GENETIC BASIS AND MOLECULAR MECHANISM

OF WHEAT DEVELOPMENT

By

MING YU

Bachelor of Science in Biotechnology
Shandong Agricultural University
Tai'an, Shandong, China
2009

Master of Science in Plant Science
Oklahoma State University
Stillwater, Oklahoma
2011

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2014

GENETIC BASIS AND MOLECULAR MECHANISM

OF WHEAT DEVELOPMENT

Dissertation Approved:

Dr. Liuling Yan

Dissertation Adviser

Dr. Brett Carver

Dr. Million Tadege

Dr. Ramamurthy Mahalingam

ACKNOWLEDGEMENTS

It is with immense gratitude that I acknowledge the support and help of my advisor, Dr. Liuling Yan who provided me with the opportunity to pursue a PhD degree at his lab. This dissertation would not have been possible without his continuous support and guidance throughout the project.

I would like to express my gratitude to my advisory committee members, Dr. Brett Carver, Dr. Million Tadege, and Dr. Ramamurthy Mahalingam, for their valuable suggestions and help. Their knowledge and advice have been of great help to my research project.

I would also like to thank Dr. Junpeng Deng and Dr. Yongfang Li for their support. I am indebted to my colleagues and many staff for their precious friendship and support: Dr. Genqiao Li, Dr. Ying Wang, Dr. Carol Power, Lei Lei, Forest Kan, and visiting scholars Dr. Fang Miao and Meiyan Liu.

This study was supported by Oklahoma Wheat Commissions and the Oklahoma Center of Advanced Science and Technology (OCAST). Thank you for the financial support to help me to finish my PhD study.

Last but not least, I owe my deepest gratitude to my family and friends for their love and encouragement during these years.

Name: MING YU

Date of Degree: DECEMBER, 2014

Title of Study: GENETIC BASIS AND MOLECULAR MECHANISM OF WHEAT DEVELOPMENT

Major Field: CROP SCIENCE

Abstract: Wheat is one of the most important crops in the world. Wheat can be classified according to its qualitative requirement for low temperature to accelerate flowering (*i.e.* vernalization). Winter wheat has a vernalization requirement whereas spring wheat does not. Winter wheat can be further classified according to its quantitative vernalization requirement: weak winter, semi-winter, and strong winter types. Three genes were cloned based on qualitative variation in vernalization requirement between spring and winter wheat, but genes controlling quantitative variation among winter wheat cultivars are not well characterized. In this study, the following genetic features for wheat development were addressed.

First, the micro RNA molecule *Ta*miR1123 present in the *Vrn-A1a* promoter was found to be associated with the transcriptional level of *Vrn-A1*. Hence an alternative regulatory mechanism is proposed such that *Vrn-A1* is regulated by plant age, low temperature, and short days through *Ta*miR1123.

Second, the vernalization requirement duration in winter wheat was found to be controlled by *Ta*VRN-A1 at the protein level. The mutation in the gene encoding *Ta*VRN-A1 thus impacts its interaction ability with *Ta*HOX1. This protein-protein interaction was confirmed by *in vitro* and *in vivo* assays. A SNP in the gene encoding *Ta*HOX1 was associated with flowering time variation. This study presented the first example that MADS and HOX proteins involving homeosis have a direct-binding relationship in higher plants.

Third, *TaHOX1* was mapped on chromosome 6B (*TaHOX-B1*), and homoeologues *TaHOX-A1* and *TaHOX-D1* were sequenced but showed no allelic variation. Whereas *TaHOX-A2* was mapped on chromosome 2A due to an 18-bp indel polymorphism in exon 1, *TaHOX-B2* or *TaHOX-D2* showed no allelic variation.

Lastly a wheat centromeric protein encoded by *TaCENPE1* from a yeast-2 hybrid (Y2H) library was also confirmed to directly bind with *TaVRN1* in living cells based on a transient expression system applied in tobacco leaves. Plant height of *TaCENPE1::RNAi* transgenic wheat was reduced and more florets were produced relative to non-transformed wheat.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

GENERAL INTRODUCTION

Abstract

All wheat cultivars grown in Oklahoma and surrounding states are of winter type, which require more or less vernalization for timely flowering. A unique characteristic of the winter wheat grown in this region is that wheat cultivars are utilized in the dual purpose agricultural system. These cultivars are required to have a longer vegetative growth phase, during which more plant biomass is produced for cattle grazing. In addition, proper timing of flowering and maturity are required to ensure that high grain yield is also achieved after grazing. This unique developmental pattern is genetically controlled by multiple genes responding to temperature and photoperiod. Three vernalization genes, *VRN1*, *VRN2* and *VRN3*, have been cloned, but all of them were cloned based on qualitative variation between winter wheat with a vernalization requirement and spring wheat without this requirement. A tri-loci model (*vrn-A1*, *PPD-D1*, and *vrn-D3*) has been established to select for the development pattern of most winter wheat cultivars, but the

current model does not match up with phenotypes in some of the tested cultivars, suggesting that more genes are involved in developmental trait of winter wheat.

In this study, we found that *Vrn-A1* is regulated at the transcriptional level via *Ta*miRNA1123 to determine flowering time in spring wheat cultivars, and the same gene is regulated at the protein level to determine the quantitative variation in vernalization requirement duration in winter wheat cultivars.

1.1 The importance of wheat development

Wheat is one of the most important crops

Winter wheat (*Triticum aestivum* L. 2n=6x=42, AABBDD) is one of the most important crops for human nutrition in the world. U.S wheat occupies approximately 11% of the world supply, and nearly 35% of world exports in recent years. A substantial increase in the productivity of wheat is needed to meet the demands for food due to the projected increase in world population from current 6 billion to 9 billion by 2050 (McMichael, 2001). Wheat grown the Great Plains accounts for 40% of the total 20 million hectares U.S. harvested wheat and more than 50% of the total 48 million metric tons U.S. wheat production .

In addition to biotic stresses of insects and diseases on wheat production, abiotic stresses such as adverse temperature and light are main factors that inhibit further increase of wheat productivity. A better understanding of major mechanisms underlying adaptation of wheat will enable us to understand why wheat is the most widely grown crop worldwide. The adaptive mechanisms can be explained by a few known genes and many more unknown genes controlling wheat development.

1.2 Genetic basis of wheat development

Three pathways are established for wheat development

Bread wheat is a hexaploid species and it could have as many as over 200 genes affecting flowering time, based on the prediction of flowering gene number in diploid *Arabidopsis* (Yan, 2009). Three pathways have been revealed involved in the regulation of flowering time in wheat, including vernalization, photoperiod, and earliness *per se* (Valárik *et al.*, 2006).

Vernalization is a major pathway controlling wheat development

The temperate grasses, which include economically important species such as wheat, barley, rye, and oats, are well-adapted to cold winters. Most of these species require a prolonged period of cold treatment for timely flowering, a process referred to as vernalization. Wheat is divided into two types, based on its qualitative requirement for low temperature to accelerate flowering (*i.e.* vernalization), winter wheat with a vernalization requirement and spring wheat without this requirement (Law, 1967; Pugsley, 1971). Vernalization is an adaptive mechanism of the most importance, allowing winter wheat to synchronize its developmental transition with changes in seasonal climate (Flood and Halloran 1986; Kirby *et al*. 1999).

This requirement delays the initiation of the reproductive phase and protects the sensitive floral meristems from frost damage during the winter. It also contributes to the precise adjustment of flowering time to seasonal changes, which is important to maximize seed production. Therefore, a better understanding of the mechanisms involved in the regulation of wheat flowering can contribute to the engineering of high yielding varieties adapted to changing environments.

In contrast with adverse effects of higher temperature on the reproductive developmental transition, increasing temperature in any season will increase growth rate resulting in precocious maturity due to a shorter phyllochron (McMaster, 2009).

Photoperiod interacts with vernalization in regulation of wheat development

Photoperiod also has complex and paradoxical effects on winter wheat development. Wheat is sensitive to long day (LD, 14 h or more light) and flowers earlier when exposed to LD; therefore, wheat is a LD plant. Winter wheat is required to meet vernalization under long days (Law, 1967; Law and Wolfe, 1966; Pugsley, 1971). Winter wheat, for which any requirement for vernalization has been fully satisfied, will have similar responses to ambient temperature as spring wheat (Snape et al. 2001).

Wheat cultivars are classified into sensitive and insensitive types, based on their responses to photoperiod. A cultivar with photoperiod insensitivity is a mutant from sensitivity to LD, which enables the wheat to flower earlier under short day (SD) photoperiod conditions (Law and Worland, 1997; Snape *et al*., 2001). Short days can partially replace low temperature to promote the developmental transition of cultivars insensitive to photoperiod (Evans et al. 1987; Heide 1994; Snape et al. 2001), but short days would repress plant growth. An earlier arrival of long days in spring in the temperate region would allow photoperiod-insensitive wheat to flower sooner (Turner et al. 2005).

Earliness pathway

Earliness *per se* (*EPS*) is another characteristic of wheat. When a cultivar is not regulated by vernalization or photoperiod for flowering, this cultivar flowers early independently of the vernalization and photoperiod or under any environmental condition (Snape *et al*., 2001; Valárik *et al*., 2006). The flowering of the winter wheat cultivars without treatment with vernalization or photoperiod is due to plant age.

Three genes for qualitative variation in vernalization requirement between spring wheat and winter wheat

Three genes in the vernalization pathway controlling spring wheat development have been cloned, *VRN1* (Yan *et al*. 2003), *VRN2* (Yan *et al*. 2004a), and *VRN3* (Yan *et al*. 2006). The three vernalization genes were cloned, based on clear segregation in three independent populations for vernalization requirement as a qualitative trait between winter wheat and spring wheat.

VRN1 is an orthologue of the *Arabidopsis* meristem identity gene *AP1* (Yan et al. 2003) that encodes a MADS-box protein and is responsible for the initiation of the transition from vegetative to reproductive apices (Mandel et al. 1992). *VRN1* transcripts can be detected in spring wheat grown under natural conditions but not in winter wheat without vernalization. This *VRN1* is further enhanced by long-day photoperiods (Petersen *et al*., 2004). A dominant *Vrn1* allele originated from mutations in the promoter or first intron of a recessive wild type *vrn1* gene in diploid wheat and barley, tetraploid, and hexaploid wheat (Dubcovsky et al. 2006; Fu et al. 2005; Yan et al. 2004a; Pidal et al. 2009).

VRN2 is a transcription factor containing a conserved region encoding a CCT domain in its protein (Yan et al. 2004a). A recessive *vrn2* allele was caused by a point mutation resulting in an alteration of an amino acid at the conserved CCT domain of the VRN2 protein or complete deletion of this gene in diploid wheat (Yan et al. 2004b; Distelfeld et al 2009a).

VRN3 is an orthologue of the *Arabidopsis* flowering gene *FT* (Abe et al. 2005; Corbesier et al. 2007; Tamaki et al. 2007; Wigge et al. 2005) and the rice *Hd3a* (Yano et al. 2001) gene, and allelic variation at *VRN3* is related with mutations in its promoter in wheat and its first intron in barley (Yan et al. 2006).

All known vernalization genes are cloned according to qualitative variation in vernalization requirement between spring and winter wheat, but genes controlling quantitative variation for more or less vernalization requirement among winter wheat cultivars remain unknown. It could be reasonable to speculate that genetic factors or mechanisms which account for the qualitative trait cannot be used to explain the quantitative trait for various durations of vernalization requirements among winter wheat cultivars.

Complex and paradoxical effects of altered temperature and photoperiod on development and growth of winter wheat

Genes which confer photoperiod response in wheat were isolated, according to the sequence of the orthologous photoperiod gene *PPD-H1* in barley (Turner et al. 2005). A PCR marker for *PPD-D1* was developed to distinguish between wheat varieties sensitive and insensitive to photoperiod, based on a 2-kb deletion upstream from its coding region that causes mis-expression in cultivars insensitive to photoperiod (Beales et al., 2007). *PPD-D1* had a large effect on heading date in spring wheat cultivars grown at high latitudes (Snape et al., 2001;

Beales et al., 2007), but it was also mapped in the center of a major QTL for segregation of heading date in two winter wheat populations grown at low latitudes (Chen et al. 2010; Wang et al. 2009), demonstrating the extensive utilization of *PPD-D1* regardless of growth habit.

No regulatory site has been determined for *PPD-B1* due to a lack of polymorphism in the genic and promoter regions in spring wheat (Beales et al. 2007). Although *PPD-B1* was mapped in the center of a major QTL for heading date in a winter wheat population, *PPD-B1* showed no allelic variation in sequence (Wang et al. 2009). No allelic variation in *PPD-A1* is known in hexaploid wheat, but two independent deletions of *PPD-A1* are associated with photoperiod insensitivity in tetraploid wheat (Wilhelm et al. 2009)**.**

In addition to the independent role of each of the genes in vernalization and photoperiod pathways, interactions between the genes in the two pathways will cause greater complexity in understanding wheat development. For example, *Vrn1* and *Vrn3* are dominant for spring growth habit, whereas *Vrn2* is dominant for winter growth habit; but the dominance of *Vrn2* is masked in the presence of a dominant *Vrn1* or *Vrn3* gene, due to epistatic interaction between *VRN2* and *VRN1* or *VRN3* (Tranquilli and Dubcovsky, 2000; Takahashi and Yasuda, 1971). *Vrn2* is repressed not only by low temperature (Yan et al. 2004b) but also by photoperiod (Dubcovsky et al. 2006; Trevaskis et al. 2006). *PPD1* genes have interactions with *FT* (=*VRN3*) at the transcriptional level (Turner et al. 2005; Wilhelm et al. 2009). These previous studies were conducted in diploid wheat or barley.

1.3 Genes controlling development of dual purpose wheat

Winter wheat cultivars have a quantitative difference in vernalization requirement duration

Winter wheat varieties are reported to have various vernalization requirement durations to attain a vernalization saturation point, from which further exposure to low temperature will not result in any more acceleration of flowering time or reduction of final leaf number (Berry et al. 1980; Wang et al. 1995a; Wang et al. 1995b). It has been reported that as many as 12 weeks of low temperature was required for winter wheat cultivar Yeoman to attain a vernalization saturation point (Baloch et al. 2003; Berry et al. 1980; Crofts 1989). Low temperature confers typically 2-10°C, which has significant effects, but vernalization rate will decline at temperatures above 11°C and it will be apparently ineffective above 18°C (Brooking 1996). Winter wheat is sub-divided into three types, based on its quantitative requirement for low temperature duration to satisfy vernalization, including weak winter wheat, semi-winter wheat, and strong winter wheat (Berry *et al*. 1980; Crofts 1989).

Dual purpose wheat has a unique development pattern

All of wheat grown in Oklahoma is winter wheat, which is used for dual purpose: biomass at seedling stage is used for cattle to graze and grains from the plants are harvested later. Wheat cultivars utilized in the dual purpose agricultural system are required to have a longer vegetative growth phase, during which more plant biomass is produced for cattle grazing (Chen *et al*. 2009). In addition, proper timing of flowering and maturity are required to ensure that high grain yield is also achieved after grazing. This unique developmental pattern is genetically controlled by multiple genes responding to temperature and photoperiod.

The timing of a developmental stage is important in production for the dual purpose wheat. Based on morphological or agronomic changes, the wheat life cycle can be recorded at several critical stages at development, including seed emergency, stem elongation prior to

jointing, booting, heading, flowering, grain filling, and maturity (Zadoks *et al*., 1974; McMaster, 2009). The dual purpose wheat is required to have a longer phase from germination to stem elongation (GE-STE), but a not too shortened phase from stem elongation to heading time (STE-HD), or from heading time to physiological maturity (HD-PM) (Chen *et al*. 2010).

Three genes have genetic effects for the unique development pattern of dual purpose wheat

Winter wheat development is regulated by genes that are identified by mapping winter wheat population. Three QTLs for variation in the developmental process from stem elongation to physiological maturity were found in a winter wheat population of recombinant inbred lines (RILs) that are generated from two winter wheat cultivars, 'Jagger' and '2174' (Chen *et al*. 2010). Each of the three major QTLs is tightly associated with a known flowering gene, *vrn-A1* on chromosome 5A, *PPD-D1* on chromosome 2D, and *vrn-D3* on chromosome 7D. The effect of *vrn-A1* decreased from stem elongation through heading to physiological maturity, the effect of *PPD-D1* was minor at stem elongation but increased up to heading then decreased at physiological maturity, and the effect of *vrn-D3* was not detected at stem elongation but increased at heading up to physiological maturity. Therefore, *vrn-A1*, *PPD-D1*, and *vrn-D3* had greatest impact on development at stem elongation, heading date, and physiological maturity, respectively.

The allelic variation in *PPD-D1* between the Jagger allele and the 2174 allele relies on a mutation in the promoter region of this gene. Jagger has an allele sensitive to photoperiod for

late development, whereas 2174 has an allele insensitive to photoperiod for early development (Chen *et al*. 2009; 2010; Wang *et al*. 2009).

The genetic and molecular mechanisms which account for the qualitative difference between the two divergent types of wheat are not expected to explain the quantitative variation in development between winter wheat cultivars. For instance, allelic variation in *vrn-D3* in winter wheat does not occur in the promoter or intron regions where the variation is characterized at the DNA level for spring wheat, but it does occur as a mutation in the coding region that alters vrn-D3 protein sequences (Bonnin *et al*. 2008; Chen *et al*. 2009; 2010; Wang *et al*. 2009). This newly discovered functional mechanism can be manipulated to regulate adult plant development of winter wheat.

vrn-A1 also has mechanism similar to that of *vrn-D3*, because the point mutation in the coding region in *vrn-A1* results in an alteration of a critical amino acid in the conserved K-box of the vrn-A1 protein (preliminary results). This mechanism is revealed by cloning the QTL centered on *vrn-A1* for quantitative vernalization duration in winter wheat (Li *et al*., 2013).

Unidentified genetic parts in the development pathways in wheat

The tri-loci model (*vrn-A1*, *PPD-D1*, and *vrn-D3*) of selection for winter wheat development is not conclusive, because the current model does not match up with phenotypes in some of the tested cultivars, such as Fannin that is extremely early and Trego that is extremely late in development, suggesting that more genes are involved in developmental trait of winter wheat. The missing part in the winter wheat development pathway could be found using two

different research strategies: one through genetics method and the other through biochemistry method.

On the genetic part, it is possible that missing part in the winter wheat development pathway is in gaps between the linkage groups mapped in the winter wheat population, since the previous SSR (simple sequence repeat) markers did not sufficiently cover the whole genome. Our lab is using genome-wide SNP (single nucleotide polymorphism) markers to construct higher density genetic maps for the mapping population. It is also possible that some genetic factors controlling winter wheat development were not detected in the Jagger x 2174 population because the two parental lines may have the same allele for those unknown gene loci. A new doubled haploid (DH) population using two winter wheat cultivars, Duster and Billings that have the same allele at each of the three known genes but have a significant difference in developmental processes, has been generated to identify new genes/QTLs in the Duster and Billings DH population.

The missing part in the winter wheat development pathway can also be found by protein and protein interactions. In comparison with research on a functional gene at the DNA level which characterizes transcription and translation of genes, protein studies have received less attention. Protein-protein interaction has been considered as increasingly important for understanding the molecular mechanism of biological process (Pawson and Nash, 2003), such as plant vernalization. Several methods have been proposed for studying the protein-protein interaction. Yeast two-hybrid (Y2H) approaches have been very successful in identifying the protein interaction partners these years. Y2H analysis is a robust method for detecting pair-wise protein-protein interactions in a cellular setting (Parrish *et al*. 2006). The proteins known for winter wheat development can be used as bait to screen the available

Y2H libraries and find interacting proteins. Mapping interacting proteins has emerged as a new technique in application of functional genes. In rice, 116 representative rice kinases and 255 of their interacting proteins have been linked to disease resistance pathways established via large-scale mass spectrometry and Y2H screening (Ding *et al*. 2009). However, no similar information is available for wheat. This project is to establish the development pathways via micro RNA and proteins interact with VRN-A1.

1.4 Pleiotropic effects of genes on vernalziation requirement, drought resistance, and maturity time in wheat

Pleiotropic effects of cloned vernalization genes

In addition to the independent role of each of these known vernalization genes in controlling the difference in vernalization requirement, the vernalization genes were found to have pleiotropic effects on agronomic traits. In a recent study on winter wheat, the *vrn-A1* locus was found associated with variation in the stem elongation in the winter wheat Jagger x 2174 RIL population (Chen et al. 2009), and this locus influenced subsequent timing of heading and physiological maturity when characterized in the field for three years (Chen et al. 2010). Allelic variation in the dominant *Vrn-A1* locus also indicated pleiotropic genetic effects in spring wheat cultivars (Baga et al. 2009; Blake et al. 2009; Distelfeld et al. 2010; Kuchel et al. 2006; Li et al. 2008; Santra et al. 2009; Shimada et al. 2009; Shitsukawa et al. 2007; Zhang et al. 2008), supporting that the *VRN-A1* locus has pleiotropic effects on wheat development.

VRN2 was cloned in diploid wheat, but this gene has not been either mapped or utilized for breeding purposes in hexaploid wheat, probably due to a rare occurrence of spring growth

habit that is caused by a null or nonfunctional allele at all three *VRN-2* genes in a given cultivar in this species. *VRN2* is indeed functional in hexaploid wheat, as confirmed by previous reports that reduction in RNA level of *VRN2* by RNAi accelerated the flowering time of transgenic plants by more than one month (Yan et al. 2004b) and that *VRN2* expression was down-regulated by vernalization in hexaploid winter wheat cv. Jagger (Yan et al. 2004b) and Triple Dirk lines (Loukoianov et al. 2005). Particularly important is that *VRN2* has been found to have pleiotropic effects in plants. A closest orthologue (*Ghd2*) of *VRN2* contributes significantly to adaptation in drought stress environment and to yield productivity in rice (Xue et al. 2008). Heterologous expression of the wheat *VRN2* gene in *Arabidopsis* had a significant effect on adaptability to the environmental stresses including dehydration, salt, heat shock, wounding, and abscissic acid (Diallo et al. 2010). The conserved CCT domain present in VRN2 proteins interacts with HEME ACTIVATOR PROTEIN (HAP), for which transcript levels are regulated by drought and osmotic stress in *Arabidopsis* (Distelfeld et al 2009b). These studies lead to the exciting possibility that *VRN2* may contribute to drought resistance, in addition to its more widely recognized effect on vernalization requirement.

VRN3 was found to have significant effects on heading date and physiological maturity in two winter wheat populations, Jagger x 2174 (Chen et al. 2010) and Intrada x Cimarron (Wang et al. 2009), and extensive wheat cultivars (Bonnin et al. 2008). Allelic variation in *vrn-D3* does not occur in the promoter or intron regions but as a mutation in coding region that alters vrn-D3 protein sequence. This new functional mechanism can be used to regulate adult plant development to avoid heat damage in later stages of grain filling.

Multiple traits are controlled by genetic factors on chromosome 5A

In winter wheat, vernalization and cold hardiness were reported to be correlated with several morphological traits, including prostrate or rosette growth habit (Salmon 1917; Klages 1926; Taylor and Olsen 1976; Zelenski and Remeslo 1977; Robert 1982; McIntosh 1983; Taylor 1983; Roberts and Larson 1985; Roberts 1986; Chaudhry 1986), plant height (Fowler and Gusta 1977; Fowler et al. 1981), and leaf length (Roberts and MacDonald 1984, 1988). Allelic variation in the dominant *Vrn-A1* locus also indicated pleiotropic genetic effects in spring wheat cultivars (Baga et al. 2009; Blake et al. 2009; Distelfeld et al. 2010; Kuchel et al. 2006; Li et al. 2008; Santra et al. 2009; Shimada et al. 2009; Shitsukawa et al. 2007; Zhang et al. 2008). These previous studies supported that the *VRN-A1* locus has different mechanisms in controlling wheat development. It was very recent reported that *VRN-A1* was associated with spike development (Pearce et al. 2013) and froest tolerance (Zhu et al., 2014). The complete deletion of the *VRN1* gene in diploid wheat enables the mutant wheat not to flower forever under any conditions, suggesting that signals from different developmental pathways converge at *VRN1*. It is postulated that *VRN1* would code for production of a temperature-sensitive protein, which undergoes conformational changes with changes in temperature (Robert 1989). In this study, we will identify proteins that interact with VRN1 and test if the interacting proteins have any genetic effects on multiple traits as reported.

CHAPTER II

*Ta*miR1123 ORIGINATED FROM A FAMILY OF MINIATURE INVERTED-REPEAT TRANSPOSABLE ELEMENTS (MITE) INCLUDING ONE INSERTED IN THE Vrn-A1a PROMOTER IN WHEAT*

This study has been published in Plant Sciences, **Ming Yu, Brett F. Carver, Liuling Yan, 2014. 117-123*

Abstract

More than half of spring wheat cultivars have a *Vrn-A1a* allele that has an insertion of a miniature inverted-repeat transposable element (MITE) in its promoter. In this study, we found that the MITE present in the *Vrn-A1a* gene (MITE_VRN) is a nearly perfect palindrome and it can form highly stable hairpin loop when expressed as RNA. MITE_VRN also possessed sequences of a microRNA in *T. aestivum* (*Ta*miR1123). The P^{32} labeled *Ta*miR1123 probe detected two RNA molecules on a small RNA gel blot, one expected for MITE_VRN, and the other expected for *Ta*miR1123. These results demonstrated that MITE_VRN was expressed as RNAs and *Ta*miR1123 was originated from the MITE_VRN family. The isogenic line TDD carrying the dominant *Vrn-A1a*

allele with MITE_VRN showed higher *Ta*miR1123 and *Vrn-A1a* transcript levels than the isogenic line TDE carrying the recessive *vrn-A1a* allele without MITE_VRN.

*Ta*miR1123 was greatly up-regulated by plant age but slightly down-regulated by low temperature and short days. These findings have pointed to a specific mechanism for the origin and evolution of *Ta*miR1123 and additional regulatory mechanisms for plant development governed by *Vrn-A1a* in spring wheat.

