------
LOC_Os08g43120.1      ---CMQILGLSECADTLVGDEMRRGISGGQKKRATIGEMLVGLARCFFMDDISTGLDSST 264
ABCG40                TDYILKILGLEVCADTMVGDDMLRGISGGQKKRVTTGEMLVGPSRALFMDEISTGLDSST 360

MLOC_11283.1          ------------------------------------------------------------
LOC_Os08g43120.1      TFEIMKFLQQMAHLMDLTMVISLLQPPPETLELFDDIILLCEGQIVYHGPRENATDFFET 324
ABCG40                TYQIVNSLRNYVHIFNGTALISLLQPAPETFNLFDDIILIAEGEIIYEGPRDHVVEFFET 420

MLOC_11283.1          ------------------------------------------------------------
LOC_Os08g43120.1      MGFKCPSRKNVADFLQEVTSKMDQKQYWIGNANKYQYHSIEKFAESFRTSYLPRLVENDH 384
ABCG40                MGFKCPPRKGVADFLQEVTSKKDQMQYWARRDEPYRFIRVREFAEAFQSFHVGRRIGDEL 480

MLOC_11283.1          ------------------------------------------------------------
LOC_Os08g43120.1      FESTNAGKSKEVKTSTSRM-ISSWNIFKACFSREVLLLKRNSPVHIFKTIQITVLALVIS 443
ABCG40                ALPFDKTKSHPAALTTKKYGVGIKELVKTSFSREYLLMKRNSFVYYFKFGQLLVMAFLTM 540

MLOC_11283.1          ------------------------------------------------------------
LOC_Os08g43120.1      TLFLRTNMRHDTVLDANKYMGALFMAVVIVNFNGMTEIAMTIKRLPIFYKQREILALPGW 503
ABCG40                TLFFRTEMQKKTEVDGSLYTGALFFILMMLMFNGMSELSMTIAKLPVFYKQRDLLFYPAW 600

MLOC_11283.1          ------------------------------------------------------------
LOC_Os08g43120.1      ALLSSVFLLSLPISFVETGLWTGLTYYVIGYAPSFVRFIQHFVVLFAMHQMSMSLYRFLA 563
ABCG40                VYSLPPWLLKIPISFMEAALTTFITYYVIGFDPNVGRLFKQYILLVLMNQMASALFKMVA 660

MLOC_11283.1          ------------------------------------------------------------
LOC_Os08g43120.1      AIGRTQVMANMLGTAALIAIYILGGFVISKDNLQPWLRWGYWTSPFTYAQNAVALNEFLD 623
ABCG40                ALGRNMIVANTFGAFAMLVFFALGGVVLSRDDIKKWWIWGYWISPIMYGQNAILANEFFG 720

MLOC_11283.1          ------------------------------------------------------MNPP 4
LOC_Os08g43120.1      DRWATEFHFANANTVGETILKVRGLLTEWHWYWICVSILFGFSLVFNILSIFALQYMRSP 683
ABCG40                HSWSRAVENS-SETLGVTFLKSRGFLPHAYWYWIGTGALLGFVVLFNFGFTLALTFLNSL 779
                                                                      :..

MLOC_11283.1          HKHQVNK------RDMQVITEFTNKIVGIGSASIPPGVLPFRPLSFVFDHINYFVEMPKE 58
LOC_Os08g43120.1      HKHQVNI------NATKVKVDYNSQIVGNGTASTDQVILPFQPLSLVFDHINYFVDMPKE 737
ABCG40                GKPQAVIAEEPASDETELQSARSEGVVEAGANKKRGMVLPFEPHSITFDNVVYSVDMPQE 839
                       * *.          ::    .. :*  *: .    :***.* *:.**:: * *:**:*
                                                 Walker Box A
MLOC_11283.1          MMKYGVTDKRLQLLQDVSGAFRPGVLTALMGITGAGKTTLLDVLAGRKTEGYIEGTITIE 118
LOC_Os08g43120.1      MTKYGVTDKKLQLLQDVSGAFRPGVLTALMGITGAGKTTLLDVLAGRKTGGYIEGTVKIA 797
ABCG40                MIEQGTQEDRLVLLKGVNGAFRPGVLTALMGVSGAGKTTLMDVLAGRKTGGYIDGNITIS 899
```

```
                    * : *. :.:* **:.*.*************::*******:******** ***:*.:.*

MLOC_11283.1        GYAKKQETFSRISGYCEQSDIHSPNLTVYESLQFSAWLRLPSEVKSCQRDMFINEVMDLV 178
LOC_Os08g43120.1    GYPKKQETFSRISGYCEQSDIHSPNLTVYESLQFSAWLRLPSNVKSHQRNMFIDEVMDLV 857
ABCG40              GYPKNQQTFARISGYCEQTDIHSPHVTVYESLVYSAWLRLPKEVDKNKRKIFIEEVMELV 959
                    **.*:*:**:********:*****::****** :*******.:*.. :*.:**:***:**
                              Walker Box B
MLOC_11283.1        ELTGLKNAMVGLAGATGLSAEQRKRLTMAVELVASPSIIFMDEPTTGLDARAAAIVMRTI 238
LOC_Os08g43120.1    ELTGLKNAMVGLAGATGLSAEQRKRLTIAVELVASPSIIFMDEPTTGLDARAAAIVMRTV 917
ABCG40              ELTPLRQALVGLPGFSGLSTEQRKRLTIAVELVANPSIIFMDEPTSGLDARAAAIVMRTV 1019
                    *** *:::*:***.* :***:*******:******.**********:**************

MLOC_11283.1        RKTVDTGRTIVCTIHQPSIGIFEAFDELLLMKRGGQTIYSGSLGRLSSNMSNYFEDIPGV 298
LOC_Os08g43120.1    RKTVDTGRTVVCTIHQPSIEIFESFDELLLMKRGGQLIYSGSLGPLSSNMIKYFEAIPGV 977
ABCG40              RNTVDTGRTVVCTIHQPSIDIFEAFDELFLLKRGGEEIYVGPLGHESTHLINYFESIQGI 1079
                    *:*******:********* ***:****:*:****: ** *.**  *::: :*** * *:

