Genetics of Pooideae freezing tolerance – an evolutionary perspective

Genetisk kontroll av frosttoleranse i Pooideae-gress - et evolusjonært perspektiv

Philosophiae Doctor (PhD) Thesis

Simen Rød Sandve

Dept. of Plant and Environmental Sciences

Norwegian University of Life Sciences

Ås 2010

Thesis number 2010: 37

ISSN 1503-1667

ISBN 978-82-575-0947-7

Table of contents

Acknowledgements

The work presented here has been carried out in the time period 2006-2010, at the Department of Plant and Environmental Sciences (IPM) at the Norwegian University of Life Sciences (UMB). It has been co-funded by the IPM/UMB, the Research Council of Norway and Graminor AS.

First, and foremost, I would like to thank my main supervisor, Professor Odd Arne Rognli, for always being enthusiastic and for giving me freedom to follow up my own ideas. Without this freedom, I do not think I would have managed to maintain my motivation, creativity, and passion for biology and genetics. Also, a big “thanks” to my co-supervisors Dr. Heidi Rudi and Dr. Torben Asp for always being helpful, positive, and interested. Thanks!

Thanks to my dear friend, colleague, and work-out companion, Vegard Martinsen, soon to be Dr. Martinsen. Our bi-weekly sessions of weight lifting, combined with discussions regarding ecology, biology, chemistry, university politics, cars, love, wine and soap-series on TV have been immensely important, both for my physical and mental health.

I also have tons of gratitude for many other colleagues. Dr. Siri Fjellheim deserves much appreciation for her work and inspiration in the process of publishing the last paper (IV) in my thesis. Dr. Magnus D. Vigeland, my childhood friend, who unfortunately was lost in the world of abstract mathematics for some years; however, he has now finally got it! It is all about genes and biology. Thanks Magnus, for always being incredibly enthusiastic, inspiring, critical, and helpful. Dr. Terje C. Ahlquist; my many discussions with you have been great! Even though Dr. Terje has his focus on very childish aspects of genetics, namely the genetics of the rectum, his creativity and competence as a researcher and geneticists has been very helpful. Other fantastically supportive colleagues which deserve their name on print are Anna, Anne Line, Bjørge, Guro, Miriam, Bruno, and all those i couldn’t fit in, you know who you are.

And now, to my parents, little sister Ine, Erlend, extra-siblings (Martin, Frida, Alex, Nora), extra-parents (Helge and Ann), and all other good friends: I am lucky to have such great and supportive people surrounding me.

Simen September, 2010

Preface: A desperate, silent battle for survival

Every autumn, to our amazement, ice crystals turn the rims of grey pavement ponds into diamond necklaces, and at night time under the street lights even the dullest, greyest road side plant looks as if it has been sprinkled with star dust. However, for plants and other living creatures the beautifying ice diamonds come with a cost - a cost that can result in death. Beneath the frosty sheaths that enclose the plant, an ice invasion of the living tissue has begun, and as the army of ice crystals penetrates deeper, the life sustaining water molecules are greedily seized and cells are torn open. Plants have no alternative but to deal with it, there and then – no escape. Every autumn, without any sound, a grand battle for survival commences – it is the plants against the wrath of winter. This is a desperate but totally silent battle for survival.

This overly dramatic first paragraph could perhaps have been taken out of a scene from a National Geographic episode. However, it is not. I wrote this paragraph with the intention of helping you, the reader, to understand the bigger picture. This thesis will mostly focus on molecular processes too small for us to see, and how they have been shaped by complex evolutionary processes too slow for us to measure. But as we embark on the journey through the maze of scientific details presented herein, we must remember that what takes place inside a plant during winter is exceptionally important for all life on this planet as we know it. Understanding the molecular mechanisms which underlie freezing tolerance and allow plants to survive inhospitable winter conditions is essential to understand how plants win the battle against winter and ensure persistence of northern ecosystems.

List of papers

I. RUDI, H., S.R. SANDVE, L. OPSETH and O.A. ROGNLI, 2010 **Identification of candidate genes important for frost tolerance in *Festuca pratensis* Huds. by transcriptional profiling.** Plant Science In press

II. BARTOŠ, J., S.R. SANDVE, R. KÖLLIKER, D. KOPECKÝ, P. NĚMCOVÁ, Š. STOČES, L. ØSTREM, A. LARSEN, A. KILIAN, O.A. ROGNLI, J. DOLEŽEL. **Genetic mapping of DArT markers in the *Festuca-Lolium* complex and their use in marker-trait (freezing tolerance) association analysis**. Manuscript

III. SANDVE, S.R, H. RUDI, T. ASP and O.A. ROGNLI, 2008. **Tracking the evolution of a cold stress associated gene family in cold tolerant grasses.** BMC Evolutionary Biology **8:** 245.

IV. SANDVE, S.R., and S. FJELLHEIM, 2010. **Did gene family expansions during the Eocene-Oligocene boundary climate cooling play a role in Pooideae adaptation to cool climates?** Molecular Ecology **19:** 2075-2088.

Associated papers:

- SANDVE, S., H. RUDI, G. DØRUM, P. BERG and O. A. ROGNLI, High-throughput genotyping of unknown genomic terrain in complex plant genomes: lessons from a case study. Molecular Breeding In press
- SANDVE, S. R., A. KOSMALA, H. RUDI, S. FJELLHEIM, M. RAPACZ *et al.*, Molecular mechanisms underlying frost tolerance in perennial grasses adapted to cold climates. Plant Science In Press

Abbreviations:

ABA	- Abscisic acid
ABRE	- ABA responsive elements
ABF	- ABRE binding factors
AFP	- Anti freeze protein
bHLH	- Basic helix-loop-helix
bp	- Base pair
CA	- Cold acclimation
CAMTA	- Calmodulin binding transcriptional activators
CBF	- C-repeat binding factor
DArT	- Diversity Array Technology
E-O	- Eocene-Oligocene
EST	- Expressed sequence tags
FST	- Fructosyl transferase
ICE	- inducer of CBF expression
IRI	- Ice re-crystallization inhibition
IRIP	- Ice re-crystallization inhibition protein
LEA	- Late embryogenesis associated
LG	- Linkage group
My	- Million years
ROS	- Reactive oxygen species
TH	- Thermal hysteresis

Summary

Freezing tolerance is an important agricultural trait for forage and cereal grasses belonging to the Pooideae sub-family. Many plants, including the Pooideae grasses, go through a cold acclimation (CA) process when exposed to low temperatures to acquire elevated levels of freezing tolerance. In four papers presented in this thesis we have investigated different genetic and evolutionary aspects of cold acclimation (CA) and freezing tolerance in Pooideae grasses.

The two first papers (I and II) deal with transcriptional responses to low temperature exposure and association of genetic variation with freezing tolerance levels. We generated sequences from 372 transcribed genes putatively involved in CA in the perennial forage grass *Festuca pratensis* Huds.. Seven of these genes were differentially regulated between plants with high- and low freezing tolerance following CA, representing good candidates for further investigations into the mechanisms underlying phenotypic differences in freezing tolerance. We also analysed differences in genomic constitution of plants with divergent freezing tolerance in a Lolium x Festuca hybrid grass population using Diversity Array Technology (DArT) markers. Three regions on the Festuca and Lolium chromosomes 2, 4, and 7 contained DArT markers associated with the freezing tolerance phenotype.

In Papers III and IV we take on a different angle and investigate patterns of molecular evolution in genes and genetic pathways involved in freezing tolerance in the entire Pooideae lineage. Evolution of a novel ice-binding domain and multiple Pooideae specific duplication of genes involved in freezing tolerance occurred in early Pooideae evolution. The initial duplication events giving birth to these Pooideae specific genes are clustered in the Eocene-Oligocene (E-O) global climate cooling period, ~34-26 million years (My) ago. We hypothesize that the E-O paleoclimatic shift was important in adaptive evolution of the molecular mechanisms fundamental for Pooideae freezing tolerance and adaptation to cold climates.

Sammendrag

Frosttoleranse er en viktig karakter for jordbruksgress (enggras og korn) tilhørende Pooideae-underfamilien. Mange planter, inkludert Pooideae-grasartene, er avhengig av å gå igjennom en kuldeakklimatiseringsperiode for å øke nivået av frosttoleranse og dermed sikre god vinteroverlevelse. I de fire artiklene presentert i denne oppgaven studerer vi ulike aspekt av de genetiske systemene involvert i kuldeakklimatisering (KA) og frosttoleranse i Pooideae-arter.

I de to første artiklene (I og II) studerer vi gener som blir regulert under KA-prosessen og assosierer ulike typer genetisk variasjon med frosttoleranse i populasjoner av fôrgress. Vi genererte 372 sekvenser fra gener involvert i KA i engsvingel (*Festuca pratensis* Huds.). Av disse genene var 7 ulikt regulert under KA-prosessen mellom planter med høy- og lav frosttoleranse og er derfor meget gode kandidater for funksjonell genetisk variasjon som ligger til grunn for ulikt nivå av frosttoleranse. Vi undersøkte også sammenhengen mellom genomisk variasjon og frosttoleranse i en Festuca x Lolium hybridpopulasjon og fant tre kromosomområder, henholdsvis på kromosom 2, 4 og 7, som var assosiert med frysetoleranse.

I de to siste artiklene er fokuset molekylære evolusjonsprosesser som har vært med å forme genene og de molekylære prosesser involvert i frosttoleranse i alle Pooideae-artene. Evolusjon av et nytt is-bindende proteindomene og flere duplikasjoner av gener involvert i frosttoleranse forekom tidlig i Pooideae-evolusjonen. De første av disse Pooideae-spesifikke duplikasjonene skjedde under en global kuldeperiode ved overgangen mellom Eocene-Oligocene (E-O) for ~34-26 millioner år siden. Vi foreslår derfor en hypotese om at klimaet under E-O kuldeperioden var viktig for evolusjon av molekylære prosesser involvert i frosttoleranse og Pooideae-artenes tilpasning til kalde klima

Introduction

Why study freezing tolerance in Pooideae grasses?

True grasses (Poaceae) are a large and diverse plant family, consisting of approximately 10.000 species belonging to 700 genera (RENVOIZE and CLAYTON 1992), which contain many of the world's most important agricultural crops, e.g. rice, cereals, forage grasses, maize, and sugarcane. The cereals (Triticeae tribe) and forage grasses (Poeae tribe) belong to the Pooideae sub-family which encompass major agricultural species in European farming; e.g. wheat, barley, ryegrass, amongst others. Crop damage caused by freezing is a common problem in northern Europe, and limits the production and cultivation range zone of elite lines of forage grasses and cereals (Fig. 1). Identification of genes and genetic markers associated with natural genetic differences in freezing tolerance, and improved understanding of molecular mechanisms underlying the freezing tolerance trait can facilitate the development of agricultural grasses with increased tolerance to freezing stress and increase agricultural production.

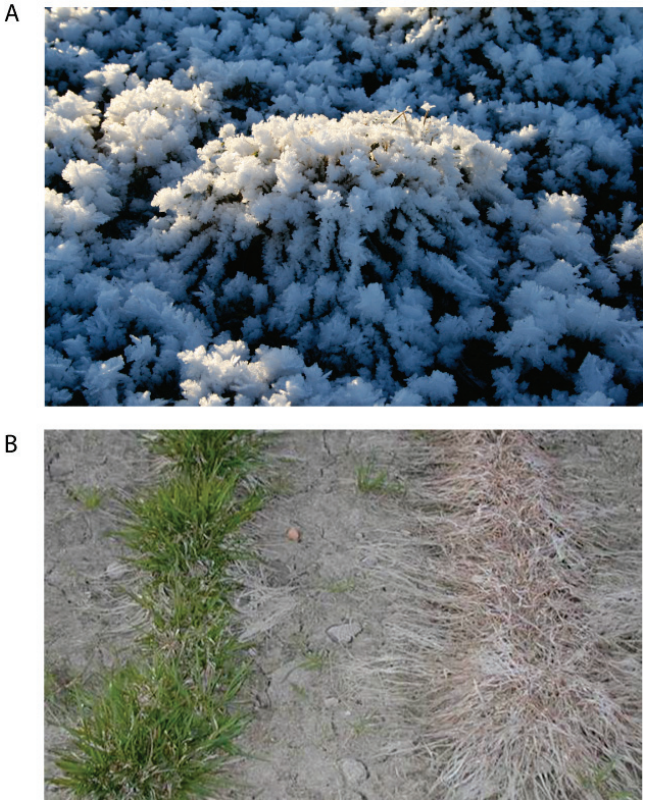


Figure 1. Forage grasses during winter and in spring. A) *Festuca pratensis* plants covered in ice crystals during early winter at Ås, Norway. B) Differences in winter survival between two forage grasses.

All plants have many common genetic pathways that are involved in freezing tolerance. Therefore, a fair question to ask is: why study freezing tolerance in Pooideae grass species and not a fully sequenced model plant like *Arabidopsis thaliana* for which enormous genetic resources already exist? The short but long answer to this question is the 150-200 million years (My) of evolution separating the monocots grasses and *A. thaliana* and other dicots (SANDERSON *et al.* 2004). This is for example more than ten times as long as it took whales to evolve from a land living animal to a bona fide sea creature (BAJPAI *et al.* 2009)! Focussing all research efforts on model species, could make us ignorant to products of evolution central for freezing tolerance in agriculturally important grasses. Cereals and forage grasses belong to different Pooideae tribes but they have the same basal number of seven chromosomes, and more importantly, comparative genomics have demonstrated large similarities with respect to the genomic loci underlying differences in freezing tolerance (ALM *et al.* submitted). The transfer of genetic knowledge between Pooideae species is therefore likely to be successful, hence understanding genetic and molecular mechanisms for freezing tolerance in any Pooideae species could be valuable for all agriculturally important cereals and forage grasses.

Apart from the obvious practical and economic aspects of studying freezing tolerance in Pooideae grasses, a second reason (and a much nerdier one) is the deep biological curiosity some of us have concerning how plants manage to adapt to cold (or other extreme) climates. The Pooideae group have an apparent skewed geographical distribution towards cold habitats in the high- northern and southern latitudes (GRASS PHYLOGENY WORKING GROUP 2001; HARTLEY 1973). A majority of their closest grass relatives on the other hand, e.g. rice, bamboo, and maize, are typical warm climate adapted grasses. It is therefore certain that after the divergence from a common ancestor with rice, novel products of evolution arose in a common Pooideae ancestor and provided adaptations to cold climates (Fig. 2). To add to the story, a global climate cooling during the Eocene-Oligocene boundary ~34-26 My ago (ZACHOS *et al.* 2001) coincides with what is believed to be the period leading up to the divergence of Triticeae and Poeae (SANDVE and FJELLHEIM 2010; VINCENTINI *et al.* 2008), the two main cool climate adapted grass groups in the Pooideae sub-family. Pooideae grasses are therefore a particularly interesting model system for studying adaptive evolution to cold climates and freezing tolerance.