2.1 Introduction

Wheat is the most widely grown crop, occupying 17% of all cultivated land and providing approximately 55% of the world's carbohydrates, and serving as a major staple to approximately 35% of the world population (Gill et al., 2004). Wheat cultivars are generally classified into two general types: winter wheat with variable vernalization requirement for a proper flowering time and thus successful grain reproduction and spring wheat without this requirement (Pugsley, 1971; Chouard, 1960; Amasino, 2004). The decoding of vernalization genes can facilitate understanding of the vernalization phenomenon that extensively exists in the plant kingdom.

The cloning of a gene, via a positional cloning approach from hexaploid common wheat (*Triticum aestivum* L, 2*n*=6*x*=42, AABBDD), seemed impossible a decade ago due to the large genome size (16,000 Mb), the complex structure of homoeologous genomes, and the highly repetitive genomic sequences of wheat (80%) (Bennett and Leitch, 1995; SanMiguel et al., 2002; Wicker et al., 2003). Diploid wheat *T. monoccoccum* (2*n*=2*x*=14, A^mA^m) has a single genome and was selected to clone the first vernalization gene *VRN-*A^m*1*, in which its genetic effect segregated according to a one-gene model in a diploid

wheat population (Tranquilli and Dubcovsky, 2000). *VRN-A^{m}1* was cloned (Yan et al., 2003), and it is an orthologue of the meristem identity gene *AP1* encoding a MADS-box protein for the initiation of the transition from vegetative to reproductive apices in *Arabidopsis* (Mandel et al., 1992). Allelic variation in *VRN-A^{m}1* between the dominant *Vrn-A^{m}1* allele for spring wheat and the recessive *vrn-A^{m}1* allele for winter wheat relies on mutations in the promoter in diploid wheat (Yan et al., 2004a; Dubcovsky et al., 2006; Pidal et al., 2009; Fu et al., 2005). Mutations in the promoter region of the wild type *vrn-A^{m}1* in diploid winter wheat are believed to result in multiple spring *Vrn-A^{m}1* alleles varying in lengths of deletions (alleles *Vrn-A^{m}1a, Vrn-A^{m}1b, Vrn-A^{m}1g, Vrn-A^{m}1f*) involved in a so-called CArG-box recognition site (Yan et al., 2004a; Pidal et al., 2009; Golovnina et al., 2010).

The availability of these *VRN-A^{m}1* sequences has facilitated identifying allelic variation in orthologous *VRN1* genes in diverse wheat species, including two homoeologous genes *VRN-A1* and *VRN-B1* in tetraploid wheat *T. turgidum* (2*n*=4*x*=28, AABB) and *VRN-A1* and *VRN-G1* in *T. timopheevii* (2*n*=4*x*=28, AAGG), and three homoeologous genes *VRN-A1*, *VRN-B1* and *VRN-D1* in hexaploid *T. aestivum*. The *Vrn-D1* gene in hexaploid wheat has a single form of mutations in intron one due to the presence of a large deletion (Fu et al., 2005); the *Vrn-B1* gene has the deletion in intron one in tetraploid *T. turgidum* ssp. *durum* and hexaploid wheat (Fu et al., 2005; Golovnina et al., 2010; Yan et al., 2004b). In a recent study, it was found that the *Vrn-B1* gene has a 5.6 kb retrotransposable element (Retrotrans_VRN) in the 5'-untranslated region (UTR) in tetraploid wheat, which is prevalent among *T. turgidum* subsp. *carthlicum* (Chu et al., 2011). In addition to small deletions in the promoter (*Vrn-A1d* and *Vrn-A1e*) observed in tetraploid wheat or a large

deletion in intron one (*Vrn-A1c*) observed in hexaploid wheat, the *Vrn-A1* gene has a miniature inverted-repeat transposable element (MITE) that is inserted in its promoter (*Vrn-A1a*) (MITE_VRN) in more than half of all hexaploid wheat varieties (Yan et al., 2004b).

Without vernalization, the winter recessive *vrn-A1a* allele was not expressed and plant flowering time was delayed, whereas the spring dominant *Vrn-A1a* allele was expressed and the plant flowered (Yan et al., 2003; Danyluk et al., 2003; Murai et al., 2003; Trevaskis et al., 2003). A model was proposed to explain mechanisms underlying growth habit by *VRN-A1a*. In this model, without vernalization the winter *vrn-A1a* allele cannot be expressed due to the presence of repressors in the binding site in the *vrn-A1a* promoter, and the repressors can be removed by vernalization. Further studies have indicated that VRT2 (Kane et al., 2005) or TaFD1 (Li and Dubcovsky, 2008) can bind to the *vrn-A1a* promoter. On the other hand, the spring *Vrn-A1a* allele is expressed without vernalization requirement, because the spring *Vrn-A1a* promoters are impaired due to the insertion of the MITE or other retrotransposal elements or deletions; therefore, the flowering repressor cannot bind to the *Vrn-A1a* promoters. However, the previous model cannot explain why those plants that carry MITE_VRN produce more *Vrn-A1a* transcripts and flowered earlier than those plants that carry other mutant alleles with deletions in their promoters or intron 1 (Yan et al., 2004b; Loukoianov et al., 2003).

In our further analysis on MITE_VRN, we found that MITE_VRN possessed sequences of a microRNA in *T. aestivum* (*Ta*miR1123). This finding has encouraged us to test if there are additional regulatory mechanisms through *Ta*miR1123 in spring wheat.

2.2 Materials and Methods

Plant materials

We investigated the spring wheat near-isogenic Triple Dirk D line (TDD), which has a dominant *Vrn-A1a* allele with MITE_VRN but a recessive allele at each of *vrn-B1* and *vrn-D1*. As a control, we used the near-isogenic Triple Dirk E line (TDE), which has a recessive *vrn-A1a* allele without MITE_VRN but is also a spring type. The spring growth habit was determined by the dominant *Vrn-D1* gene (Fu et al., 2005). These near-isogenic lines have provided a useful tool to study the effect of different *VRN1* genes in hexaploid wheat without the confounding effect of other genes affecting flowering time (Pugsley, 1971).

The TDD and TDE lines were initially grown in a greenhouse at 20-25°C and with a long day photoperiod (LD, 16/8 hours light/dark). At the 3rd-leaf stage, the first set of the two lines were moved into a cold room with 4°C and the same LD photoperiod, and the second set of the two lines were moved into a growth chamber with 20-25°C but with a short day photoperiod (SD, 8/16 hours light/dark). After 5 days, these temperature-photoperiod treated plants and the 3rd set plants that were continuously kept in the greenhouse were collected for leaf samples for analyses of small RNA blot and *VRN-A1a* transcript levels.

RNA gel blot analysis with a P^{32} labeled *Ta*mi1123 probe

Small RNAs were isolated from leaves of the TDD and TDE plants growing under different conditions. Total RNA was extracted using Trizol regents (Invitrogen). RNA samples were size-fractioned on a 15% denaturing polyacrylamide gel and then

electrophoretically transferred to a Hybond-N^+ membrane (Amersham Biosciences). The membrane was UV cross-linked to fix RNA on the membrane. The small RNA membrane was hybridized with two probes. One probe was *Ta*miR1123 (5'-TCCGTGAGACCTGGTCTCATAGA-3') that has the same sequence as observed in MITE_VRN. The other probe was U6 (5'-TCATCCTTGCGCAGGGGCCA-3') that was used as a control to ensure that the membrane had small RNA molecules. The hybridization was performed using the protocol described previously (Sunkar, 2008). Briefly, a DNA fragment was end-labeled with γ-^{32}P-ATP using T_4 polynucleotide kinase (New England Biolabs), which was used to test if the probe has sequence complementary to small RNA or micro RNA molecules that were size-fractionated on the small RNA blot. The blot membrane was pre-hybridized in Perfect-Hyb Plus buff (Sigma) for 2 hours, and the ^{32}P probe was then added to hybridize with the membrane at 38°C for 16 hours. The hybridized membrane was washed with 2X SSC buffer plus 1% SDS for three times at 50°C, 10 minutes for each. The filter was exposed for 24 hours and images were taken using a scanner (Typhoon).

ImageJ 1.32 software (National institutions of Health, Bethesda, MD. http://rsb.info.nih.gov/ij) was used to quantify the density of different signals on blots. The signal of MITE_VRN or *Ta*mR1123 was compared between different samples by converting to the ratio of its own signal density over the density of its respective U6 signal density.

VRN-A1a expression pattern

The RNA samples from the TDD and TDE plants were treated with Deoxyribonuclease I and first-strand cDNA was synthesized using a SuperScript™ II Reverse Transcriptase kit (Invitrogen). Primers vrn-A1-Exp-F1 (5'-GAATAAAGTTCTCCAGAAGGAACTCGTG-3') and vrn-A1-Exp-R2 (5'-GCATGAAGGAAGAAGATGAAGAGCTG-3') that are specific to *Vrn-A1a* (Loukoianov et al., 2005) were used to determine its transcript levels in leaves of plants. Primers actin-F1 (5'-ATGGAAGCTGCTGGAATCCAT-3') and actin-R1 (5'-CCTTGCTCATACGGTCAGCAATAC-3') were used to amplify transcripts of *actin* as endogenous control. A quantitative RT-PCR was performed using a 7500 Real-Time PCR System (Applied Biosystems) and iQ™ SYBR Green Supermix kit (BIO-RAD) and the Applied Biosystems 7500 Real-Time PCR Systems.

Sequence data analyses

It was reported that the TDD line has a duplicated copy including the promoter, exon 1 and partial intron 1 (Yan et al., 2004b), but it was not known which copy is original or duplicated. The only difference observed between the two copies is the size of a MITE_VRN in the promoters. The MITE-VRN structure of each of two *Vrn-A1a* copies was predicted using DNA folding form program at http://mfold.rit.albany.edu/. The MITE-VRN sequences were used to search the miRNA database at http://www.mirbase.org.

The MITE_VRN sequences were used to search the wheat genomic sequence database at http://www.cerealsdb.uk.net to determine its copy number. The MITE_VRN flanking

sequences of targeted fragments were searched in GenBank for potential genes. The *Ta*miR1123 sequences found in the sequences were also used to search in GenBank expressed sequence tags (EST) databases for potential target genes.

2.3 Results

The insertion of transposable elements in multiple *Vrn-A1* genes

The previously reported foldback element (MITE_VRN) with a target site duplication (TSD= TTAAAAACC) in the dominant *Vrn-A1a* allele in the TDD line was inserted in a CG-rich region, where there are 14 G/C in 16 bp (CCTCCCCCCCTGCCGG) at the 3' downstream of the TSD element (Fig.1C). A member of the MITE_VRN family was found to insert in Vrn-G1, with a TSD (CTCCGCCC), where there are 9 G/C in 11 bp (TCCCCTCCCCG) at the 3' downstream of the TSD element (Fig.1D). Interestingly, a 5.6 kb retrotransposal element (Retrotrans_VRN) (Chu et al., 2011)was inserted in *Vrn-B1* at the exactly same site as *Vrn-G1*, though *Vrn-B1* has a different TSD sequence (CTCCG) (Fig.1E).

These observations indicated that different MITE and TE inserted in *Vrn-A1* genes in the upstream region from the start codon ATG, where it is GC-rich. This is characteristic of a MITE or TE insertion preference (Ferguson et al., 2011). The difference in the insertion site between *Vrn-A1a* and the other two *Vrn-1* genes is that MITE_VRN was placed on the upstream side of the transcriptional site (or the promoter region) in *Vrn-A1a* but MITE_VRN or Retrotrans_VRN was located on the downstream side of the transcriptional site or the 5'-UTR in the other two *Vrn-1* genes.

A nearly perfect palindrome of the MITE_VRN containing *Ta*miR1123

A previous study reported that there were two copies of *Vrn-A1a* in the TDD line; one has the insertion of a 222-bp MITE and the other has a 131-bp MITE (Yan et al., 2004b). The two MITE_VRN sequences have the same inverted repeat sequences (GGAAAAAATT) but different lengths between the flanking sequences, suggesting that they were originated from a duplication event. Using DNA folding form program, it was predicted that the 222-bp MITE_VRN possessed a nearly perfect palindrome (Fig.1B), whereas the shorter 131-bp MITE cannot form a perfect palindrome due to the absence of a 91-bp deletion (data not shown). It is likely that the 91-bp section in the duplicated *Vrn-A1a* copy lost during or after the gene duplication. Only the original *Vrn-A1a* containing MITE_VRN was further analyzed in this study.

The search of MITE_VRN sequences in the miRNA database at http://www.mirbase.org found that MITE_VRN possessed a 23 bp sequence, TCCGTGAGACCTGGTCTCATAGA (Fig.1A), which was complementary to *Ta*miR1123 (Fig.1A). The *Ta*miRNA1123 was a microRNA that was found expressed in wheat, but it was mistakenly named as miR507 (Yao et al., 2007). There are 4 mismatches between *Ta*miR1123 and the element (TATATGAGACCAGGTCTCATATA) (Fig.1B). These observations suggested that if MITE_VRN with highly stable hairpin loops was expressed as RNA molecules, it could be recognized by RNA interference enzymatic machinery and processed to form mature *Ta*miR1123.

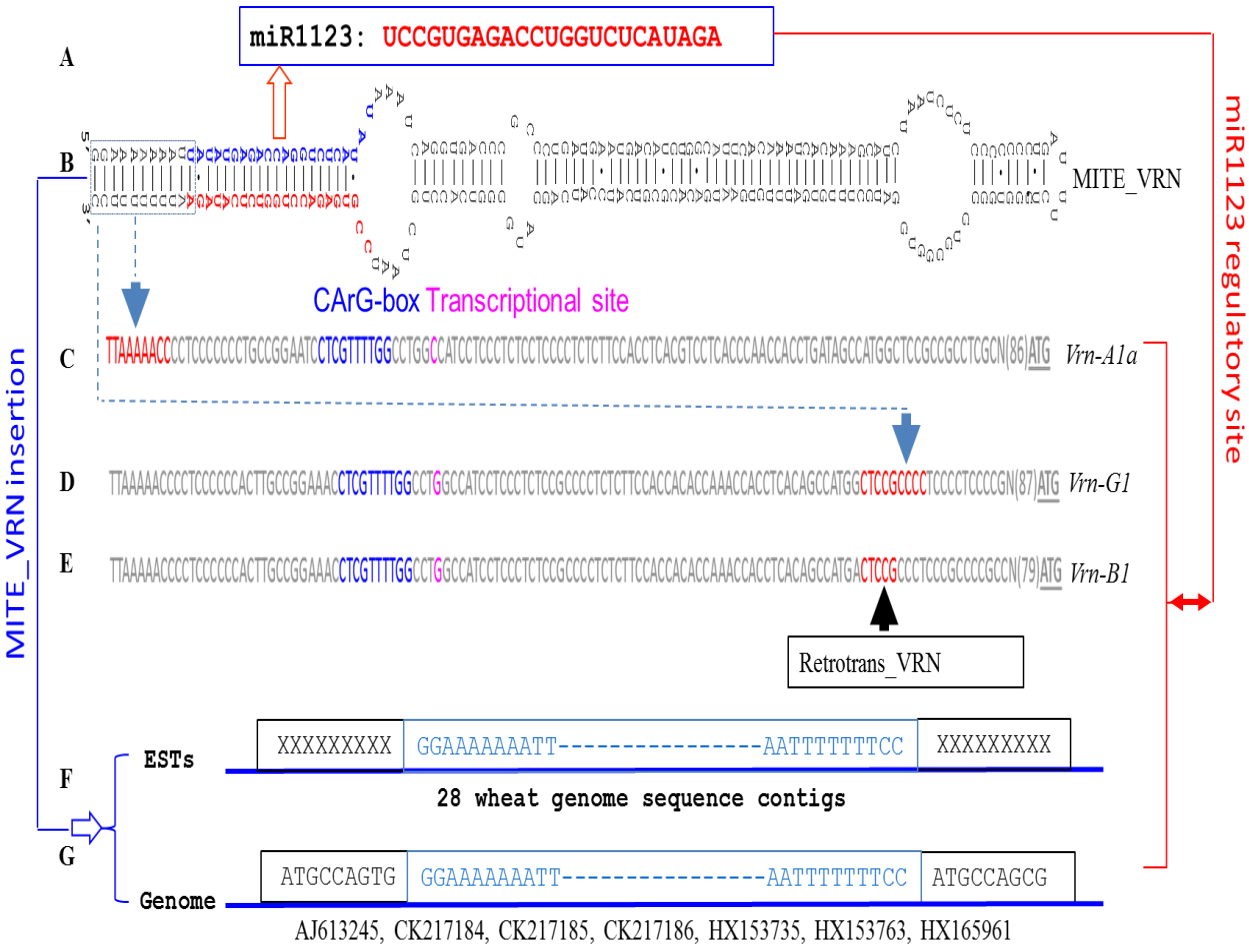


Fig.1 MITE_VRN in wheat genomes. (A) *Ta*miR1123 is derived from MITE_VRN. (B) MITE_VRN forms a stem loop structure. (C) MITE_VRN is inserted in *Vrn-A1a* at TTAAAAACC. (D) MITE_VRN is inserted in *Vrn-G1* at CTCCGCCC. (E) Retrotrans_Vrn in Vrn-B1 is inserted at CTCCG in the same site as MITE_VRN in *Vrn-G1*. (C)-(E), CArG-box and transcriptional site in Vrn-A1a is indicated to show that MITE_VRN is in the Vrn-A1a promoter region, whereas MITE_VRN in Vrn-G1 and Retrotrans_VRN are in the 5'-UTR. (F**)** MITE_VRN is in 28 wheat genome sequence contigs with various host direct duplication. (G) MITE_VRN is in 8 EST sequences with the same host direct duplication. A work model for the loop between MITE_VRN, miR1123, and targeted genes/genomic sites is diagrammed with lines.

The existence of MITE_VRN and *Ta*miR1123 in small RNAs

To confirm that MITE_VRN was expressed as RNAs, a ^{32}P labeled *Ta*miR1123 probe was used to analyze a small RNA blot that was generated from the leaves of the TDD and TDE plants under different conditions. The ^{32}P labeled small RNA U6 probe was used as a control to ensure that small RNAs from each samples were transferred on blot membranes. As expected, the ^{32}P probes of both *Ta*miR1123 and U6 detected positive hybridization signals on the small RNA blot.

As shown in Fig.2A, the U6 probe detected RNAs as a single band in all samples. When the same membrane hybridized with the U6 probe was washed and then hybridized with the *Ta*miR1123 probe, three bands appeared in each sample as shown in Fig.2B. The highest band was from U6 signal rescues, and the middle band and the lowest band were two new signals that should be from the *Ta*miR1123 probe. When a brand new membrane was hybridized with the *Ta*miR1123 probe, only two bands with the same sized as shown for the middle band and the lowest band were observed (data not shown), confirming that the *Ta*miR1123 probe detected the two small RNA molecules.

The middle band size (Fig.2B) was expected for MITE_VRN, compared with the U6 band (250-300 bp). The lowest band was expected for *Ta*miR1123. These results demonstrated that both MITE_VRN and *Ta*miR1123 existed in the RNA samples. The appearance of the positive and discrete small RNA molecules on the small RNA blot indicated that they were not degradation products of RNA but from MITE_VRN of a hairpin structure. The expressed MITE_VRN was also processed to mature *Ta*miR1123.

Therefore, the member of MITE_VRN family might be one of the precursors of *Ta*miR1123.

The regulation of MITE_VRN and *Ta*miR1123 by internal and external factors

Both of the isogenic TDD line (lanes 1 and 3-5, Fig.2B) and the isogenic TDE line (lanes 2 and 6-8, Fig.2B) were detected to have the middle hybridization signal representing MITE_VRN and the lower signal representing *Ta*miR1123 molecules. Using the ImageJ software, the amount of the MITE_VRN and *Ta*miR1123 molecules in each sample was converted to percentage of U6 hybridization signals in the same sample. U6 was used as an endogenous control.

The MITE_VRN in the TDD line (graph 1, Fig.2C) was not higher but slightly lower than that in the TDE line (graph 2, Fig.2C). It was possible that the difference in MITE_VRN transcription between the TDD line and the TDE line was masked due to the presence of multiple MITE copies in wheat, as shown in later results. In both TDD (graphs 3-5, Fig.2C) and TDE (graphs 6-8, Fig.2C), MITE_VRN signals were decreased with plant age and treatments with low temperature and short days.

A significant difference in the *Ta*miR1123 molecules was detected between the two isogenic lines. The TDD line (graph 1, Fig.2D) showed much more *Ta*miR1123 molecules than the TDE line (graph 2, Fig.2D). When the TDD line was continuously kept in a greenhouse with long day and room temperature, the plants were observed to have more *Ta*miR1123 molecules after 5 days (graph 3, Fig.2D). When the plant was treated with low temperature (graphs 4, Fig.2D) or short day (graph 5, Fig.2B),

*Ta*miR1123 molecules were significantly increased. The *Ta*miR1123 at TDE line showed a low level under these conditions (graphs 6-8, Fig.2D).

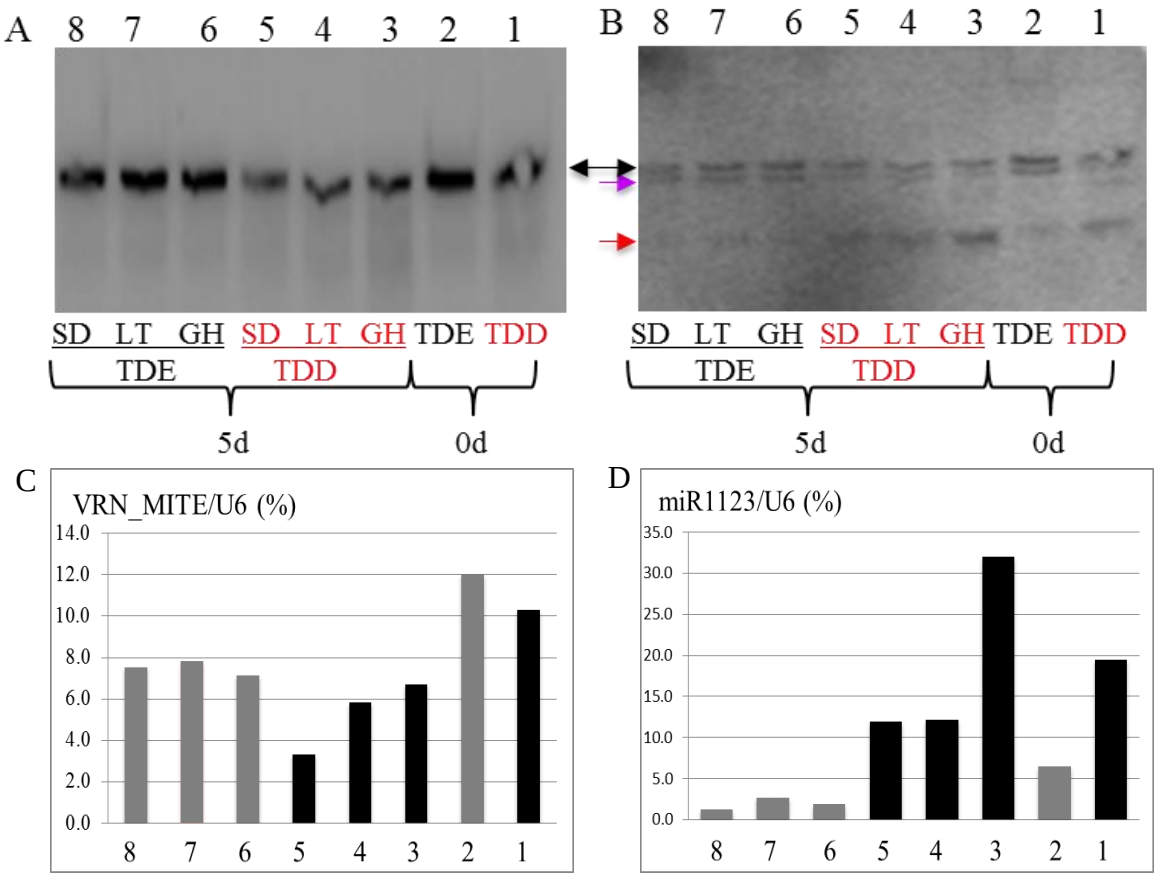


Fig.2 Expression of MITE_VRN and *Ta*miR1123. (A) Expression patterns of micro RNA U6. (B) Expression patterns of *Ta*miR1123. RNA gel blots of small RNAs from isogenic TDD line with MITE_VRN at *Vrn-A1a* and the TDE line without MITE_VRN at *Vrn-A1a***.** The blot was probed with labeled oligonucleotides. The small RNA samples were collected from leaves of the plants at the beginning of experiments (CK) and the plants grown at greenhouse (GH) or treated with low temperature (LT) or short days (SD) for 5 days. (C) Expression amount of MITE_VRN relative to U6. (D) Expression amount of *Ta*miR1123 relative to U6.

Association of *Ta*miR1123 molecules with *Vrn-A1a* transcript levels

The *VRN-A1a* transcript levels in the TDD and TDE lines were determined using a quantitative RT-PCR (Fig.3). The expression of dominant *Vrn-A1a* in the TDD line was detected in seedlings grown in the greenhouse (TDD-CK). The *Vrn-A1a* transcript level was 31.7 before the plants were treated with different conditions, and it was increased to 52.5 after 5 days (TDD-GH). *Vrn-A1a* transcript level was also increased in plants treated with low temperature (TDD-LT) or short day (TDD-SD) compared with TDD-CK but slightly decreased compared with the plants that were continuously kept in the greenhouse (TDD-GH). The *vrn-A1a* transcript level in the TDE seedlings grown in the greenhouse (TDE-CK) was 4.8, which was 40% in the TDE seedlings grown under the same condition after 5 days (Fig.3). A similar result was shown in the TDE line treated with low temperature and short days as the TDD line. Overall, *Vrn-A1a* transcripts were 5.8 folds as *vrn-A1a* transcripts in the isogenic lines. Under any condition, the *Vrn-A1a* transcript level in the TDD line was significantly higher than the *vrn-A1a* transcript level in the TDE line.