MLOC_11283.1        PRIKDGQNPASWLLDVTSHTTEYKIGVDYAEIYRNSSLYRENMLLIEELGKPAPNIEDLH 358
LOC_Os08g43120.1    PRIKEGQNPAAWMLDISSRTAEYEIGVDYAEIYQRSSLYWENRQLIDDLGKPEPNTEDLH 1037
ABCG40              NKITEGYNPATWMLEVSTTSQEAALGVDFAQVYKNSELYKRNKELIKELSQPAPGSKDLY 1139
                     :*.:* ***:*:*:::: : *  :***:*::*:.*.** .*  **.:*.:* *. :**:

MLOC_11283.1        FPPKYWVNFRAQCIACLWKQRCAYWKNPEHNVVRFLNTFVLSIMFGIVFWQTGSTIKRQQ 418
LOC_Os08g43120.1    FPPKYWQDFRAQCMACLWKQNCAYWKNSEHNVVRFINTFAVSIMFGIVFWKIGSTIKDEQ 1097
ABCG40              FPTQYSQSFLTQCMASLWKQHWSYWRNPPYTAVRFLFTIGIALMFGTMFWDLGGKTKTRQ 1199
                    **.:*  .* :**:*.****. :**:*. :..***: *: :::*** :**. *.. * .*

MLOC_11283.1        DIFNILGLIYGSSLFLGFNNCCILQPVVATERVVLYREKAAGTYSTSAYAIAQVAIELPY 478
LOC_Os08g43120.1    DVFNILGVVYGSALFLGFMNCSILQPVVGMERVVLYREKAAGMYSTMAYAIAQVAVELPY 1157
ABCG40              DLSNAMGSMYTAVLFLGLQNAASVQPVVNVERTVFYREQAAGMYSAMPYAFAQVFIEIPY 1259
                    *: * :* :* : ****: *.. :****  **.*:***:*** **: .**:*** :*:**

MLOC_11283.1        MLVQSFIFATIVYTMIGFQMAANKFFWFLLYMVLSYMYYTLFGMMTVALTPNIEIASGLS 538
LOC_Os08g43120.1    MFVQVFIFSAIVYPMIGFQMTATKFFWFALYMVLSFLYYTLYGMMTVALTPNIEIAAGLS 1217
ABCG40              VLVQAIVYGLIVYAMIGFEWTAVKFFWYLFFMYGSFLTFTFYGMMAVAMTPNHHIASVVS 1319
                    ::** :::. ***.****: :* ****: ::*  *:: :*::***:**:*** .**: :*

MLOC_11283.1        FLIFIFWNIFSGFIIGREMIPVWWRWVYWANPAAWTVYGLMFSQLGDRTEPILVPGQPNQ 598
LOC_Os08g43120.1    FLIFIFWNVFSGFIIGRQMIPVWWRWVYWANPAAWTVYGLMFSQLGDRTELIQVPGQPEQ 1277
ABCG40              SAFYGIWNLFSGFLIPRPSMPVWWEWYYWLCPVAWTLYGLIASQFGDITEPMADS---NM 1376
                      :: :**:****:* *  :****.* **  *.***:***: **:** ** :  .   :

MLOC_11283.1        TVREFLEGYLGLEDHYFNLVTYLHLIIIMLFAVIFFISLKYLNFQRR 645
LOC_Os08g43120.1    TVKEFLEGYLGLQDRYFNLVTSLHVAIIALFTFLFFLSIKHLKFQRR 1324
ABCG40              SVKQFIREFYGYREGFLGVVAAMNVIFPLLFAVIFAIGIKSFNFQKR 1423
                    :*::*:. : * .: ::.:*: ::: :  **:.:* :.:* ::**:*



LOC_Os08g43120.1    ------------------------------------------------------MAAA 4
MLOC_54794.6        ------------------------------------------------------------
ABCG40              MEGTSFHQASNSMRRNSSVWKKDSGREIFSRSSREEDDEEALRWAALEKLPTFDRLRKGI 60


LOC_Os08g43120.1    VELTGDGGTTAETRWLSP-------------PLTHDDNRGFLQMLREKKERLGVGAAKVE 51
MLOC_54794.6        ------------------------------------------------------------
ABCG40              LTASHAGGPINEIDIQKLGFQDTKKLLERLIKVGDDEHEKLLWKLKKRIDRVGIDLPTIE 120


LOC_Os08g43120.1    VRLEKLTVEADVRVGRRAVPTLLNCAINAAQELAACAHMCTTRKKPMKIINEATGTIRPS 111
MLOC_54794.6        -----------------------------------MCTTRKKPIKIINGASGTIRPS 22
ABCG40              VRFDHLKVEAEVHVGGRALPTFVNFISNFADKFLNTLHLVPNRKKKFTILNDVSGIVKPG 180
                                                      : ..*** :.*:* .:* ::*.
                              Walker Box A
LOC_Os08g43120.1    RMTLLLGAPGSGKTTLLKALAGKLDSSLKMKGKVTYNGEEVNSSTPQYLHAYVSQYDLHH 171
MLOC_54794.6        RMTLLLGAPGSGKTTFLKALAGKLDSSLKLKGKVMYNGEEVKSSTPQYLHAYISQYDLHH 82
ABCG40              RMALLLGPPSSGKTTLLLALAGKLDQELKQTGRVTYNGHGMNEFVPQRTAAYIGQNDVHI 240
                    **:****.*.*****:* *******..** .*:* ***. ::. .**   **:.* *:*

LOC_Os08g43120.1    AEMTLSS----LKKLICT-------GHNLWRRKQPYNKLYY----------------QAI 204
MLOC_54794.6        AEMTVRETIDFSSKMLGTNNEFEMLGEAIRRKKGVINKVDQDLDSFIKAT-TFGEGSNLT 141
ABCG40              GEMTVRETFAYAARFQGVGSRYDMLTELARREKEANIKPDPDIDIFMKAMSTAGEKTNVM 300
                    .***: .      ::  .        .   *.*    *                    :
```