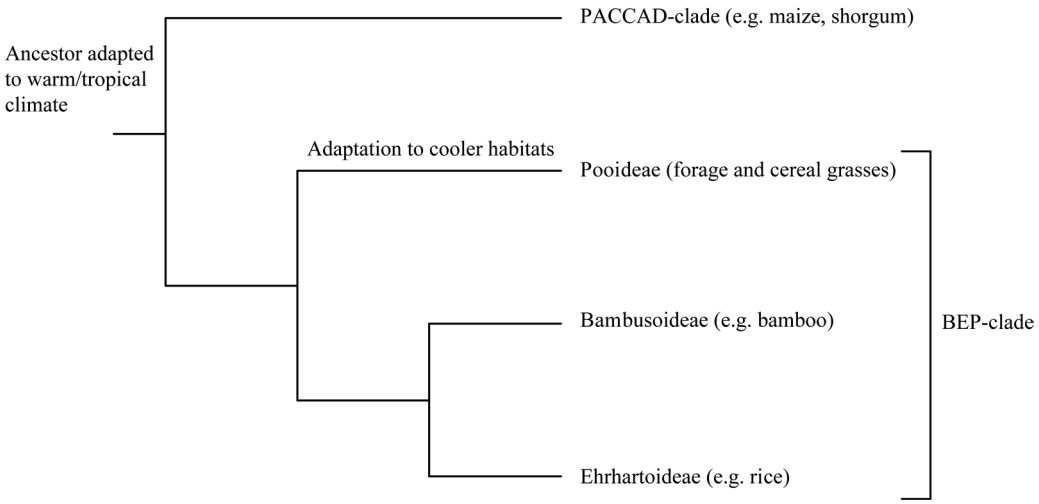


Figure 2. Evolutionary relationship between the main groups of extant grasses and major shift in climate adaptation. The evolutionary relationships between the two major clades of extant grasses, the PACCAD (Panicoideae, Aristidoideae, Centothecoideae, Chloridoideae, Arundinoideae, and Danthonioideae) and BEP (Bambusoideae, Ehrhartoideae, and Pooideae) clades.

What is freezing tolerance really?

Living in cold environments is a challenge to all organisms. Low temperatures affect molecular processes in the cells by slowing down their rate, and even sometimes disrupt them completely (Fig. 3). This comes as a result of low temperature-induced changes in the stability, shape or function of biomolecules. One example from the plant kingdom is how low temperature induces changes in functionality of biomolecules involved in the photosynthesis, the process that converts solar radiation into chemical energy in plants. Plants that do not possess adaptations to low temperatures may be incapable of balancing the energy input/output during cold periods (Fig. 3). Producing too much energy relative to the plant needs at low temperature, generates higher levels of reactive oxygen species (ROS) molecules (TAKAHASHI and MURATA 2008) which can severely damage plant cells (APEL and HIRT 2004).

A second effect of low temperature on living cells is related to the physiochemical properties of water molecules. Below 0°C water molecules crystallizes, a process referred to as freezing. This phase transition from liquid (or vapour) to crystallized water has multiple adverse effects on living organism. Firstly, when water freezes it becomes unavailable for normal water-dependent biochemical processes in cells. Secondly, because ice has a lower water potential

than that of liquid water, the formation of ice in extracellular spaces results in water being drawn out of the cells. This process is called freeze-induced dehydration and will affect plant cells in a similar fashion as drought stress (Fig. 3) (PEARCE 2001). Lastly, during prolonged freezing ice re-crystallization occurs, i.e. the formation of larger ice crystals at the expense of small crystals. This process is thought to lead to cell damaged due to ice crystal volume expansion (Fig. 3).

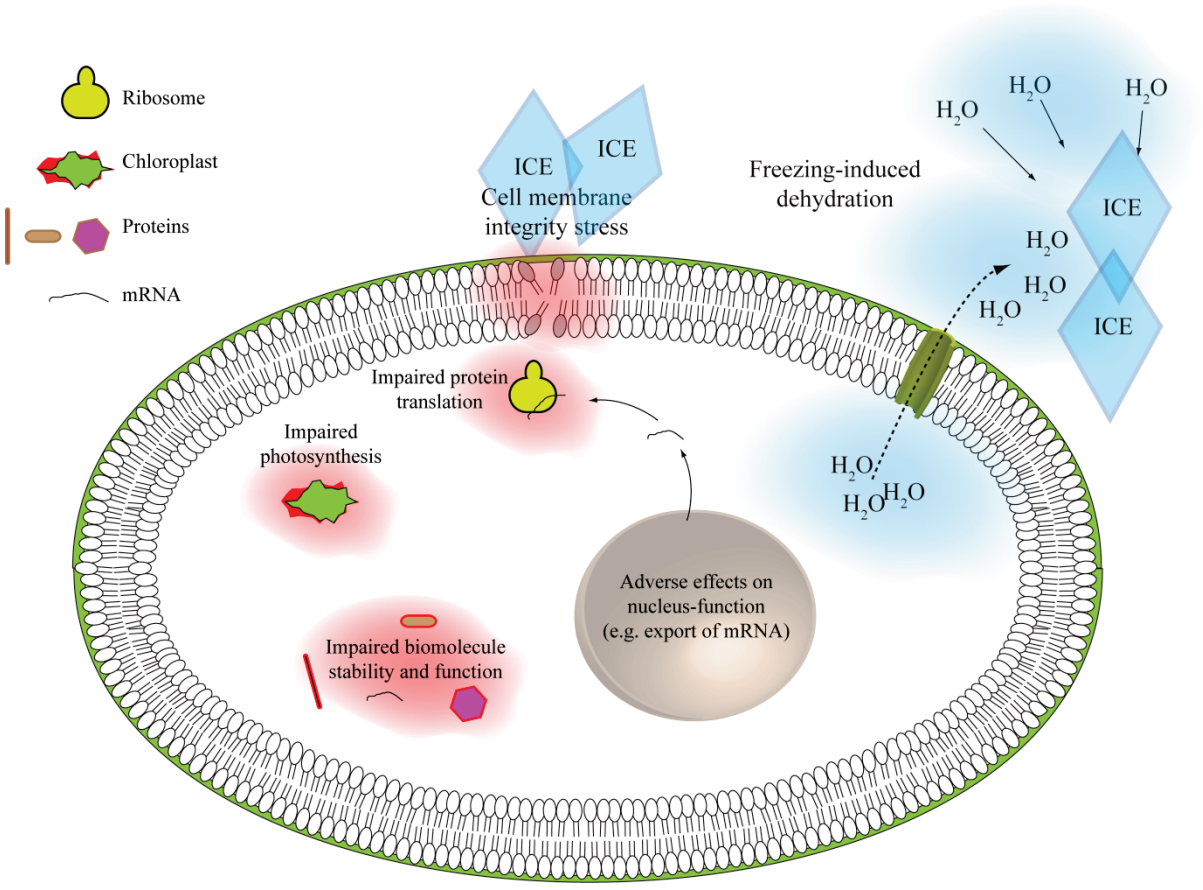


Figure 3. Harmful effects of freezing on plant cells.

It is evident that low temperatures have adverse effects on very many aspects of the plant physiology; hence freezing tolerance is a trait that is under the control of many genes. These genes are involved in regulating physiological processes to accommodate biochemical changes that occur at low temperature (e.g. biomolecule function, metabolism, photosynthesis rate) and protection against the direct harmful effects of freezing (dehydration and ice crystal expansion) (SMALLWOOD and BOWLES 2002). To add to the complexity, in nature,

interactions between temperature dependent stress (abiotic) and winter-associated pathogen attacks (biotic stress) also play a big role (GAUDET 1994). Thus the winter survival trait includes pathogen defence as well as all aspects of low temperature stress tolerance. Furthermore, plants have different life history strategies and as a consequence also different challenges in regards to freezing stress. Annual plants overwinter as seeds (or are sown in spring in agricultural systems) and their biggest threat is episodic freezing in spring time subsequent to germination. Perennial plants survive the winter in a vegetative growth stage and are in need of a long-term system for defence against freezing stress (winter-associated freezing stress) on top of the ability to withstand sporadic spring freezing.

In this thesis, I restrict my focus to the genes and molecular systems involved in the abiotic aspects of winter associated freezing stress in the perennial grasses. This encompasses plant stress related to ice formation (dehydration, ice crystal formation) in addition to all non-ice related effects of low temperature (biochemical/metabolic effects of low temperature) (Fig. 3).

Freezing tolerance – It is about being prepared

The importance of being well prepared for a difficult task is something all of us are familiar with. Giving a lecture on an unfamiliar topic without researching the subject matter, or running a marathon without months of hard training, are tasks that very likely result in utter failure. Similarly, in nature preparation for tough impending challenges is essential for both animals and plants. Perennial plants in temperate and polar climates like the Pooideae grasses have molecular mechanisms that help prepare for the low temperatures and freezing stress these plants must overcome every winter. During autumn, temperature and day length gradually drops in synchrony and function as critical heads up for the coming winter. Plants translate these external environmental cues into molecular responses that increase the level of freezing tolerance and ensure winter survival. This preparation process is called cold acclimation (CA).

Some authors have defined CA as: plant responses to low but non-freezing temperatures resulting in increased freezing tolerance (FOWLER and THOMASHOW 2002). This definition

however only accounts for the temperature dependent processes. Many plant molecular responses to low temperatures are in fact dependent on an interaction between light (quantity and quality) and low temperature (CROSATTI *et al.* 1999), hence the latter definition is artificial and restricted to laboratory settings. In nature, the seasonal changes in ambient temperature co-vary with changes in day length and light quality. A more realistic definition of CA is therefore; plant responses to the natural conditions of autumn, i.e. low but non-freezing temperature in combination with changing day length and light quality, resulting in increased freezing tolerance. It is important to note that most studies on molecular mechanisms in CA are biased towards responses triggered by temperature only, and many light-dependent molecular responses might therefore have been overlooked.

There are numerous ways of structuring the content of a general introduction on molecular basis for freezing tolerance in plants. But because freezing tolerance of perennial plants in northern ecosystems is inextricably entwined with the molecular changes that occur during CA, I feel that molecular basis for freezing tolerance best can be understood in the light of the CA process. The CA process can be divided into three stages; (1) initial sensing and signalling of low temperature, (2) remodelling of gene expression levels, and (3) an output in the form of cellular changes resulting in an increased level of freezing tolerance. With basis in these three stages of CA, I will give a short introduction to the major genes and molecular mechanisms involved in freezing tolerance in plants, with a special emphasis on Pooideae grasses.

Sensing cold

It is obvious that plants sense changing temperatures, but the details on how low temperature is sensed and then signalled to the cell nucleus to initiate changes in gene transcription is not well understood. The best studied mechanism of plant temperature sensing is membrane fluidity. The membranes that surrounds cells, mitochondria, and chloroplasts consist of a lipid bi-layer and under low temperature this lipid membrane changes into a more rigid state (ALONSO *et al.* 1997). It has been suggested that membrane rigidification is the first physiological effect of decreasing temperatures on the plant, and the earliest signal input into the CA pathways (Fig. 4). This notion is supported by experiments which have shown that

chemically triggered membrane rigidification result in transcription of cold responsive genes, and that this artificially induced gene expression mimicked CA and improved freezing tolerance of plants (SANGWAN *et al.* 2001; ÖRVAR *et al.* 2000). Associated with membrane rigidification are secondary effects like cytoskeleton changes, increased Ca^{2+} influx, and changes in activity of certain protein kinases (Fig. 4) (SANGWAN *et al.* 2001; ÖRVAR *et al.* 2000). The disruption of any of these secondary responses during CA is sufficient to prevent normal CA-mediated improvement of freezing tolerance. Conversely, triggering Ca^{2+} influx or certain cytoskeleton changes at warm temperatures initiate the CA-processes and increase freezing tolerance (SANGWAN *et al.* 2001; ÖRVAR *et al.* 2000).

In cyanobacteria, two-component systems are involved in low-temperature sensing and signalling (SUZUKI *et al.* 2000). Two-component systems are made up by a membrane bound kinase which upon an environmental signal phosphorylates a response regulator that mediates the signal from the membrane to the genes. Homologous genes to those involved in two-component temperature sensing system in bacteria does not exist in plants, however plants do contain two-component systems, thus it is possible that similar mechanisms also are involved in low temperature sensing in plants (BROWSE and XIN 2001).

Secondary signals which are not related to low temperature *per se*, e.g. reactive oxygen species (ROS) or abscisic acid (ABA), may also trigger signalling cascades similar to those created by cell membrane rigidification (Fig. 4) (CHINNUSAMY *et al.* 2007). Furthermore, impairment of normal production of both ABA (LLORENTE *et al.* 2000; XIONG *et al.* 2001) and ROS (LEE *et al.* 2002) have been shown to affect the regulation of gene expression during CA and disrupt the acquirement of freezing tolerance.

Low temperature modulation of gene expression

The initial sensing and cellular response during CA, including membrane rigidification with the associated cellular changes, and most likely also several other unknown mechanisms, initiate remodelling of gene transcription activity. Whole genome microarrays have provided detailed genome-wide snap-shots of how gene expression is modulated during CA. This work,

supplemented with studies of loss-of-function mutants and transgenic plants over-expressing specific CA-associated genes, have been instrumental in understanding the genetic basis for CA in plants. The majority of studies on genes and genetic pathways involved in CA have been carried out on *A. thaliana*; even so, these results are most often of relevance to monocot grass species due to the conservation of the major genetic pathways involved in CA in all land plant lineages. When data from Pooideae grass species exists, this will be highlighted.

The estimated numbers of cold responsive genes involved in CA will depend on many factors; laboratory conditions, statistical tools, the species studied, plant tissue, and so on, hence these estimates vary substantially between studies. Lee et al. (2005) identified 939 (4%) cold regulated genes in the *A. thaliana* genome, 655 that were up-regulated and 284 that were down-regulated. Vogel et al. (2005) used a more stringent criteria for classifying a gene as cold responsive and found 514 (2%) genes regulated by cold, 302 up-regulated and 212 down-regulated, respectively. In a study on barley, 2735 (12%) genes were found to be significantly regulated by low temperature (SVENSSON *et al.* 2006). In general, transcription factors and genes involved in signal transduction are regulated as an initial CA response, while CA-associated genes involved in metabolism and protection against freezing related cell injuries are regulated later in the CA process (LEE *et al.* 2005).