Both *Vrn-A1a* and *Ta*miR1123 transcripts were regulated positively by plant age and negatively by low temperature and short days. The concomitant expression of *Vrn-A1a* in the normal RNA samples with *Ta*miR1123 molecules in the small RNA samples suggested that *Vrn-A1a* could be induced by *Ta*miR1123.

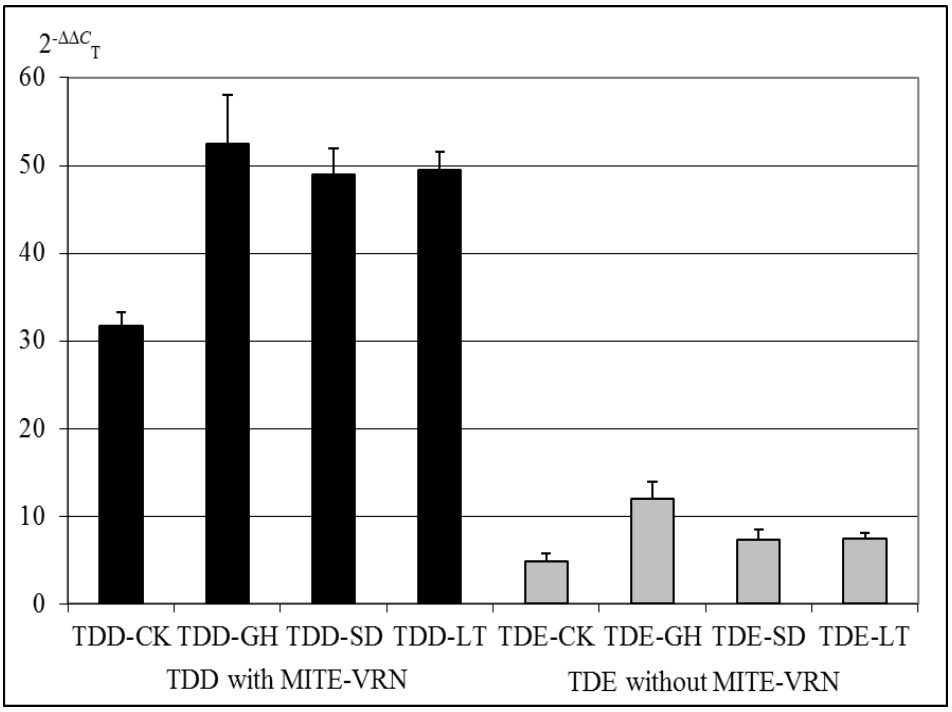


Fig.3 Expression profiles of *Vrn-A11*. Transcript levels of *Vrn-A1a* in TDD and *vrn-A11a in* TDE are shown using the values calculated by the $2^{(-\Delta\Delta CT)}$ method, where CT is the threshold cycle, and *actin* was used as an endogenous control. RNA samples were collected were collected from leaves of the plants at the beginning of experiments (CK) and the plants grown at greenhouse (GH) or treated with low temperature (LT) or short days (SD) for 5 days plants. The values represent mean expression levels (n=12), and the bar indicates standard error.

The copy number of MITE_VRN in hexaploid wheat

The availability of 5-fold coverage of the whole-wheat genome (http://www.cerealsdb.uk.net) has facilitated identification of the copy number of MITE_VRN in the entire wheat genome. A total of 124 wheat genome contigs (as of March 30, 2013) were hit in 'Draft assembly of gene-rich regions' (E value<1.0e-05). However, only 28 hits were found to have similar sequences in the complete MITE_VRN

region (Fig.1F), and the remaining hits included no or one end of the MITE_VRN. Further analyses showed that each member of the MITE_VRN family was inserted with a different host direct duplication of 3 or 7 bp sequences (Table 1). It was predicted to have approximately 375 copies of the MITE_VRN present in the tetraploid wheat genome based on the clone number in the Langdon BAC library (Yan et al., 2004b). The dramatic difference in the MITE-VRN copy number between hexaploid wheat and tetraploid wheat was because some positive BACs hit by the MITE_VRN probe should be false or not belong to the MITE_VRN family. In this study, only the genome sequence contigs that have intact MITE_VRN sequences were analyzed.

Target genes of MITE_VRN and *Ta*miR1123

The search of MITE_VRN sequences in the wheat EST databases deposited in GenBank with BLAST algorithms found a total of 16 EST accessions (E<0.01). After removing the EST sequences with more than 2 mismatches, seven EST accessions were found to have either MITE_VRN or miR1123 or both in sequence and structure., including three from cultivar 'Norstar' (CK217184, CK217185, and CK217186), two from cultivar 'Halberd' HX153735 and HX153763), and the other two from 'Ofanto' (AJ613245 and 'Cranbrook' (HX165961). These ESTs were inserted with the same host direct repeat (ATGCCAGTG) (Fig.1G). The EST MITE sequences in various cultivars showed more than 85% identity to each other (Fig.4). The conservations of insertion sites of MITEs in various cultivars and their sequence identity to MITE-VRN suggested that these expressed ESTs belong to the same MITE_VRN family.

*Ta*miR1123 sequence alone was searched in the GenBank wheat EST databases. Expect for the seven ESTs that had MITE_VRN and thus *Ta*miR1123, no new wheat EST was found to have identical sequences to *Ta*miR1123 only.

Fig.4 Multiple sequence alignment of MITE_VRN. EST sequences are derived from GenBank. Wheat genome sequence contigs are derived at http://www.cerealsdb.uk.net. Except for host direct duplication, the sequences flanking MITE_VRN are included in the alignment.

2.4 Discussion

Many miRNA molecules have been found to exist in different plant species and they may play important roles in plant responses to abiotic and biotic stresses as well as signal transduction (Chen, 2010). Tens of miRNA genes have been identified in wheat (Yao et al., 2007). However, the origin and evolution of these non-coding regulatory sequences remain largely unknown. No miRNA has been characterized with a functional gene in wheat.

In this study, we demonstrated that *Ta*miR1123 was derived from a MITE that was inserted in the promoter of a dominant *Vrn-A1a* gene. First, MITE_VRN was nearly perfect palindromes. When expressed as RNA it can form highly stable hairpin loops. This structure can be processed to form mature *Ta*miR1123 sequences if MITE_VRN was recognized by the RNA interference enzymatic machinery as described in plants or animals (Chen, 2010; Ambros, 2004; Bartel et al., 2004). It is possible that the MITE_VRN hairpin stem is cleaved and then degraded to produce a *Ta*miR1123. Second, the ^{32}P labeled *Ta*miR1123 detected RNAs in the same size as expected for MITE_RNA, indicating that the small RNA contained *Ta*miR1123. The probe also detected *Ta*miR1123 molecules, provided experimental evidence that *Ta*miR1123 could be released from MITE_RNA. The association of *Vrn-A1a* and *Ta*miR1123 at the transcript level and the concomitant regulation of their expression by plant age also supported that *Ta*miR1123 was derived from MITE_VRN in the TDD line. The MITE_VRN in the *TaVRN-A1* promoter could be the direct target of TamiR1123 because they have complementary sequences. However, the MITE_VRN in the VRN-A1 promoter could be one of the targeted sites of *Ta*miR1123, since an average miRNA has approximately 100

target sites (Brennecke et al., 2005). No other target site was found in the *TaVRN-A1* gene. The *Ta*miR1123 sequence was identified in the previous study (Yao et al., 2007), but its function is unknown. Our results clearly showed that *Ta*miR1123 was increased with plant age, which was concomitant with *Vrn-A1a* transcript levels during development. In this study, we provided a depth understanding of the MITE inserted in the promoter of the dominant *Vrn-A1* allele in hexpaloid wheat that was previously described (Yan et al., 2004b). We provided experimental evidence that the level of *Ta*miR1123 was positively correlated with of *Vrn-A1a*; therefore, it is possible that *Vrn-A1a* could be induced by TamiR1123. However, it is also possible that *Vrn-A1a* was not induced directly by *Ta*miR1123 and *Ta*miR1123 could be the by-product of transcription of *Vrn-A1a*. Based on characteristics of the sequence and target site duplication of MITE_VRN, it is a Mutator-like element (MULE). A MULE can harbor the promoters of a gene for transcription (Ferguson and Jiang, 2012; Lisch, 2002), and the function of such a MULE was reported in the promoter of the hcf106-mum1 gene (Mutator transposons (Das and Martienssen, 1995). The potential mechanism could explain why *Vrn-A1a* is linked to a much stronger expression than other *Vrn-A1* alleles. The two mechanisms presented in this study are different from the previous hypothesis that the insertion of MITE_VRN in the promoter of *Vrn-A1a* resulted in a damage of recognition site by a flowering repressor (Yan et al., 2003; Yan et al., 2004b). The recognition of the promoter of *Vrn-A1a* by a *Ta*miR1123 resulted in *Vrn-A1a* expression without vernalization requirement. This study provided additional mechanisms for developmental regulation by MITE_VRN through *Ta*miR1123 in spring wheat.

MITE_VRN in *Vrn-A1a* was expressed as RNA, which released *Ta*miR1123 that in return induced the expression by a direct or indirect manner. This formed a loop that regulated the expression of *Vrn-A1a* (Fig.1). This loop could involve in plant development and other phenotypes through both MITE_VRN and *Ta*miR1123. On the one hand, MITE_VRN is active and movable, which can insert in a functional gene or a regulatory site like *Vrn-A1a* promoter. On the other hand, *Ta*miR1123 has many regulatory sites throughout the genome. The expression of these targeted genes could be regulated due to the insertion of MITE_VRN or recognition of their regulatory sites by *Ta*miR1123, which would form a dynamic gene regulatory network governed by *Vrn-A1a* in plant development. The *VRN1* region was reported to have association with multiple traits including vernalization, cold hardening, and the development of rosette (Roberts, 1990). This study provided a machinery explanation for the complex association among some of the multiple phenotypes.

The MITE_VRN derived *Ta*miR1123 characteristics can be applied in wheat breeding. Firstly, the genomic and EST sequences flanking MITE_VRN can be used to design specific primers to map members of the MITE_VRN family. The characteristics that such a MITE has a few hundred nucleotides will facilitate development of a PCR marker for mapping the MITE_VRN dispersed throughout the genome. Any phenotypic variation that shows association with a MITE_VRN marker can be suggested to link with development controlled by *VRN-A1a* or *Ta*miR1123 donated from MITE_VRN in the *VRN-A1a*. These potential applications of MITE_VRN and *Ta*miR1123 in wheat need to be investigated in the future studies.

Acknowledgements

This project was supported by the Oklahoma Agricultural Experiment Station and the Oklahoma Wheat Research Foundation.

CHAPTER III

*TA*VRN1 AND *TA*HOX1 PROTEINS INTERACTIVELY REGULATE VERNALIZATION REQUIREMENT DURATION*

** The study has been published in Plant Journal, Genqiao[1], **Ming Yu**[1], Tilin Fang[1], Brett F. Carver, Liuling Yan, 2014. 76:742-753 ([1] these authors contributed equally to the project). Only the research work that was mainly conducted by Ming Yu is presented herein.*

Abstract

Winter wheat requires a period of low temperatures to accelerate flowering (vernalization). This requirement could make winter wheat more vulnerable to elevated global temperature via insufficient vernalization. All known vernalization genes are cloned according to qualitative variation in vernalization requirement between spring and winter wheat, but genes controlling quantitative variation for more or less vernalization requirement among winter wheat cultivars remain unknown. A major QTL on chromosome 5A was found associated with quantitative vernalization requirement duration (*QVrd.osu-5A*) in the winter wheat population generated from Jagger and 2174, this QTL was cloned by using a positional cloning approach, and the allelic variation at *vrn-A1* at the protein level was finally found to be responsible for *QVrd.osu-5A* (Li and Yu et al., 2013). The Ala^{180} in vrn-A1a encoded by the dominant allele for 3-week

vernalization was mutated to Val^{180} in vrn-A1b encoded by the recessive allele for 6-week vernalization. Further studies indicated that the mutated Val^{180} in vrn-A1b protein decreased the ability to bind with *Ta*HOX1 (the first homeobox protein in *T. aestivum*) by *in vitro* protein pulldown assays and immunoprecipitation analyses. The direct binding of *Ta*VRN-A1 and *Ta*HOX1 proteins was confirmed in the nucleus of living plant cells by bimolecular fluorescence complementation (BiFC) analyses. The *TaHOX1* gene was found to be up-regulated by low temperature and have significant genetic effect on heading date, suggesting that *Ta*HOX1 functions in the flowering pathway in winter wheat.

3.1 Introduction

Wheat (*Triticum aestivum*, 2n=6x=42, AABBDD) is cultivated across more land area than any other grain crop. Wheat cultivars are qualitively classified as two general types: winter wheat with variable low temperature requirement for a proper flowering time (vernalization) and thus successful grain reproduction and spring wheat without this requirement (Amasino, 2004; Chouard, 1960; Pugsley, 1971). Winter wheat cultivars are quantitatively classified as three types according to the low temperature duration required to reach a vernalization saturation point or achieve the maximum vernalization effect: a weak winter type that is stimulated to flower by brief exposure to low temperature (less than two weeks), a semi-winter type that requires 2 to 4 weeks of cold exposure for flowering, and a strong winter type that requires more than 4 weeks of cold exposure for timely flowering (Crofts, 1989).

Vernalization usually occurs at temperatures less than 8°C (Amasino, 2004; Chouard, 1960; Crofts, 1989; Pugsley, 1971). Recent studies showed that average global surface air temperature rose 0.5 °C in the 20^{th} century and is projected to continue its increase by

roughly 3°C or 5 °C by the end of the 21st century (Kerr, 2007; Semenov and Halford, 2009). As various simulation models have shown, winter wheat is more vulnerable to increasing temperatures during winter seasons due to its low temperature requirement to ensure proper flowering time and successful seed reproduction (Humphreys et al., 2006; Miglietta et al., 1995). A shortened duration at low temperature, due to global warming, could result in a failed or insufficient vernalization in winter wheat.

Three vernalization genes have been cloned from wheat by using a positional cloning approach, *VRN1* (Yan et al., 2003), *VRN2* (Yan et al., 2004a), and *VRN3* (Yan et al., 2006), and each of them was cloned in the context of growth habit as a discrete trait of winter and spring types in diploid wheat and barley. *VRN1* (=*AP1*) is a central promoter for spring cultivars to flower without vernalization (Danyluk et al., 2003; Murai et al., 2003; Trevaskis et al., 2003; Yan et al., 2003), and dominant *Vrn-A1* alleles originated from mutations in the promoter or the first intron of the wild recessive *vrn-A1* in hexaploid wheat (Yan et al., 2004b; Fu et al., 2005). A recent study reported that the increased copy number of *TaVRN-A1* resulted in an increased requirement for vernalization in winter wheat (Diaz et al 2012). *VRN2* (=*ZCCT1*) is a flowering repressor, and a recessive *vrn2* allele was caused by a point mutation at the conserved CCT domain or complete deletion of the wild dominant *Vrn2* in diploid wheat (Yan et al., 2004). *VRN3* (=*FT1*) is another flowering promoter, and the early flowering plants carry *Vrn3* alleles originating from mutations in the promoter or the first intron in wheat and barley (Yan et al., 2006). The genetic and molecular mechanisms to account for the qualitative difference in the three genes between the two divergent types of wheat are not expected to explain the quantitative variation in vernalization requirement duration among winter

wheat cultivars, because all winter wheat cultivars are supposed to have the same winter allele for each of the three cloned vernalization genes, recessive *vrn1*, recessive *vrn3*, and dominant *Vrn2* alleles (Pugsley, 1971; Tranquilli and Dubcovsky, 2000). No exception was reported to invalidate this genetic model, when numerous cultivars/germplasm from different wheat species varying in ploidy level were screened using molecular markers for each of the three genes (Yan et al., 2003; Yan et al., 2004a; Yan et al., 2004b; Fu et al., 2005; Bonnin et al., 2008; Zhang et al., 2008; Santra et al., 2009; Chen et al., 2010).

The molecular mechanism underlying quantitative vernalization requirement in a winter type of *Arabidopsis* is unknown yet, but could be explained by three models. The first model is that vernalization results in a quantitative reduction in *FLC* mRNA levels, which negatively correlates with flowering time (Michaels and Amasino, 1999; Sheldon et al., 1999). The second model is that the transcript level of *FLC* is up-regulated or down-regulated by multiple genes including *VIN3* and *FRI* in the vernalization pathway, as well as *GI*, *CO*, and *FT* in the photoperiod pathway (Corbesier et al., 2007; Heo and Sung 2011; Levy et al., 2002; Reeves and Coupland, 2001). The last model is that quantitative vernalization requirement is modulated by accumulation of the Polycomb-based epigenetic-silencing complexes and histone modifications at the *FLC* gene (Angel et al., 2011) and the quantitative modulation of Polycomb silencing is associated with natural variation in the sequence of *FLC* (Coustham et al., 2012). The signals from these pathways are integrated at *FLC* to induce *AP1* for flowering. Previous studies have demonstrated that vernalization has evolved different mechanisms between winter wheat and *Arabidopsis* (Yan et al., 2003). The gene(s) responsible for various vernalization

requirement durations in winter wheat should be cloned from a population generated using two winter wheat cultivars having different vernalization saturation points.

In previous studies, we developed a population of recombinant inbred lines (RILs) generated from a cross between two winter wheat cultivars, 'Jagger' and '2174'. When different sets of the same population were tested in field, the RIL population was segregated in stem elongation, heading date, and physiological maturity (Chen et al., 2009). In this previous study, a PCR marker developed for allelic variation in exon 4 of *TaVRN-A1* was found to have genetic association with a major QTL for the phenotypes, but it was not known whether *TaVRN-A1* caused the QTL in winter wheat. If *TaVRN-A1* was indeed responsible for this QTL, it could have different mechanisms in regulating the developmental process in winter wheat (as opposed to spring wheat). Otherwise, a novel gene at the *TaVRN-A1* locus should be responsible for the QTL regulating developmental process in winter wheat. In the present study, we tested this RIL population under thermal-photo-controlled greenhouse conditions and found a major QTL for vernalization requirement duration, and we then generated a large backcross population and cloned the first gene for vernalization requirement duration in winter wheat.

3.2 Materials and Methods

Plant materials and vernalization experiments

Jagger and 2174 were initially grown in a greenhouse at 20-25°C and with a long day photoperiod (LD, 16/8 hours light/dark). At the 5^{th}-leaf stage, the parental plants were moved into a cold room with 4°C and the LD photoperiod. The LD photoperiod was applied throughout this study to avoid disruption of photoperiod effects on vernalization

(Wang et al., 2009). The two parental lines were vernalized for varying weeks. The vernalized plants were returned to the greenhouse, whereas non-vernalized plants that continuously remained in the greenhouse were used as controls. Heading date for 8 plants of each treatment was scored.

Three populations from the same Jagger x 2174 RILs were initially grown in the same greenhouse as for the parental lines. At the 5th-leaf stage, two of these populations were vernalized for 3 weeks and 6 weeks. The vernalized populations were returned to the greenhouse for comparison with the 3rd population as non-vernalized control. Heading date was scored for 3 plants of each line in these 3 populations.

Positional cloning of *QVrd.osu-5A*

A major QTL on chromosome 5A was found associated with quantitative vernalization requirement duration (*QVrd.osu-5A***)** in the winter wheat population generated from Jagger and 2174, and this QTL was cloned using a positional cloning approach (Li and Yu et al., 2013). The allelic variation at *vrn-A1* at the protein level was finally found to be responsible for *QVrd.osu-5A* (Li and Yu et al., 2013).

Identification of proteins interacting with vrn-A1

The 'Matchmaker™' Library Construction & Screening System (Clontech, USA) was used to construct an Y2H 'prey' library for Jagger**.** RNA was extracted from pooled samples of young leaves and apices from vernalized plants for 1, 2, and 3 weeks, and the vrn-A1a was used as 'bait' to screen the 'prey' library. The constructing and screening procedures were described in a previous study in which 2174 was used as a host plant (Cao and Yan, 2013).

The *in vivo* interaction of *Ta*VRN-A1 and *Ta*HOX1 proteins

The complete *Ta*VRN-A1 cDNA was cloned into pDONR207 with the BP cloning kit (Invitrogen), and then transferred to pEarleygate 101 (pEG101) using the LR cloning kit for subcellular localization of *Ta*VRN-A1. *TaVRN-A1* in pDONR207 was fused to the N-terminal 174 amino acid portion (1-174) of YFP in the pEarleyGate201-YN vector (pEG201-YN) to test *in vivo* interaction with *Ta*HOX1 fused to the C-terminal amino acid portion (175-239) of YFP in the pEarleyGate202-YC vector (pEG202-YC). Empty vectors were also used as negative controls for interaction with *Ta*VRN-A1 or *Ta*HOX1 proteins. *A. tumefaciens* strains (GV3101) carrying the BiFC constructs were used together with the p19 strain for infiltration of *N. benthamiana* leaves (5 weeks old). Leaf discs were cut for BiFC for imaging 3 days after infiltration. The images were taken with a bright filter (BF) to indicate the background of the leaves infiltrated with *A. tumefaciens* carrying constructs, or with an ultraviolet filter (DAPI) to indicate the position of the nucleus stained with 4′, 6-diamidino-2-phenylindole. The overlay images align the locations of YFP with the DAPI-stained nucleus. Images were taken under a fluorescent microscope (Olympus BX51) with GFP filter to indicate the presence of fluorescent proteins.

3.3 Results

The discovery of a major gene for vernalization requirement duration

We tested the two parental lines, Jagger and 2174, for variation in vernalization requirement duration to reach the vernalization saturation point or gain the maximum vernalization effect on heading under a long day condition (16/8 hrs for day/light). Jagger

required 3 weeks at 4°C to reach the maximum vernalization effect on heading, whereas 2174 required 6 weeks under the same condition (Li and Yu et al., 2013). When vernalized for 3 weeks, Jagger flowered 101 days after planting but 2174 flowered 136 days after plants. When Jagger and 2174 were tested in the field, they showed a difference of 3-5 days in heading date and 1-3 days in physiological maturity (Fig.5). The 3 weeks' vernalization produced phenotypic difference in flowering time for up to one month between the two alleles. The enlarged difference could be easily scored for mapping and cloning of genes segregated in the Jagger x 2174 population.

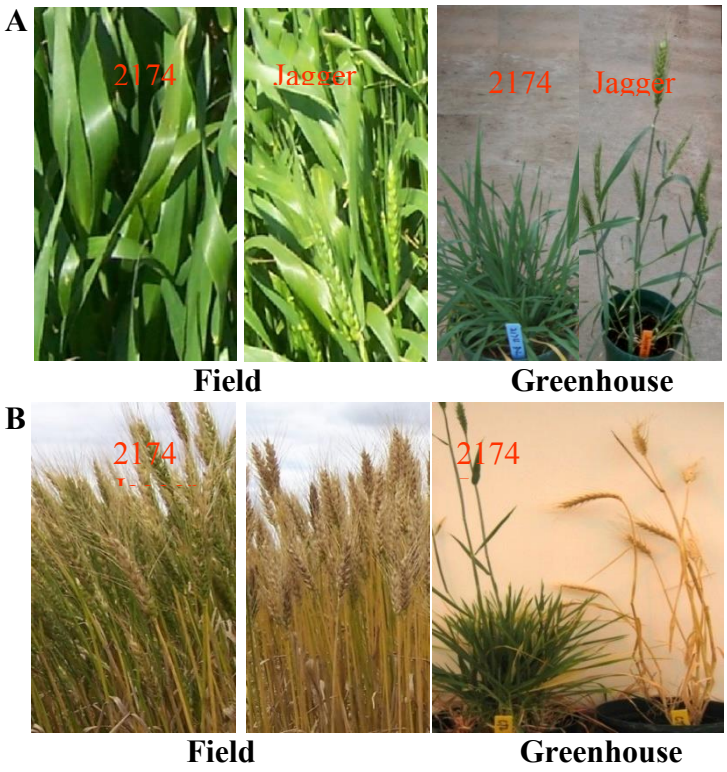


Fig.5 Phenotypic comparison of Jagger and 2174. (A) Phenotypic comparison of heading date between Jagger and 2174. Jagger and 2174 showed a difference in heading date for only 3-5 days when tested in the field but up to one month when tested in a greenhouse and vernalized for 3 weeks. (B) Phenotypic comparison of physiological maturity between Jagger and 2174. Jagger and 2174 showed a difference in heading date for only 1-3 days when tested in the field but up to one month when tested in a greenhouse and vernalized for 3 weeks.

A major QTL for vernalization requirement duration that segregated in the mapping population of Jagger x 2174 RILs was mapped to the long arm of chromosome 5A in a genomic region encompassing the *vrn-A1* locus (Figure 6), and this QTL was designated *QVrd.osu-5A*. The *QVrd.osu-5A* locus explained 63.4% (LOD=18.8) of the total phenotypic variation in the population vernalized for 3 weeks, and 20.9% (LOD=4.3) in the control population with no vernalization (Fig.6).