```
                                  Signature motif Walker box B
LOC_Os08g43120.1      KIECMQILGLSECADTLVGDEMRRGISGGQKKRATIGEMLVGLARCFFMDDISTGLDSST 264
MLOC_54794.6          TNYIIKILGLSECADTMVGDELRRGISGGQKKRATIGEMLVGLARCFFMDDISTGLDSST 201
ABCG40                TDYILKILGLEVCADTMVGDDMLRGISGGQKKRVTTGEMLVGPSRALFMDEISTGLDSST 360
                      .   ::****. ****:***:: **********.* ****** :*.:***:*********

LOC_Os08g43120.1      TFEIMKFLQQMAHLMDLTMVISLLQPPPETLELFDDIILLCEGQIVYHGPRENATDFFET 324
MLOC_54794.6          TYEIVKFLQQMAHLMDLTMVISLLQPPPETLELFDDIILLCEGQIIYHGPRENATNFFEI 261
ABCG40                TYQIVNSLRNYVHIFNGTALISLLQPAPETFNLFDDIILIAEGEIIYEGPRDHVVEFFET 420
                      *::*:: *:: .*::: * :******.***::*******:.**:*:*.***::..:***

LOC_Os08g43120.1      MGFKCPSRKNVADFLQEVTSKMDQKQYWIGNANKYQYHSIEKFAESFRTSYLPRLVENDH 384
MLOC_54794.6          MGFKCPSRKNVADFLQEVTSKMDQKQYWIGDENKYQYRPIEKFAESFRSSYLPRFAKDDL 321
ABCG40                MGFKCPPRKGVADFLQEVTSKKDQMQYWARRDEPYRFIRVREFAEAFQSFHVGRRIGDEL 480
                      ******.**.*********** ** ***    : *::  :.:***:*:: :: *   ::

LOC_Os08g43120.1      FESTNAGKSK-EVKTSTSRMISSWNIFKACFSREVLLLKRNSPVHIFKTIQITVLALVIS 443
MLOC_54794.6          CRTNNTGKSK-EIITSATRRISRWNIFKACFSREVLLLKRNSPLHIFKTVQITVMALVIS 380
ABCG40                ALPFDKTKSHPAALTTKKYGVGIKELVKTSFSREYLLMKRNSFVYYFKFGQLLVMAFLTM 540
                        . :  **:    *: .  :.  ::.*:.**** **:**** :: **  *: *:*::

LOC_Os08g43120.1      TLFLRTNMRHDTVLDANKYMGALFMAVVIVNFNGMTEIAMTIKRLPIFYKQREILALPGW 503
MLOC_54794.6          TIFLRTNMNHKTVLDANKYMGSLFMAVVIVNFNGMTEIAMTIKRLPTFYKQRELLALPGW 440
ABCG40                TLFFRTEMQKKTEVDGSLYTGALFFILMMLMFNGMSELSMTIAKLPVFYKQRDLLFYPAW 600
                      *:*:**:*.:.* :*.. * *:**: :::: ****:*::*** :** *****::*  *.*

LOC_Os08g43120.1      ALLSSVFLLSLPISFVETGLWTGLTYYVIGYAPSFVRFIQHFVVLFAMHQMSMSLYRFLA 563
MLOC_54794.6          ALLSSVFLISLPMSLLETGLWTSLTYYVIGYAPSFLRFIQQFLVLFAMHQMSMGLYRFLA 500
ABCG40                VYSLPPWLLKIPISFMEAALTTFITYYVIGFDPNVGRLFKQYILLVLMNQMASALFKMVA 660
                      .   . :*:.:*:*::*:.* * :******: *.. *::::::::*. *:**: .*::::*

LOC_Os08g43120.1      AIGRTQVMANMLGTAALIAIYILGGFVISKDNLQPWLRWGYWTSPFTYAQNAVALNEFLD 623
MLOC_54794.6          AIGRTQVMANMLGTAALIAIYIFGGFVISKDNLQPWLQWGYWTSPFTYAQNAVALNEFLD 560
ABCG40                ALGRNMIVANTFGAFAMLVFFALGGVVLSRDDIKKWWIWGYWISPIMYGQNAILANEFFG 720
                      *:**. ::** :*: *::.:: :**.*:*:*::: *  **** **: *.***:  ***:.

LOC_Os08g43120.1      DRWATEFHFANANTVGETILKVRGLLTEWHWYWICVSILFGFSLVFNILSIFALQYMRSP 683
MLOC_54794.6          ERWAIEFHYANAKTVGEAILKIRGLLMEWHWYWICVGILFGFSLVFNILTIFALEFMKSP 620
ABCG40                HSWSRAVENS-SETLGVTFLKSRGFLPHAYWYWIGTGALLGFVVLFNFGFTLALTFLNSL 779
                      . *:  .. : ::*:* ::** **:* . :**** .. *:** ::**:   :** ::.*

LOC_Os08g43120.1      HKHQVNINATKVKVDYNSQ------IVGNGTASTDQVILPFQPLSLVFDHINYFVDMPKE 737
MLOC_54794.6          HKHQVNIDSTKTKTECKKQ------KVGTGNASTGQVVLPFQPLSLVFDHINYFVDMPKE 674
ABCG40                GKPQAVIAEEPASDETELQSARSEGVVEAGANKKRGMVLPFEPHSITFDNVVYSVDMPQE 839
                       * *. *    .. : : *       *  *  ..  ::***:* *:.**:: * ****:*
                                                      Walker Box A
LOC_Os08g43120.1      MTKYGVTDKKLQLLQDVSGAFRPGVLTALMGITGAGKTTLLDVLAGRKTGGYIEGTVKIA 797
MLOC_54794.6          MMKYGVTEKKLQLLQDVSGVFRPGVLTALMGVTGAGKTTLLDVLAGRKTGGYIEGTIRIA 734
ABCG40                MIEQGTQEDRLVLLKGVNGAFRPGVLTALMGVSGAGKTTLMDVLAGRKTGGYIDGNITIS 899
                      * : *. :.:* **:.*.*.**********::*******:************:*.: *:

LOC_Os08g43120.1      GYPKKQETFSRISGYCEQSDIHSPNLTVYESLQFSAWLRLPSNVKSHQRNMFIDEVMDLV 857
MLOC_54794.6          GYPKKQDTFSRISGYCEQSDIHSPNLTVHESLQFSAWLRLPSNVNSRQRDMFIDEVMDLV 794
ABCG40                GYPKNQQTFARISGYCEQTDIHSPHVTVYESLVYSAWLRLPKEVDKNKRKIFIEEVMELV 959
                      ****:*:**:********:*****::**:*** :*******.:*...:*.:**:***:**
                              Walker Box B
LOC_Os08g43120.1      ELTGLKNAMVGLAGATGLSAEQRKRLTIAVELVASPSIIFMDEPTTGLDARAAAIVMRTV 917
MLOC_54794.6          ELTGLKNAMVGIAGATGLSAEQRKRLTIAVELVASPSIIFMDEPTTGLDARAAAIVMRTV 854
ABCG40                ELTPLRQALVGLPGESGLSTEQRKRLTIAVELVANPSIIFMDEPTSGLDARAAAIVMRTV 1019
                      *** *::*:**:.* :***:**************.**********:**************

LOC_Os08g43120.1      RKTVDTGRTVVCTIHQPSIEIFESFDELLLMKRGGQLIYSGSLGPLSSNMIKYFEAIPGV 977
MLOC_54794.6          RKTVDTGRTVVCTIHQPSIGIFESFDELLLMKRGGQIIYSGPLGPLSSNMIKYFEAIPGV 914
ABCG40                RNTVDTGRTVVCTIHQPSIDIFEAFDELFLLKRGGEEIYVGPLGHESTHLINYFESIQGI 1079
                      *:***************** ***:****:*:****: ** *.**  *:::*:***:* *:

LOC_Os08g43120.1      PRIKEGQNPAAWMLDISSRTAEYEIGVDYAEIYQRSSLYWENRQLIDDLGKPEPNTEDLH 1037
MLOC_54794.6          PRIKEGQNPAAWVLDISSHITEYEIGVDYAEIYRSSSLYRENMLLIDELGQPAPNTEDLH 974
ABCG40                NKITEGYNPATWMLEVSTTSQEAALGVDFAQVYKNSELYKRNKELIKELSQPAPGSKDLY 1139
                       :*.** ***:*:*::*:   *  :***:*::*: *.** .*  **.:*.:* *.::**:

LOC_Os08g43120.1      FPPKYWQDFRAQCMACLWKQNCAYWKNSEHNVVRFINTFAVSIMFGIVFWKIGSTIKDEQ 1097
MLOC_54794.6          FPPGYWQNFRAQCMACLWKQRCAYWKNSEHNVVRFLNTFAVSIMFGIVFWKIGSIIKGQQ 1034
ABCG40                FPTQYSQSFLTQCMASLWKQHWSYWRNPPYTAVRFLFTIGIALMFGTMFWDLGGKTKTRQ 1199
                      **. * *.* :****.****. :**:*. :..***: *:.:::*** :**.:*.  * .*
```

```
LOC_Os08g43120.1      DVFNILGVVYGSALFLGFMNCSILQPVVGMERVVLYREKAAGMYSTMAYAIAQVAVELPY 1157
MLOC_54794.6          DVFNILGVVYGSALFLGFMNCSILQPVVTMERVVLYREKAAGMYCTLAYAIAQMAIELPY 1094
ABCG40                DLSNAMGSMYTAVLFLGLQNAASVQPVVNVERTVFYREQAAGMYSAMPYAFAQVFIEIPY 1259
                      *: * :* :* :.****: *.: :**** :**.*:***:*****.::.**:**: :*:**

LOC_Os08g43120.1      MFVQVFIFSAIVYPMIGFQMTATKFFWFALYMVLSFLYYTLYGMMTVALTPNIEIAAGLS 1217
MLOC_54794.6          MLVQVLIFASIVYPMIGFEMTAVKFFWFVLYMVLSFMYYTLYGMMTVALTPNLEIAAGLS 1154
ABCG40                VLVQAIVYGLIVYAMIGFEWTAVKFFWYLFFMYGSFLTFTFYGMMAVAMTPNHHIASVVS 1319
                      ::**.::::. ***.****: **.****: ::*  **: :*:****:**:*** .**: :*

LOC_Os08g43120.1      FLIFIFWNVFSGFIIGRQMIPVWWRWVYWANPAAWTVYGLMFSQLGDRTELIQVPGQPEQ 1277
MLOC_54794.6          FLIFIFWNVFSGFIIGRELIPIWWRWVYWANPAAWTVYGLMFSQLGDRTEIIHVPGQPDQ 1214
ABCG40                SAFYGIWNLFSGFLIPRPSMPVWWEWYYWLCPVAWTLYGLIASQFGDITEPMADSN---M 1376
                        :: :**:****:* *  :*:**.* **  *.***:***: **:** ** :  ..

LOC_Os08g43120.1      TVKEFLEGYLGLQDRYFNLVTSLHVAIIALFTFLFFLSIKHLKFQRR 1324
MLOC_54794.6          TVREFLEGYLGLENRYFNLVTCLHLAIIALFALFFIFIKHLKFQRR 1261
ABCG40                SVKQFIREFYGYREGFLGVVAAMNVIFPLLFAVIFAIGIKSFNFQKR 1423
                      :*::*:. : * .: ::.:*:.::: :  **:.:* : ** ::**:*



MLOC_68581.1          ------------------------------------------------------------
LOC_Os08g43120.1      ------------------------------------------------------------
ABCG40                MEGTSFHQASNSMRRNSSVWKKDSGREIFSRSSREEDDEEALRWAALEKLPTFDRLRKGI 60


MLOC_68581.1          -----------MTEYSRGAEQAPETRARPPPPLTHEDNRRFLEMLREKKERLGVGVEKVE 49
LOC_Os08g43120.1      --------MAAAVELTGDGGTTAETRWLSPP-LTHDDNRGFLQMLREKKERLGVGAAKVE 51
ABCG40                LTASHAGGPINEIDIQKLGFQDTKKLLERLIKVGDDEHEKLLWKLKKRIDRVGIDLPTIE 120
                                :    .   .:.       : .:::. :*  *::: :*:*:.  .:*