The CBF genes; key transcriptional regulators early in the cold acclimation process

The best studied genetic component of the CA pathway involves the early expressed DREB1/CBF transcription factors (from now on referred to as CBF). CBFs transcript levels accumulate rapidly (<15 minutes) in plants after exposure to low non-freezing temperatures (SKINNER *et al.* 2005; VOGEL *et al.* 2005) and regulate downstream target genes by binding to the CBF binding element A/GCCGAC in the promoter regions (SAKUMA *et al.* 2002). The majority of the functional studies on CBFs have been carried out in *A. thaliana,* but the role of CBF genes as key regulators of cold responsive transcription is conserved in species across the entire land plant lineage (e.g. LIU *et al.* 2007; PENNYCOOKE *et al.* 2008; SKINNER *et al.* 2005; XIONG and FEI 2006). It is important to be aware that although general conclusions on CBF functions from a the *A. thaliana* dicot model system seems valid for many other species including the monocot grasses (XIONG and FEI 2006), species-specific differences in the

number CBF genes and their functional spectre is probably the rule rather than the exception (AGARWAL *et al.* 2006; PENNYCOOKE *et al.* 2008; ZHANG *et al.* 2004).

In *A. thaliana* three CBF genes (*CBF1*, *CBF2*, *CBF3*) are involved in the transcriptional regulation of about 30% of the cold responsive genes (Fig. 4), of which 85% is up-regulated during CA (VOGEL *et al.* 2005). The importance of this CBF controlled genetic pathway in CA has been rigorously demonstrated in numerous studies with loss-of-function mutants and CBF over-expressing transgenic plants (JAGLO-OTTOSEN *et al.* 1998; KASUGA *et al.* 1999; LIU *et al.* 1998). Moreover, almost identical changes in the plant cell metabolites occur in normal cold acclimated plants and plants over-expressing CBF genes at warm temperatures (COOK *et al.* 2004; MARUYAMA *et al.* 2009). Two studies by Novillo and co-workers have demonstrated an intricate interplay between CBF genes in the CA process. First, mutants with non-functional *CBF2* genes provided evidence that *CBF2* act as a negative regulator of *CBF1* and *CBF3* genes (NOVILLO *et al.* 2004). In a later study, using RNAi experiments, *CBF1* and *CBF3* were shown to have overlapping functions in controlling downstream gene transcription, however expression of *CBF1* or *CBF3* alone were not sufficient for wild type levels of target gene up-regulation, resulting in a sub-optimal CA process and lower freezing tolerance (NOVILLO *et al.* 2007).

As mentioned, the role of CBFs in CA is conserved in the Pooideae grasses (XIONG and FEI 2006) but the Pooideae CBF-family has undergone substantial expansion compared to the *A. thaliana* CBF family (BADAWI *et al.* 2007). At least 13 and 17 CBF members exist in wheat (MILLER *et al.* 2006) and barley (SKINNER *et al.* 2005), respectively. Despite this, the size of the CBF-dependent regulon in barley were suggested to be similar to that of *A. thaliana* (SVENSSON *et al.* 2006). In the forage grasses little is known about the number of CBF genes, however several Pooideae-specific CBF gene duplication events occurred in the cereal lineage after the Triticeae-Poeae split (SANDVE and FJELLHEIM 2010), thus it is possible that Poeae species contains fewer CBFs than Triticeae species.

CBF genes themselves are regulated at exposure to low temperatures. The inducer of CBF expression 1 (*ICE1)*, a MYC-like bHLH transcriptional activator, were first reported in *A.*

thaliana (CHINNUSAMY *et al.* 2003). Two homologs with similar function to the *A. thaliana ICE1* gene have subsequently been identified in Pooideae grasses (in wheat) (BADAWI *et al.* 2008). Over-expression of *ICE1* increased expression of downstream genes involved in CA and provided improved freezing tolerance in *A. thaliana* (CHINNUSAMY *et al.* 2003). Interestingly, *ICE1* itself is not cold regulated, but is constitutively expressed in the plant (BADAWI *et al.* 2008; CHINNUSAMY *et al.* 2003; ZARKA *et al.* 2003). Transcriptional activation of the CBF pathway by *ICE1* therefore requires a different level of regulation. Sumoylation is the attachment of small ubiquitin-related modifier proteins onto another protein (JOHNSON 2004). Miura *et al.* (2007) demonstrated that sumoylation of ICE1 proteins by SIZ1 takes place under Low temperatures and is essential for ICE1 to become a transcriptional activator of CBF genes (Fig. 4). *ICE1* also regulates the expression of numerous other transcription factors involved in the CA process, albeit little is known about their particular regulons (LEE *et al.* 2005). A negative regulator of *ICE1* have also been identified, the RING-finger-protein HOS1. HOS1 post-translationally modifies ICE1 by ubiquitination and thereby increases the rate of ICE1degradation (Fig. 4) (DONG *et al.* 2006a).

Other modifiers of CBF function are the *SFR6* and *ZAT12* genes (KNIGHT *et al.* 2009; KNIGHT *et al.* 1999; VOGEL *et al.* 2005). The *SFR6* gene encodes a large nuclear located protein with unknown structure and no defined domains (KNIGHT *et al.* 2009). Very little is known about the functional role(s) of SFR6 in plants, but studies of transgenic lines indicate that SFR6 is involved in post translational modifications of CBF genes and that *SFR6* loss-of-function mutants have impaired CBF function and cannot undergo CA (KNIGHT *et al.* 2009). *ZAT12*, a Zn-finger protein encoding transcription factor involved in regulating CA transcriptional responses independent of the CBF pathway, has also been implicated as a negative regulator of CBF genes (Fig. 4) (VOGEL *et al.* 2005).

CBF-independent pathways involved in CA

Of the 514 cold regulated genes in *A. thaliana* defined by Vogel et al. (2005), the CBF pathway can only be attributed the transcriptional control of ~30% of downstream target genes. Apart from the transcription factor ZAT12, which controls >20 cold responsive genes (VOGEL *et al.* 2005), no other non-CBF transcription factors have so far been assigned larger

cold responsive regulons. Nevertheless, there is not a lack of candidate transcription factors for being involved in the CBF-independent CA pathways; at least ~50-60 other transcription factors have been implicated in CA, including members of AP2, MYB, MYC, bZIP, Zn-FINGER transcription factors families (Fig. 4), but little is known about their downstream targets (FOWLER and THOMASHOW 2002; LEE *et al.* 2005; VOGEL *et al.* 2005).

The plant hormone abscisic acid (ABA) is a major player in regulating genes involved in plant stress response (e.g. drought and salt stress) through the transcriptional activation of ABA-dependent transcription factors (SHINOZAKI and YAMAGUCHI-SHINOZAKI 2000). The importance of ABA in CA and freezing tolerance is debated (GUSTA *et al.* 2005), but much evidence supports a ABA role in CA under natural conditions. First, endogenous ABA-levels have been shown to increase in *A. thaliana* and wheat during low temperature exposure (CUEVAS *et al.* 2008; SHAKIROVA *et al.* 2009). Second, application of exogenous ABA have been shown to enhance freezing tolerance in whole plants (CHEN *et al.* 1983; MANTYLA *et al.* 1995) and calli (DALLAIRE *et al.* 1994). Third, many genes expressed during CA are regulated by ABA (Fig. 4) (AGARWAL and JHA 2010; CUEVAS *et al.* 2008; HOTH *et al.* 2002; KOBAYASHI *et al.* 2008; SHAKIROVA *et al.* 2009). ABA-dependent transcription factors bind to an abscisic acid responsive element (ABRE) in the promoters of target genes. Different classes of ABREs are targeted by different ABRE binding factors (ABFs). For example, ABA induces the expression of many bZIP transcription factors that bind to the (C/G)ACGT ABRE in promoters of abiotic stress induced genes (CHOI *et al.* 2000; YAMAGUCHI-SHINOZAKI and SHINOZAKI 1994), while other ABREs are targets for MYC and MYB transcription factors (ABE *et al.* 2003).

It is important to be aware that many transcription factors with ABRE/ABRE-like (ABREL) elements are not necessarily ABA-responsive. Evening Element (EE)-like (EEL) elements coupled with an ABREL motifs have been shown to be enriched in promoters of cold regulated genes (MIKKELSEN and THOMASHOW 2009). In two genes regulated by CBF-independent cold regulated pathways differences in the presence of EEL-ABREL motifs were shown to be important for transcriptional control during CA. The transcriptions factors that bind EEL-ABREL motifs are, however, not identified (MIKKELSEN and THOMASHOW 2009). In conclusion, even though the CBF-pathway is well studied there are massive gaps in our

understanding of the CBF-independent transcriptional control of downstream genes in the CA-process, and future research should address these gaps.

Crosstalk between drought and cold responsive pathways

As dehydration is a common consequence of drought and freezing, it is evident that there must be crosstalk between the genetic pathways involved in the two different stress responses. This is supported by identification of many genes that are both drought- and cold-responsive (SEKI *et al.* 2002). One link between genetic systems involved in drought and cold stress is mediated by ABA-dependent gene regulation (KNIGHT and KNIGHT 2001; SHINOZAKI and YAMAGUCHI-SHINOZAKI 2000). As mentioned above, ABA is not likely involved in initial CA transcriptional response due to the fact that ABA synthesis during cold stress does not commence until after the CBF-pathway is activated (CUEVAS *et al.* 2008; SHAKIROVA *et al.* 2009; THOMASHOW 1999). However, in the later stages of CA, endogenous ABA levels has been shown to rise (SHAKIROVA *et al.* 2009) and this may influence expression levels of many stress related genes, including those common for all cellular dehydration responses.

Downstream transcriptional responses in the CA process

After a rapid up-regulation of genes involved in controlling the initial transcriptional responses in the CA process, the downstream CA-transcriptional changes commences. This includes down-regulation of many genes, particularly genes involved in metabolism (FOWLER and THOMASHOW 2002; LEE *et al.* 2005), and up-regulation of genes involved in protective roles during freezing stress, e.g. hydrophilic late embryogenesis abundant (LEA) –like proteins and sugar metabolism genes. Similar trends have been observed in late transcriptional responses during CA in Pooideae grasses (RUDI *et al.* 2010; WINFIELD *et al.* 2010).

Many different genes encode proteins with putative protective functions during freezing but it is worth spending some time on two particularly interesting groups, the LEA (including the dehydrins) and genes encoding anti freeze proteins (AFPs). LEA proteins are a diverse group, only united by (i) being highly hydrophilic (GARAY-ARROYO *et al.* 2000) and (ii) lacking well defined secondary structure in the hydrated state (MOUILLON *et al.* 2006). Apart from this,

LEA proteins are highly divergent at the amino acid level, show diverse expression pattern, and different sub-cellular localization. Some LEA genes are transcribed constitutively, others only during seed development, and yet others only under abiotic stress (TUNNACLIFFE and WISE 2007). Moreover, LEA proteins can be localized in the nucleus, cytoplasm, cell membranes, endoplasmatic reticulum, and mitochondria (HUNDERTMARK and HINCHA 2008; TUNNACLIFFE and WISE 2007).

In *A. thaliana* 11 LEA genes are induce upon cold exposure, 5 belonging to the dehydrin sub-group, while the other 6 belong to four other sub-groups (HUNDERTMARK and HINCHA 2008). One of them, *COR15a*, which belongs to the sub-group LEA_4 (also called "group 3") have been shown to affect the stability of cell membranes during freeze-induced dehydration (STEPONKUS *et al.* 1998). However, functional roles of LEA genes in CA and freezing tolerance are not restricted to cell membrane stabilization. Other studies point to potential roles of LEA proteins in osmolyte action (retention of water inside cells), antioxidant function, and as molecular shields (Reviewed in RORAT 2006; TUNNACLIFFE and WISE 2007). Furthermore there is accumulating evidence that LEA proteins have important functions as chaperones during abiotic stress. As mentioned above, LEA proteins are characterized by intrinsically disordered proteins (IDP), i.e. proteins with no/little defined three dimensional structure (TOMPA and KOVACS 2010). Function of IDPs in plants is associated with chaperone activity, i.e. stabilization and assistance of folding of mRNA or proteins and prevention of protein aggregation during stress (CHAKRABORTEE *et al.* 2007; REYES *et al.* 2008; TOMPA and KOVACS 2010).

One of the most studied sub-groups of LEA homologs with regard to freezing tolerance and CA are the dehydrins. These proteins are characterised by a conserved 15-amino acid domain called the K-segment, (EKKGIMDKIKEKLPG) and two other motifs referred to as the S- and Y-segments, however, like other LEAs the dehydrins do not have a conserved secondary structure or one specific molecular function (RORAT 2006). In dicots allelic variation in dehydrin sequences have been associated with differences in freezing tolerance (RÉMUS-BOREL *et al.* 2010). Over-expression and accumulation of dehydrins to levels similar (and higher) to those reached during CA also improved plant freezing tolerance (PUHAKAINEN *et al.* 2004). Because the dehydrins were located at the cell membranes, Puhakainen *et al.* (2004)

suggested that improved freezing tolerance was a result of dehydrin membrane stabilization. Many dehydrins have been identified and studied extensively in Pooideae species (see for example CHOI *et al.* 1999; DANYLUK *et al.* 1998; ZHU *et al.* 2000). Dehydrins have been suggested as candidate genes underlying QTLs for freezing tolerance (ALM *et al.* submitted) and recently a study demonstrated that increased levels of the dehydrin DHN5 during CA correlated well with freezing tolerance in barley (KOSOVÁ *et al.* 2010). Despite the uncertainties of LEA protein functions *in planta*, there *is* scientific consensus that LEA gene transcription and protein accumulation is important for CA and the development of freezing tolerance (FOWLER and THOMASHOW 2002; STEPONKUS *et al.* 1998; TUNNACLIFFE and WISE 2007).

The other group of genes with protective properties during freezing that I want to highlight are the genes encoding AFPs. AFPs can act protectively during freezing stress through two different mechanisms; (i) by thermal-hysteresis (TH), or (ii) ice re-crystallization inhibition (IRI). TH depresses the freezing point at which ice crystallization initiates, decreasing the likelihood of intracellular ice formation. IRI function is simply a manipulation of the growth of the ice crystals such that small ice crystals grow at the expense of larger ice crystals. Even though IRI function does not repress ice formation, the IRI mechanism is thought to minimize cellular damage in plants (GRIFFITH *et al.* 2005; SMALLWOOD and BOWLES 2002). Plant AFPs usually exhibit low TH-activity, but strong IRI activity *in vitro* (GRIFFITH and YAISH 2004). The first report of ice-binding proteins in plants was in 1992 from experiments with rye (*Secale cereale*) (GRIFFITH *et al.* 1992). Subsequently many plant AFPs of different evolutionary origin have been discovered in monocots and dicots (GRIFFITH and YAISH 2004; SIDEBOTTOM *et al.* 2000).