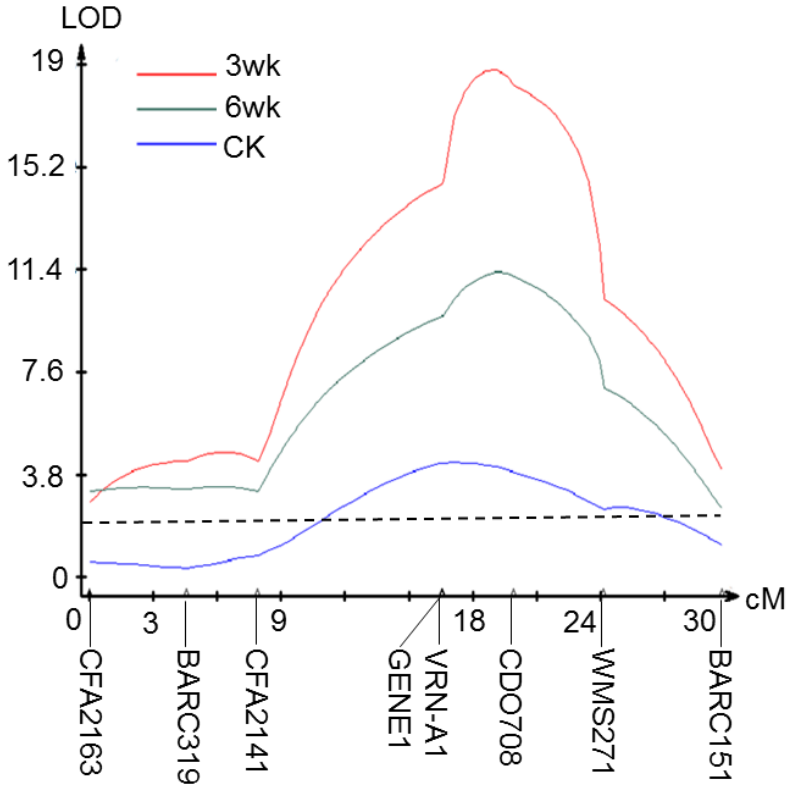


Fig.6 Genetic map of *QVrd.osu-5A*. Two sets of the Jagger x 2174 RIL population were vernalized for 3 weeks (3wk) and 6 weeks (6wk) and one set of the same population was not vernalized for control (CK). The horizontal dotted line represents a threshold value of 2.5 LOD.

A recombinant inbred line, RIL23 carrying the Jagger *vrn-A1* allele was backcrossed with the parental line 2174 to generate a BC_1F_2 population, in which the *QVrd.osu-5A* locus

was heterozygous but each of *PPD-D1* and *VRN-D3* was fixed at the homozygous allele. The latter two genes showed significant effects on heading date in the Jagger x 2174 RIL population when tested in the field (Chen et al., 2009). The resulting BC_1F_2 population was used to test genetic effects of the *QVrd.osu-5A* locus on vernalization requirement in winter wheat. The 70 plants either homozygous or heterozygous for the Jagger *vrn-A1a* allele for early heading showed a significant difference from the 20 plants homozygous for the 2174 *vrn-A1b* allele for late heading ($p<0.001$). The observed segregation ratio between the earlier heading and later heading groups was not significantly different from a 3:1 ratio ($X^2=0.37$, df=1, $p=0.54$) and fit a one-gene model. Finally, *vrn-A1* was identified being the gene that caused the difference in vernalization requirement duration between Jagger and 2174 (Li and Yu, et al., 2013).

A modified PCR marker for allelic variation in *vrn-A1*

No difference was observed between the Jagger and 2174 *vrn-A1* alleles in the previously identified regulatory sites that accounted for allelic variation between the winter *vrn-A1* allele and the spring *Vrn-1* allele. No difference was observed in gene expression between the Jagger and 2174 *vrn-A1* alleles (data not shown). However, two point mutations were found in coding regions, resulting in alteration of two amino acids in the conserved domain between the Jagger and 2174 alleles. A PCR marker for the mutated site corresponding to amino acid residue (L^{117}/F^{117}) between the Jagger *vrn-A1a* allele and 2174 *vrn-A1b* allele was developed (Chen et al., 2009). Primers vrn-A1F7B (5'-GTGGAGAAGCAGAAGGCGCATG-3') and vrn-A1R7 (5'-CCGACAGAACTGCATAGAGACC-3') were designed in this study to map A^{180}/V^{180}

between the Jagger *vrn-A1a* allele and 2174 *vrn-A1b* allele (Fig.7A) (Li and Yu, et al., 2013).

The previously pusbilished PCR marker was used for gene cloning. However, the PCR marker cannot be used to show that 2174 has two copies of vrn-A1b as found in the previous study and reported in another study (Díaz et al., 2012). A new PCR marker for the polymorphism between the Jagger *vrn-A1a* allele and the 2174 *vrn-A1b* allele, as well as for the polymorphism between the duplicated *vrn-A1b1* and *vrn-A1b2* copies in 2174, by using one-shoot PCR following digestion with two restriction enzymes (Fig.7B). The PCR marker will facilitate identification of the point mutation at L^{117}/F^{117} or A^{180}/V^{180} or at both of them in diverse wheat genetic germplasm.

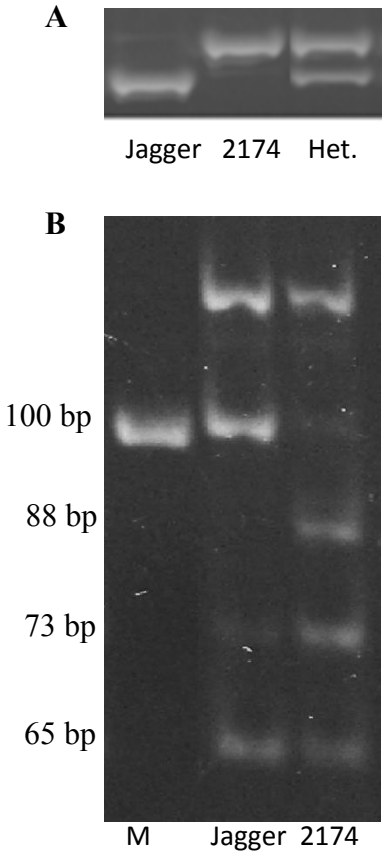


Fig.7 Comparison of PCR markers for the Jagger and 2174 *vrn-A1* alleles. (A) A PCR marker for the SNP in exon 7. The two primers amplified a 221bp fragment using an annealing temperature of 55°C and extension time for 1 min. The PCR products digested with restriction enzyme *Sph* I were run on a 2% agarose gel, showing polymorphic bands between the *vrn-A1a* allele (199 bp) and *vrn-A1b* allele (221 bp). (B) A PCR marker for both of two SNPs in exon 4 and exon 7. The PCR products were digested with *Mbo*I + *Aci*I + *Hpa*II. The digested PCR products were 342 bp + 73 bp + 470 bp + 88 bp + 9 bp + 65 bp + 123 bp in size for the Jagger allele and 415 bp + 470 bp + 97 bp + 65 bp + 123 bp for duplicated vrn-A1b1 and 342 bp + 73 bp + 470 bp + 97 bp + 65 bp + 123 bp for vrn-A1b2 of the 2174 allele. The digested PCR products were run on a 9% acrylamide gel.

The interaction site of *Ta*VRN-A1 and *Ta*HOX1 proteins

A protein that has the same structure of the homeodomain (HD) and the leucine zipper domain (Zip) as HOX proteins reported in animals (Ariel et al., 2007) was identified from the Y2H library using vrn-A1 as a probe. *Ta*HOX1 and *Ta*VRN-A1 shared 5 leucine residues present in the Zip domain (Fig.8). The interaction of *Ta*HOX1and *Ta*VRN-A1 is attributable to the presence of the zipper in the two proteins. However, the Ala^{180} in vrn-A1b represented the same amino acid residue as *Ta*HOX1, but Val^{180} caused a mismatch between vrn-A1b and *Ta*HOX1. This mismatch explained the decreased ability of vrn-A1b to interact with *Ta*HOX1 (Li and Yu et al. 2013).

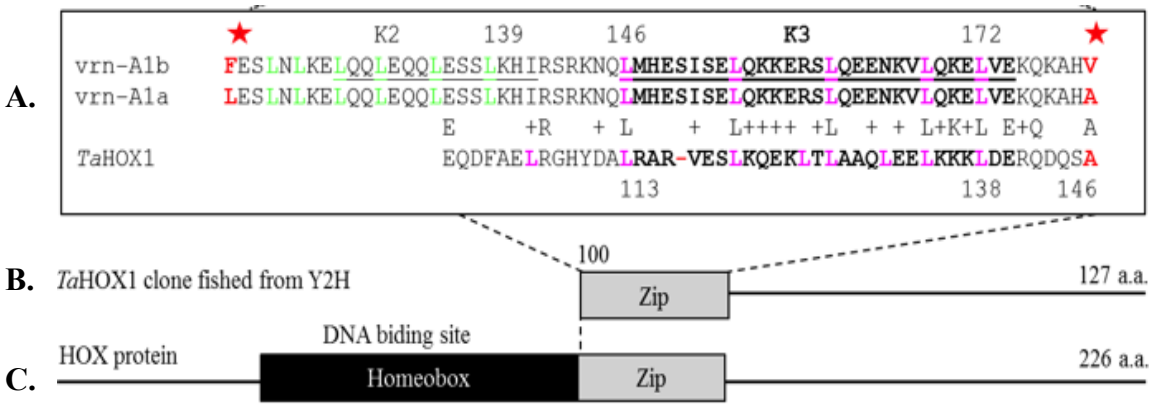


Fig.8. Interacting site of *Ta*VRN-A1 and *Ta*HOX1 proteins. (A) Sequence comparison between the *Ta*VRN-A1 and *Ta*HOX1 proteins. Locations of conserved domains in *Ta*VRN-A1, and two altered sites in amino acid sequence are indicated with a red star. Leu^{117}/Phe^{117} or Ala^{180}/Val^{180} residues are highlighted in red, and the conserved leucine residues are highlighted in pink. (B) The cDNA of *TaHOX1* fished out from the Y2H library was expressed to test protein interactions with *Ta*VRN-A1. (C) Positions of conserved HD and Zip domains in *Ta*HOX1 proteins in plants.

The *in vivo* interaction of *Ta*VRN-A1 and *Ta*HOX1 proteins

The direct binding between *Ta*VRN-A1 and *Ta*HOX1 proteins was confirmed by *in vitro* experiments (Li and Yu, et al., 2013). The physical interaction of *Ta*VRN-A1 and *Ta*HOX1 proteins was also confirmed by *in vivo* experiments (Fig.9). When *Ta*VRN-A1-YN and *Ta*HOX1(1-180)-YC were simultaneously expressed in the same cell, yellow fluorescence was observed in the nucleus with a confocal microscope (Fig.9). The interaction happened on cell nucleus that was confirmed using DAPI to stain the cell nucleus (Fig.9).

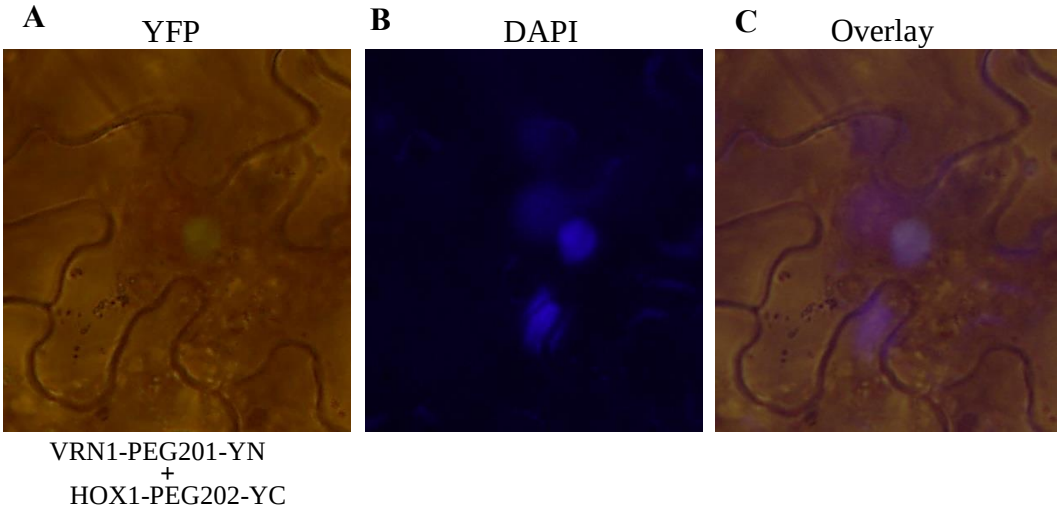


Fig.9. The *in vivo* interaction of *Ta*VRN-A1 and *Ta*HOX1 proteins. (A) The fluorescent proteins resulted from the *in vivo* interaction between *Ta*VRN-A1-YN and *Ta*HOX1(1-180)-YC in the nucleus are shown in green. by 'N'. No YFP was observed in a negative control of the co-transformation of pEG202-YC with *Ta*VRN-A1-pEG201-YN or pEG201-YN with *Ta*HOX1-pEG202-YC (Figures not shown). YN, YFP fragment at the N-terminal end expressed from pEG201-YN vector; YC, YFP fragment at the C-terminal end expressed from pEG202-YC vector. Images were taken under a fluorescent microscope. *A. tumefaciens* strains (GV3101) carrying the BiFC constructs were used together with the p19 strain for infiltration of *N. benthamiana* leaves (5 weeks old). Leaf discs were cut for BiFC for imaging 3 days after infiltration. (B) The image was also taken with an ultraviolet filter (DAPI) to indicate the position of the nucleus stained with 4′, 6-diamidino-2-phenylindole. (C) The overlay images align the locations of YFP with the DAPI-stained nucleus.

Allelic variation in *TaHOX1*

The complete *TaHOX1* gene including 956-bp upstream from the start codon and the region from the start codon to the stop codon. The only difference in *TaHOX1* between Jagger and 2174 alleles was one SNP that occurred in exon 1. This DNA SNP resulted in an alteration Leu^{99}/Pol^{99}. This site was not included in the protein fragment that was used for protein interaction with *Ta*VRD-A1, but could be included in the suggested DNA binding site. This SNP facilitated mapping of the *TaHOX1* gene. A PCR marker was developed for a SNP between the Jagger *TaHOX1a* allele and the 2174 *TaHOX1b* allele (Fig.10A). Primers *Ta*HOX1-6BF1 (5'-GCGGCGCGCCAAGCTGGAC-3') and *Ta*HOX1-R2M (5'- CAGCTGCACATCGAGCAGACAC-3') were used to map the SNP in *TaHOX1*. The *TaHOX1* marker was mapped into a genetic linkage group including 11 SSR markers located on chromosome 6B (Fig.10B).

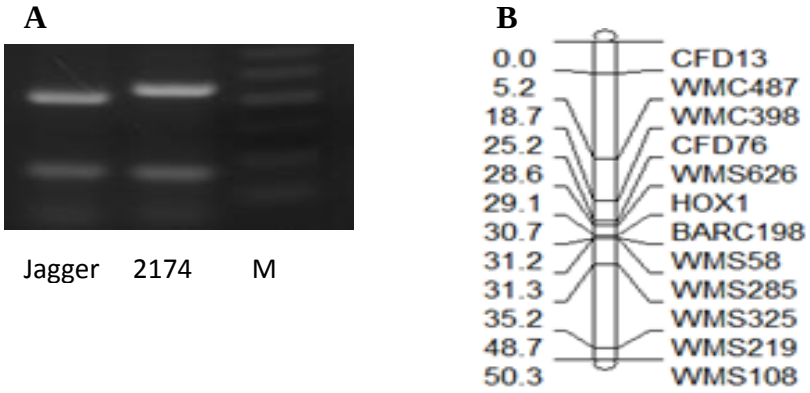


Fig.10. Mapping of *TaHOX1*. (A) The specific primers amplified a 317 bp gDNA fragment using an annealing temperature of 55°C and extension time for 1 minute. The PCR products digested with restriction enzyme *Msp* I were run on a 2% agarose gel, showing polymorphic bands between the *TaHOX1a* allele (198 bp) and *TaHOX1b* allele (217 bp). (B) Location of *TaHOX1*. *TaHOX1* was mapped in the Jagger×2174 RIL population.

The up-regulation of *TaHOX1* by low temperature and its genetic effects on heading date

We isolated the complete *TaHOX1* gene from Jagger (JQ915061) and 2174 (JQ915062) for functional characterization. Using RT-PCR, we found that the *TaHOX1* transcripts in leaves were greatly up-regulated and remained at a high level during vernalization (Figure 11A). This result indicated that *TaHOX1* had a similar expression pattern as the recessive *vrn1* allele characterized in previous studies (Danyluk et al., 2003; Murai et al., 2003; Trevaskis et al., 2003; Yan et al., 2003). *TaHOX1* and *TaVRN-A1* were concomitantly up-regulated by low temperature, indicating that they functioned in the flowering pathway.

To determine if *TaHOX1* was involved in the flowering pathway, the PCR marker was used as a single marker to analyze the phenotypic data that were collected from field and greenhouse experiments. A significant difference in heading date was observed between the Jagger *TaHOX1a* and the 2174 *TaHOX1b* alleles (120 days vs. 109 days, respectively) in the RIL population vernalized for 3 weeks ($P<0.05$) but not in the RIL population vernalized for 6 weeks or the control population (Figure 11B). The SNP in exon 1 resulted in a 'leucine' residue at position 99 in *Ta*HOX1b and a 'proline' residue at the same position in *Ta*HOX1a. This point mutation in *Ta*HOX1 did not affect its interaction with vrn-A1. In contrast, the higher *TaHOX1* transcriptional level was associated with earlier heading, suggesting that the regulation of heading date by *TaHOX1* was at the transcriptional level.

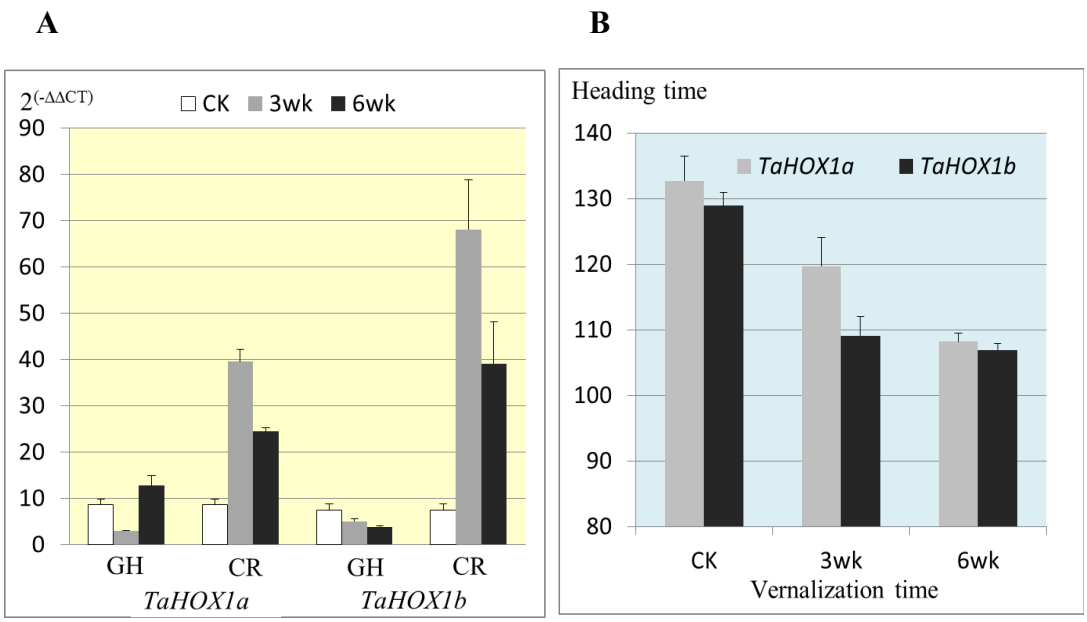


Fig.11. Expression profiles of *TaHOX1* and its genetic effect on flowering time. (A) Expression pattern of *TaHOX1* in greenhouse (GH) and cold room (CR). Transcript levels of *TaHOX1a* (the Jagger allele) and *TaHOX1b* (the 2174 allele) in leaves of vernalized plants for 3 weeks (3wk) and 6 weeks (6wk) as well as non-vernalized plants (CK) are shown using the values calculated by the $2^{(-\Delta\Delta CT)}$ method, where CT is the threshold cycle, and *actin* was used as an endogenous control. The values represent mean expression levels (*n*=15-20), and the bar indicates standard error. (B) Genetic effect of *TaHOX1* on heading date. The heading date was from each line of the population (n=96) that was vernalized for 3 weeks (3 wk) or 6 weeks (6 wk) or used for control in a greenhouse (CK). The significant effects of the *TaHOX1* gene on heading date were determined using one-way analysis of variance (ANOVA). Bar indicates standard error.

VRN1 was not necessary for the expression of *HOX1*

In an *mvp* plant in which *TmVRN1* was deleted, transcripts of *TmHOX1* were detected in both flowered plants and non-flowered plants (Fig.12), indicating *TmVRN1* is not necessary for the expression of *TmHOX1* in diploid wheat *T. monococcum*.

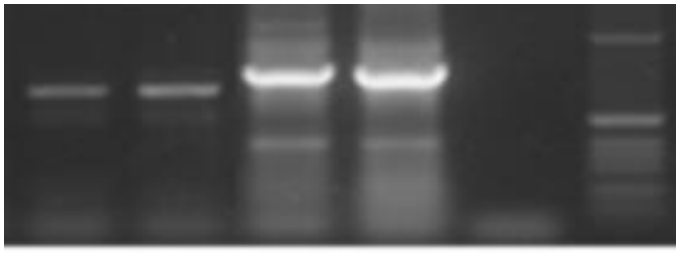


Fig.12 HOX expression in MVP plants. Primers TmHOX1-F1 (5'-CGCGACGGCTCCGAAATG-3') and TmHOX-R1 (5'-TCATGCCACTGCGTTCCACT-3') were used to test if *HOX1* is expressed in MVP plants, with the wild type plant as control. The gDNA was also used as control to ensure that the primers work for diploid wheat.
The expected size of cDNA PCR products was 696 bp, while the expected size of gDNA PCR products was larger than 696 bp due to the presence of an intron.

3.4 Discussion

A C/T polymorphism in exon 4 that is responsible for the amino acid change at Leu^{117}/Phe^{117} in *Ta*VRN-A1 was found between the Jagger allele and the 2174 allele, and a PCR marker for this polymorphism was developed to show an association with developmental variation in the Jagger x 2174 RIL population tested in the field (Chen et al 2009). A new PCR marker was developed for the polymorphism in exon 7 that is responsible for the amino acid change at A^{180}/V^{180} (Li and Yu et al., 2013). In this study,

we developed a PCR marker that can be used to identify the point mutation at L^{117}/F^{117} or A^{180}/V^{180} or at both of them in diverse wheat genetic germplasm. This new PCR marker can also be used to distinguish the polymorphisms between the Jagger *vrn-A1a* allele and the 2174 *vrn-A1b* allele and between the duplicated *vrn-A1b1* and *vrn-A1b2* copies in 2174.

It was recently reported that *TaVRN-A1* was duplicated in winter wheat cv. 'Hereward' (JF965397) but not in 'Claire' (JF965395) (Díaz et al., 2012). In the population of Claire x Hereward, plants homozygous for the Claire allele flowered earlier and plants homozygous for the Hereward allele flowered later, while heterozygotes had an intermediate flowering time; therefore, the authors concluded that the increased copy number of *TaVRN-A1* in Hereward resulted in an increased requirement for vernalization (thus late flowering). We found that at the polymorphic sites of both exon 4 and exon 7, Jagger has the same allele as Claire and 2174 has the same allele as Hereward. We also observed the duplication event of *vrn-A1b* in 2174 for more vernalization but not in Jagger for less vernalization. Thus on the surface, it seems that the two studies have consistent results and that the phenomenon the Diaz et al. described in their paper that plants with an increased copy number of *TaVRN-A1* have an increased requirement for vernalization was correct. However, we affirm that a greater vernalization requirement in 2174 was not because 2174 has one more *vrn-A1* copy than Jagger but because Jagger produces a different vrn-A1 protein from 2174. Our results from several independent progeny populations clearly and consistently demonstrated the Jagger *vrn-A1a* allele for early flowering was dominant over the 2174 *vrn-A1b* allele for late flowering, which phenotypic segregation was 3:1. The two studies have made different conclusions for

several reasons. First, in the Diaz et al. study, vernalization was performed under a short day condition (7°C, 8 h light), and then the vernalized plants were grown under a long day condition (18 h light) in a greenhouse (no temperature condition was provided). In our study, however, vernalization was performed under 4°C, 16 h light and then the vernalized plants were moved to 25°C, 16 h light. The same long day was set throughout the experiment in our study to avoid disruption of photoperiod that was well known to have significant effects on flowering in winter wheat (Dubcovsky et al., 2006; Wang et al., 2009). Second, in the Diaz et al. study, a small mapping population of 96 F_2 lines was used to test the association between the *vrn-A1* copy number and flowering time. However, we generated a backcross population and screened 6,500 plants to find several lines that had crossovers at the *vrn-A1* locus and were segregated in progeny for flowering time by a single gene *vrn-A1a* to avoid disruption of other genetic factors for this phenotype. Third, *vrn-A1* is the promoter of flowering as shown in wheat and barley (Danyluk et al., 2003; Murai et al., 2003; Trevaskis et al., 2003; Yan et al., 2003), and *Arabidopsis* (Ng and Yanofsky, 2000). The dosage effect of two copies from such a flowering promoter should not result in a phenotype of a later flowering time as suggested in the previous study (Díaz et al., 2012). In this study, we used the positional cloning strategy to prove that *vrn-A1a* in Jagger was dominant for early flowering, regardless of the function of the duplicated *vrn-A1b* in 2174. Therefore, we concluded that less vernalization requirement by Jagger was due to its vrn-A1a protein form or that more vernalization requirement in 2174 was not due to the presence of its two *vrn-A1b* copies.

Acknowledgements

We thank J. Deng (Oklahoma State University) for his assistance with the development of the *in vitro* protein interaction systems, and Y Cui (Agriculture and Agri-Food Canada) for providing pEarleygate201-YN and pEarleygate202-YC vectors, F. Zhang (Oklahoma State University) for his assistance with the development of the *in vivo* protein interaction systems in tobacco. We thank Y. Chen for technical assistance in generating backcross populations. This project was supported by the National Research Initiative Competitive Grants CAP project 2011-68002-30029 from the USDA National Institute of Food and Agriculture and from the Oklahoma Center of Advanced Science and Technology (OCAST, PS07-032). This research project was partially funded by the Oklahoma Agricultural Experiment Station and the Oklahoma Wheat Research Foundation.