MLOC_68581.1          VRFEELTVEADVRVGRRTLPTLLNCVLNAAQELATYSHMCTTRKKPIKIINGVSGTIRPS 109
LOC_Os08g43120.1      VRLEKLTVEADVRVGRRAVPTLLNCAINAAQELAACAHMCTTRKKPMKIINEATGTIRPS 111
ABCG40                VRFDHLKVEAEVHVGGRALPTFVNFISNFADKFLNTLHLVPNRKKKFTILNDVSGIVKPG 180
                      **::.*.***:*:** *::**::*   * *:::    *: ..*** :.*:* .:* ::*.
                                Walker Box A
MLOC_68581.1          RMTLLLGAPGSGKTTLLKALAGKLDSSLKLKGKVMYNGEEVNFSTPQYMHAYISQYDLHN 169
LOC_Os08g43120.1      RMTLLLGAPGSGKTTLLKALAGKLDSSLKMKGKVTYNGEEVNSSTPQYLHAYVSQYDLHH 171
ABCG40                RMALLLGPPSSGKTTLLLALAGKLDQELKQTGRVTYNGHGMNEFVPQRTAAYIGQNDVHI 240
                      **:****.*.*****:* *******..** .*:* ***. :*  .**   **:.* *:*

MLOC_68581.1          AEMTVRETIDFSSKMLGTNNAFGMLGEGIRRKYGLINKVDEDIDSFTKDT-TFGEASNLT 228
LOC_Os08g43120.1      AEMTLSS----LKKLICT-------GHNLWRRKQPYNKL-------------YYQAIKIE 207
ABCG40                GEMTVRETFAYAARFQGVGSRYDMLTELARREKEANIKPDPDIDIFMKAMSTAGEKTNVM 300
                      .***: .      ::  .        .   *.     *                :  ::
                                Signature motif  Walker Box B
MLOC_68581.1          TNYIIKILGLYDCADTLVGDELRRGISGGQKKRATIGEMLVGLARCFFMDDISTGLDSST 288
LOC_Os08g43120.1      C---MQILGLSECADTLVGDEMRRGISGGQKKRATIGEMLVGLARCFFMDDISTGLDSST 264
ABCG40                TDYILKILGLEVCADTMVGDDMLRGISGGQKKRVTTGEMLVGPSRALFMDEISTGLDSST 360
                          ::****  ****:***:: **********.* ****** :*.:***:*********

MLOC_68581.1          TYEILKFVQQMTHLMDLTMVISLLQPPPETLELFDDIILLCEGQIVYHGPRENAIHFFGI 348
LOC_Os08g43120.1      TFEIMKFLQQMAHLMDLTMVISLLQPPPETLELFDDIILLCEGQIVYHGPRENATDFFET 324
ABCG40                TYQIVNSLRNYVHIFNGTALISLLQPAPETFNLFDDIILIAEGEIIYEGPRDHVVEFFET 420
                      *::*:: ::: .*::: * :******.***::*******:.**:*:*.***::. .**

MLOC_68581.1          MGFKCPERKNVADFLQEVTSKMDQKQYWMGDENKYQYRPIEKFVESFHSSYLPQLIEDNH 408
LOC_Os08g43120.1      MGFKCPSRKNVADFLQEVTSKMDQKQYWIGNANKYQYHSIEKFAESFRTSYLPRLVENDH 384
ABCG40                MGFKCPPRKGVADFLQEVTSKKDQMQYWARRDEPYRFIRVREFAEAFQSFHVGRRIGDEL 480
                      ****** **.*********** ** ***    : *::  :.:*.*:*:: :: : : ::

MLOC_68581.1          YMTNNTRENK-VIKTSETHRISRWDIFKACFLREVLLLKRNSPLHIFKTVQTVVLAFVIS 467
LOC_Os08g43120.1      FESTNAGKSK-EVKTSTSRMISSWNIFKACFSREVLLLKRNSPVHIFKTIQITVLALVIS 443
ABCG40                ALPFDKTKSHPAALTTKKYGVGIKELVKTSFSREYLLMKRNSFVYYFKFGQLLVMAFLTM 540
                        . :  :.:    *: .  :.  ::.*:.* ** **:**** :: **  *  *:*::

MLOC_68581.1          TIFLRTNMNHKNVLDANKYMGSLFIAIVIVNMNGMTEIAMTIKRLPTFYKQRELLALPGW 527
LOC_Os08g43120.1      TLFLRTNMRHDTVLDANKYMGALFMAVVIVNFNGMTEIAMTIKRLPIFYKQREILALPGW 503
ABCG40                TLFFRTEMQKKTEVDGSLYTGALFFILMMLMFNGMSELSMTIAKLPVFYKQRDLLFYPAW 600
                      *:*:**:*.:.. :*.. * *:**: :::: :***:*::*** :** *****::*  *.*

MLOC_68581.1          ALLSSIFLISIPMSLVETFLWTSLTYYVIGYAPSFLRFFQQFMVLFATHQMSMGLYRFLA 587
LOC_Os08g43120.1      ALLSSVFLLSLPISFVETGLWTGLTYYVIGYAPSFVRFIQHFVVLFAMHQMSMSLYRFLA 563
ABCG40                VYSLPPWLLKIPISFMEAALTTFITYYVIGFDPNVGRLFKQYILLVLMNQMASALFKMVA 660
```

```
                    .   . :*:.:*:*::*: * * :******: *.. *::::::::*.  :**: .*::::*

MLOC_68581.1        AIGRTQVMANMLSTTAVIAMYIFGGFVISKDNLQPWLQWGCWISPFTYAQNAVALNEFLD 647
LOC_Os08g43120.1    AIGRTQVMANMLGTAALIAIYILGGFVISKDNLQPWLRWGYWTSPFTYAQNAVALNEFLD 623
ABCG40              ALGRNMIVANTFGAFAMLVFFALGGVVLSRDDIKKWWIWGYWISPIMYGQNAILANEFFG 720
                    *:**. ::** :.: *::.:: :**.*:*:*::: *  ** * **: *.***:  ***:.