It is the Pooideae specific IRI-proteins (IRIPs), initially identified by Sidebottom and colleagues (2000), which have been most extensively studied among the plant AFPs. Subsequent to the first discovery, several IRIP homologs were isolated from Pooideae grass species and it is now clear that the genes encoding Pooideae IRIPs belong to a multi-gene family (JOHN *et al.* 2009; KUMBLE *et al.* 2008; SANDVE *et al.* 2008; TREMBLAY *et al.* 2005). Studies of protein structure have shown that a repeated amino acid motif (NxVxG/NxVxxG) forms the ice-binding face of the proteins and is responsible for the IRI function *in vitro* (MIDDLETON *et al.* 2009). The Pooideae IRIP genes also encode a conserved signal peptide

that targets the proteins to the apoplast (SANDVE *et al.* 2008). Such apoplastic export of IRIPs have been demonstrated *in vivo* in transgenic *Arabidopsis* plants (JOHN *et al.* 2009). Synthetic AFPs have been shown to enhance FT (GRIFFITH and YAISH 2004), but the importance of IRIPs for freezing tolerance in Pooideae grasses have yet to be demonstrated. Recently however, *Arabidopsis* plants expressing *Lolium perenne* IRIP genes were shown to have improved cell membrane stability during freezing and increased freezing tolerance (ZHANG *et al.* 2010). This strongly supports the notion that IRIP genes are important for Pooideae survival in frost exposed environments.

RNA-processing and transport

During stress, cells adapt to the new biotic or abiotic challenges by making biochemical adjustments. These adjustments ultimately rely on RNA-processing and RNA-export to the cytoplasm. Genes involved in RNA-processing and RNA-export have therefore been studied in the context of CA and freezing tolerance. Extreme temperatures are known to change the stability and/or conformation of both RNA and proteins and disrupt their normal functions (ZHU *et al.* 2007). An early response to low temperature stress is the expression of cold shock protein (CSP) encoding genes. Some of these CSPs have been shown to function as chaperones, i.e. molecules that stabilize and maintain the correct conformation of other RNAs and proteins during stress (KARLSON *et al.* 2002; NAKAMINAMI *et al.* 2006; PARK *et al.* 2010).

Before translation can take place, cells must export the mRNA from the nucleus to the ribosomal translation machinery in the cytoplasm. Two mutants, the *los4* and *atnup160* are affected in RNA-nucleus export and have been shown to have altered freezing tolerance. *los4* is a DEAD-box RNA helicase and interestingly two different *los4* mutants have the opposite effects on the freezing tolerance. A loss-of-function mutation in the *los4* gene (*los4-1*) impaired the export of RNA from the nucleus, resulting in a non-functional CBF-pathway and freezing sensitive plants (GONG *et al.* 2002). A second *los4* mutant, *los4-2,* exhibit an up-regulated CBF-pathway, is hypersensitive to ABA, and has increased freezing tolerance (GONG *et al.* 2005). The *los4* protein is located at the nuclear rim and is thought to affect freezing tolerance by regulating export of CA-involved mRNA to the cytoplasm under low temperatures (GONG *et al.* 2005). Similar to *los4,* the nucleoporin *atnup160* is localized at the nuclear rim, and the *atnup160-1* mutant has a defective RNA-export and is impaired in CBF-

protein accumulation during CA, resulting in a freezing sensitive phenotype compared to wild-type plants (DONG *et al.* 2006b).

A possible secondary function of *los4* during cold stress has been proposed. DEAD-box RNA helicases are known to control the conformation of RNA molecules (GONG *et al.* 2005), and it is known that RNAs can act as molecular thermometers through temperature mediated conformational changes (JOHANSSON *et al.* 2002). Hence it was suggested that *los4* could control RNA-mediated thermometer function in plants (GONG *et al.* 2005), albeit this remains a mere speculation at this point.

Involvement of endogenous small RNAs and DNA-methylation in CA

The importance of endogenous small RNA and their role as transcriptional regulators have only recently been acknowledged. This novel insight into the world of small RNAs has revolutionized our view on the complexity of gene regulation. Small RNA molecules affect gene regulation directly by interacting with and down-regulating target mRNA (SUNKAR *et al.* 2007). But in addition, small RNAs can also affect transcription indirectly by directing methylation of genomic loci (MOLNAR *et al.* 2010). Small RNAs involved in transcriptional regulation belongs to two groups; (1) micro RNAs (miRNA), and (2) small interfering RNAs (siRNA) (RANA 2007). Endogenous miRNA and siRNA are 21-24 nucleotides long and are transcribed from the genome; miRNA is derived from short double stranded (ds) hairpin structures while endogenous siRNA is derived from cleaved fragments of longer dsRNAs (RANA 2007). Further processing of miRNA and siRNA are dependent on dicer enzymes to produce single stranded (ss) RNAs which are subsequently incorporated into the RISC protein complex (KROL *et al.* 2010). The RISC complex then uses the ssRNA as a guide to silence complementary target mRNA or genomic loci (KROL *et al.* 2010).

siRNA work as transcriptional repressors by cleaving and degrading target mRNA while miRNA usually bind to the target mRNA and suppresses its translation (RANA 2007). It should be noted that certain siRNA have been shown to function in a similar manner as miRNA, by suppressing the target mRNA without cleaving and degrading it (DOENCH *et al.*

2003). siRNA can also regulate transcription through DNA methylation, i.e. the attachment of a methyl group on nucleotides. DNA-methylation occurs in plants at CG, CHG, or CHH (H=A, C, or T) sites, and in many cases methylation of a transcribed genomic locus is inversely related to the transcriptional activity (TEIXEIRA and COLOT 2009). The methylation status of a gene is not static, but change in response to internal and external stimuli, for example during abiotic stress (LUKENS and ZHAN 2007), and siRNAs have been shown to direct such dynamic modifications of DNA methylation (MOLNAR *et al.* 2010).

Our understanding of the roles miRNA and siRNA play in CA and freezing tolerance is still very limited. Up-regulation of miRNA transcription under low temperatures have been demonstrated in several species such as *A. thaliana* (SUNKAR *et al.* 2007), rice (JIAN *et al.* 2010), and *Brachypodium* (ZHANG *et al.* 2009). It is therefore likely that small RNAs are important in transcriptional regulation during CA. To my knowledge there are to this date no known examples of endogenous siRNA involved in CA. The importance of miRNA or siRNA in transcriptional silencing during CA could potentially be large. One reason for this is that very few transcription factors have been implicated in gene silencing during CA (VOGEL *et al.* 2005). Even though this is a mere speculation, perhaps these "missing" transcriptional silencers could be explained by yet undiscovered roles of miRNA and siRNA in the CA-process? Recent experiments have in fact shown that low temperature stress induces silencing of genes through DNA-methylation (BOYKO *et al.* 2010) but further research is needed to assess the quantitative importance of small RNAs in CA transcriptional regulation.

The CA-output; changes in the proteome and metabolome

The end points in the cascades of CA transcriptional re-modelling are changes in the biochemical and physiological state of the cell. New types of proteins and enzymes are made, others are removed, and this results in changes to the metabolome. There is to some extent a deterministic relationship between transcriptional responses and the output in the form of proteomic changes during CA. However, gene expression and protein levels is not necessarily strictly correlated (e.g. MOONEY *et al.* 2006). This can be due to effects of post-transcriptional or posttranslational processes influencing the protein accumulation independent of the transcription level of the gene. Hence analyses of gene transcription do not offer a complete

picture of plant molecular responses during CA and cold stress, and studying proteomic CA responses are thus important to complement gene expression studies.

The CA proteome has been studied extensively in *A. thaliana* and many proteins which are differentially displayed between normal and cold acclimated plants have been identified. Not surprisingly, the major finding from these studies is that the there is a strong correlation between the transcriptional response during CA and the changes in corresponding protein levels (AMME *et al.* 2006; BAE *et al.* 2003). Proteins that were found to be regulated during CA includes photosynthesis related proteins, dehydrins, chaperons, transcription factors, and calcium binding proteins (AMME *et al.* 2006; BAE *et al.* 2003; GOULAS *et al.* 2006). An effect of CA on the proteome was almost absent shortly after cold shock (1 day) but increased rapidly thereafter (GOULAS *et al.* 2006). A recent study investigated the difference in the proteome between *Festuca pratensis* Huds. genotypes of differing freezing tolerance levels (KOSMALA *et al.* 2009). Forty-one (5.1%) proteins showed > 1.5-fold difference in abundance between high and low freezing tolerant plants during the CA process, and the majority of these differentially displayed proteins were involved in photosynthesis. Several of the proteins had not previously been linked to CA and freezing tolerance levels. This includes chloroplast-localized Ptr ToxA binding protein 1, globulin 2, 50S ribosomal protein L10 from chloroplasts, 30S ribosomal protein S10 from chloroplasts, ADP (adenosine diphosphate) - glucose pyrophosphatase, and ADP-ribosylation factor 1 (KOSMALA *et al.* 2009).

As a consequence of the re-modelling of cellular protein content, and low temperature effects on the biochemical processes, the cellular content of many hundreds, if not thousands, of metabolites change during CA. Out of 434 metabolites monitored, 325 (75%) were found to increase significantly in *A. thaliana* during CA (COOK *et al.* 2004). Of these 325 low temperature responsive metabolites, 256 (79%) were also found to increase when *CBF3* was over-expressed at warm temperatures (COOK *et al.* 2004), emphasizing the importance of the CBF-pathway in the CA-process. Among the major metabolite groups that increased were amino acids and carbohydrates, however, the biggest group were un-identified metabolites (COOK *et al.* 2004). In a different study on *A. thaliana,* 60% (155) of the monitored metabolites increased during CA, many of which were carbohydrates (MARUYAMA *et al.* 2009). Other important cellular changes during CA are changes to the lipid composition of

cell membranes (UEMURA *et al.* 1995), and modifications of biochemical processes related to the adverse effects of performing photosynthesis at low temperatures (photoinhibition) (HUNER *et al.* 1993; MURATA *et al.* 2007).

The cryptic role of carbohydrates in freezing tolerance

The physiological role of changes in carbohydrate content during CA have historically been assumed to be related to (i) the osmolyte action which can help retain water during freeze-induced dehydration (YANCEY *et al.* 1982) or (ii) the properties of carbohydrates as easily accessible energy reserves during winter and spring (VIJN and SMEEKENS 1999). Although the roles of carbohydrates in freezing tolerance could at least partly be related to these historically assumed functions, recent research efforts have painted a much more complex picture of carbohydrate function in CA.

One example is the increase in raffinose during CA, which has been thought to be related to osmolyte function (TAJI *et al.* 2002). However, when raffinose levels was increase or removed completely in transgenic *A. thaliana* plants, no changes in freezing tolerance were observed (ZUTHER *et al.* 2004). On the other hand, manipulating α-Galactosidase levels in Petunia resulted in an elevated raffinose level and increased freezing tolerance (PENNYCOOKE *et al.* 2003). It is therefore apparent that a raffinose level increase during CA is not critical for developing freezing tolerance in all plants. Moreover, raffinose and galactinol, which earlier were assumed to act as osmolytes, have more recently been prescribed roles as antioxidants that protect plants from oxidative damage (NISHIZAWA *et al.* 2008).

Another group of CA-associated carbohydrates that has received much attention is fructans. Convergent evolution have independently resulted in the capacity of fructan biosynthesis in monocots and dicots, and the distribution of fructan synthesising plants is skewed towards ecosystems characterized by intermittent drought and low temperature stress (HENDRY 1993). It is recognized from field and laboratory studies that Pooideae grass species accumulate fructans during CA (KEREPESI *et al.* 2004; LIVINGSTON and HENSON 1998; POLLOCK and JONES 1979). Several studies have found correlative evidence between high levels of fructans

and increased freezing tolerance (LIVINGSTON and HENSON 1998; PUEBLA *et al.* 1997). The role of fructans in CA has historically been assumed to be an easily accessible energy reserve (VIJN and SMEEKENS 1999). This idea has however been challenged by results from functional studies on fructans and freezing stress. Studies on model cell membranes have showed that fructans are incorporated into the lipid bi-layer of the cell membrane and that increased levels of membrane bound fructans provides higher cell membrane stability and improved freezing tolerance (DEMEL *et al.* 1998; HINCHA *et al.* 2000; VEREYKEN *et al.* 2001). These results have been reproduced *in vivo*. Transgenic *L. perenne* plants with an increased level of endogenous fructans production had improved cell membrane stability and superior freezing tolerance compared to wild type plants (HISANO *et al.* 2004). Similarly, transgenic tobacco and rice, species which naturally cannot synthesize fructans, engineered to synthesize fructans, had improved low temperature stress tolerance (KAWAKAMI *et al.* 2008; LI *et al.* 2007; PARVANOVA *et al.* 2004).

It is about what type of fat

During freezing temperatures, it is the adverse consequences of freeze-induced dehydration that cause most damage to cell membranes (DOWGERT *et al.* 1987). Freeze-induced dehydration causes formation of endocytotic vesicles that decreases the surface area of cell membranes in an irreversible way (UEMURA *et al.* 1995). Decreased cell membrane surface can directly result in an osmotically un-responsive membrane and cell death, or alternatively it can affect cell death indirectly by increasing the rate of cell death caused by over-expansion during the re-hydration phase (DOWGERT and STEPONKUS 1984; DOWGERT *et al.* 1987). The CA process lowers the temperature at which these membrane injuries occur (UEMURA and STEPONKUS 1999) and this CA-triggered improvement of freezing tolerance is partly dependent on the reorganization of the lipid composition in the cell membrane during CA (UEMURA and STEPONKUS 1999).

Comparisons between non-acclimated and cold acclimated plants have shown that cell membrane stability is associated with membrane lipid composition changes during CA (UEMURA *et al.* 1995; UEMURA and STEPONKUS 1999). Expression of the CA-involved gene *COR15a* in warm temperatures improved the stability of *A. thaliana* chloroplasts membranes

in vivo and protoplasts *in vitro* during freeze-dehydration (ARTUS *et al.* 1996). It was later shown that *COR15a* expression lowered the temperature at which destructive lipid structure changes occurred during freezing-induced dehydration, and thereby reduced cell death (STEPONKUS *et al.* 1998). Manipulations of lipid metabolism and membrane lipid composition in transgenic plants have improved freezing tolerance in tobacco (KHODAKOVSKAYA *et al.* 2006) and poplar (ZHOU *et al.* 2009), and improved chilling tolerance in tomato (DOMINGUEZ *et al.* 2010).