CHAPTER IV

GENETIC ANALYSES OF HOMOWOLOGOUS AND HOMOLOGOUS GENES OF *TaHOX1*

Abstract

Bread wheat (*Triticum aestivum*) is an allohexaploid species that possesses three genomes A, B and D; therefore, each gene should have three homoeologous genes. The availability of *TaHOX1* sequences has facilitated identifying allelic variation in its homoeologous and homologous genes. *TaHOX1* presented in the previous chapter was mapped on chromosome 6B (*TaHOX-B1*), and the complete gene of homoeologous *TaHOX-A1* on chromosome 6A and *TaHOX-D1* on chromosome 6D were sequenced but showed no allelic variation and mapped using Chinese Spring nullisomic-tetrasomic lines. A set of three homoeologous *TaHOX2* genes similar to *TaHOX1* was observed in the wheat genome. While *TaHOX-A2* was mapped on chromosome 2A due to an 18 bp indel polymorphism in exon 1, *TaHOX-B2* on chromosome 2B or *TaHOX-D2* on chromosome 2D showed no allelic variation but mapped using Chinese Spring nullisomic-tetrasomic lines. Using subcellular localization, *Ta*HOX-A2 protein was

characterized. The expression patterns of these homoeologous and homologous *TaHOX1 and TaHOX2* in different developmental stages and tissues were characterized in wheat.

4.1 Introduction

Bread wheat (*Triticum aestivum*) is an allohexaploid species that possesses three genomes, A, B and D, each of which is originated from diploid wheat. These three genomes are closely related in genome structure, chromosome size and shape, and gene content and order, but genome sequences are divergent particularly in repetitive DNA content. Each of most genes in bread wheat is present as three similar sequences of homoeologous loci with high exonic homology and lower homology in introns. Based on specific sequences, three homoeologous loci can be distinguished and designed gene A, B, and D.

VRN1 was originally cloned from diploid wheat, *T. monococcum* that is a relative of genome A in hexaploid wheat (Yan et al., 2003). The availability of the *VRN1* gene sequence from diploid wheat has enabled three homoeologous genes *VRN-A1*, *VRN-B1*, and *VRN-D1* in bread wheat to be isolated using a relatively readily PCR approach. Allelic variation in each of the three *VRN1* genes in extensively collected germplasm in different geographical areas of the world has been identified and functionally characterized (Kamran et, 2014). *VRN-A1*, *VRN-B1*, and *VRN-D1* in bread wheat also play important roles in regulating developmental processes. The first objective of this study is to isolate two homoeologous genes, *TaHOX-A1* on chromosome 6A and *TaHOX-D1* chromosome 6D, based on the *TaHOX-B1* on chromosome 6B that was isolated in the

previous study. Each of the two homoeologous genes was isolated from each of Jagger and 2174 to identify allelic variation for mapping in the available Jagger x 2174 RIL population.

In addition to homoeologous genes in hexaploid wheat, this species may have many genes that have been duplicated during evolution and domestication. Early studies with molecular markers indicated the presence of duplicated loci or regions on the genetic maps in wheat, revealing ancestral genome duplications and polyploidization events in the history of this species. During the processing of isolating the three *TaHOX1* genes, we found that *TaHOX1* was duplicated, thus producing three homoeologous *TaHOX2* genes. The second objective of this study is to isolate and map homoeologous *TaHOX2* genes in bread wheat.

The co-existence of six homologous genes including three homoeologous *TaHOX1* and three *TaHOX2* in the same wheat plant makes it not easy to obtain the complete gene of each of these genes. However, the release of the wheat genome sequences enables us to design specific to each gene. In this study, we have developed a method allowing respective isolation of each gene by a single PCR reaction and provided important new insights into developmental mechanism, showing that wheat development involves the differential expression of *HOX* family genes.

4.2 Materials and Methods

Isolation of homoeologous and homologous *TaHOX* genes

The complete gene from the translational start codon and stop codon, the 5'-UTR and part of the promoter, as well as part of the 3'-UTR were isolated for each of 6 homoeologous and homologous *TaHOX* genes.

Specific primers for each *TaHOX* gene were designed, based on multiple sequence alignment. Chinese Spring (CS) nullisomic-tetrasomic lines were used to confirm determine which chromosome a *HOX* is located. The N6A-T6B, N6B-T6D, and N6D-T6A deletion lines were used to determine chromosomal assignment in homoeologous group 6, and the N2A-T2B, N2B-T2A, and N2D-T2A deletion lines were used to determine chromosomal assignment in homoeologous group 2. The use of deletion lines enables markers to be localized to a certain chromosome. Once specific primers were determined, Jagger and 2174 were used to amplify each gene, and polymorphic sequences were used to map in the Jagger x 2174 RIL population. Sequences of primers used for isolation of these genes are provided in Table 1. PCR was performed with annealing temperature of 55°C and extension time of 1 minute for approximate 1 kb PCR products of these *TaHOX* genes. The PCR products were purified and sequenced directly.

Expression patterns of homoeologous and homologous *HOX* genes

Expression pattern of each *TaHOX* gene was investigated by using specific primers to test cDNA samples from leaves and roots of seedling plants plus spikes when adult plants were tested. Total RNA was extracted using Trizol regents (Invitrogen). The first-strand cDNA was synthesized using a SuperScript^{TM} II Reverse Transcriptase kit (Invitrogen). Sequences of primers used for gene expression are provided in Table 2.

Subcellular localization of *Ta*HOX-A2

The *TaHOX-A2* cDNA was amplified by specific primers HOXC4-BiFC2-F1-2 (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGATTACCACCACCATCAC-3') and HOXC4(124/130)-BiFC2-R (5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCGAAGTCGTGCTCCAGCT-3'). The cDNA fragment encodes 1-124 amino acids for the Jagger allele and 1-130 amino acids for the 2174 allele. The difference of 6 amino acids between the two proteins was due to an 18-bp deletion in the Jagger allele. DMSO was added into the PCR system for higher amplification efficiency. The cDNA fragments were cloned into pDONOR207 with the BP cloning kit, and then transferred to pEarleyGate 101 (pEG101) using the LR cloning kit. *TaHOX-A2* in pEG101 was used for subcellular localization. Agrobacterium tumefaciens strains (GV3101) carrying the BiFC constructs were used for infiltration of tobacco leaves (at 5 weeks old). Three days after agrobacterium infiltration, leaf discs were cut for BiFC for imaging. Images were taken using a fluorescent microscope. Images taken under GFP filter or ultraviolet filter indicate the presence of fluorescent proteins, or the position of the nucleus stained separately.

4.3 Results

Isolation of homoeologous genes of *TaHOX1*

The sequences of the complete gene of *TaHOX-B1* are deposited in GenBank for the Jagger allele (JQ915061) and the 2174 allele (JQ915061). The isolated genes from the two alleles have the same length of 1,737 bp including 956 bp upstream from the translational start codon and 781 bp from the start codon to the stop codon. The two

alleles showed one SNP in exon 1 that resulted in alteration of an amino acid (Chapter III). The gene consists of two exons and one intron, and which is the same as all of the other *TaHOX* genes isolated in this study.

The isolated *TaHOX-A1* gene was 2,736 bp in length, including 987 bp before the translational start codon, 798 bp between the start codon and the stop codon, 951 bp after the translational stop codon. No difference was found in the complete *TaHOX-A1* gene between the Jagger and 2174 alleles. The chromosomal location of *TaHOX-A1* was determined based on identification of this gene in CS deletion lines (Fig.13A).

The isolated *TaHOX-D1* gene was 2,239 bp in length, including 1,141 bp before the translational start codon, 774 bp between the start codon and the stop codon, 324 bp after the translational stop codon, and. No difference was found in the complete *TaHOX-D1* gene between the Jagger and 2174 alleles. The chromosomal location of *TaHOX-D1* was determined based on identification of this gene in CS deletion lines (Fig.13B).

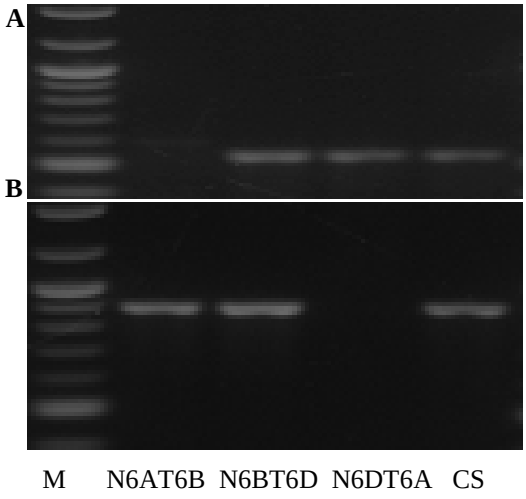


Fig.13 Chromosomal locations of *TaHOX-A1* and *TaHOX-D1*. Primers were designed specific to *TaHOX-A1* (A) and *TaHOX-D1* (B) respectively. PCR were performed with Chinese Spring homoeologous group 6 nulli-tetrasomics. M: PCR 100 bp ladder marker. PCR was performed using standard programs except annealing temperature of 55°C and extension time for 1 min.

Isolation of homoeologous genes of *TaHOX2*

TaHOX1 sequences were used to search GenBank nucleotide collection databases, a cDNA (AK335335.1) was found to have very high identity with *TaHOX1*. This cDNA was different from all of *TaHOX-A1*, *TaHOX-B1*, and *TaHOX-D1*; therefore, this cDNA could be from a novel gene. An effort was made to isolate homoeologous genes that are different from *TaHOX1* and thus were *TaHOX2*. We have also found three new *HOX* genes that have high identity to *TaHOX1* genes, and these new *HOX* genes are hereafter referred to as *TaHOX2*.

The first *TaHOX2* gene was 1,522 bp in length, including 306 bp before the translational start codon, 905 bp between the start codon and the stop codon, 301 bp after the translational stop codon. An 18 bp indel (insertion/deletion) polymorphism was observed in exon 1 of this gene between the Jagger and 2174 alleles. A PCR marker was developed for 18 bp indel using specific primers HOXC4-MF2 (5'-CATCCAGCAGAGCAGAGGAGAGC -3') and HoxC4-MR2 (5'-GAACATGGACTCCAGCGACCGTG -3') (Fig.14).

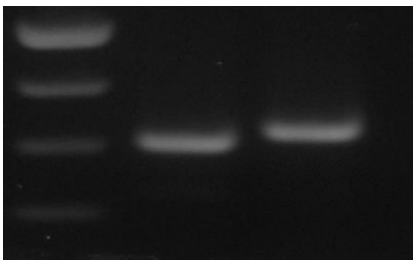


Fig.14 PCR marker for *TaHOX-A2*. Specific primers HOXC4-MF2 and HoxC4-MR2 were used to map the 18 bp indel in *TaHOX-2A*. The specific primers amplified a 310 bp gDNA fragment in Jagger and a 328 bp gDNA fragment in 2174, using PCR programs with annealing temperature of 56°C and extension time for 30 seconds. The PCR products were run in a 2% agarose gel directly. M: PCR 100 bp DNA ladder.

The marker for the first *TaHOX2* gene was developed by one shoot PCR without requirement for digestion with restriction enzymes. This marker was mapped on the short arm of chromosome 2A in the Jagger x 2174 RIL population (Fig.15); therefore, this *TaHOX2* was designated *TaHOX-A2*.

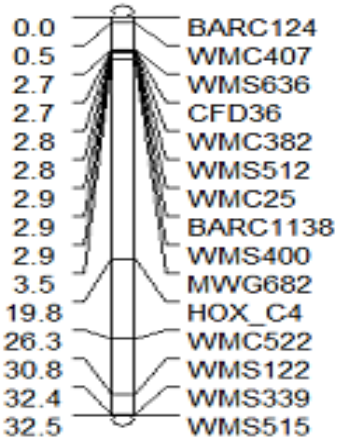


Fig.15 Chromosomal locations of *TaHOX-A2*. Primers specific to *TaHOX-A2* were used for mapping of the gene. SSR markers linked with *TaHOX-A2* are reported to locate on chromosome 2A.

The isolated *TaHOX-B2* gene was 1,536 bp in length, including 426 bp before the translational start codon, 876 bp between the start codon and the stop codon, 234 bp after the translational stop codon. No difference was found in the complete *TaHOX-B2* gene between the Jagger and 2174 alleles. The chromosomal location of *TaHOX-B2* was determined based on identification of this gene in CS deletion lines (Fig.16A).

The isolated *TaHOX-D1* gene was 1,219 bp in length, including 124 bp before the translational start codon, 789 bp between the start codon and the stop codon, 306 bp after the translational stop codon. No difference was found in the complete *TaHOX-D2* gene between the Jagger and 2174 alleles. The chromosomal location of *TaHOX-D2* was determined based on identification of this gene in CS deletion lines (Fig.16B).

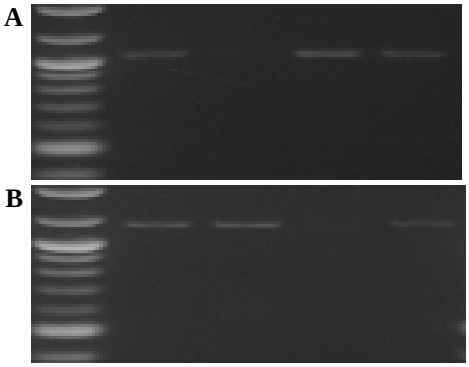


Fig.16 Chromosomal locations of *TaHOX-B2* and *TaHOX-D2*. Primers were designed specific to *TaHOX-B2* (A) and *TaHOX-D2* (B) respectively. PCR were performed with Chinese Spring homoeologous group 2 nulli-tetrasomics. M: PCR 100 bp ladder marker. PCR was performed using standard programs except annealing temperature of 55°C and extension time for 1 min.

Expression of *TaHOX1* and *TaHOX2*

As shown in Fig.17, all of the five genes were observed expression in leaves of seedling plants. Except *TaHOX-A2*, all other genes were observed expression in spikes. *TaHOX-D1* expression was detectable in roots of both seedling and adult plants. Another gene that was found to express in roots was *TaHOX-D2*, but it was at seedling only. All of the five genes showed higher transcriptional levels in leaves of seeding plants in the 2174 allele than the Jagger allele.

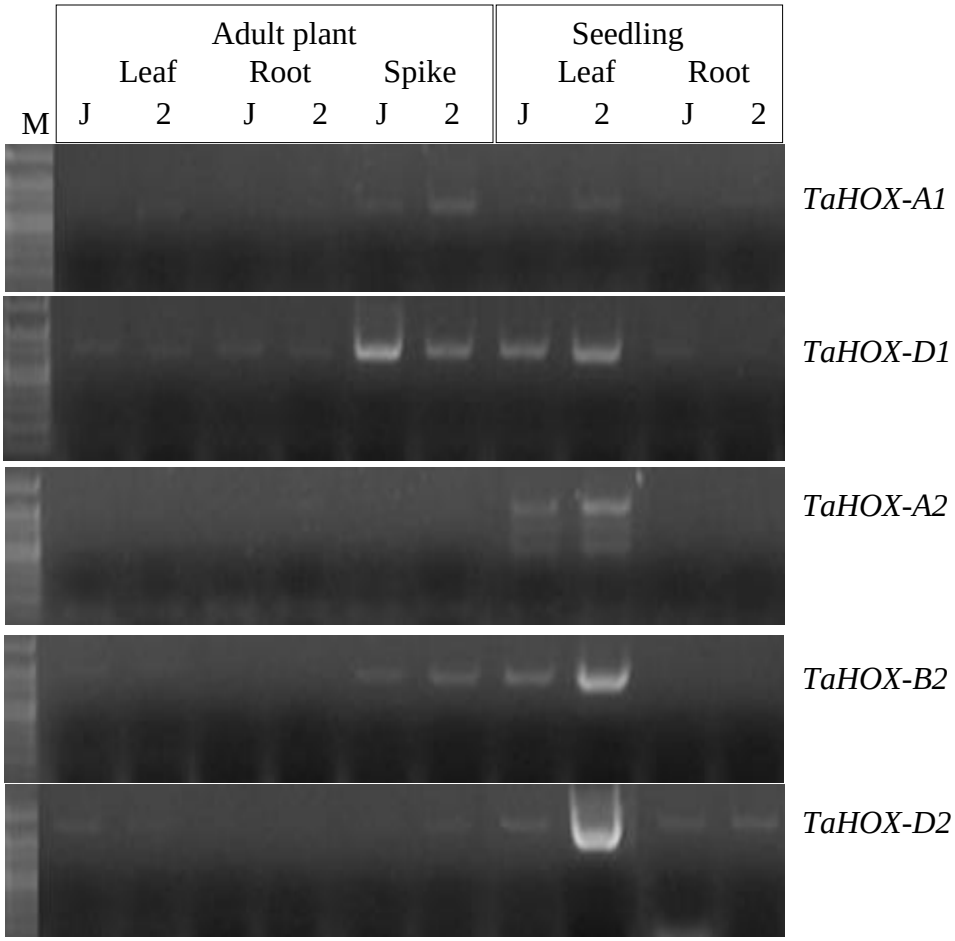


Fig.17 Chromosomal locations of *TaHOX-B2* and *TaHOX-D2*. Primers specific to five different *HOX* genes were used to test the same cDNA samples. (These cDNAs were tested using actin gene data not shown here).

Subcellular localization of *Ta*HOX-A2 proteins

To test if there is any difference in the expression location of the *Ta*HOX-A2 protein between Jagger and 2174, *Ta*HOX-A2 protein in pEG101 was expressed in living cells. As shown in Fig.18, the Jagger *Ta*HOX-A2 protein was expressed predominantly in the nucleus and partly on the cytoplasm membrane. No detectable difference was found in

the subcellular location of *Ta*HOX-A2 between Jagger and 2174, though Jagger *Ta*HOX-A2 lost 6 amino acids in the nucleotide binding site.

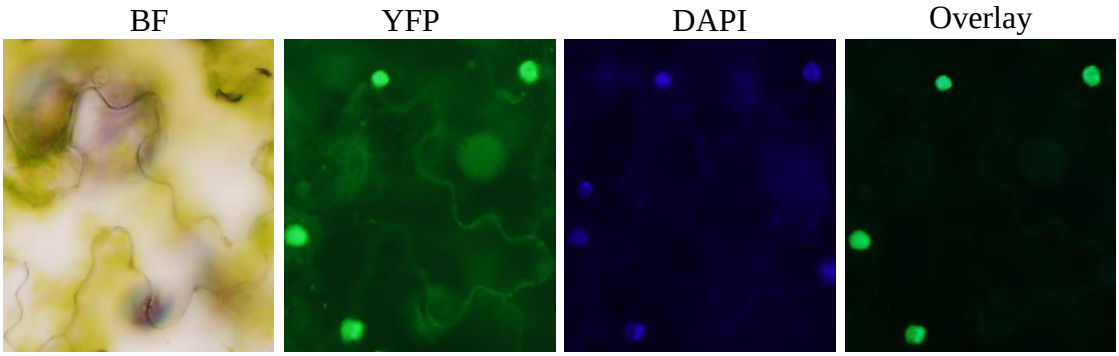


Fig.18 Subcellular localization of *TaHOX-A2*. *A.tumefaciens* strains (GV3101) carrying the pEG101 construct was used with the P19 strain for infiltration of *N. benthamiana* leaves (5 weeks old). Leaf discs were cut for imaging 3 days after infiltration, with an ultraviolet filter (DAPI) to indicate the position of the nucleus stained with 4′, 6-diamidino-2-phenylindole. The overlay images align the locations of YFP with the DAPI-stained nucleus.

4.4 Discussion

There are 42 genes in the HDZip gene family in *Arabidopsis* (*Arabidopsis* Genome Initiative, 2000). These HDZip proteins interact as dimers by their leucinezippers domains, and bind DNA sequence specifically via their homeodomains (Sessa et al., 1993; Aoyama et al., 1995; Meijer et al., 1997; Frank et al., 1998; Johannesson et al., 2001). Based on sequence blast, both *TaHOX-A1* and *TaHOX-A2* showed the highest identity to *Arabidopsis* thaliana homeobox 7 (AtHB7) (NP_182191.1), suggesting that *TaHOX-A1* and *TaHOX-A2* are two genes that were duplicated during evolution.

AtHB proteins in *Arabidopsis* have been grouped into four classes (Sessa et al., 1994). AtHB7 was characterized as a member in HD-Zip class I that involves responses to water and light stresses. AtHB7 could be induced by water deficit conditions or ABA treatment; hence, it is a gene that had response to water deficit (Soderman et al., 1996 and 1999; Lee and Chun, 1998). Over-expression of *AtHB7* resulted in early flowering in transgenic *Arabidopsis* plants (Soderman et al., 2000), and reduced elongation of the inflorescence stem and rosette leaves (Hjellstro¨m et al., 2003). These studies in *Arabidopsis* suggest that the orthologous *TaHOX-A1* and *TaHOX-A2* may play a role in regulation of flowering time and drought tolerance in wheat.

TaHOX-B1 has been demonstrated that it does have genetic effects on flowering time in winter wheat when vernalized for 3 weeks (Chapter 3). *TaHOX-B1* was expressed in seedling leaves, and all of the five homoeologous and homologous *HOX* genes presented in this study were also observed expression in leaves of seedling plants, suggesting they may have similar role in regulation of flowering time. However, only *TaHOX-D1* and *TaHOX-D2* were observed expression in roots of seedling, suggesting that if any of these wheat *HOX* genes play a similar role in plant response to drought stress as AtHB7, *TaHOX-D1* and *TaHOX-D2* should be the better candidates.

Like *TaHOX-B1*, both *TaHOX-A1* and *TaHOX-D1* were expressed in spikes in wheat. However, only *TaHOX-B2* in the *TaHOX2* homoeologous genes was were expressed in spikes. These genes expressed in spikes may be involved in spike development, as over-expression of orthologous *AtHB7* resulted in reduced elongation of the inflorescence stem. The functional validation of these HOX genes needs to be tested in transgenic

wheat. It was unfortunate that no transgenic wheat was successfully generated in this study.

Homeotic protein is one of the HOX proteins known to function as an on-off switch in controlling development, including specialization of regional identities along the anterio/posterior axis in a wide range of phyla in animals (Manak et al. 1994), whereas vrn-A1 (=AP1) is one of the MADS-box proteins known to act as floral switches for the transition from the vegetative to reproductive development in plants (Ng and Yanofsky, 2001). This study presented the first example that MADS and HOX proteins involving homeosis have a direct binding in higher plants.

CHAPTER V

TaVRN1 AND *TaCENPE1* PROTEINS INTERACTIVELY REGULATE PLANT HEIGHT AND SPIKE DEVELOPMENT IN WHEAT

Abstract

Using *Ta*VRN-A1 as bait, a wheat kinase interacting protein from a yeast-2 hybrid (Y2H) library was gained. The protein showed very high similarity to the CENP-E like kinetochore protein in cereals and it thus designated *Ta*CENPE1. *Ta*CENPE1 protein has a long region of sequence that was similar to conserved domains including leucine residues in *Ta*VRN-A1, suggesting that this region includes interaction sites. *Ta*CENPE1 was localized in the same pattern as *Ta*VRN-A1, indicating that they had the interaction site in living cells. *Ta*CENPE1 and *Ta*VRN-A1 were also confirmed to have a direct binding in living cells by using a transient expression system in tobacco leaves. The interference of *TaCENPE1* through RNAi approach showed the plant height was reduced and more florets were produced in transgenic wheat. These results suggest that *Ta*VRN1 and *Ta*CENPE1 proteins interactively regulate cell division and agronomic traits in wheat.

5.1 Introduction

*Ta*VRN-A1 and *Ta*HOX1 proteins were found to have the physical interaction in wheat, though such an interaction was not reported in any plants or other organisms. A single marker analysis showed that *TaHOX-B1* had a significant effect on flowering time of the plants with 3 weeks' vernalization, supporting that *Ta*HOX1 and *Ta*VRN-A1 function in the same flowering pathway in wheat. The previous study encouraged us to test if any other proteins identified from the Y2H library have interactions with *Ta*VRN-A1.

The Y2H system can generate a significant number of both false-positive and false-negative interactions. False clones may be identified for two reasons. One reason is that the two proteins show interactions with each other in yeast cells actually may not be expressed in the same tissue at the same time. Another reason is that the two proteins in yeast cell may interact indirectly, due to other proteins or factors that are also involved in the same pathway. As discussed in previous chapters, in addition to SOC1 and VRT2 belonging to the MADS family that have interactions with *Ta*VRN-A1, HOX1 proteins with a leucine rich region have a direct binding with *Ta*VRN-A1. The previous study suggests that the proteins that have conserved leucine rich domain could have interactions in plants. In this chapter, a CENP-E like kinetochore protein (CENP-E) was selected to test if it has interactions with *Ta*VRN1, based on a sequence analysis result that the *Ta*CENPE1 and *Ta*VRN1 have a conserved region including leucine rich domain. We first confirmed the interaction between *Ta*VRN-A1 and *Ta*CENPE1 proteins in an *in vivo* protein interaction system as previously described for *Ta*VRN-A1 and *Ta*HOX1.