MLOC_68581.1        ERWATEFYYTNANTVGEAILKIRGLLTNWHWYWICVSILFAFSMVFNVLTIFALEFMKSP 707
LOC_Os08g43120.1    DRWATEFHFANANTVGETILKVRGLLTEWHWYWICVSILFGFSLVFNILSIFALQYMRSP 683
ABCG40              HSWSRAVENS-SETLGVTFLKSRGFLPHAYWYWIGTGALLGFVVLFNFGFTLALTFLNSL 779
                    . *:  .  : ::*:* ::** **:*.. :**** .. *:.* ::**.   :** ::.*

MLOC_68581.1        HKHQVNTS------ATKSKLICKNYKIGTGNASSGQVVLPFQPLSIVFDHINYFVDMPKE 761
LOC_Os08g43120.1    HKHQVNIN------ATKVKVDYNSQIVGNGTASTDQVILPFQPLSLVFDHINYFVDMPKE 737
ABCG40              GKPQAVIAEEPASDETELQSARSEGVVEAGANKKRGMVLPFEPHSITFDNVVYSVDMPQE 839
                     * *.          *: :   ..  :  *  ..  ::***:* *:.**:: * ****:*
                                                Walker Box A
MLOC_68581.1        MMKYGVTEKKLQLLQDVSGVFRPGVLTALMGITGAGKTTLLDVLAGRKTGGYTEGSIRIA 821
LOC_Os08g43120.1    MTKYGVTDKKLQLLQDVSGAFRPGVLTALMGITGAGKTTLLDVLAGRKTGGYIEGTVKIA 797
ABCG40              MIEQGTQEDRLVLLKGVNGAFRPGVLTALMGVSGAGKTTLMDVLAGRKTGGYIDGNITIS 899
                    * : *. :.:* **:.*.*.**********::*******:********** :*.: *:

MLOC_68581.1        GYPKKQDTFSRISGYCEQNDIHSPNLTVHESLQFSAWLRLPSNVKSHERNMFIDEIMDLV 881
LOC_Os08g43120.1    GYPKKQETFSRISGYCEQSDIHSPNLTVYESLQFSAWLRLPSNVKSHQRNMFIDEVMDLV 857
ABCG40              GYPKNQQTFARISGYCEQTDIHSPHVTVYESLVYSAWLRLPKEVDKNKRKIFIEEVMELV 959
                    ****:*:**:********.*****::**:*** :*******.:*..::*::**:*:*:**
                            Walker box B
MLOC_68581.1        ELTGLKNAMVGQAGATGLSAEQRKRLTIAMELVASPSIIFMDEPTTGLDARAAAIVMRTI 941
LOC_Os08g43120.1    ELTGLKNAMVGLAGATGLSAEQRKRLTIAVELVASPSIIFMDEPTTGLDARAAAIVMRTV 917
ABCG40              ELTPLRQALVGLPGESGLSTEQRKRLTIAVELVANPSIIFMDEPTSGLDARAAAIVMRTV 1019
                    *** *::*:** .* :***:*********:****.**********:*************:

MLOC_68581.1        RKTVDTGRTVVCTIHQPSIEIFESFDELLLMKRGGQIIYNGSLGPLSSNMIKYFEDIPSV 1001
LOC_Os08g43120.1    RKTVDTGRTVVCTIHQPSIEIFESFDELLLMKRGGQLIYSGSLGPLSSNMIKYFEAIPGV 977
ABCG40              RNTVDTGRTVVCTIHQPSIDIFEAFDELFLLKRGGEEIYVGPLGHESTHLINYFESIQGI 1079
                    *:*****************:***:****:*:****: ** *.**  *:::*:*** * .:

MLOC_68581.1        PRIKEGQNPASWVLDISSHTMEYEIGVDYAEIYRSSYLYRDKMLLIDELGHPTPNTEDLY 1061
LOC_Os08g43120.1    PRIKEGQNPAAWMLDISSRTAEYEIGVDYAEIYQRSSLYWENRQLIDDLGKPEPNTEDLH 1037
ABCG40              NKITEGYNPATWMLEVSTTSQEAALGVDFAQVYKNSELYKRNKELIKELSQPAPGSKDLY 1139
                     :*.** ***:*:*::*: : *  :***:*::*: * **  :  **.:*.:* *.::**:

MLOC_68581.1        FPPGYWQNFMAQCMACLWKQRCAYWKNLEHNGVRFLNTFVVSIMFGVVFWKIGSTITQEQ 1121
LOC_Os08g43120.1    FPPKYWQDFRAQCMACLWKQNCAYWKNSEHNVVRFINTFAVSIMFGIVFWKIGSTIKDEQ 1097
ABCG40              FPTQYSQSFLTQCMASLWKQHWSYWRNPPYTAVRFLFTIGIALMFGTMFWDLGGKTKTRQ 1199
                    **. * *.* :****.****. :**:*  :. ***: *: :::*** :**.:*.. . .*

MLOC_68581.1        DVFNILGVVFGSALFLGFNNCIGLQPVVIMERVVLYREKASGMYSTLAYTIAQMAIELPY 1181
LOC_Os08g43120.1    DVFNILGVVYGSALFLGFMNCSILQPVVGMERVVLYREKAAGMYSTMAYAIAQVAVELPY 1157
ABCG40              DLSNAMGSMYTAVLFLGLQNAASVQPVVNVERTVFYREQAAGMYSAMPYAFAQVFIEIPY 1259
                    *: * :* :: :.****: *.  :**** :**.*:***:*:****::.*::**: :*:**

MLOC_68581.1        ILVQVLVFSSIVYPMTGFQMTTVKFFWFVLYMVLSFMYYTLYGMMTVALTPNLEMAAGLS 1241
LOC_Os08g43120.1    MFVQVFIFSAIVYPMIGFQMTATKFFWFALYMVLSFLYYTLYGMMTVALTPNIEIAAGLS 1217
ABCG40              VLVQAIVYGLIVYAMIGFEWTAVKFFWYLFFMYGSFLTFTFYGMMAVAMTPNHHIASVVS 1319
                    ::**.:::. ***.* **: *:.****: ::*  **: :*:****:**:*** .:*: :*

MLOC_68581.1        FLIYVFWTVFSGFIIGRELIPIWWRWVYWANPAAWTVYGLMFSQLGDRTELISVPGQPDQ 1301
LOC_Os08g43120.1    FLIFIFWNVFSGFIIGRQMIPVWWRWVYWANPAAWTVYGLMFSQLGDRTELIQVPGQPEQ 1277
ABCG40              SAFYGIWNLFSGFLIPRPSMPVWWEWYYWLCPVAWTLYGLIASQFGDITEPMADSN---M 1376
                      :: :*.:****:* *  :*:**.* **  *.***:***: **:** ** :  ..