Metabolic changes as a signalling system during cold stress

CA-associated changes in the metabolome do not only play a direct role as biochemical modifications providing protection against freeze-associated cellular damage. Some metabolites produced during CA are also involved in signalling and regulation of further transcriptional and metabolic responses. Soluble sugars, amino acids, and ROS have all been implicated in low temperature stress signalling (Fig. 4) (ZHU *et al.* 2007). For example, when photosynthesis is carried out at low temperature this generates elevated ROS levels (see next section for details) (TAKAHASHI and MURATA 2008). ROS accumulation can affect transcriptional signalling cascades by influencing protein kinase activity, oxidization, or TF activity (Fig. 4) (Reviewed in APEL and HIRT 2004). Similarly, an increase in soluble sugars during CA can affect regulation of protein kinase activity in plants during low temperatures (Reviewed in HEY *et al.* 2010). Amino acids, e.g. proline, is also known to accumulate under different abiotic stress, including cold (SAVOURÉ *et al.* 1997). The role of amino acids in plant cold stress has been thought to be related to the osmoprotectant property (as shown in bacteria), however the relationship between proline accumulation and water stress tolerance is not unambiguous in plants (SZABADOS and SAVOURÉ 2010). Increasing evidence suggests that proline is a multifunctional amino acid playing a role in many processes during CA, including signalling (SZABADOS and SAVOURÉ 2010).

Photoinhibition and photosynthetic acclimation

From numerous studies on gene transcription, protein and metabolite level changes during CA it is apparent that regulation of photosynthesis-related processes is important during cold stress. At low temperatures, CO_2 fixation is limited and the light energy absorbed by the

plants exceeds the energy demand in light-independent processes (MURATA *et al.* 2007; TAKAHASHI and MURATA 2008). This energy imbalance leads to subsequent photo system (PS) II over-reduction and inhibition of the photosynthetic capacity, a process referred to as photoinhibition (TAKAHASHI and MURATA 2008). Not only does photoinhibition damage the photosynthetic machinery (MURATA *et al.* 2007), it is also detrimental to the entire cell due to the increased production of ROS which accompanies PSII over-reduction (KRAUSE 1988). The plants' tolerance to sub-zero temperatures is therefore inherently linked to how they deal with low temperature-induced photoinhibition (HUNER *et al.* 1993).

To avoid photoinhibition related cell damage, plants make biochemical adjustments to avoid PSII over-reduction (MULLER *et al.* 2001; POWLES 1984), a process called photosynthetic acclimation (HUNER *et al.* 1993). Two main types of CA-associated photosynthetic acclimation have been described in higher plants (HUNER *et al.* 1993; POLLOCK and JONES 1979); (i) a photochemical mechanism and (ii) non-photochemical mechanisms. The photochemical mechanism works by increasing the energy demand through increased carbon assimilation and carbon metabolism (HUNER *et al.* 1993). Non-photochemical mechanisms depend on dissipation of excess excitation energy as heat (DALL'OSTO *et al.* 2005).

Variation in the capacity for photosynthetic acclimation is shown to be correlated with genotypic differences in winter survival and freezing tolerance (RAPACZ *et al.* 2004). Moreover, there seems to be a link between freezing tolerance levels and the type of photosynthetic acclimation mechanism a plant exploits. In Pooideae grasses, a high level of winter hardiness and freezing tolerance is associated with a photosynthetic acclimation by non-photochemical mechanisms (HUMPHREYS *et al.* 2007; RAPACZ *et al.* 2004). Recently, a link between the CBF-pathway and photoinhibition has also been demonstrated; transgenic tobacco plants over-expressing *CBF1* experienced less photoinhibition during chilling stress (YANG *et al.* 2010).

There is evidence that selection for improved photosynthetic acclimation is involved in natural adaptation to cold environments in *Poeae* species. In the Babia Gora Mountain in Poland, photosynthetic acclimation to cold conditions is thought to play a role in the

altitudinal range distribution of *Festuca* species. *Festuca rubra* is limited to lower altitudes compared to species of mountainous fescues (*Festuca versicolor* and *Festuca supina*) which have more efficient non-photochemical mechanisms of photosynthetic acclimation compared to *F. rubra* (SENDERSKA *et al.* 2005).

A molecular model linking cold sensing with transcriptional responses

Although we understand some aspects of how plants sense cold and we have a good overview of major molecular responses during CA, fairly little is known about the molecular mechanisms linking cold sensing and CA transcriptional changes. Recently, a model linking cold sensing and transcription has been put forward. The model includes four components; Ca^{2+}, calcium modulated proteins (calmodulins), calmodulin binding transcriptional activators (CAMTAs), and cold responsive transcription factors (DOHERTY *et al.* 2009) (Fig. 4). Low temperature exposure increases Ca^{2+} influx (probably as an indirect response to membrane rigidification). Increased Ca^{2+} levels can activate calmodulin proteins which have the capability to activate other Ca^{2+} unresponsive CAMTA proteins (SNEDDEN and FROMM 2001). The *A. thaliana* genome contain six CAMTA genes with calmodulin binding- and DNA binding CG-1 (CGCG) domains (BOUCHÉ *et al.* 2002). Promoter studies have shown that the absence of the CG-1 element in promoters of important low temperature responsive transcription factors (both CBF and non-CBF (*ZAT12*)), lead to a decrease in their transcript levels of up to 40-50% (DOHERTY *et al.* 2009). Moreover, loss of function CAMTA gene mutants have impaired CA and are freezing-sensitive (DOHERTY *et al.* 2009). Since CAMTAs are involved in transcriptional control of TFs in the CA-process, calmodulins and CAMTAs could link low temperature sensing (Ca^{2+} influx) and downstream CA transcriptional responses (DOHERTY *et al.* 2009). However, it should be noted that the validity of this model with respect to the role of calmodulin as Ca^{2+} signal sensors during CA is not experimentally tested.

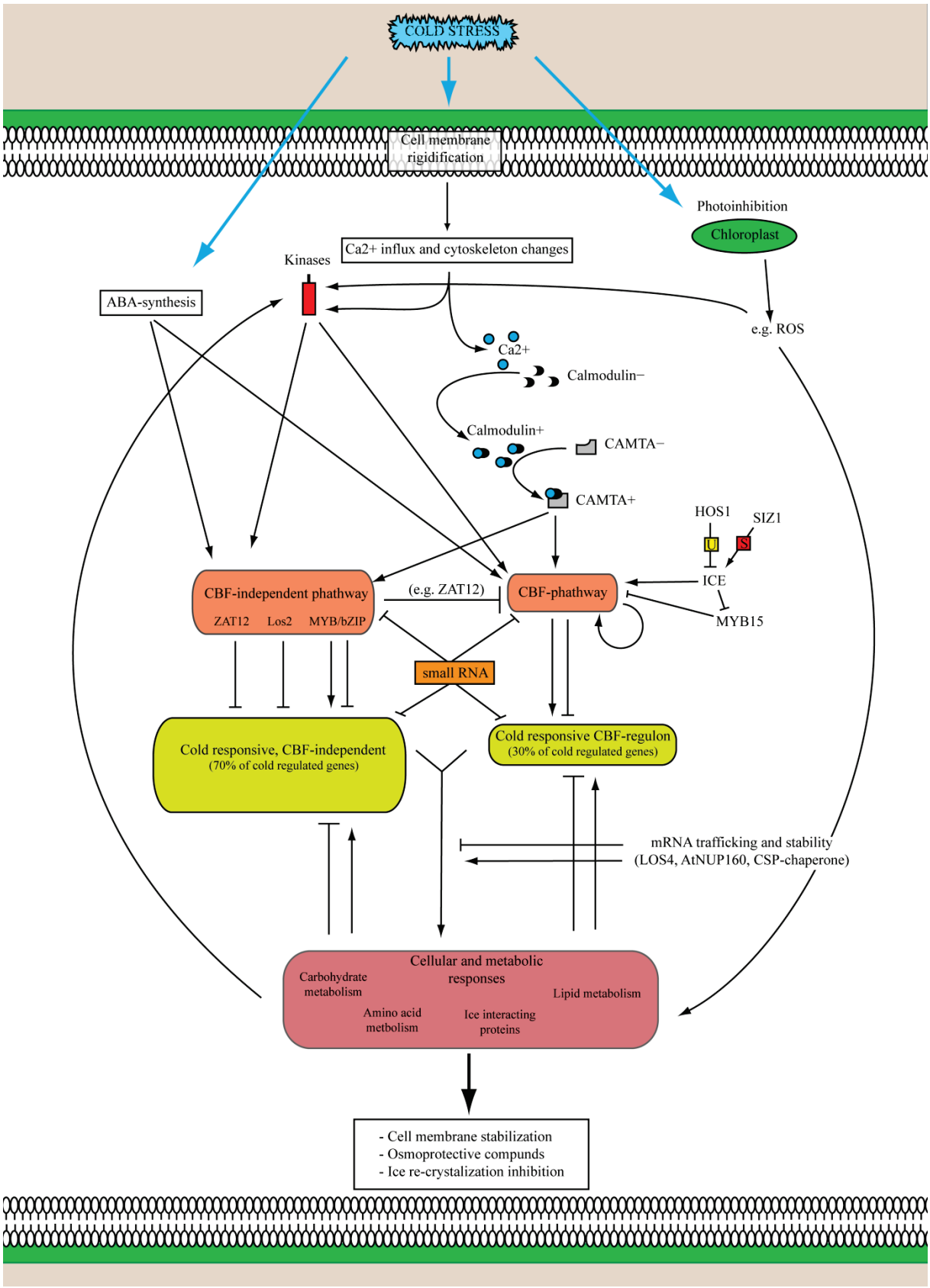


Figure 4. Schematic representation of major molecular pathways involved in the CA-process. U and S are ubiquitination and sumoylation, respectively.

Evolution of molecular adaptations

Now that the major molecular mechanisms underlying plant freezing tolerance have been outlined, it is time to change the focus from the function of small molecules in the cell to the function of big processes in nature – namely evolution. Since the time plants conquered land ~850-550 My ago (BECKER and MARIN 2009), land plants have been under selection pressure from a constantly changing climate. The result is a dazzling display of plant adaptations to virtually any type of climate - from extremely hot and dry desert climates, to the blistering cold climates of the Arctic and Antarctic. In the following part I aim to outline the evolutionary mechanisms that shape molecular adaptations in response to natural selection.

Adaptive evolution from standing and novel variation

Evolution can be defined as transgenerational phenotypic change caused by shift in the frequencies of heritable components in a population. The heritable components have traditionally been thought of as only the DNA-sequence itself, but modern biology also includes epigenetic marks, i.e. DNA-methylation and histone modifications (MARGUERON and REINBERG 2010). There are four driving forces of evolution; mutation, drift, migration and natural selection, all of which can alter the composition of genetic and epigenetic variation. Mutation, drift and migration act mainly in an environment non-dependent manner (note: we come back to the degree of environmental non-dependence of mutations later in this section, as this is the focus of an interesting debate). The process of natural selection on the other hand is solely dependent on the environment and can be defined as differential capability for survival and reproduction of individuals in a population in a given environment. Natural selection works in two major ways; (i) to conserve those genetic and epigenetic variants that are already successful in a particular environment (purifying selection), or (ii) to increase the frequency of genetic and epigenetic variants that improve adaptation, i.e. increase survival and reproductive success in the environment (directional selection). Adaptation through directional selection is often referred to as adaptive evolution or positive selection, while the relative contribution of genetic and epigenetic variants to the next generation is referred to as the fitness level of an organism.

The foundation for adaptive evolution is genetic and epigenetic variation. Without existing variation, no change can occur. Natural selection can utilize two types of variation; the standing variation in the population (pre-existing) and novel genetic or epigenetic variants that arise spontaneously (ROWAN and SCHLUTER 2008). In a given environment, the standing genetic and epigenetic variation in the population confers a range of different fitness levels. Since the fitness level of particular genetic or epigenetic variants is relative to the environment the organism live in, the standing genetic and epigenetic variation in a population could be considered a stockpile of secret weapons ready to use if the environment brings on new challenges. The mutation process constantly adds novel variants to the stockpile of potential adaptive variation, but the birth of novel and fitness enhancing mutations is an extremely rare event. Thus adaptive evolution from novel variation is of relatively little importance for rapid adaptation on a short time scale (a few generations), but plays a bigger role in adaptation over the course of millions of years (ROWAN and SCHLUTER 2008).

Genetic and epigenetic variation comes in a multitude of different forms. Some genetic variants are distinguished by the substitution of a single base pair, while others comes in the form of a gene copy number differences, or even differences in copy numbers of whole chromosomes. The different types of genetic and epigenetic variation can have different effects on the phenotype and therefore also play different roles in adaptive evolution. Smaller changes to a gene might for example change the enzyme efficiency, alter the properties of protein-substrate binding, or in rare cases create a novel protein function. Larger changes to the genome, like duplications of genes or chromosomes, create novel and redundant genetic material for evolution to work with.

Small mutations affecting gene function: SNPs, INDELs, micro- and minisatellites

Mutations in the form of single nucleotide polymorphisms (SNPs), small insertions or deletions (INDELS), and length variation of small repeated sequences (micro- and minisatellites) are responsible for constant generation of new variation at the gene level. These mutations can affect the function of a gene, and provide new molecular adaptations in many different ways. If the substitutions of a single base pair (bp) occur at a non-synonymous

site in a gene it affects the amino acid content of the protein and may result in a change in protein function (TISHKOFF *et al.* 2007). Alternatively, if a SNP is situated in a regulatory gene region it could affect variation in gene transcription patterns (STRANGER *et al.* 2007; ZHOU *et al.* 2008). INDELs and micro-/minisatellite mutations can affect gene function in a similar manner as SNPs, by affecting their function or regulation (JOHANSON *et al.* 2000; LI *et al.* 2004; ZHANG *et al.* 2008) but they differ from SNP variation in some aspects. INDELs and micro-/minisatellites change the length of the DNA and hence have different properties than SNPs. Firstly, length variations can easily result in a frame shift mutation and loss-of-function with dramatic consequences for the phenotype. Although most loss-of-function mutations has adverse effects on an organism, some have actually been shown to play a role in adaptive evolution (JOHANSON *et al.* 2000). Secondly, even if DNA-sequence length mutations do not cause a frame shift, it will likely have a large impact on protein structure and function. SNP variation on the other hand might not impact the amino acid content. Lastly, the repetitive nature of micro- (1-6 bp) and minisatellite (~10-60 bp) loci results in a high mutation rate and consequently high allelic variability in the population.