In winter wheat, vernalization and cold hardiness were reported to be correlated with several morphological traits, including prostrate or rosette growth habit (Salmon 1917; Klages 1926; Taylor and Olsen 1976; Zelenski and Remeslo 1977; Robert 1982b; McIntosh 1983; Taylor 1983; Roberts and Larson 1985; Roberts 1986; Chaudhry 1986), plant height (Fowler and Gusta 1977; Fowler et al. 1981), and leaf length (Roberts and MacDonald 1984, 1988). Allelic variation in the dominant *Vrn-A1* locus also indicated pleiotropic genetic effects in spring wheat cultivars (Baga et al. 2009; Blake et al. 2009; Distelfeld et al. 2010; Kuchel et al. 2006; Li et al. 2008; Santra et al. 2009; Shimada et al. 2009; Shitsukawa et al. 2007; Zhang et al. 2008). These previous studies supported that the *VRN-A1* locus has different mechanisms in controlling wheat development. It was very recently reported that *VRN-A1* was associated with spike development (Perrce et al. 2013) and forest tolerance (Zhu et al., 2014). The complete deletion of the *VRN1* gene in diploid wheat enables the mutant wheat not to flower forever under any conditions, suggesting that signals from different developmental pathways converge at *VRN1*. It is postulated that *VRN1* would code for production of a temperature-sensitive protein, which undergoes conformational changes with changes in temperature (Robert 1989). In this study, we identified proteins that interact with *Ta*VRN1 and then test if the interacting proteins have any genetic effects on multiple traits as reported.

5.2 Materials and Methods

Identification of *TaCENPE1* clones from Y2H library

Three clones from the Y2H library were found to have a *TaCENPE1* cDNA fragment. The wheat genomic sequences database (http://www.cerealsdb.uk.net) and GenBank were used to determine three homoeologous *TaCENPE1* genes in hexaploid wheat. Chinese Spring deletion lines were used to identify specific primers that can be used to distinguish each of the three homoeologous genes in common wheat. Each of three homoeologous genes from parental lines were amplified and sequenced.

Subcellular localization of *Ta*CENPE1

*Ta*CENPE1 and *Ta*VRN-A1 proteins were tested if they have a direct binding in living cells by using a transient expression system in tobacco leaves (Lu *et al*., 2010). The full-length *Ta*VRN-A1 was expressed by pEG101-YFP vector for subcellular localization.

The complete *TaCENPE1* cDNA was amplified (KIP-5-F1, 5'-GGGCGGATCCCCATCTCC-3'; KIP-5-R2, 5'-ATTCGATTTTGAAAAGGAATGCCATAATGGTCCAT-3'; HF Taq, 60°C for annealing temperature, 1 minute extension, with 40 cycles) and cloned into TA vector using ligase kit. Another pair of primers (KIP-BiFC2-F1, 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAAGCGCATGAGTAG -3'; KIP-BiFC2-R1r2, 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCCAGCAAACCAGACACCTTATTA -3') was used to add the attB-flanked DNA fragment at both end of *TaCENPE1* so that it can be cloned into the donor vector pDONOR207 with the BP cloning kit. Then the full

length of *TaCENPE1* was transferred to pEarley-gate 101 (pEG101) using the LR cloning kit for subcellular localization of *TaCENPE1*.

In vivo interaction of *Ta*CENPE1-*Ta*VRN-A1

For *in-vivo* interaction between *Ta*VRNA1 and *Ta*CENPE1, only partial *TaCENPE1* cDNA was cloned into vector pEarley-gate 202 (pEG202) since the full gene did not express in the pEG201 vector. Similarly, the attB-flanked DNA fragment was added at both end of partial *TaCENPE1* (KIP-BiFC2-F1, KIP-BiFC2-R2r) and then cloned into pDONOR207 with the BP cloning kit. Then, the partial length of *TaCENPE1* was transferred to pEarley-gate 202 (pEG202) using the LR cloning kit for *in-vivo* protein interaction.

TaVRN-A1 in pDONR207 was fused to the N-terminal amino acid portion of YFP in the pEarleyGate201-YN vector (pEG201-YN) to test *in-vivo* interaction with *Ta*CENPE1 (1-533 aa) fused to the C-terminal amino acid portion of YFP in the pEarleyGate202-YC vector (pEG202-YC).

Empty vectors were used as negative controls for interaction with *Ta*VRN-A1 or *Ta*CENPE1 proteins. Agrobacterium tumefaciens strains (GV3101) carrying the BiFC constructs were used together with the p19 strain for infiltration of tobacco leaves (at 5 weeks old). Leaf discs were cut for BiFC for imaging 3 days after infiltration. Images were taken using a fluorescent microscope with different filters to indicate different signals/background. The images taken with a bright filter (BF) are used to indicate the GFP filter to indicate the background of the leaves infiltrated with A. tumefaciens carrying constructs. Images taken under filter GFP indicate the presence of fluorescent

proteins, and ones taken under ultraviolet filter indicate the position of the nucleus stained with the stain DAPI.

Interference of *TaCENPE1* using RNAi

The gene encoding *Ta*CENPE1 was cloned into the vector pMCG161 and the construct was transformed into wheat to characterize the function of *Ta*CENPE1. The pMCG161 is an RNAi vector used for RNA interference. When the synthetic dsRNA is introduced into plants, endogenous RNAs of *Ta*CENPE1 could be degraded. The plasmid pMCG161 contains the bar gene resistance to the herbicide and also the Bacterial chloramphenicol resistance gene to facilitate the selection of transgenic plants.

A two-step method was used to clone *Ta*CENPE1 covering 309 bp into the pMCG161 RNAi vector. In the first cloning step, the PCR fragment was amplified using primers KIP-RNAi-F1 (5'- ACTAGTGGCGCGCCGTCTGGTTTGCTGAGTGATCTCACA-3') and KIP-RNAi-R1 (5'- GCGATCGCCCTAGGACCATCGGCCTCTTGTGGCCCTGA-3'), which were cleaved at the 'inner' restriction sites *Asc*I and *Avr*II and then ligated to pMCG161 cleaved with the same restriction enzymes. In the second cloning step, the plasmid DNA resulting from the first cloning step served as a template for a second amplification using the original primers. The resulting PCR product was cleaved at the 'outer' restriction sites *Spe*I and *Sgf*I and then ligated to pMCG161 cleaved with the same restriction enzymes. This second ligation inserted the *Ta*CENPE1 fragment in pMCG161 in an inverted orientation with respect to the first cloned fragment. The construct plasmid DNA was extracted in a large scale using a commercial kit (OMEGA) for plant transformation.

Generation of transgenic wheat using *TaCENPE1*-RNAi

Transgenic wheat was generated using Bobwhite, which is a spring wheat advanced linewidely used in wheat transformation, because of its relatively high transformation efficiency. Brief procedures include:

***Constructs of transformation*:** A two-step method was used to clone *Ta*CENPE1 covering 309 bp into the pMCG161 RNAi vector. In the first cloning step, the PCR fragment was amplified using primers KIP-RNAi-F1 (5'-ACTAGTGGCGCGCCGTCTGGTTTGCTGAGTGATCTCACA-3') and KIP-RNAi-R1 (5'- GCGATCGCCCTAGGACCATCGGCCTCTTGTGGCCCTGA-3'), which were cleaved at the 'inner' restriction sites *Asc*I and *Avr*II and then ligated to pMCG161 cleaved with the same restriction enzymes. In the second cloning step, the plasmid DNA resulting from the first cloning step served as a template for a second amplification using the original primers. The resulting PCR product was cleaved at the 'outer' restriction sites *Spe*I and *Sgf* I and then ligated to pMCG161 cleaved with the same restriction enzymes.

Immature embryo used to host transgene: Immature embryos were collected from the spikes of plants approximately 14 days after anthesis. At this stage the embryos were about 0.1-1.0 mm in length and had great differentiation ability suitable for gun bombardment gene transformation.

The collected seeds were sterilized with 70% ethanol for 5 min, 20% bleach for 30 min, and then rinsed 5 times with sterile and distilled water. The sterilized seeds were excised using aseptic technique in a hood, and the excised embryos were cultured in dissection

medium (Table 3) at room temperature (22-23 °C) in the dark for 4-6 days, before they were used for bombardment.

Gene-gun bombardment: The dark-cultured embryos were transferred to high osmoticum media (bomb medium) (Table 4) four hours before bombardment. The embryos, with scutellum up, were placed in a circle in the middle of the plate. The *TaCENPE1::RNAi* construct plasmid DNAs were coated with gold particles (BioRad) and shot into the embryos by helium pressure.

Callus incubation: The bombed calli were cultured at recovery medium (Table 5) in dark for 4 weeks, then regeneration medium (Table 6) in lighted growth chamber (22-23 °C, 16 hours light/ 8 hours dark) for 6-7 weeks, followed by rooting medium (Table 7). Seedlings with healthy roots were transferred to soil in pots until plant matured.

Identification of transgenic plants: A PCR method was used to check if the transgene was in the plants that survived. A pair of primers in the regions before and after the inner restriction enzyme digestion sites was designed to check the presence of the sense fragment of *TaCENPE1*, forward sense primer 5'-ATATCCCCTAGCCACCCAAG-3', and reverse sense primer 5'-CCCCTGGGTGTGTTTCTCTA-3'. Another pair of primers in the regions before and after the outer restriction enzyme digestion sites was designed to check the presence of the antisense fragment of *TaCENPE1*, forward antisense primer 5'- GTAAGGTGTTGGGCTGGAAA-3', and reverse antisense primer 5'-CGCTCGGTGTGTCGTAGATA-3'. The PCRs were conducted in a reaction at 94 °C for 3 min, followed by 40 cycles of 94 °C for 30 sec, 53 °C for 30 sec, as well as a final

step of 72 °C for 3 min. Construct plasmid DNA was used as positive control, and the wild type plant DNAs and water were used for negative control.

Gene expression of transgenic wheat

T_1 transgenic wheat populations were tested to see if endogenous *TaCENPE1* is repressed by RNAi. The positive transgenic plants were identified using primers that designed based on sequences on vector pMCG161 for RNAi experiments. RNA will be extracted with the Trizol method, and RT-PCR approach was used to determine transcript level with SYBER Green® systems (Loukoianov *et al*. 2005). Non-transgenic plants were be used as controls.

5.3 Results

Discovery of *TaCENPE1*

Of these positive cDNA clones from the Y2H library using *Ta*VRN-A1 as bait, 3 independent ones are predicted to encode part of the protein *Ta*CENPE1, which is first defined for kinase interacting protein in *T. aestivum*.

A BLAST search using the *TaCENPE1* sequence revealed the existence of many related genes in two families, one is CENP-E1 in cereals including *Hv*CENP-E in barley, ADH94798 with 95% similarity at the protein level; the other is kinase interacting protein in *Arabidopsis* (*At*KIP1), NP_180608 with 34% identity or 51% similarity at the protein level. The original name of the gene for research work was *TaKIP1* but it was renamed as *TaCENPE1* hereafter.

Interaction sites between *Ta*VRN-A1 and *Ta*CENPE1

*Ta*VRN-A1 is a MADS box protein that contains a conserved modular structure consisting of four different functional domains: the MADS-box, the intervening region (I-region), the coiled-coil keratin-like (K-box), and the C-terminal region (Kane et al., 2005; Mandel et al., 1992; Yang et al., 2004). MADS box proteins often homodimerize or heterodimerize with other MADS box protein in the same family to gain functional diversity (Immink and Angenent, 2002; Riechmann *et al.*, 1996). Intriguingly, *Ta*CENPE1 was found to have a long region with amino acids conserved with *Ta*VRN-A1 including several leucine residues (Fig.19), suggesting that this region might be interaction sites of the two proteins.

```
TaVRN-A1    56   EFSTESCMDKILERYERYSYAEKVLVSSESEIQ-------GNWC---HEYRKLKAKVETI  105
                 +   E C  ++ E  ERY+Y +  L +   E+Q       GN      E  +L+A+V  +
TaCENPE1   271   DLEIEKCKRELEEVSERYTYGKSTLETEIGELQEVVKNLEGNLAKLSEEKLQLEAQVMDL  330

TaVRN-A1    106  QKCQKHLMGEDFESLNLKELQQLEQQLESSLKHIRSRKNQLMHESISELQKKERSLQE  163
                 ++    L   D  S  +K+LQ++ + L++ L++  S + +++ E   E ++  R L++
TaCENPE1    331  EQTSHSL---DDSSAEIKKLQKVIKDLQARLEN-DSNEKRVLEERAIEFEQVHRELED  384
```

Fig.19 Sequence alignment between *Ta*VRN-A1 and *Ta*CENPE1 proteins. *Ta*VRN-A1 is encoded by the Jagger allele. An amino acid in blue indicates a point mutation that occurred between Jagger and 2174 *Ta*VRN-A1 proteins. Conserved leucine residues are highlighted in red.

Subcellular localization and *in vivo* interaction of *Ta*VRN-A1 and *Ta*CENP1 proteins

The full length *Ta*CENPE1 was expressed by pEG101-YFP vector, and enriched yellow fluorescent signals were detected predominantly in the nucleus (Fig.20A) and partly on the cytoplasm membrane (Fig.20B). The expression pattern of *Ta*CENPE1 was exactly the same as *Ta*VRN-A1 observed in the previous study. This result indicated that *Ta*VRN-A1 and *Ta*CENPE1 have interaction sites in the cell.

*Ta*VRN-A1 into pEG201-YN vector and *Ta*CENPE1(1–513) into pEG202-YC vector and analyzed in vivo protein interactions by BiFC. When *Ta*VRN-A1-YN and *Ta*CENPE1-YC were simultaneously expressed in the same cell, yellow fluorescence was observed in the membrane with a fluorescent microscope (Fig.20C-F), but no signal was observed in the negative controls using *Ta*CENPE1 and empty vectors to express in the same cell (data not shown). These results indicated that *Ta*VRN-A1 and *Ta*CENPE1 proteins had a physical interaction in plants.

The *in vivo* protein interactions are usually conducted using agrobacterium tumefaciens strains (GV3101) to carry the BiFC constructs to infiltrate tobacco leaves that are complete at 5 weeks old. Interestingly, when the bacteria was used to infiltrated with leaves that were young and not yet complete, interaction signals were observed on two dividing nuclei in the same cell (Fig.20G), dotted interaction signals were observed on spindle fibers (Fig.20H) and two divided cells (Fig.20I).

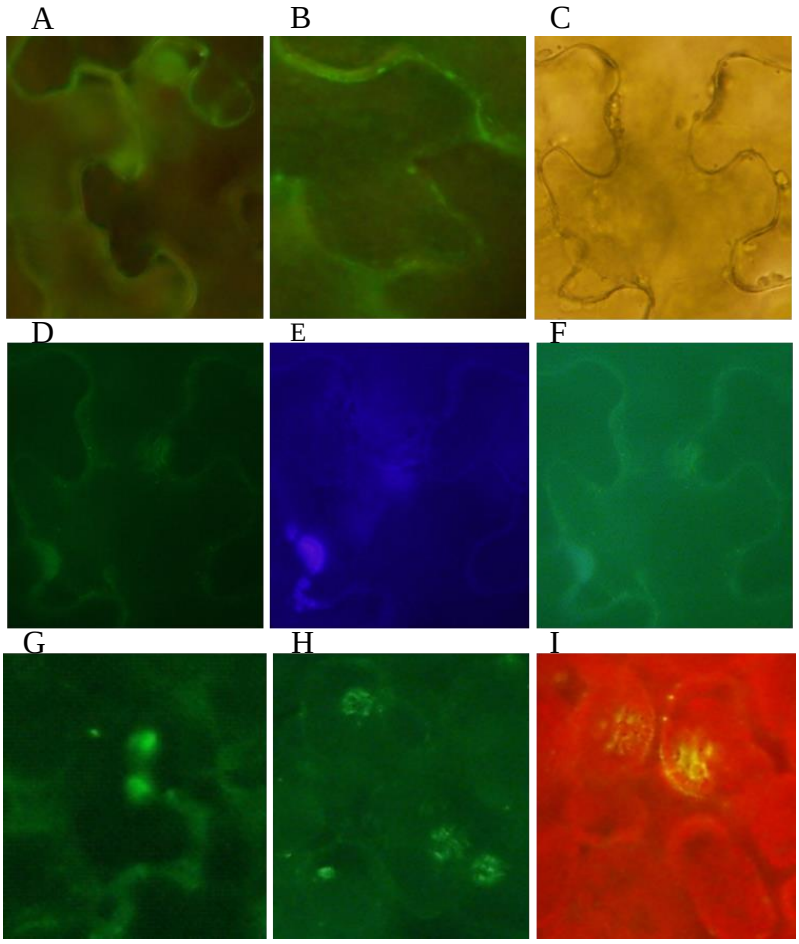
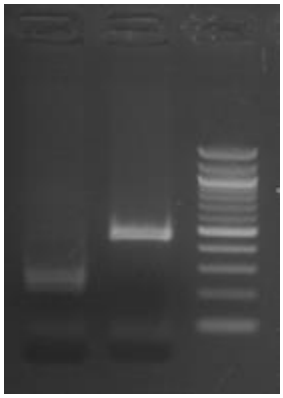


Fig.20 Subcellular localization and *in vivo* interaction of *Ta*VRN-A1 and *Ta*CENP1 proteins. (A) Circle signals representing *Ta*CENPE1-YFP protein expressed by pEG101 on nucleus. (B) Dot signals representing *Ta*CENPE1 on the cytoplasm membrane. (C)-(I) *In vivo* interaction of *Ta*VRN-A1 and *Ta*CENPE1 when the two proteins were simultaneously expressed in the same living cell in Nicotiana tabacum (tobacco) leaves. (C) No signals in cells when imaged with bright filter (BF). (D) The green fluorescent proteins resulted from the *in vivo* interaction between *Ta*VRN-A1-YN and *Ta*CENPE1-YC. YN, YFP fragment at the N-terminal end expressed from pEG201-YN vector; YC, YFP fragment at the C-terminal end expressed from pEG202-YC vector. (E) Leaf discs were cut for imaging 3 days after infiltration, with an ultraviolet filter (DAPI) to indicate the position of the nucleus stained with 4′, 6-diamidino-2-phenylindole. (F) Overlay of image D and image E to indicate that interaction signals on the nucleus. (G) Interaction signals of *Ta*VRN-A1-YN and *Ta*CENPE1-YC on two dividing nuclei in the same dividing cell. (H) Interaction signals of *Ta*VRN-A1-YN and *Ta*CENPE1-YC on spindle fibers. (I) Interaction signals in two divided young cells.

Isolation and mapping of *TaCENPE1* genes

To map the single nucleotide polymorphism in *TaCENPE1*, forward primer KIP2-F2 5'-GGGTAAAAAGGGAAGTGATGGATCG-3' and reverse primer KIP2-R2 5'-AGTTCTTGACCACTTCTTGGAGA-3' were designed to amplify a 530 bp fragment containing this SNP. Jagger PCR DNA product contained two *Aci* I digestion sites while 2174 PCR contained only one. After digestion, the Jagger PCR showed 184-, 255- and 91-bp bands, whereas the 2174 PCR showed 439- and 91-bp bands. The digested PCR products of 184-bp and 255-bp for the Jagger *TaCENPE2a* allele and 439-bp for the 2174 *TaCENPE2b* allele were distinguishable on a 2% agarose gel (Fig. 21).



Jagger 2174 M

Fig.21 A PCR marker for *TaCENPE-A1* in wheat. *TaCENPE-A1a* has 'GCGG'which can be digested with *Aci*I, whereas *TaCENPE-A1b* has 'GCTG' which cannot be digested with *Aci*I. The expected Jagger PCR product has a lower band than 2174.

The mapping of the newly developed PCR marker for *TaCENPE2* in 96 RILs of the Jagger x 2174 population showed that *TaCENPE1* was on chromosome 3A, therefore, this gene is designated *TaCENPE-A1* (Fig.22)

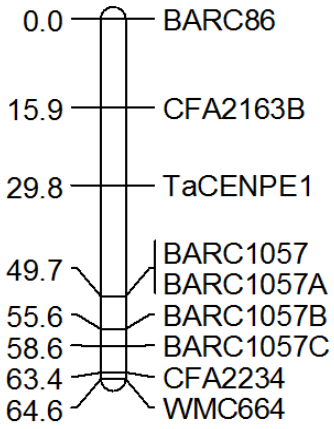


Fig.22 Chromosomal location of *TaCENPE1* in wheat. Primers specific to *TaCENPE1* were used for mapping of the gene. SSR markers linked with *TaCENPE1* are reported to locate on chromosome 3A.The genetic linkage map was constructed using MapMaker 3.0.

TaCENPE1::*RNAi* transgenic plant

In order to investigate the function of *TaCENPE1*, we used a vector of RNAi pMCG161 (RNA interference) to transform this gene into the hexaploid spring wheat variety Bobwhite. RNAi is a biological process in which RNA molecules inhibit gene expression, typically by causing the destruction of specific mRNA molecules. This method is widely used to silence gene expression in the study of gene function. The phenotype of transgenic plant of *TaCENPE1*::*RNAi* would provide solid evidence to the function of *TaCENPE1*. Two transgenic wheat plants (T20 and T33) showed the expected PCR products of transgenes for both sense and anti-sense orientation (Fig.23)

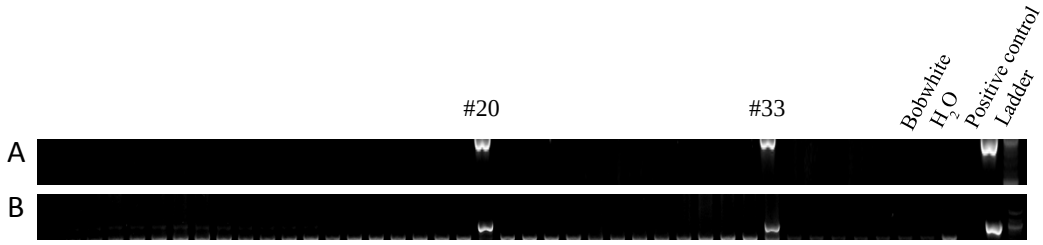


Fig.23 Screening of *TaCENPE1*::*RNAi* transgenic plants. (A) PCRs used to amplify the transgene in sense. (B) PCRs used to amplify the transgene in antisense. Plant #20 and #33 are two positive plants. Bobwhite gDNA and H_2O as negative controls, *TaCENPE1*::RNAi construct as positive control.

Functional understanding of *TaCENPE1* in transgenic wheat

Since only a few seeds were harvested, no T_1 population was generated to test phenotypes but the phenotypes of a few T_1 plants were presented herein. A T_2 population generated from a few T_1 plants is growing and will be tested for phenotypes. Positive plants in the T_2 populations have been identified for gene expression.

For T_2 populations of both T20-3 and T33-4, a ratio of 3:1 segregation was observed for plants carrying the *TaCENPE1*::RNAi construct and plants without this construct. The *TaCENPE1* transcript levels in leaf samples of the seedling stage plants were determined using a quantitative RT-PCR (Fig.24). The transcriptional levels of endogenous *TaCENPE1* in the plants carrying *TaCENPE1*::RNAi construct were significantly reduced compared with the plants without the construct in the same population. For T20-3 population, positive plants showed only 2/5 of the transcriptional level of the controls, and T33-4 positive plants showed only 1/2 of transcriptional level of the corresponding

controls. This result indicates that the *TaCENPE1*::RNAi construct efficiently decreased the *TaCENPE1* transcriptional level.

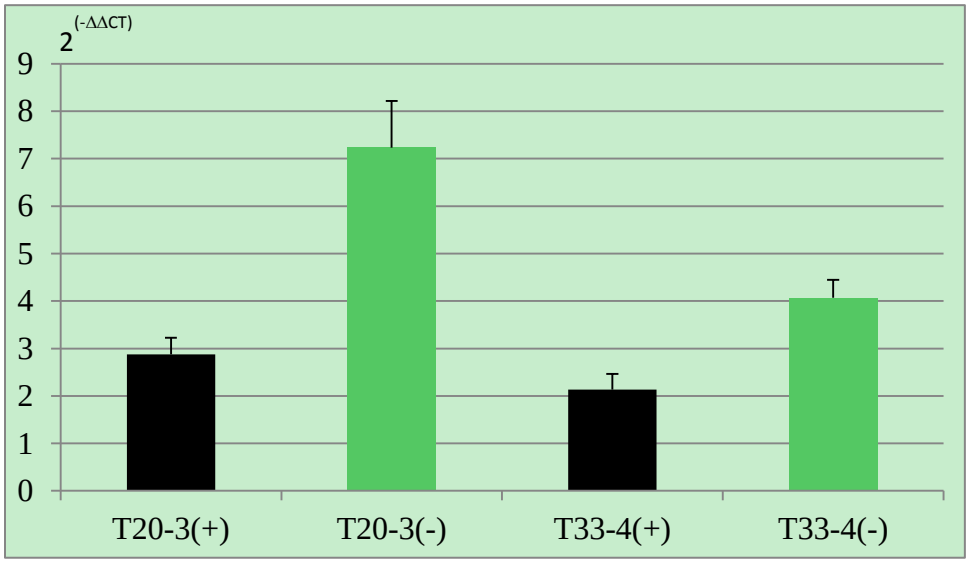


Fig.24 *TaCENPE1* expression level in T2 population of *TaCENPE1*::*RNAi* transgenic wheat. Transcript levels of *TaCENPE1* in T_2 populationare shown using the values calculated by the $2^{(-\Delta\Delta CT)}$ method, where C^T is the threshold cycle, and actin was used as an endogenous control. For T_2 population of T20-3: plants carrying RNAi construct (#1, 5, and 8) have lower expression level than plants without the construct (#12, 34, and 35). For T_2 population of T33-4: most plants carrying RNAi construct (#1 and 7) have lower expression level than plants without the construct (#6, 9, and 13). Only the expression level for plant #3 is not depressed obviously, maybe the RNAi construct breakdown during cell division and lost function as a result.

As shown in Fig.25A, the plant height of the transgenic wheat was significantly reduced. Several positive transgenic plants showed similar phenotypes. Another striking phenotype in transgenic wheat plants was that more flowerets per spikelet were

developed (Fig.25B and 25C) compared with wild type (Fig.25D). There usually are 6-10 flowerets each spikelet on the middle-lower part of a spike and 1-5 of them are visible and can be developed to seed. However, in the transgenic plants, up to 10 flowerets were visible (Fig.25B and 25C), and more grains per spikelet were produced in the *TaCENPE1*::*RNAi* transgenic wheat.