MLOC_68581.1        TVQEFLEGYLGIEGRYFYLITYLHLAIIAIFAFLFFIFVKHLKFEWR 1348
LOC_Os08g43120.1    TVKEFLEGYLGLQDRYFNLVTSLHVAIIALFTFLFFLSIKHLKFQRR 1324
ABCG40              SVKQFIREFYGYREGFLGVVAAMNVIFPLLFAVIFAIGIKSFNFQKR 1423
                    :*::*:. : * .  :: ::: ::: :  :*:.:* : :* ::*: *



LOC_Os11g07600.1    ---------------------------MDCFLSSVCTPLNLQFIDVAYRVKVSTTAAAAK 33
MLOC_62985.1        MPPNGQSVHGGGGLVSPATAAPAPPSKMDCFLTSVCTPLNLQFIDVAYRVKVERTAKE-- 58
ABCG25              MSAFDGVENQMNGPDSSPRLS-QDPREPRSLLSSSCFPITLKFVDVCYRVKIHGMSND-- 57
                                        .:*:* * *:.*:*:**.****:   :
```

```
LOC_Os11g07600.1      GAPPGRISHAGGTGGGGGAQ------EERTILKGITGEARPGEVLAVLGPSGSGKSTLLS 87
MLOC_62985.1          --PPGRISHSGGGGGGGAAGGLSASVEERTILKGITGEARPGEVLAVLGPSGSGKSTLLS 116
ABCG25                ---SCNIKKLLGLKQKPSDE--TRSTEERTILSGVTGMISPGEFMAVLGPSGSGKSTLLN 112
                         . .*.:  *     .        ******.*:**   ***.:***************.

LOC_Os11g07600.1      ILGGRLAG-RHAGTVLAGGRAPCRAVQRRTGFVAQDDVLHPHLTVRETLLFCAMLRLPAS 146
MLOC_62985.1          ILGGRLAG-RYSGTVLTGGRAPCRAVQRRTGFVAQDDILHPHLTVRETLAFCAMLRLPTS 175
ABCG25                AVAGRLHGSNLTGKILINDGKITKQTLKRTGFVAQDDLLYPHLTVRETLVFVALLRLPRS 172
                       :.*** * . :*.:* ..    : . :*********:*:******** * *:**** *
                                                 Walker box A
LOC_Os11g07600.1      SPAAAKAAAAEAVIAELGLAPCADTIVGNAFVRGVSGGERKRVSIGHELLVNPSLLVLDE 206
MLOC_62985.1          APTSAKLAAADAVIAELGLGTCADTIVGNAFVRGVSGGERKRVSIGHELLVNPSLLVLDE 235
ABCG25                LTRDVKLRAAESVISELGLTKCENTVVGNTFIRGISGGERKRVSIAHELLINPSLLVLDE 232
                       .  .*  **::**:****  * :*:***:*:**:**********.****:*********

LOC_Os11g07600.1      PTSGLDSTAAARLVATLSSLAR-RGRTVVMSVHQPSTRVYRMFDSVLLLAEGTCLYFGAG 265
MLOC_62985.1          PTSGLDSTAASRLVATLSALAR-KGRTVVLSVHQPSSRVYRAFDSVLLLSEGSCMYHGPG 294
ABCG25                PTSGLDATAALRLVQTLAGLAHGKGKTVVTSIHQPSSRVFQMFDTVLLLSEGKCLFVGKG 292
                      ******:*** *** **:.**: :*:*** *:****:**:: **:****:**.*:: * *

LOC_Os11g07600.1      RDAMDYFAAVGFSPAFHVNPADFMLDLANGFAQTEYDNCNAADGGNVKQSLISSYNRVLA 325
MLOC_62985.1          RDAMDYFASVGFAPGFHVNPADFMLDLANGFAQAEYSDR-AAEGGSVKQSLVASYGRVLA 353
ABCG25                RDAMAYFESVGFSPAFPMNPADFLLDLANGVCQTDGVTE--REKPNVRQTLVTAYDTLLA 350
                      **** ** :***:*.* :*****:******..*::       :  .*:*:*:::*. :**

LOC_Os11g07600.1      PRVKASINAG-DVHGGEQQPPPPAAES--CSGCTSWSNQFAILLRRSLKERRHEAFTSLR 382
MLOC_62985.1          PKVKAAIATGAHTDDGHGHAAEQPLES--CSGCTSWTNQFAILLRRSLKERRHETFTSLR 411
ABCG25                PQVKTCIEVSHFPQDNARFVKTRVNGGGITTCIATWFSQLCILLHRLLKERRHESFDLLR 410
                      *:**:.* ..   ...          .   :  ::* .*:.***:* *******:*  **

LOC_Os11g07600.1      LFQIIAPALVAGAMWWRSSPAAVGDRMGLLFFVSIFWGVFASFNAVFAFPQERPVLARER 442
MLOC_62985.1          LFQIIAPAVVAGAMWWRSTPLEVQDRMGLLFFISIFWGVFASFNAVFAFPQERPVLAREL 471
ABCG25                IFQVVAASILCGLMWWHSDYRDVHDRLGLLFFISIFWGVLPSFNAVFTFPQERAIFTRER 470
                      :**::*.::::.* ***:*    * **:*****:******:.******:*****.:::**