Some putative examples of the role of allelic variation in molecular evolution of freezing tolerance have been reported. Allelic variation In *L. perenne* have been correlated with both natural adaptation to winter conditions and freezing tolerance in the laboratory (SKØT *et al.* 2002). Two studies on the CBF genes have demonstrated a link between sequence variation and freezing tolerance. In an association study in barley SNP variation in a CBF gene were statistically associated with freezing tolerance levels (FRICANO *et al.* 2009). Another study on *A. thaliana* demonstrates that populations from southern climatic zones with no selection pressure for freezing tolerance had elevated rates of nucleotide substitutions in their CBF genes, consistent with relaxed selection compared to northern populations (ZHEN and UNGERER 2008).

A putative illustration of adaptive evolution of novel gene function for improved freezing tolerance is the evolution of FST enzymes. FST genes are up-regulated during CA, take part in converting sucrose molecules into fructans sugars, and improve freezing tolerance (HISANO *et al.* 2004; LIVINGSTON *et al.* 2009). The most likely ancestral function of the FST encoding genes is a vacuolar invertase (VI) function and it has been demonstrated that VI genes can

gain FST function by the substitution of a few critical nucleotides (JI *et al.* 2007; SCHROEVEN *et al.* 2008). It is therefore possible that positive selection for novel mutations transforming a VI gene function to a FST function allowed fructan accumulation during cold stress, rendering plants more tolerant to freezing (HISANO *et al.* 2004; KAWAKAMI *et al.* 2008; LI *et al.* 2007; PARVANOVA *et al.* 2004).

Function and importance of gene duplication in adaptive evolution

The ultimate source for generation of novel genetic material on which evolution can act on is DNA duplication events. These duplication events can range in size from whole genome duplication (polyploidization), duplication of whole or partial chromosomes, to duplication of single genes. Even smaller DNA duplication events occur frequently, duplications spanning only a few nucleotides, however this type of duplication, which does not create a full copy of an existing gene is not of relevance in this part. Gene duplication is hereby referred to as any process that creates one or several new redundant gene copies. The molecular mechanisms generating gene duplications are diverse, including polyploidization, duplicative transposition (e.g. non-homologous recombination), retrotransposition, and unequal crossing over (HAHN 2009) but these generative mechanisms will not be treated in further detail. I will focus on the role of gene duplication in adaptive evolution.

Birth of a novel and functionally redundant gene will usually have no or very little effect on plant fitness and the novel gene copies will accumulate mutations as if they were neutral DNA sequences. The evolutionary fate of a gene copy is therefore most often silencing (non-functionalization) within a few million years from the duplication event (LYNCH and CONERY 2000). In some instances however, gene duplication has a positive effect on fitness, resulting in adaptive evolution through preservation of both gene copies. We can classify these positive fitness effects of gene duplication into (1) proximate and (2) post-duplication fitness effects.

A proximate positive fitness effect is any fitness gain that comes as a direct consequence of having two functional genes instead of one. Theoretically, proximate positive fitness effects could be caused by either an improved mutational buffer or a gene dosage effect (HAHN

2009). If a gene plays an important role in one particular environment (for example during freezing), acquirement of an extra gene copy will act as a buffer against detrimental effects of a future mutation in this gene. Such mutational buffers improve fitness in theory, but because deleterious mutation rates are very low, the importance of gene redundancy in adaptive evolution is questioned (HAHN 2009). The gene dosage effect is likely more important in adaptive evolution. Given that variation in the quantity of a gene product can influence the phenotype, for example by providing stronger transcriptional responses, or perhaps increase the capacity of biochemical synthesis, a double gene dose can provide an instant fitness boost.

Even if there is no proximal fitness gain of having two gene copies, gene duplication can still be important in adaptive evolution. Having redundant gene copies can lead to relaxed selection pressure for maintenance of ancestral functions in both the copies and this might result in post-duplication evolution causing subfunctionalization or neofunctionalization (LYNCH and CONERY 2000; LYNCH and KATJU 2004). If relaxation of selection pressure is symmetrical on both gene copies this may result in subfunctionalization by equal partitioning of ancestral function(s). The adaptive advantage of subfunctionalization is thought to be related to an adaptive conflict between different functions of a gene (DES MARAIS and RAUSHER 2008; HITTINGER and CARROLL 2007; LYNCH and KATJU 2004). A multifunctional gene encoding a protein that carries out different functions could be under adaptive conflict if the improvement of one protein function is hindered by a conflict between alternative functions of the protein. However, if the genes are duplicated, this cancels the adaptive conflict and may facilitate molecular evolution which improves both molecular functions.

When the relaxation of selection pressure on gene copies is asymmetrical, one gene copy maintains all or most of the ancestral function(s) while the other copy can evolve freely. If then the freely evolving copy acquires a new function through post-duplication mutations, this is referred to as neofunctionalization; the evolution of a completely novel gene function (HAHN 2009). Genomic data from yeast and humans support that the most common pattern of post-duplication evolution of gene copies is rapid subfunctionalization followed by prolonged neofunctionalization, a model referred to as sub-neofunctionalization (HE and ZHANG 2005).

Gene duplication seems to be important in adaptive evolution of stress tolerance in plants. Gene ontology distribution of duplicated genes reveals overrepresentation of tandem duplicated genes in plant stress response (HANADA *et al.* 2008). An analysis of >10.000 genes involved in stress response in *A. thaliana* found that most duplicated genes undergo highly asymmetric partitioning of ancestral function (ZOU *et al.* 2009). This was further supported by data that showed that DNA-regulatory elements in the promoters of duplicated genes are retained in a highly asymmetric manner (ZOU *et al.* 2009). Extreme asymmetry does not fit predictions for the classical subfunctionalization model, but fits better with the model of sub-neofunctionalization proposed by HE and ZHANG (2005).

Comparative genome analysis has provided further evidence for the importance of gene duplication in adaptive evolution of freezing tolerance. In highly cold-tolerant species of the Brassicaceae family duplication of *COR15* genes and subsequent rapid amino acid replacements suggest that sub- or neofunctionalization of this gene family could be involved in adaptation to low temperature stress. In Pooideae grasses, lineage-specific gene duplication events of CBF, FST, and IRIP genes, have been implicated in adaptation to cool climates (BADAWI *et al.* 2007; HISANO *et al.* 2008; SANDVE *et al.* 2008; SANDVE and FJELLHEIM 2010). In addition, copy number variation in the CBF genes in temperate cereals has recently been associated with within-species differences in freezing tolerance (KNOX *et al.* 2010).

Novel miRNA and siRNA loci can arise from inverted duplications of genes (DE WIT *et al.* 2009; VAZQUEZ *et al.* 2008), and there are indications that evolution of novel miRNA is important for lineage-specific evolution of low temperature stress tolerance. Recently, the presence of many cold regulated non-conserved Pooideae specific miRNAs were identified (ZHANG *et al.* 2009). Zhang and colleagues (2009) suggested that selection for acquirement of novel miRNA involved in regulation of low temperature gene expression could have been important in adaptation to winter habits and low temperatures in the Pooideae lineage.

Epigenetic mechanisms and adaptive evolution

In the traditional model of adaptive evolution the environment has one role; it dictates the natural selection pressure which act on the genetic variation underlying the phenotypic variation. In recent models of adaptive evolution, the environment has an additional role in creating epigenetic variation on which natural selection can act on (JABLONKA and LAMB 2005; ZEH *et al.* 2009). Epigenetic variation arises spontaneously as a response to environmental stimuli (independent of DNA-sequence variation) and influence gene transcription (BOYKO and KOVALCHUK 2008). Many recent studies have demonstrated that variation in DNA-methylation is strongly associated with adaptations to specific environments in many plant species (GAO *et al.* 2010; HERRERA and BAZAGA 2010; LIRA-MEDEIROS *et al.* 2010) and both DNA-methylation and histone modifications have now been shown to cause transgenerational adaptation to new environments in the laboratory, including adaptation to low temperature stress (BOYKO *et al.* 2010; LANG-MLADEK *et al.* 2010).

Even though adaptive advantage of transgenerational epigenetic inheritance has been demonstrated in nature (GALLOWAY and ETTERSON 2007), the importance of epigenetic mechanisms in adaptation to freezing stress is still largely unknown. One rare example has been reported in Norway spruce (*Piceae abies*). Studies of clonal plants of Norway spruce revealed that low temperature tolerance is dependent on epigenetic changes that take place in the developing embryo (JOHNSEN *et al.* 2005). Albeit the molecular mechanisms responsible for this epigenetic memory are not understood, there are indications that miRNA mediated transcriptional regulation could be involved (YAKOVLEV *et al.* 2010).

Apart from the direct effect of epigenetic variation on the phenotype, theoretical biologists also speculate if the environment also triggers generation of novel DNA-sequence variation mediated by epigenetic mechanisms. Zeh et al. (2009) proposed that "environmental shock" due to climate change can repress the epigenetic control of transposable element (TE) activity, which in turn leads to bursts of TE activity. At the molecular level TE activity can contribute to adaptive evolution by producing novel gene/genome structures, changing gene regulation, and promoting gene duplication (OLIVER and GREENE 2009). The consequences of epigenetic mechanisms in adaptive evolution are large. First, environmentally triggered adaptive changes

in epigenetic variation imply that adaptation to novel climatic conditions can happen with exceptional rapidity, within one generation. Secondly, if abiotic stress triggers an increased rate of TE-related mutations, this will boost the level of standing genetic variation for natural selection to act upon. Compared to the traditional model of adaptive evolution, a model of adaptive evolution which includes epigenetic mechanisms brings about highly efficient adaptive responses.

The never ending story of adaptive evolution

In an evolutionary time perspective the environment is a constantly changing factor. We can therefore view adaptive evolution to the climate as a never ending story of repeated cycles, starting and ending with sexual reproduction. The standing genetic and epigenetic variation present in the population after sexual reproduction in the spring/summer provides the basis for natural selection to act on. Then, the environment dictates the terms of survival and success in sexual reproduction the subsequent year, but at the same time affects the generation of new genetic and epigenetic variation.

Figure 5 shows a schematic and simplified model of the processes involved in adaptive evolution of freezing tolerance in perennial Pooideae grasses. When autumn makes an entrance and the process of CA initiates, plants with differing CA-capacity acquire different levels of freezing tolerance. In addition, molecular processes induced by environmental stress, such as epigenetic modifications and increased TE-activity, can generate even more phenotypic variation. Differential survival through the winter and spring results in changes in the frequencies of the genetic and epigenetic variants in the population, and the cycle can start all over again. As long as the environment favours plants with the best freezing tolerance, and genetic/epigenetic variation for freezing tolerance exists in the population, adaptive evolution of improved freezing tolerance will resume.

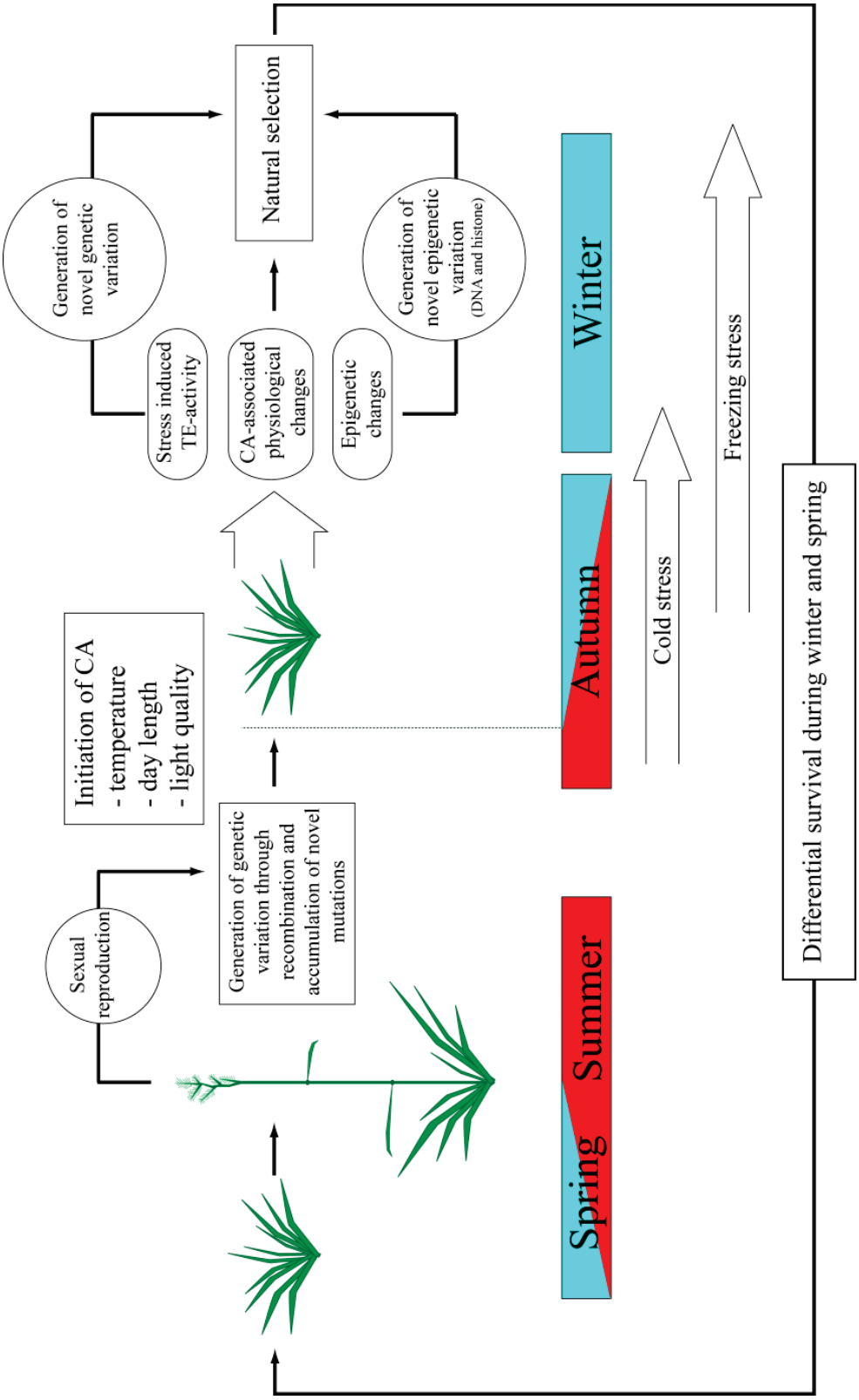


Figure 5. Adaptive evolution of improved freezing tolerance. Schematic model of the processes involved in adaptive evolution of freezing tolerance in perennial Pooideae grasses.