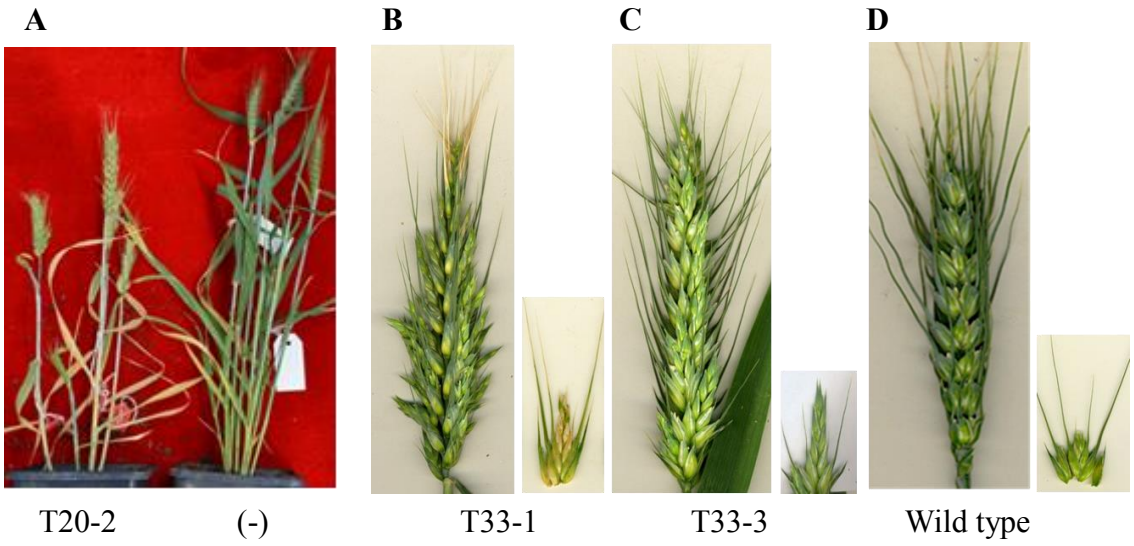


Fig.25 Phenotypes of *TaCENPE1*::*RNAi* transgenic wheat. (A)The plant height of positive transgenic wheat T20-2 is significantly reduced caompared with the wild type. (B) and (C) Spike and floweret structure for positive transgenic wheat T33-1 and T33-3. (D) Spike and floweret structure for wild type wheat as control.

5.4 Discussion

CENP-E is a kinesin-like protein that binds to kinetochores during mitosis in organisms (Yen et al., 1991, 1992; Cooke et al., 1997; Yao et al., 1997). Mitosis and meiosis and two ways to divide and reproduce cells. Meiosis is a special type of cell division necessary for sexual reproduction in plants, in which homologous chromosome pairing,

synapse and recombination occur. Mitosis is the process in which the chromosomes in the cell nucleus are separated into two identical sets of sister chromosomes, each in its own nucleus. Mitosis is followed immediately by cytokinesis, which divides the cytoplasm, organelles, and cell membrane, and later karyokinesis, which divides the nucleus, dividing the cell into two new cells containing roughly equal shares of these cellular components. Mitosis and cytokinesis together define the mitotic (M) phase of the cell cycle—the division of the mother cell into two daughter cells, genetically identical to each other and to their parent cell. During mitosis, the two identical chromatids or sister chromatids are held together by a specialized region of the chromosome: a DNA sequence called the centromere. A kinetochore is a complex protein structure that assembles on the centromere and links the chromosome to microtubule polymers from the mitotic spindle during mitosis and meiosis. In human cells, the centromeric proteins include constitutive proteins, such as CENP-A, -B, -C, and -H, that are present at the centromere throughout the cell cycle, and transient proteins that appear after the onset of mitotic phase, such as CENP-E and -F.

CENP-E of a genomic sequence similar to ZW10 of Drosophila was identified in *Arabidopsis* (Starr et al 1997) and in barley (Hoopen et al. 2002). The previous evidence suggests that CENP-E like protein may also occur at plant centromeres (cross reactivity of antibodies agaisnt CENP-E (Yen 1991) with the kinetochores of Vicia faba and/or Hordeum vulgare. *Arabidopsis* genome revealed 61 sequences encoding proteins with a kinesin motor domain (Teddy and Day 2001). Some of the Arabidipsis kinesin sequences grouped with CENP-E according to the maximun likelihood method (Lawrence et al 2002).

By the sequence alignment, *Ta*CENP-E1 could have a direct physical binding with *Ta*VRN-A1 protein. The *in vivo* interaction approach also demonstrated that TaCENPE-A1 had interaction with *Ta*VRN-A1 protein. In view of sequence, the interaction of the two proteins in different families is due to the presence of conserved sequences including leucine rich domain between them.

It was reported that interference of CENP-E function reduces tension across the centromere, increases the incidence of spindle pole fragmentation, and results in monooriented chromosomes approaching abnormally close to the spindle pole. Recent efforts to eliminate CENP-E expression by an antisense strategy yielded a similar phenotype (Yao et al., 2000). These studies have been interpreted to mean that CENP-E is required for stable attachment of kinetochore microtubules (kMts) and chromosome congression to the spindle equator (Schaar et al., 1997; Yao et al., 2000). The presence of bipolar aligned chromosomes in the absence of CENP-E function has been attributed to chromosomes located near the center of the forming spindle during early prometaphase. The locations of the interaction signals in nuclei and spindle suggested that both *Ta*CENPE1 and *Ta*VRN-A1 move with chromosomes during cell division.

Plant growth rate at the cellular level is regulated by the combined activity of two processes: cell proliferation and expansion, which involves the integration of signals from the intrinsic genetic programs with environmental cues. Gibberellins (GA) are one of endogenous hormones that play a central role in regulating responses of plant growth and development to environment conditions (Olszewski et al. 2002, Achard et al. 2006). It is well known that GA promote cell expansion and proliferation by repressing DELLA proteins restraining cell division activity in the shoot meristematic zone and early

developing leaves, which enhance the level of kinase interacting protein *Kip-related protein 2* (*KRP2*), an inhibitor of the cell production (Olszewski et al. 2002; Peng et al., 1997; Achard et al. 2009). GA signaling, by modulating the expression levels of KRP2, controls the cell proliferation rate, resulting in dwarf and branching spikelets. A recent study on wheat reported both *VRN1* and GA are required in the wheat shoot apical meristem for the up-regulation of SOC1-1 and LFY and for the acceleration of spike development (Pearce et al., 2013). Put these studies together, we establish a model for floweret development, in which *Ta*CENPE1, which may play a functional role like KRP2, regulate plant height and spike development through VRN1 and GA pathway in wheat. Further functional understanding of these proteins is of importance because the pathway is related to grain yield.

CHAPTER VI

DISCUSSION AND CONCLUSION

6.1 New approaches in identifying genes for important traits

A genetic approach is usually utilized to identify genes associated with important traits by using genome-wide markers to construct higher density genetic linkage groups for the mapping population. For developmental genes in winter wheat, three major QTLs (*vrn-A1*, *PPD-D1*, and *vrn-D3*) have been genetically associated with developmental phases (Chen et al., 2010). However, The tri-loci model of selection for winter wheat development is not conclusive, because the current model does not match up with phenotypes in some of the tested cultivars, such as Fannin that is extremely early and Trego that is extremely late in development, suggesting that more genes are involved in developmental trait of winter wheat. The missing part in the winter wheat development pathway could be found using further genetic research. It is possible that some genes is in gaps between the linkage groups mapped in the winter wheat population, since the previous SSR (simple sequence repeat) markers did not sufficiently cover the whole genome. It is also possible that some genetic factors controlling winter wheat

development were not detected in the previous Jagger x 2174 population because the two parental lines may have the same allele for those unknown gene loci. A new doubled haploid (DH) population using two winter wheat cultivars, Duster and Billings that have the same allele at each of the three known genes but have a significant difference in developmental processes, has been generated to identify new genes/QTLs in the Duster and Billings DH population.

The missing part in the winter wheat development pathway can also be found by using new research strategies. In comparison with research on a functional gene at the DNA level which characterizes transcription and promotion of genes, micro RNA and protein studies have received less attention. These new research areas have been considered as increasingly important for understanding the molecular mechanism of biological process (Pawson and Nash, 2003), such as plant development.

6.2 A flowering pathway via *Ta*miR1123

A new model was proposed to explain mechanisms underlying spring growth habit by *Ta*miR1123 present in the promoter of the *Vrn-A1a* allele without vernalization requirement. In this model, without vernalization the spring *Vrn-A1a* allele can be expressed much earlier than other spring alleles due to the presence of *Ta*miR1123 in the *Vrn-A1a* promoter. In previous studies, all of spring *Vrn-A1* alleles, regardless of mutation types in the promoters of the insertion of the MITE or retrotransposal elements or deletions, were believed to lose the binding sites of repressors, but the previous model cannot explain why those plants that carry MITE_VRN produce more *Vrn-A1a*

transcripts and flowered earlier than those plants that carry other mutant alleles with deletions in their promoters or intron 1 (Yan *et al*. 2004, Fu *et al*., 2005).

In this study, we have experimentally demonstrated that the level of *Ta*miR1123 was positively correlated with the transcriptional levels of *Vrn-A1a*. Therefore, it is likely that *Ta*miR1123 binds to the same sequence present in MITE_VRN in the promoter of the *Vrn-A1a* to induce expression of *Vrn-A1a*. However, the possibility that *Ta*miR1123 is the by-product of transcription of *Vrn-A1a* that is used to enhance the expression of *Vrn-A1a* cannot be excluded. A MULE (Mutator-like element) could harbor the promoters of *Vrn-A1a* for transcription, based on characteristics of the sequence and target site duplication of MITE_VRN. The potential mechanism could explain why *Vrn-A1a* is linked to a much stronger expression than other *Vrn-A1* alleles.

*Ta*miR1123 can be released from MITE_VRN that is expressed as RNA. The released *Ta*miR1123 can induce the expression of the genes that have a target of *Ta*miR1123. Moreover, MITE_VRN is active and movable, which can insert in the promoter or regulatory site of a functional gene. In this study, we have found several cDNA clones or sequence contigs in the available databases that have identical or similar sequences as *Ta*miR1123 or MITE_VRN. Many more regulatory sites with such a sequence would be found as the wheat genome is completely sequenced. The expression patterns or coding sequences of these targeted genes could be altered due to the insertion of MITE_VRN or the cognition of *Ta*miR1123, which would form a dynamic gene regulatory network governed by *Vrn-A1a* in plant development. The *VRN1* region was reported to have association with multiple traits including vernalization, cold hardening, and the

development of rosette. This study provided a machinery explanation for the complex association among some of the multiple phenotypes.

6.3 Protein interaction pathways for flowering time in wheat

Using VRN1 as bait, we identified HOX and CENPE1 proteins from the yeast-two-hybrid (Y2H) library and confirmed the interaction using *in vitro* approach, in addition to other proteins such as VRT2 and SOC1that have been reported to involve in development processes. We have used pull down assay for *in vitro* interaction and the bimolecular fluorescence complementation (BiFC) system for *in vivo* interaction to confirm that these proteins have real interaction. These interacting proteins may be a necessary component in different biological function pathways.

The HOX proteins are known to function as an on-off switch in controlling development including specialization of regional identities along the anterio/posterior axis in a wide range of phyla in animals, and MADS-box proteins, such as wheat *VRN1* and *Arabidopsis FLC*, are known to act as the floral switches for the transition from the vegetative organ to reproductive organs in plants. The striking parallels in function involving homeosis between animal HOX proteins and plant MADS-box proteins has provided the possibility that plant HOX may direct floral architecture (Ng and Yanofsky, 2001). This study presented the first example that plant HOX regulated flowering time by directing homeosis for organ identity from the apical meristem to the stem.

6.4 Future research perspectives

With recent releases of wheat survey genome sequences, we have entered a post-genomic era that should focus on identification of functions of genes and gene products (proteins)

and application of multiple functional genes/proteins in wheat. In comparison with previous research on a marker associated with a QTL, future researches will focus on functional characterization of the candidate genes associated with QTLs.

Many miRNA molecules have been found to play important roles in plant responses to abiotic and biotic stresses as well as signal transduction. In this study, several gene sequences have been identified to be direct targets of MITE_VRN or *Ta*miR1123, but functions of these potential targets are not known. A future research will be to establish network via *Ta*miR1123 and MITE_VRN governing multiple traits in wheat. In addition, the MITE_VRN derived *Ta*miR1123 characteristics needs to utilize in wheat breeding. Firstly, the genomic and EST sequences flanking MITE_VRN can be used to design specific primers to map members of the MITE_VRN family. Any phenotypic variation associated with a MITE_VRN marker can be suggested to link with development regulated by *Ta*miR1123 donated from MITE_VRN in the *VRN-A1a*. These potential applications of MITE_VRN and *Ta*miR1123 in wheat need to be developed in the future studies.

*Ta*HOX1 protein was suggested to function in the flowering pathway, based on its direct interaction with *Ta*VRN-A1 and genetic association with flowering time. However, the role of *Ta*HOX1 in regulating flowering time needs to confirm in transgenic plants, and further experiments are also needed to test if the *Ta*VRN-A1-*Ta*HOX1 protein complex is sensed by low temperature in winter wheat. We hypothesize that when a protein like VRN1 interacts with multiple proteins such as HOX1, CENPE1, SOC1, and VRT2, the protein interactions have resulted in a competition for signals from different pathways in

plants. It would be intriguing that a model is established to understand how these interacting proteins are balanced by binding strength or anything else.

The mechanism how plants measure the vernalization duration in *Arabidopsis* is unknown yet, but a model is proposed to understand this mechanism. In the model, some gene(s) is gradually up-regulated or down-regulated during the vernalization process to reach a threshold at which the vernalized plants acquit the competence to flower. The vernalization duration depends on when the plant reach the maximum vernalization effects or any further vernalization no longer makes the plant to flower earlier but may delay flowering due to paradox effects of low temperature on development and growth. In the threshold model, the cold sensor could simply be an enzyme that measures the duration of the cold. In *Arabidopsis*, *VIN3* could be such an enzyme based on its unique expression pattern. One challenge research for the future is to understand the nature of the measurement of cold duration at a protein or enzyme level in wheat.

Current scientific consensus shows that average global temperature rose in the 20th century and will continue its increase by 3 to 5 °C by the end of the 21st century. Winter wheat requiring a few weeks of low temperature at lower than 8 °C for proper flowering time is more vulnerable to elevated temperature, because a shortened low-temperature duration may result in a failed or incomplete vernalization. The precise characterization of interacting molecules at micro RNA or protein intersection will lead to gene networks that can be used to regulate development processes in wheat. This knowledge, in turn, has a direct impact on our ability to develop new plant cultivars that are responsive to the increasing challenges of climate change.

REFERENCES

Abe, M., Kobayashi, Y., Yamamoto, S., Daimon, Y., Yamaguchi, A., Ikeda, Y., and Araki, T. (2005) FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science*, **309**, 1052-1056.

Achard, P., Cheng, H., De Grauwe, L., Decat, J., Schoutteten, H., Moritz, T., and Harberd, N. P. (2006) Integration of plant responses to environmentally activated phytohormonal signals. *Science*, **311**, 91-94.

Achard, P., Gusti, A., Cheminant, S., Alioua, M., Dhondt, S., Coppens, F., and Genschik, P. (2009) Gibberellin Signaling Controls Cell Proliferation Rate in *Arabidopsis*. *Current biology*, **19**, 1188-1193.

Amasino, R. (2004) Vernalization, competence, and the epigenetic memory of winter. *Plant Cell*, **16**, 2553-2559.

Ambros, V. (2004) The functions of animal microRNAs. *Nature*, **431**, 350-355.

Angel, A., Song, J., Dean, C. and Howard, M. (2011) A polycomb-based switch underlying quantitative epigenetic memory. *Nature*, **476**, 105-108.

Aoyama, T., Dong, C.-H., Wu, Y., Carabelli, M., Sessa, G., Ruberti, I., Morelli, G. and Chua, N. H. (1995) Ectopic expression of the *Arabidopsis* transcription activator Athb-1 alters leaf cell fate in tobacco. *Plant Ceil*, **7**, 1773-1785.

Ariel, F.D., Manavella, P.A., Dezar, C.A. and Chan, R.L. (2007) The true story of the HD-Zip family. *TRENDS in Plant Sci*ence, **12**, 419-426.

Båga, M., Fowler, D. B., and Chibbar, R. N. (2009) Identification of genomic regions determining the phenological development leading to floral transition in wheat (Triticum aestivum L.). *Journal of experimental botany*, **60**, 3575-3585.

Baloch, D. M., Karow, R. S., Marx, E., Kling, J. G., and Witt, M. D. (2003) Vernalization studies with Pacific Northwest wheat. *Agronomy journal*, **95**, 1201-1208.

Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281-297.

Beales, J., Turner, A., Griffiths, S., Snape, J. W., and Laurie, D. A. (2007) A pseudo-response regulator is misexpressed in the photoperiod insensitive Ppd-D1a mutant of wheat (Triticum aestivum L.). *Theoretical and Applied Genetics*, **115**, 721-733.

Bennett, M.D. and Leitch, I.J. (1995) Nuclear-DNA amounts in angiosperms. *Annals. Bot.*, **76**, 113-176.

Berry, G. J., Salisbury, P. A., and Halloran, G. M. (1980) Expression of vernalization genes in near-isogenic wheat lines: duration of vernalization period. *Ann. Bot*, **46**, 235-241.

Blake, N. K., Lanning, S. P., Martin, J. M., Doyle, M., Sherman, J. D., Naruoka, Y., and Talbert, L. E. (2009) Effect of variation for major growth habit genes on maturity and yield in five spring wheat populations. *Crop science*, **49**, 1211-1220.

Bonnin, I., Rousset, M., Madur, D., Sourdille, P., Dupuits, C., Brunel, D., and Goldringer, I. (2008). *FT* genome A and D polymorphisms are associated with the variation of earliness components in hexaploid wheat. *TAG Theoretical and Applied Genetics*, **116,** 383-394.

Brennecke, J., Stark, A., Russell, R.B. and Cohen, S.M. (2005) Principles of MicroRNA–Target Recognition. *PLoS Biol*, **3**, e85.

Brooking, I. R. (1996). Temperature response of vernalization in wheat: a developmental analysis. *Annals of Botany*, **78**, 507-512.

Cao, S. and Yan, L. (2013) Construction of a high-quality Yeast two-hybrid (Y2H) library and its application in identification of interacting proteins with key vernalization regulator *Ta*VRN-A1 in wheat. *BMC Res. Note*, **6**, 81-86.

Chaudhry, S. H. (1986) Initiation of the reproductive stage in wheats carrying in winterhardiness. *Dissertation Abstracts International* [Section B], **46**, 2129-2130.

Chen, Y., Carver, B., Wang, S., Cao, S., and Yan, L. (2010) Genetic regulation of developmental phases in winter wheat. *Molecular Breeding*, **26**, 573-582.

Chouard, P. (1960) Vernalization and its relation to dormancy. Annu. Rev. *Plant Physiol.*, **11**, 191-238.

Chu, C., Tan, C., Yu, G., Zhong, S., Xu, S.S., and Yan, L. (2011) A novel retrotransposon inserted in the dominant Vrn-B1 allele confers spring growth habit in tetraploid wheat (*Triticum turgidum* L.). *G3: Genes, Genomes, Genetics*, **1**, 637-645.

Cooke, C. A., Schaar, B., Yen, T. J., and Earnshaw, W. C. (1997) Localization of CENP-E in the fibrous corona and outer plate of mammalian kinetochores from prometaphase through anaphase. *Chromosoma*, **106**, 446-455.

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakountis, A., Farrona, S., Gissot, L., Turnbull, C. and Coupland, G. (2007) FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. *Science*, **316**, 1030-1033.

Coustham, V., Li, P., Strange, A., Lister, C., Song, J. And Dean, C. (2012) Quantitative modulation of Polycomb silencing underlies natural variation in vernalization. *Science*, **337**, 584-587.

Crofts, H.J. (1989) On defining a winter wheat. *Euphytica*, **44**, 225-234.

Danyluk, J., Kane, N. A., Breton, G., Limin, A. E., Fowler, D. B., and Sarhan, F. (2003) *TaVRT-1*, a putative transcription factor associated with vegetative to reproductive transition in cereals. *Plant Physiology*, **132**, 1849-1860.

Das, L. and Martienssen, R. (1995) Site-selected transposon mutagenesis at the hcf106 locus in maize. *Plant Cell*, **7**, 287-294.

Diallo, A., Kane, N., Agharbaoui, Z., Badawi, M., and Sarhan, F. (2010) Heterologous expression of wheat VERNALIZATION 2 (TaVRN2) gene in *Arabidopsis* delays flowering and enhances freezing tolerance. *PloS one*, **5**, e8690.

Díaz, A., Zikhali, M., Turner, A. S., Isaac, P., and Laurie, D. A. (2012) Copy number variation affecting the Photoperiod-B1 and Vernalization-A1 genes is associated with altered flowering time in wheat (Triticum aestivum). *PLoS One*, **7**, e33234.

Ding, X., Richter, T., Chen, M., Fujii, H., Seo, Y. S., Xie, M., Zheng, X., Kanrar, S., Stevenson, R. A., Dardick, C., Li, Y., Jiang, H., Zhang, Y., Yu, F., Bartley, L. E., Chern, M., Bart, R., Chen, X., Zhu, L., Farmerie, W. G., Gribskov, M., Zhu, J.-K., Fromm, M. E., Ronald, P. C., and Song, W. Y. (2009) A rice kinase-protein interaction map. *Plant Physiology*, **149**, 1478-1492.

Distelfeld, A., and Dubcovsky, J. (2010) Characterization of the maintained vegetative phase deletions from diploid wheat and their effect on VRN2 and FT transcript levels. *Molecular Genetics and Genomics*, **283**, 223-232.

Distelfeld, A., Tranquilli, G., Li, C., Yan, L., and Dubcovsky, J. (2009a) Genetic and molecular characterization of the *VRN2* loci in tetraploid wheat. *Plant physiology*, **149**, 245-257.

Distelfeld, A., Li, C., and Dubcovsky, J. (2009b) Regulation of flowering in temperate cereals. *Current opinion in plant biology*, **12,** 178-184.

Dubcovsky, J., Loukoianov, A., Fu, D., Valarik, M., Sanchez, A. and Yan, L. (2006) Effect of photoperiod on the regulation of wheat vernalization genes *VRN1* and *VRN2*. *Plant Mol. Biol*., **60**, 469-480.

Evans, L. T. (1987) Short day induction of inflorescence initiation in some winter wheat varieties. *Functional Plant Biology*, **14**, 277-286.

Ferguson, A. A., and Jiang, N. (2012) Mutator-like elements with multiple long terminal inverted repeats in plants. *Comparative and functional genomics*, **2012**, 695827-695827.

Ferguson, A.A. and Jiang, N. (2011) Pack-MULEs: Recycling and reshaping genes through GC-biased acquisition. *Mob. Genet. Elements*, **1**, 135-138.

Flood, R. G., and G. M. Halloran. (1986) The influence of genes for vernalization response on development and growth in wheat. *Annals of botany*, **58.4**, 505-513.

Fowler, D. B., and Gusta, L. V. (1977) Influence of fall growth and development on cold tolerance of rye and wheat. *Canadian Journal of Plant Science*, **57**, 751-755.

Fowler, D. B., Gusta, L. V., and Tyler, N. J. (1981) Selection for winterhardiness in wheat. Screening methods. *Crop Science*, **21**, 896-901.

Frank, W., Phillips, J., Salamini, F., Bartels, D. (1998) Two dehydration-inducible transcripts from the resurrection plant Craterostigma plantagineumencode interacting homeodomain-leucine zipper proteins. *Plant Journal*, **15**, 413–421.

Fu, D., Szucs, P., Yan, L., Helguera, M., Skinner, J.S., von Zitzewitz, J., Hayes, P.M. and Dubcovsky, J. (2005) Large deletions within the first intron in VRN-1 are associated with spring growth habit in barley and wheat. *Mol. Genet. Genomics*, **273**, 54-65.

Gill, B.S., Appels, R., Botha-Oberholster, A.M., Buell, C.R., Bennetzen, J.L., Chalhoub, B., Chumley, F., Dvořák, J., Iwanaga, M., Keller, B., Li, W., McCombie, W.R., Ogihara, Y., Quetier, F. and Sasaki, T. (2004) A workshop report on wheat genome sequencing: International Genome Research on Wheat Consortium. *Genetics*, **168**, 1087-1096.

Golovnina, K.A., Kondratenko, E.Y., Blinov, A.G. and Goncharov, N.P. (2010) Molecular characterization of vernalization loci *VRN1* in wild and cultivated wheats. *BMC Plant Biol.*, **10**, 168-182.

Heo, J.B. and Sung, S. (2011) Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science**, 331**, 76-79.

Hjellström, M., Olsson, A. S. B., Engström, P., and Söderman, E. (2003) Constitutive expression of the water deficit-inducible homeobox gene *ATHB7* in transgenic *Arabidopsis* causes a suppression of stem elongation growth. *Plant Cell Environ,* **26,** 1127–1136.

Humphreys, M.W., Yadav, R.S., Cairns, A.J., Turner, L.B., Humphreys, J. and Skot, L. (2006) A changing climate for grassland research. *New Phytologist*, **169**, 9-26.

Immink, R. G., Gadella, T. W., Ferrario, S., Busscher, M., and Angenent, G. C. (2002) Analysis of MADS box protein–protein interactions in living plant cells. *Proceedings of the National Academy of Sciences*, **99**, 2416-2421.

Johannesson, H., Wang, Y., & Engström, P. (2001). DNA-binding and dimerization preferences of *Arabidopsis* homeodomain-leucine zipper transcription factors in vitro. *Plant molecular biology*, **45**, 63-73.

Kamran, A., Randhawa, H. S., Yang R. C., and Spaner, D. (2014) The effect of *VRN1* genes on important agronomic traits in high-yielding Canadian soft white spring wheat. *Plant Breeding*, **133**, 321-326.