LOC_Os11g07600.1      ASGMYALSSYFMSRMAGDLPMELALPAAFTVIVYLMAGLNPSPAAFALTLAVILSYVLVA 502
MLOC_62985.1          ASGMYSLSSYFMSRMAGDLPMELALPTAFTLIVYLMAGLNPAPAAFALTLLVILSYVLVA 531
ABCG25                ASGMYTLSSYFMAHVLGSLSMELVLPASFLTFTYWMVYLRPGIVPFLLTLSVLLLYVLAS 530
                      *****:******::: *.*.***.**::*  :.* *. *.*. ..* *** *:* ***.:

LOC_Os11g07600.1      EGLGLAVGAVMMDAKRASTLVTVVMLAYLLTGGFYVHNVPGFMAWAKYTSFTYYCYRLLI 562
MLOC_62985.1          EGLGLAIGALMMDAKRASTLATVIMLAYLLTGGFYVHNVPVFMIWAKYSSFTYYCYRLLI 591
ABCG25                QGLGLALGAAIMDAKKASTIVTVTMLAFVLTGGYYVNKVPSGMVWMKYVSTTFYCYRLLV 590
                      :*****:** :****:***:.** ***::****:**::**  * * ** * *:******:

LOC_Os11g07600.1      AVQYSG--RLARLLPPEE--------------------ARGEASPAACVAALVAMFFAYR 600
MLOC_62985.1          AVQYSG--HLAQLLPPDS--------------------THGEASTWTCVAALVVMFFGYR 629
ABCG25                AIQYGSGEEILRMLGCDSKGKQGASAATSAGCRFVEEEVIGDVGMWTSVGVLFLMFFGYR 650
                      *:**..  .: ::*  :.                    . *:..  :.*..*. ***.**

LOC_Os11g07600.1      LLAYLALRRVRK 612
MLOC_62985.1          LLAYFALRRVRT 641
ABCG25                VLAYLALRRIKH 662
```

Asterisk indicate the conservation of residue in all the three species and dashes indicate the gaps or absence of sequence data; alignments were made using CLUSTALW2 (http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=clustalw2; last accessed May 2014) and then bioedit was used to align the conserved functional motifs; bold: indicate ABC signature motif; box: indicate Walker box A and B; rice sequences were derived from the MSU rice genomic annotation project (http://rice.plantbiology.msu.edu/; last accessed May 2014) and Arabidopsis sequences were derived from TAIR (http://www.arabidopsis.org/index.jsp); **E.** motif sequence and the position were derived from Brule and Smart (2002).

Appendix XVI Analysis of Sanger sequencing data.

NTC	Gene family	% identity of sequenced data to NGS sequence	BLAST results of sequenced data (NCBI)	BLAST results of sequenced data (TIGR)
NTC10767	PP2CI	1 of 913 bp not matched	-	-
NTC14333	PP2CD	100%	-	-
NTC40607	PP2CE	1 of 957 bp not matched	-	-
NTC2524	PYR	50% to NTC 2524	99% to AK365182.1 having 97% coverage (unknown barley gene	98% with protein|actin, putative, expressed
NTC9836	ABC transporter	93% 1200 bp	90% with Brachypodium distachyon pleiotropic drug resistance	91% ABC transporter
NTC4445	ABC transporter	100%	-	-

Appendix XVII K^+ ion concentration based stress tolerance of 16 varieties of barley.

	Calmariout	Hindmarsh	Mundah	ELB14	Golden promise	Buloke	Maythorpe	Vlamingh
K^+ control (root)	8.69±0.34	8.81±0.46	9.33±0.13	9.84±0.49	10.57±0.44	11±0.76	13.9±0.29	13.35±0.33
K^+ stress (root)	7.82±0.14	7.50±0.45	7.44±0.37	6.04±0.36	5.57±0.77	5.38±0.48	5.73±0.39	5.477±0.25
K^+ Index (root)	**90.03**	**85.19**	**79.73**	**61.42**	**52.67**	**48.91**	**41.26**	**41.01**
K^+ control (shoot)	54.33±0.43	56±0.56	75.87±0.47	58.9±0.44	58.5±0.32	64.22±0.16	63.55±0.09	62.66±0.18
K^+ stress (shoot)	41.56±0.76	41.02±0.48	42.67±0.24	40.95±0.43	44.38±0.12	45.83±0.07	46.44±0.37	45.62±0.46
K^+ Index (shoot)	**76.50**	**73.25**	**56.24**	**69.53**	**75.86**	**71.37**	**73.07**	**72.79**

Na^+ ion Index = (Na^+ (stress)/Na^+ (control))*100; Na^+/K^+ ion Index = (Na^+/K^+ (stress)/Na^+/K^+ (control))*100; ±= standard error

Appendix XVII K^+ ion concentration based stress tolerance of 16 varieties of barley (cont).

	Beecher	Clipper	Skiff	CM72	Numar	Arivat	Gairdner	Franklin
K^+ control (root)	13.81±0.47	14.40±0.65	15.90±0.23	15.28±0.15	16.56±0.19	13.20±0.47	12.43±0.26	12.92±0.22
K^+ stress (root)	5.37±0.13	5.27±0.25	5.30±0.43	5.36±0.37	5.37±0.24	4.63±0.07	4.483±0.28	4.38±0.11
K^+ Index (root)	**38.92**	**36.63**	**33.36**	**35.12**	**32.42**	**35.07**	**36.04**	**33.91**
K^+ control (shoot)	50.81±0.10	59.11±0.24	57.88±0.27	50.90±0.49	48.70±0.67	52.11±0.46	50.11±0.37	52.75±0.33
K^+ stress (shoot)	44.89±0.14	43.17±0.12	42.09±0.32	42.01±0.27	40.65±0.25	40.69±0.52	38.52±0.48	36.08±0.09
K^+ Index (shoot)	**88.35**	**73.04**	**72.71**	**82.55**	**83.47**	**78.09**	**76.87**	**68.41**

Na^+ ion Index = (Na^+ (stress)/Na^+ (control))*100; Na^+/K^+ ion Index = (Na^+/K^+ (stress)/Na^+/K^+ (control))*100; ±= standard error