Results in brief

Paper I: *Identification of candidate genes important for frost tolerance in Festuca pratensis Huds. by transcriptional profiling.*

We identified candidate genes involved in CA with putative importance for freezing tolerance in the forage grass *F. pratensis* by a combination of suppressive subtraction hybridization (SSH) library of expressed sequence tags (ESTs) and microarray analysis. A SSH-EST library from non-cold acclimated and cold acclimated crown tissues were cloned and sequenced. Cloned ESTs from the SSH library were spotted on a cDNA array and the expression of these genes was measured in crown tissue from plants with high and low freezing tolerance. The SSH-EST library produced 372 ESTs putatively involved in CA, many of which were homologous to known genes with functions in CA and abiotic stress in other plant species. The microarray analysis identified 7 differentially expressed genes between genotypes of different freezing tolerance. Out of these 7 genes, four were up-regulated, and three down-regulated in high freezing tolerance plants compared to low freezing tolerance plants. Two of the up-regulated genes, homologues of a tumor suppressor gene (QM) and a triose phosphate/phosphate translocator (TPT), are excellent candidate genes underlying the difference in freezing tolerance.

Paper II: *Genetic mapping of DArT markers in the Festuca-Lolium complex and their use in marker-trait (freezing tolerance) association analysis*

The newly developed DArTFest genotyping chip was used to saturate existing genetic maps of *Lolium multiflorum* and *F. pratensis*. We then used the DArT marker resources in an integrated analysis to identify genomic loci involved in freezing tolerance in the Furs0357 *L. perenne* x *F. pratensis* hybrid population. Furs0357 were selected for divergent levels of freezing tolerance to generate high and low freezing tolerant sub-populations and the differences in genome content between the phenotypes were studied. In total, 96 DArT markers were found to associate with freezing tolerance phenotypes, and five of these were placed on one of the genetic maps. These five mapped markers corresponded to three genomic regions on LG2, LG4, and LG7, respectively. The associated region on LG4 corresponded to a previously reported QTL for freezing tolerance in *F. pratensis* and cereals close to *Vrn1,* a key regulator in vernalization.

Paper III. *Tracking the evolution of a cold stress associated gene family in cold tolerant grasses*

Data from BAC-sequencing combined with different bioinformatics approaches was integrated to (i) demonstrate the presence of a multi-gene family of ice re-crystallization inhibition protein (IRIP) coding genes in Pooideae grasses and (ii) study the evolution of this gene family. The first IRIP gene probably evolved from a LRR phytosulfokine receptor kinase-like gene through increased copy numbers of a repeated amino acid motif (NxVxG). Subsequently, the IRIP gene underwent multiple rounds of gene duplication, both prior to- and after the split between cereals and forage grasses. Structural divergence between IRIP gene paralogs and evidence from functional studies by other research groups indicate that IRIP paralogs underwent post-duplication subfunctionalization.

Paper IV. *Did gene family expansions during the Eocene-Oligocene boundary climate cooling play a role in Pooideae adaptation to cool climates?*

In the last paper, we explored the idea that paleoclimatic conditions during the Eocene-Oligocene (E-O) boundary global cooling (~34-26 My ago) was important in Pooideae adaptation to cold environments. We did this by investigating the temporal patterns of gene duplication events giving birth to Pooideae specific gene families involved in freezing tolerance. The timing of the initial Pooideae specific duplications in CBF, FST, and IRIP gene families were estimated to have occurred inside-, or in a few instances just prior to the E-O cooling period. Hence, our data support the hypothesis that the E-O cooling selected for novel molecular adaptations to low temperature stress in a Pooideae ancestor, which could have been instrumental in subsequent ecological expansion of Pooideae grasses into cooler habitats.

Discussion

Molecular adaptation to freezing on different temporal scales

Extant land plants are thought to share a common ancestor 488-443 My ago (SANDERSON *et al.* 2004; STEEMANS *et al.* 2009). Along this ~500 My evolutionary time line, land plants lineages have been under varying selection pressure for cold and freezing tolerance. This selection pressure has been enforced by three global scale glaciations (YOUNG *et al.* 2010), numerous regional scale ice ages, and constant between-year shifts in local climatic conditions. The result is a fascinating diversity of molecular mechanisms which help plants surviving freezing temperatures and cold winters. Some of these mechanisms are ancient and shared between many land plants, and yet other mechanisms have evolved recently and are specific to certain plant groups.

Analyses of molecular adaptation on different temporal scales reveal different evolutionary patterns. At the million year perspective, larger and rarer evolutionary events are more frequently observed at the DNA level, while over the course of a few generations most molecular adaptations come from changes in the standing genetic variation (ROWAN and SCHLUTER 2008). In Papers I and II, we used artificial selection for divergent freezing tolerance and measurements of allele frequencies and gene expression levels to study rapid adaptive responses to selection. Although artificial selection only mimics selection pressure under natural conditions, these types of experiments can help us identify loci underlying natural variation in freezing tolerance and improve our understanding of molecular mechanisms important for adaptation to abiotic stress in nature. Papers III and IV, on the other hand, deal with molecular evolution of genes involved in CA and freezing tolerance in the entire Pooideae lineage, a time window of ~45 million years. The focus of these papers was the evolution of rare Pooideae specific molecular innovations and how these molecular novelties potentially impacted Pooideae adaptation to cold environments. Differences in the temporal scale of the analyses in the two first and two last papers provide different answers to the same question, and complement each other to obtain a fuller picture of how evolution works to improve freezing tolerance in grasses.

Rapid adaptation: natural selection on standing genetic variation

Standing genetic variation in gene expression associated with freezing tolerance

In Paper I, we investigated gene expression differences between plants selected for high and low freezing tolerance in a gene set enriched with genes involved in CA. After three weeks of CA seven genes (1.3%) had >2 fold difference in expression level between plants with high and low freezing tolerance. Two of these genes, the triose phosphate/phosphate translocator (TPT) -like and QM-like, have functionally characterized homologues in other plant species and are striking candidates for being involved in adaptive evolution of freezing stress tolerance.

The TPT protein is situated in the chloroplast membrane and is responsible for carbohydrate export to the cytoplasm during photosynthesis (WALTERS *et al.* 2004). Expression of TPT has been positively correlated with increased levels of TPT protein in the chloroplast and increased levels of sucrose, fructose and other carbohydrates in the cytosol of *A. thaliana* (WALTERS *et al.* 2004). TPT activity has also been positively correlated with accumulation of cytosolic sucrose, fructose and other carbohydrates in tobacco (HÄUSLER *et al.* 2000) and potato (HEINEKE *et al.* 1994). Carbohydrate accumulation is a major metabolic response during CA (COOK *et al.* 2004; MARUYAMA *et al.* 2009) and is thought to affect freezing tolerance through various molecular mechanisms, for example aiding in photosynthetic acclimation (STRAND *et al.* 2003) and cell membrane interaction and stabilization (HINCHA *et al.* 2000).

One putative mechanism in which increased TPT expression could affect freezing tolerance is by indirectly affecting fructan levels. In Pooideae grasses fructan accumulation during CA is a common response (LIVINGSTON and HENSON 1998; PUEBLA *et al.* 1997), and high fructan levels have been shown to improve freezing tolerance by stabilizing cell membranes during freezing (HISANO *et al.* 2004). The initial step in fructan synthesis is the conversion of two sucrose molecules to a fructan molecule by the FST enzyme 1-sucrose-sucrose-fructosyltransferase (LIVINGSTON *et al.* 2009). Although highly speculative at this point, it is conceivable that higher expressions of TPT in high freezing tolerance phenotypes increase

cytosolic sucrose levels and thereby cause increased rates of fructan synthesis and freezing tolerance.

The second highly interesting candidate gene identified in Paper I is the QM-like homolog. *QM* was first isolated and identified as a putative tumor suppressor gene (MONTECLARO and VOGT 1993) and has subsequently been shown to have DNA binding properties and to regulate the transcription of a proto-onco gene in human cells (OH 2002). QM homologs are highly conserved among all eukaryotic organisms, suggesting an essential role in basic cell maintenance (FARMER *et al.* 1994) but the QM function is elusive. Studies of cellular localization demonstrate co-localization with the ribosome, and QM-homologues have a conserved L10e ribosomal binding domain which is involved in organizing the architecture of the aminoacyl tRNA binding site of the ribosome (LILLICO *et al.* 2002; LOFTUS *et al.* 1997). Both animal and plant QM-homologs have been implicated in abiotic and biotic stress responses. In carp fish, a QM homolog was up-regulated subsequent to viral and bacterial infections (WEN *et al.* 2005). In plants, up-regulation of QM during cold (and also other abiotic stress) has been demonstrated in rice (XIO-CHUN 2002) and the temperate plant *Caragana jubata* that grows under extreme cold and high altitude in the Himalayas (BHARDWAJ 2010). Furthermore, expression of a tomato QM gene in yeast protected against oxidative stress by increased accumulation of proline (CHEN 2006). Since molecular biotic and abiotic stress responses ultimately depend on mRNA-translation and protein synthesis, one hypothesis is that QM up-regulation during cold stress mitigates adverse effects of abiotic stress on the translation machinery. However, ribosomal proteins are often implicated in extra-ribosomal functions (WARNER and MCINTOSH 2009). In the case of the human QM gene it has been shown to be a transcriptional regulator (OH 2002), thus another hypothesis is that the QM protein is involved in transcriptional regulation of other genes involved in CA, e.g. genes in the proline synthesis pathway (CHEN 2006).

Rapid adaptive evolution to abiotic stress through changes in gene expression patterns is common (SAUVAGE *et al.* 2010; SØRENSEN *et al.* 2007). Because abiotic stress tolerance is a highly quantitative trait, adaptive evolution from standing genetic variation is expected to happen through small changes in many genes, so called “soft sweeps” (PRITCHARD *et al.* 2010). This exact pattern was observed in *Drosophila melanogaster* during adaptation to

different abiotic stresses, where most genes only changed their expression level slightly, much less than 2-fold (SØRENSEN *et al.* 2007). Unfortunately, limitations in the production of the custom made microarrays in our study excluded the possibility of testing for small significant changes in expression levels; hence in Paper I, we used the 2-fold change threshold to classify genes with different expression patterns. Moreover, in our analysis we used pooled mRNA from multiple genotypes in the high and low freezing tolerance groups, respectively. The advantage of pooled mRNA is that it acts as a buffer to average out genotype specific expression differences not related to the freezing tolerance *per se.* But at the same time the pooling strategy also averages out small genotype specific expression differences that truly are related to adaptive changes in gene expression. Consequently, the combination of the 2-fold filter and the mRNA pooling strategy results in a highly conservative estimate of the numbers of genes differentially expressed between high- and low freezing tolerant genotypes. On the other hand, this approach increases the chance that the expression differences associated with freezing tolerance levels are causally linked to the phenotype.

Selection for freezing tolerance affected allelic composition and polyploidy levels

In Paper II 96 genomic markers were associated with freezing tolerance levels in a population of *L. perenne* x *F. pratensis* hybrids. Five of these markers had known genetic map positions in the *Lolium* or *Festuca* genomes and interestingly these five markers correspond to linkage groups (LGs) and genomic loci which have previously been implicated in freezing tolerance in cereals and forage grasses.

Three of the significant markers on LG4, two mapped in *L. multiflorum* and one in *F. pratensis*, mapped to a chromosomal region corresponding to a freezing tolerance QTL in *F. pratensis* closely linked to the *Vrn1* gene (ALM *et al.* submitted). This QTL is also found in cereals where it is referred to as FR-1 (GALIBA *et al.* 2009). Figure 6 shows a comparative map between LG4 in forage grasses and the syntenic relationship to the *B. distachyon* chromosome 1 (Bd1). The existence of *Vrn1* paralogs in grasses (PRESTON and KELLOGG 2007) complicates the syntenic relationships. In rice the two *Vrn1* paralogs (also referred to as FUL1 and FUL2) are situated on different chromosomes (3 and 7), while both copies are found on Bd1 (HIGGINS *et al.* 2010). Syntenic relationships between forage grass LG4 and

both regions on Bd1 harbouring Vrn1 copies are revealed by blast analysis of the mapped DArT markers (Fig. 6). Markers situated close to both *Vrn1* paralogs were significantly associated with freezing tolerance, however as many as 7 significant DArT markers were clustering close to the BdVrn1b/FUL2 (Fig. 6). This strongly supports the notion that the significant markers on LmLG4 and FpLG4 reflect the same underlying QTL and that this QTL is located close to a Vrn1 gene.

It has been debated whether the gene underlying FR-1 QTL is directly involved in freezing tolerance or if the FR-1 QTL is an artefact of tight linkage with the vernalization response gene *Vrn1*. *Vrn1* is a mads-box transcription factor involved in the transition between vegetative and reproductive phase in Pooideae grasses (CATTIVELLI *et al.* 2002; TREVASKIS *et al.* 2007). This transition involves down-regulation of genes involved in CA (GALIBA *et al.* 2009). Recently, polymorphisms in *Vrn1* were shown to exhibit a pleiotropic effect on freezing tolerance by directly or indirectly affecting expression of genes involved in the CA pathway (e.g. CBF genes) (DHILLON *et al.* 2010). Hence, the DArT markers associated with freezing tolerance on LG4 in forage grasses could be linked to differences in the vernalization control rather than genes involved in CA.

The significant markers on FpLG7 and LmLG2 also correspond to chromosomal regions which previously have been implicated in Pooideae freezing tolerance. The peak marker of a freezing tolerance QTL on *F. pratensis* LG7 is mapped only a few centi-Morgan (cM) from the significant DArT marker on LG7 in Paper II (ALM *et al.* submitted). LG2 has previously been shown to harbour QTL(s) for freezing tolerance in *Lolium* x *Festuca* hybrids (KOSMALA *et al.* 2006) and winter survival in *F. pratensis* (ALM *et al.* submitted). Unfortunately the syntenic relationship between the significant marker on LG2 in this study and the other above mentioned QTLs on LG2 is uncertain due to the lack of common bridging markers between the genetic maps.

A second result worth discussing in the context of adaptation to stress tolerance is the observed effect of selection on ploidy level (tetraploids vs diploids) in Paper II. A significant higher proportion of tetraploid plants were found in the low frost tolerance group. This result

was surprising, first of all because we were unaware that the hybrid population we studied segregated for ploidy levels, but mostly because polyploidization normally is assumed to improve stress tolerance in plants (CHANDRA and DUBEY 2010; CROW and WAGNER 2006; FAWCETT *et al.* 2009). The discrepancy between our findings and the general presumption on polyploidization and stress tolerance could be related to genome constitution of the hybrid population in this study. It has been shown that introgression from *F. pratensis* into a *Lolium* genomic background enhances freezing tolerance in Festulolium hybrids (KOSMALA *et al.* 2006; RAPACZ *et al.* 2004), and it is possible that the tetraploid plants in the population had fewer *Festuca* introgressions due to their ploidy levels. Limited FISH-experiments showed a higher number of introgressions in the high freezing tolerance group compared to the tetraploid plants in the low freezing tolerance group, and supports this idea.