Kane, N.A., Danyluk, J., Tardif, G., Ouellet, F., Laliberte, J.F., Limin, A.E., Fowler, D.B. and Sarhan, F. (2005) *TaVRT-2*, a member of the StMADS-11 clade of flowering repressors, is regulated by vernalization and photoperiod in wheat. *Plant Physiol.*, **138**, 2354-2363.

Kerr, R.A. (2007) Climate change: Global warming is changing the world. *Science*, **316**, 188-190.

Kirby, E. J. M., Spink, J. H., Frost, D. L., Sylvester-Bradley, R., Scott, R. K., Foulkes, M. J., Clare, R. W., and Evans, E. J. (1999) A study of wheat development in the field: analysis by phases. *European Journal of Agronomy*, **11**, 63–82.

Klages, K. H. (1926) Metrical attributes and the physiology of hardy varieties of winter wheat. *Agronomy Journal*, **18**, 529-566.

Kuchel, H., Hollamby, G., Langridge, P., Williams, K., and Jefferies, S. P. (2006) Identification of genetic loci associated with ear-emergence in bread wheat. *Theoretical and Applied Genetics*, **113**, 1103-1112.

Law, C. N. and Worland, A. J. (1997) Genetic analysis of some flowering time and adaptive traits in wheat. *New Phytologist*, **137**, 19–28.

Law, C. N., A. J. Worland, and B. Giorgi. (1976) The genetic control of ear-emergence time by chromosomes 5A and 5D of wheat. *Heredity*, **36.1**, 49-58.

Law, C. N., and M. S. Wolfe. (1966) Location of genetic factors for mildew resistance and ear emergence time on chromosome 7B of wheat. *Canadian Journal of Genetics and Cytology*, **8.3**, 462-470.

Lawrence, C. J., Malmberg, R. L., Muszynski, M. G., and Dawe, R. K. (2002) Maximum likelihood methods reveal conservation of function among closely related kinesin families. *Journal of molecular evolution*, **54**, 42-53.

Levy, Y.Y., Mesnage, S., Mylne, J.S., Gendall, A.R. and Dean, C. (2002) Multiple roles of *Arabidopsis VRN1* in vernalization and flowering time control. *Science*, **297**, 243-246.

Li, C., and Dubcovsky, J. (2008) Wheat FT protein regulates VRN1 transcription through interactions with FDL2. *The Plant Journal*, **55**, 543-554.

Li, G., Yu, M., Fang, T., Cao, S., Carver, B. F., and Yan, L. (2013) Vernalization requirement duration in winter wheat is controlled by TaVRN-A1 at the protein level. *The Plant Journal*, **76**, 742-753.

Lisch, D. (2002) Mutator transposons. *Trends Plant Sci.*, **7**, 498-504.

Liu, S., Rudd, J.C., Bai, G., Haley, S.D., Ibrahim A., and Xue, Q. (2014) Molecular markers linked to genes important for Hard Winter wheat production and marketing in the US Great Plaints. *Crop Science*

Loukoianov A, Yan L, Blechl A, Sanchez A, and Dubcovsky J (2005) *Plant Physiol* **138**, 2364–2373.

Manak, J.R. and M.P. Scott. (1994) A class act: Conservation of homeodomain protein functions. *Development* (Suppl.) 61–77.

Mandel, M. A., Gustafson-Brown, C., Savidge, B., and Yanofsky, M. F. (1992). Molecular characterization of the *Arabidopsis* floral homeotic gene APETALA1. *Nature*, **163**, 273-277.

Mcintosh, R. A. (1983) A catalogue of gene symbols for wheat (1983 edition). Proceedings of the 6th International Wheat Genetics Symposium, November 28 -

December 3, 1983, Kyoto, Japan. Edited by S. Sakamoto. Kyoto University, Kyoto, Japan. 1197-1254.

McMaster, G. S. (2009) Development of the wheat plant. In: Carver, B. F. (ed) *Wheat: science an trade*. Blackwell, IA 31-50.

McMichael, P. (2001) The impact of globalization, free trade and technology on food and nutrition in the new millennium. *Proceedings of the Nutrition Society*, **60**, 215-220.

Meijer, A. H., Scarpella, E., Dijk, E. L., Qin, L., Taal, A. J., Rueb, S., and Hoge, J. H. C. (1997). Transcriptional repression by *Oshox1*, a novel homeodomain leucine zipper protein from rice. *The Plant Journal*, **11**, 263-276.

Michaels, S.D. and Amasino, R.M. (1999) FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell*, **11**, 949-956.

Miglietta, F., Tanasescu, M. and Marica, A. (1995) The expected effects of climate change on wheat development. *Global Change Biol.*, **1**, 407-415.

Murai, K., Miyamae, M., Kato, H., Takumi, S. and Ogihara, Y. (2003) *WAP1*, a wheat *APETALA1* homolog, plays a central role in the phase transition from vegetative to reproductive growth. *Plant Cell Physiol*., **44**, 1255-1265.

Ng, M. and Yanofsky, M.F. (2001) Function and evolution of the plant MADS-box gene family. *Nature Reviews Genetics*, **2:** 186-195.

Ng, M., and Yanofsky, M. F. (2000). Three ways to learn the ABCs. *Current opinion in plant biology*, **3**, 47-52.

Olszewski, N., Sun, T. P., and Gubler, F. (2002) *Gibberellin signaling: biosyn* (Doctoral dissertation, thesis, catabolism, and response pathways. *Plant Cell*, **14** (Suppl), 61-80.

Parrish, J. R., Gulyas, K. D., and Finley Jr, R. L. (2006) Yeast two-hybrid contributions to interactome mapping. *Current Opinion in Biotechnology* **17**, 387-393.

Pawson, T., and Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains. *Science*, **300**, 445-452.

Pawson, T., and Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains. *Science* **300**,445-452.

Pearce, S., Vanzetti, L. S., and Dubcovsky, J. (2013) Exogenous gibberellins induce wheat spike development under short days only in the presence of VERNALIZATION1. *Plant physiology*, **163**, 1433-1445.

Peng, J., Carol, P., Richards, D. E., King, K. E., Cowling, R. J., Murphy, G. P., and Harberd, N. P. (1997) The *Arabidopsis* GAI gene defines a signaling pathway that negatively regulates gibberellin responses. *Genes & development*, **11**, 3194-3205.

Petersen, K., Didion, T., Andersen, C. H., and Nielsen, K. K. (2004) MADS-box genes from perennial ryegrass differentially expressed during transition from vegetative to reproductive growth. *Journal of plant physiology*, **161**, 439-447.

Pidal, B., Yan, L.,Fu, D,. Zhang, F., Tranquilli, G., Dubcoveky, J. (2009) The CArG-boxlocated upstream from the transcriptional start of wheat vernalization gene*VRN1* is not necessary for the vernalization response, *J. Hered*. **100**, 355–364.

Pugsley, A.T. (1971) A genetic analysis of the spring-winter habit of growth in wheat. *Australian Journal of Agricultural Research*, **22**, 21-31.

Reddy, A. S., and Day, I. S. (2001) Kinesins in the *Arabidopsis* genome: a comparative analysis among eukaryotes. *BMC genomics*, **2**, 2.

Reeves, P.H. and Coupland, G. (2001) Analysis of flowering time control in *Arabidopsis* by comparison of double and triple mutants. *Plant Physiol.*, **126**, 1085-1091.

Riechmann, J. L., Krizek, B. A., and Meyerowitz, E. M. (1996) Dimerization specificity of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. *Proceedings of the National Academy of Sciences*, **93**, 4793-4798.

Roberts D. W. A., and MacDonald, M. D. (1984) Evidence for the multiplicity of alleles at *Vrn1*, the winter-spring habit locus in common wheat. *Canadian journal of genetics and cytology*, **26**, 191-193.

Roberts, D. W. A. (1982) Induction of prostrate growth habit in winter wheat under artificial conditions. *Canadian Journal of Botany*, **60**, 1-6.

Roberts, D. W. A. (1986) Chromosomes in 'Cadet' and 'Rescue' wheats carrying loci for cold hardiness and vernalization response. *Canadian journal of genetics and cytology*, **28**, 991-997.

Roberts, D. W. A., and Larson, R. I. (1985) Vernalization and photoperiodic responses of selected chromosome substitution lines derived from 'Rescue', 'Cadet', and 'Cypress' wheats. *Canadian Journal of Genetics and Cytology*, **27**, 586-591.

Roberts, D.W.A. (1990) Identification of loci on chromosome 5A of wheat involved in control of cold hardiness, vernalization, leaf length, rosette growth habit, and height of hardened plants. *Genome*, **33**, 247-259.

Salmon, S. C. (1917) Why cereals winterkill. *Agronomy Journal*, **9**, 353-380.

SanMiguel, P.J, Ramakrishna, W., Bennetzen, J.L., Busso, C.S. and Dubcovsky, J. (2002) Transposable elements, genes and recombination in a 215-kb contig from wheat chromosome $5A^m$. *Funct. Integr. Genomics*, **2**, 70-80.

Santra, D.K., Santra, M., Allan, R.E., Campbell, K.G. and Kidwell, K.K. (2009) Genetic and molecular characterization of vernalization genes *Vrn-A1, Vrn-B1*, and *Vrn-D1* in spring wheat germplasm from the pacific northwest region of the USA. *Plant Breed*, **128**, 576-584.

Schaar, B. T., Chan, G. K. T., Maddox, P., Salmon, E. D., and Yen, T. J. (1997) CENP-E function at kinetochores is essential for chromosome alignment. *The Journal of cell biology*, **139**, 1373-1382.

Semenov, M. A., and Halford, N. G. (2009) Identifying target traits and molecular mechanisms for wheat breeding under a changing climate. *Journal of Experimental Botany*, **60**, 2791-2804.

Sessa, G., Carabelli, M., Ruberti, I., Lucchetti, S., Baima, S., Morelli, G. (1994) Identification of Distinct Families of HD-ZIP Proteins in *Arabidopsis* Thaliana. *Plant molecular biology*, **81**, 411-426.

Sessa, G., Morelli, G., Ruberti, I. (1993) The Athb-1 and -2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities. *EMBO J.* **12**, 3507-17.

Sheldon, C.C., Burn, J.E., Perez, P.P., Metzger, J., Edwards, J.A., Peacock, W.J. and Dennis, E.S. (1999) The *FLF* MADS box gene: A repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. *Plant Cell*, **11**, 445-458.

Shimada, S., Ogawa, T., Kitagawa, S., Suzuki, T., Ikari, C., Shitsukawa, N., and Murai, K. (2009) A genetic network of flowering-time genes in wheat leaves, in which an APETALA1/FRUITFULL-like gene, VRN1, is upstream of FLOWERING LOCUS T. *The Plant Journal*, **58**, 668-681.

Shitsukawa, N., Ikari, C., Shimada, S., Kitagawa, S., Sakamoto, K., Saito, H., and Murai, K. (2007) The einkorn wheat (Triticum monococcum) mutant, maintained vegetative phase, is caused by a deletion in the VRN1 gene. *Genes & genetic systems*, **82**, 167-170.

Snape, J. W., Butterworth, K., Whitechurch, E., and Worland, A. J. (2001) Waiting for fine times: genetics of flowering time in wheat. *Euphytica*, **27**, 377-412.

Söderman, E., Mattsson, J., & Engström, P. (1996). The *Arabidopsis* homeobox gene ATHB-7 is induced by water deficit and by abscisic acid. *The Plant Journal*, ***10***, 375-381.

Söderman, E., Mattsson, J., Svenson, M., Borkird, C., & Engström, P. (1994). Expression patterns of novel genes encoding homeodomain leucine-zipper proteins in *Arabidopsis* thaliana. *Plant molecular biology*, ***26***, 145-154.

Starr, D. A., Williams, B. C., Li, Z., Etemad-Moghadam, B., Dawe, R. K., and Goldberg, M. L. (1997) Conservation of the centromere/kinetochore protein ZW10. *The Journal of cell biology*, **138**, 1289-1301.

Sunkar, R., Zhou, X., Zheng, Y., Zhang, W. and Zhu, J.K. (2008) Identification of novel and candidate miRNAs in rice by high throughput sequencing. *BMC Plant Biol.*, **8**, 25-41.

Takahashi, R., and Yasuda, S. (1971). Genetics of earliness and growth habit in barley. In *Int Barley Genet Symp Proc*.

Tamaki, S., Matsuo, S., Wong, H. L., Yokoi, S., and Shimamoto, K. (2007) Hd3a protein is a mobile flowering signal in rice. *Science*, **316,** 1033-1036.

Taylor, G. A., and Olsen, R. A. (1976) The relationship of dessication and seedling morpho-developmental characteristics with winter survival of wheat. *Agronomical Abstract*, 64.

Taylor, R. G., Jackson, T. L., Powelson, R. L., and Christensen, N. W. (1983). Chloride, nitrogen form, lime, and planting date effects on take-all root rot of winter wheat. *Plant disease*, **67**, 1116-1120.

Ten Hoopen, R., Schleker, T., Manteuffel, R., and Schubert, I. (2002) Transient CENP-E-like kinetochore proteins in plants. *Chromosome Research*, **10**, 561-570.

Tranquilli, G.E. and Dubcovsky, J. (2000) Epistatic interactions between vernalization genes *Vrn-A^{m}1* and *Vrn-A^{m}2* in diploid wheat. *J. Hered.*, **91**, 304-306.

Trevaskis, B., Bagnall, D.J., Ellis, M.H., Peacock, W.J. and Dennis, E.S. (2003) MADS box genes control vernalization-induced flowering in cereals. *Proc. Natl. Acad. Sci. USA.*, **100**, 13099-13104.

Trevaskis, B., Hemming, M. N., Peacock, W. J., and Dennis, E. S. (2006) *HvVRN2* responds to daylength, whereas *HvVRN1* is regulated by vernalization and developmental status. *Plant Physiology*, **140**, 1397-1405.

Turner, A., Beales, J., Faure, S., Dunford, R. P., and Laurie, D. A. (2005) The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. *Science*, **310**, 1031-1034.

Turner, A., Beales, J., Faure, S., Dunford, R. P., and Laurie, D. A. (2005) The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. *Science*, **310**, 1031-1034.

Valárik, M., Linkiewicz, A., and Dubcovsky, J. (2006) A microcolinearity study at the earliness *per se* gene *Eps-A* m1 region reveals an ancient duplication that preceded the wheat-rice divergence. *TAG Theoretical and Applied Genetics*, **112**, 945-957.

Valárik, M., Linkiewicz, A., and Dubcovsky, J. (2006) A microcolinearity study at the earliness *per se* gene *Eps-A* m1 region reveals an ancient duplication that preceded the wheat-rice divergence. *TAG Theoretical and Applied Genetics*, **112**, 945-957.

Wang, S. Y., Ward, R. W., Ritchie, J. T., Fischer, and R. A., Schulthess, U. (1995a) Vernalization in wheat I. A model based on the interchangeability of plant age and vernalization duration. *Field Crops Research*, **41**, 91-100.

Wang, S. Y., Ward, R. W., Ritchie, J. T., Fischer, R. A., and Schulthess, U. (1995b) Vernalization in wheat II. Genetic variability for the interchangeability of plant age and vernalization duration. *Field Crops Research*, **44**, 67-72.

Wang, S., Carver, B., and Yan, L. (2009). Genetic loci in the photoperiod pathway interactively modulate reproductive development of winter wheat. *TAG Theoretical and Applied Genetics* **118**, 1339-1349

Wicker, T., Yahiaoui, N., Guyot, R., Schlagenhauf, E., Liu, Z.D., Dubcovsky, J. and Keller, B. (2003) Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and Am genomes of wheat. *Plant Cell*, **15**, 1186-1197.

Wigge, P. A., Kim, M. C., Jaeger, K. E., Busch, W., Schmid, M., Lohmann, J. U., and Weigel, D. (2005) Integration of spatial and temporal information during floral induction in *Arabidopsis*. *Science*, **309,** 1056-1059.

Wilhelm, E. P., Turner, A. S., and Laurie, D. A. (2009) Photoperiod insensitive Ppd-A1a mutations in tetraploid wheat (Triticum durum Desf.). *Theoretical and Applied Genetics*, **118**, 285-294.

Xue, W., Xing, Y., Weng, X., Zhao, Y., Tang, W., Wang, L., and Zhang, Q. (2008) Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. *Nature genetics*, **40**, 761-767.

Yan, L, Loukoianov, A., Tranquilli, G., Helguera, M., Fahima. T. and Dubcovsky, J. (2003) Positional cloning of wheat vernalization gene *VRN1*. *Proc. Natl. Acad. Sci. USA*, **100**, 6263-6268.

Yan, L. (2009) The flowering pathway in wheat. *In* Wheat: Science and Trade. Wiley-Blackwell Publishing, *Ames, IA.* 57-72.

Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., Sanchez, A., Valarik, M., Yasuda, S. and Dubcovsky, J. (2006) The wheat and barley vernalization gene *VRN3* is an orthologue of FT. *Proc. Natl Acad. Sci.* USA, **103**, 19581–19586.

Yan, L., Helguera, M., Kato, K., Fukuyama, S., Sherman, J. and Dubcovsky, J. (2004b) Allelic variation at the *VRN1* promoter region in polyploid wheat. *Theor. Appl. Genet.*, **109**, 1677-1686.

Yan, L., Loukoianov, A., Tranquilli, G., Blechl, A., Khan, I.A., Ramakrishna, W., SanMiguel, P., Bennetzen, J.L., Echenique, V., Lijavetzky, D. and Dubcovsky, J. (2004a) The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science*, **303**, 1640-1644.

Yang, Y., Fanning, L. and Jack, T. (2004) The K domain mediates heterodimerization of the *Arabidopsis* floral organ identity proteins, APETALA3 and PISTILLATA. *Plant Journal*, **33**, 47-59.

Yano, M., Kojima, S., Takahashi, Y., Lin, H., and Sasaki, T. (2001) Genetic control of flowering time in rice, a short-day plant. *Plant physiology*, **127,** 1425-1429.

Yao, X., Abrieu, A., Zheng, Y., Sullivan, K. F., and Cleveland, D. W. (2000) CENP-E forms a link between attachment of spindle microtubules to kinetochores and the mitotic checkpoint. *Nature Cell Biology*, **2**, 484-491.

Yao, X., Anderson, K. L., and Cleveland, D. W. (1997) The microtubule-dependent motor centromere–associated protein E (CENP-E) is an integral component of kinetochore corona fibers that link centromeres to spindle microtubules. *The Journal of cell biology*, **139**, 435-447.

Yao, Y., Guo, G., Ni, Z., Sunkar, R., Du, J., Zhu, J. K., and Sun, Q. (2007) Cloning and characterization of microRNAs from wheat (*Triticum aestivum L.*). *Genome Biology*, **8**, R96.

Yen, T. J., Compton, D. A., Wise, D., Zinkowski, R. P., Brinkley, B. R., Earnshaw, W. C., and Cleveland, D. W. (1991) CENP-E, a novel human centromere-associated

protein required for progression from metaphase to anaphase. *The EMBO journal*, **10**, 1245.

Zadoks, J. C., Chang, T. T., and Konzak, C. K. (1974) A decimal code for the growth stages of cereals. *Weed research*, **14**, 415-421.

Zelenski, M. A., and Remeslo, V. V. (1977) Inheritance of habit and related characters in winter wheat hybrids. *Plant Breeding Abstract*, **50**, 3665.

Zhang, X. K., Xiao, Y. G., Zhang, Y., Xia, X. C., Dubcovsky, J., and He, Z. H. (2008) Allelic variation at the vernalization genes, and in chinese wheat cultivars and their association with growth habit. *Crop Science*, *48*(2), 458-470.

Zhu, J., Pearce, S., Burke, A., See, D. R., Skinner, D. Z., Dubcovsky, J., and Garland-Campbell, K. (2014). Copy number and haplotype variation at the *VRN-A1* and central *FR-A2* loci are associated with frost tolerance in hexaploid wheat. *Theoretical and Applied Genetics*, **127**, 1183-1197.

Table 1. The MITE_VRN family in wheat genomes

Sequence source	Accession #	Host direct duplication sequences	MITE length (bp)	Sequence length (bp)
EST	HX153763	ATGCCAGTG	195	624
	CK217184	ATGCCAGTG	194	1049
	CK217185	ATGCCAGTG	195	1047
	CK217186	ATGCCAGTG	195	1027
Wheat genome sequence contig	Contig299482	CCAAATATAAG	219	1679
	Contig134865	ATGGTTTGAG	224	2194
	Contig80285	GTGTTTTTC	221	2554
	Contig3014060	CTATTATAC	214	525
	Contig859785	GCAGTTTAG	222	1093
	Contig311912	CATAATTAC	221	1663
	Contig4207044	CTGAATTTG	221	325
	Contig1137015	GCTCAACAC	220	955
	Contig279496	CTAGGATGC	220	1742
	Contig3124132	GCCAT	218	509
	Contig2088947	ATG	220	476

Table 2. Primers used for *TaHOX* expression in Jagger and 2174

Target	Primer name	Sequence
TaHOX-A1	TaHoxC2-F3	TTCTCGCTCTGACGATTTGATACCGCG
	TaHoxC2-R1	TGCAACGTTGGTTCATGCCACTGCG
TaHOX-D1	TaHoxC2-F2	GGAGTTCTGTCTCTGACGATTCGATTCCAAT
	TaHoxC2-R1	TGCAACGTTGGTTCATGCCACTGCG
TaHOX-A2	HoxC4-MF2	CATCCAGCAGAGCAGAGGAGAGC
	Hox-M-R4	GGCGGTCGCCTCCATGGTTTCAGGTC
TaHOX-B2	Hox-M-F6	GTAGGGAGGGGATCCATCCAGATCC
	Hox-M-R6	ACTTAAGACATTTTGAGACGGAGGGAGTATAC
TaHOX-D2	HoxC5-MF3	GATCCATCCAACAGAGCAGAGCAGAGT
	Hox-M-R5	GATTCAGGTGCATCATCCTCTATCTTACATTC

Table 3. Dissection medium (1 L)

Murashige & Skoog Salt Mixture (BRL)	4.4 g
Maltose	40 g
Thiamine-HCL (25mg/500ml)	10 ml
L-asparagine	0.15 g

Adjust pH with 1.0 M KOH to 5.85.
Add 3.5 g phytagel and autoclave in the autoclave machine.
Cool to 60 °C in water bath.
Add 2.0 ml 0.5 mg/ml filter-sterilized 2, 4-D to each 500-ml bottle.
Add 24.5µl 0.1M filter-sterilized $CuSO_4$ to each 500-ml bottle.
Pour into 100 mm × 15 mm plates (15-20 plates per 500 ml).

Table 4. High osmoticum media (bomb medium) (1L)

Murashige & Skoog Salt Mixture (BRL)	4.4 g
Maltose	40 g
Thiamine-HCL (25mg/500ml)	10 ml
L-asparagine	0.15 g
Sucrose	171.15 g

Adjust pH with 1.0 M KOH to 5.85.
Add 3.5 g phytagel and autoclave in the autoclave machine.
Cool to 60 °C in water bath.
Add 2.0 ml 0.5mg/ml filter-sterilized 2, 4-D to each 500-ml bottle.
Add 24.5 µl 0.1M filter-sterilized $CuSO_4$ to each 500-ml bottle.
Pour into 60 mm × 15 mm plates (20-30 plates per 500 ml).

Table 5. Recovery medium (1 L)

Murashige & Skoog Salt Mixture (BRL)	4.4 g
Maltose	40 g
Thiamine-HCL (25mg/500ml)	10 ml
L-asparagine	0.15 g

Adjust pH with 1.0 M KOH to 5.85.
Add 3.5 g phytagel and autoclave in the autoclave machine.
Cool to 60 °C in water bath.
Add 2.0 ml 0.5 mg/ml filter-sterilized 2, 4-D to each 500-ml bottle.
Add 24.5µl 0.1M filter-sterilized $CuSO_4$ to each 500-ml bottle.
Pour into 100 mm × 15 mm plates (15-20 plates per 500 ml).

Table 6. Regeneration medium (1 L)

Murashige & Skoog Salt Mixture (BRL)	4.4 g
Maltose	40 g
Thiamine-HCL (25mg/500ml)	10 ml
L-asparagine	0.15 g

Adjust pH with 1.0 M KOH to 5.85.
Add 3.5 g phytagel and autoclave in the autoclave machine.
Cool to 60 °C in water bath.
Add 2.0 ml 0.5mg/ml filter-sterilized 2, 4-D to each 500-ml bottle.
Add 24.5µl 0.1M filter-sterilized $CuSO_4$ to each 500-ml bottle.
Add 1.5 ml 3.0 mg/ml filter-sterilized bialaphos to each 500-ml bottle.
Add 50 µl of 6-Benzylamino purine (6-BA) to each 500-ml bottle.
Pour into 100 mm × 20 mm plates (12-15 plates per 500 ml).

Table 7. Rooting medium (1 L)

Murashige & Skoog Salt Mixture (BRL)	2.2 g
Maltose	20 g
Thiamine-HCL (25mg/500ml)	510 ml
L-asparagine	0.075 g

Adjust pH with 1.0 M KOH to 5.85.
Add 2.5 g phytagel and autoclave in the autoclave machine.
Cool to 60 °C in water bath.
Add 1.5 ml 3.0 mg/ml filter-sterilized bialaphos to each 500-ml bottle.
Pipet 18mls into each tube (13 tubes/ 500 ml)

VITA

MING YU

Candidate for the Degree of

Doctor of Philosophy

Thesis: GENETIC BASIS AND MOLECULAR MECHANISM OF WHEAT DEVELOPMENT

Major Field: Crop Science

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Crop Science at Oklahoma State University, Stillwater, Oklahoma in December, 2014.

Completed the requirements for the Master of Science in Plant Science at Oklahoma State University, Stillwater, Oklahoma in December, 2012.

Completed the requirements for the Bachelor of Science in Biotechnology at Shandong Agricultural University, Tai'an, Shandong/China in August 2009.