The allelic composition in the Festulolium plants undoubtedly differed between populations with differing selection regimes; however, there is always a chance that other evolutionary forces than natural selection contribute to allele frequency shifts. Genetic drift during artificial selection could lead to allelic divergence. Another possibility is that the genotyping technology itself could produce biased genotyping scores that mimic the pattern of phenotype-genotype correlations, for example due to the skewed distribution of ploidy levels. Although we have not experimentally controlled for drift or genotyping biases in Paper II, we find it unlikely that our results are heavily influenced by these factors. Firstly, an identical selection regime to the one used in this study has previously been tested in other experiments with no effects of drift (unpublished results). Secondly, genotyping error is expected to produce a random distribution of markers with low *P*-values along the chromosomes. In Paper II there is a clear tendency for low *P*-value markers to be closely linked on the chromosomes, reflecting the decreasing linkage disequilibrium with the distance to the significant markers (data not shown).

Lastly, I want to highlight two methodological aspects of Paper II, relevant for the interpretation of the results. First, the genetic map positions in *Lolium* were estimated in a *L. multiflorum* population, while the Festulolium population used in this study had a *L. perenne* background. In a worst case scenario, this could result in erroneous inferences of the genetic position of markers associated with freezing tolerance in this study. However, the chance of

this is negligible due to the close evolutionary relationship between *L. multiflorum* and *L. perenne* genomes (CATALÁN *et al.* 2004). Second, the linkage disequilibrium in the Furs0357 population is expected to be high as a result of the breeding scheme (few rounds of recombination and hybrid plants). Consequently, significant associations between DArT markers and phenotype can only identify chromosome regions containing QTLs, and not specific genes. This is also why as many as 96 markers were identified as significantly associated with freezing tolerance in Paper II.

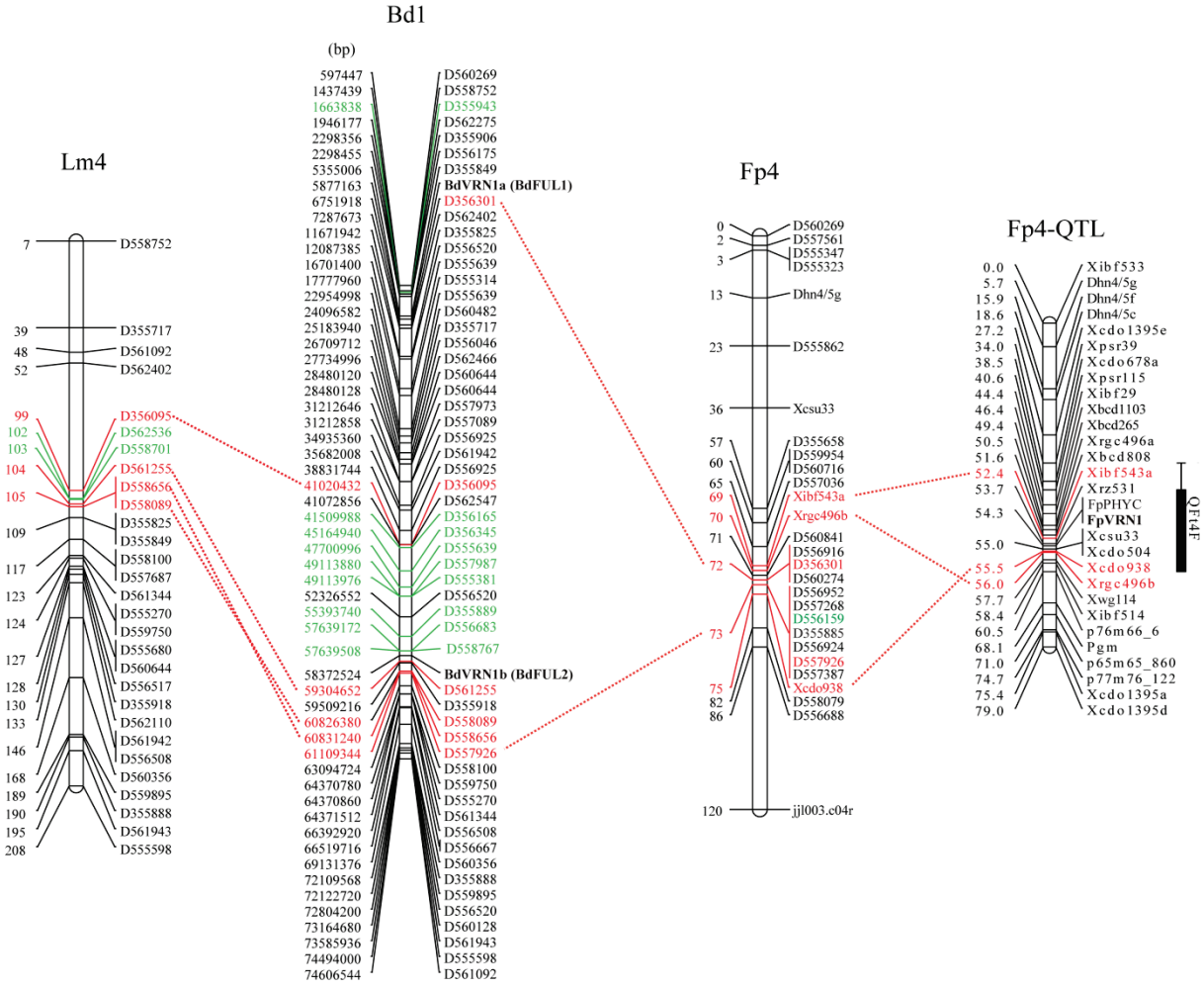


Figure 6. Syntenic relationships between *F. pratensis*, *L. multiflorum*, and *B. distachyon* chromosome. Comparative map showing the syntenic relationships between the freezing tolerance QTL on LG4 in *F. pratensis*, DArT markers associated with freezing tolerance phenotype, and the VRN1 paralogs in the *B. distachyon* genome. Red colour text denotes bridging markers. Green colour text denotes DArT markers associated with freezing tolerance phenotype. On the Bd1 chromosome the left side numbers are physical positions in base pair. Left side numbers for all other maps are genetic distance in cM. Fp4-QTL is the chromosome 4 linkage group from *F. pratensis* with the *Vrn1*-associated freezing tolerance QTL (QFt4F) from Alm et al. (submitted).

Adaptation on a geological time scale; Pooideae specific innovations

Evolution of novel ice-binding protein domains

Pooideae specific evolution of an ice-binding domain (IRI domain) took place in the IRIP gene family (SANDVE *et al.* 2008) and it is speculated if evolution of this IRI domain was important for Pooideae lineage adaptation to cold climates (SIDEBOTTOM *et al.* 2000; TREMBLAY *et al.* 2005). IRIP genes evolved from a bipartite ancestral gene consisting of a leucine rich repeat (LRR) domain and a protein kinase (PK) domain (TREMBLAY *et al.* 2005). The PK domain was later lost in the Pooideae lineage due to a frame shift mutation caused by the evolution of the IRI domain. The lack of IRI domain homologs outside the Pooideae lineage led some authors to suggest that the IRI domain could represent exogenous DNA inserted into an ancestral Pooideae genome by a TE (TREMBLAY *et al.* 2005). However, no TE signatures flanking the IRIP genes have been identified (TREMBLAY *et al.* 2005); so how did the IRI domain arise?

Even when no homology to other plant genes can be found, alternative hypotheses to the TE insertion of exogenous DNA exists. Novel endogenously derived DNA can either arise through (i) large scale duplications (whole genes, chromosomes or genomes) or (ii) small scale duplications, for example through mini- or microsatellite copy number expansion. If multiple duplications of a small DNA-sequence motif occur, this novel DNA-sequence could appear as having no homology to other genes. In Paper III, we studied the sequence similarities between the IRIP genes and their closest homolog, the rice LRR-PK gene, to identify small common IRI domain-like motifs. Analysis of the amino acid sequences demonstrated that the IRI domain most likely arose through duplication and expansion of a small 5 codon long motif (Fig. 7). Evolution of the IRI domain thus represents a putative example of adaptive evolution of minisatellite repeat variation. Although in the case of the IRI domain, the variation in repeated elements did not modify the function of an existing allele as is known from many examples (KASHI and KING 2006; LI *et al.* 2004), but gave birth to a completely novel gene function (Fig. 7).

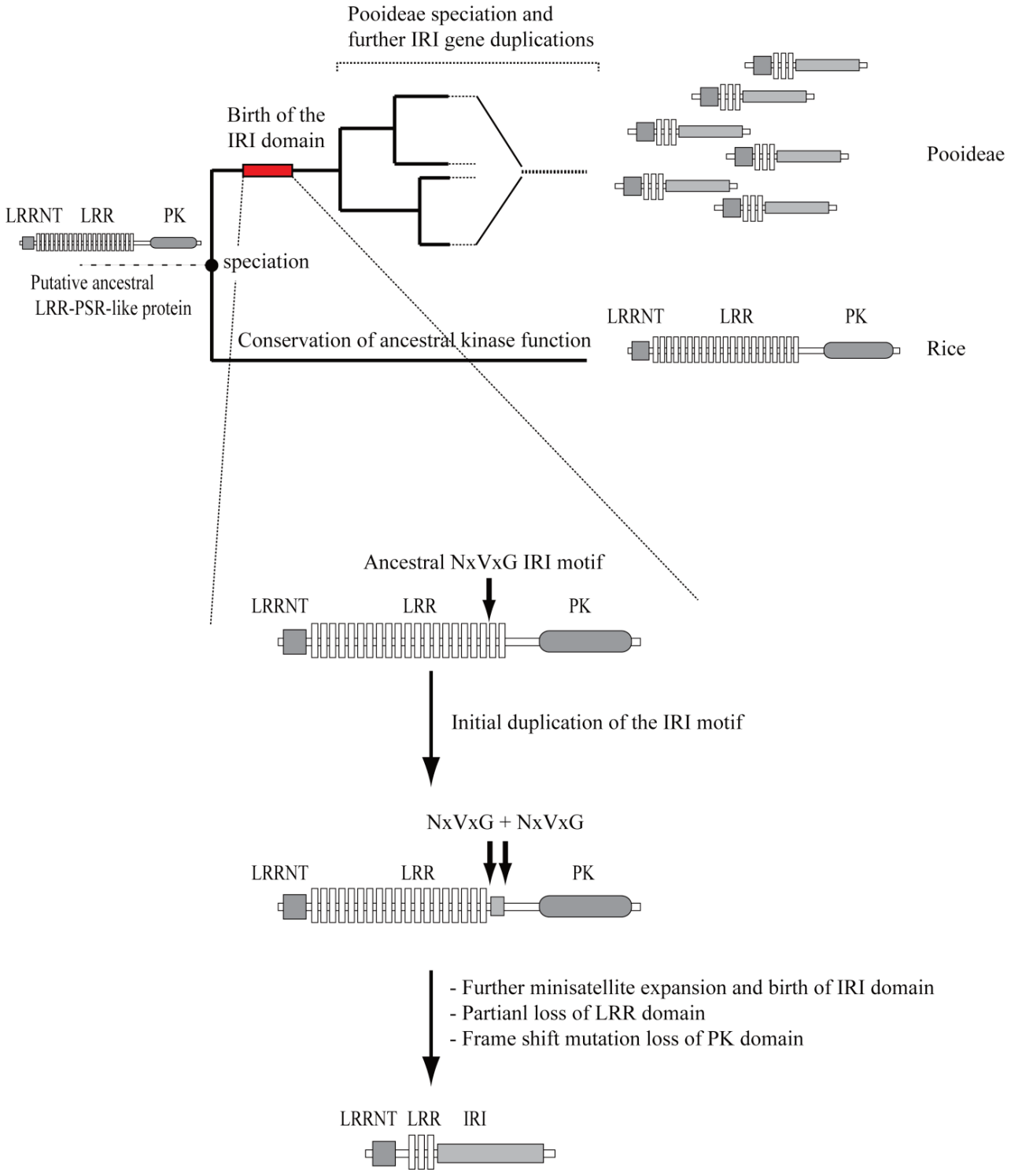


Figure 7. Schematic model of IRI domain birth and subsequent Pooideae IRI gene family expansion. After the divergence of rice and Pooideae grasses the LRR-PK homologue in the Pooideae lineage underwent substantial molecular evolution. An ancestral amino acid motif (NxVxG) duplicated and expanded in copy numbers. This resulted in a frame shift mutation and subsequent loss of the protein kinase (PK) domain. As the novel IRI genes evolved novel functions, the LRR domain was reduced. The figure is not a precise representation of the timing of duplications or IRI domain evolution. LRRNT=Leucine-Rich Repeats N-Terminal, LRR=Leucine-Rich Repeats, IRI=Ice Re-crystallization Inhibition, PSR=Phytosulfokine Receptor.

Adaptive lineage-specific gene duplications

Many observations suggest that duplicated genes are overrepresented in functional classes involved in abiotic stress responses in plants for, in particular tandem duplicated genes (RIZZON *et al.* 2006; WANG *et al.* 2008). One hypothesis is therefore that lineage-specific gene duplication is an important mechanism in adaptive evolution of abiotic stress tolerance (HANADA *et al.* 2008; LESPINET *et al.* 2002). In a study comparing gene duplication patterns between *A. thaliana*, moss, rice, and poplar, genes with a function in cold stress response were significantly enriched for linage specific gene copies (HANADA *et al.* 2008). In the Pooideae grasses, lineage-specific expansions of gene families involved in cold stress has been identified in the IRIP family in Paper III (SANDVE *et al.* 2008), the CBF family (BADAWI *et al.* 2007), and FST enzyme family (HISANO *et al.* 2008).

Inferences of the adaptive evolutionary role of the Pooideae specific gene duplications can be based on extant gene functions and DNA sequence substitution patterns. Post-duplication regulatory subfunctionalization has been demonstrated for CBF (AGARWAL *et al.* 2006; WINFIELD *et al.* 2010), IRIP (TREMBLAY *et al.* 2005), and FST genes (HISANO *et al.* 2008). Positive selection on post-duplication mutations is expected to leave footprints at the DNA sequence level in the form of elevated rates of non-synonymous to synonymous nucleotide substitutions. Such footprints were indeed found in Paper IV in the CBFIIIc/d clade where several amino acid sites were predicted to have been under positive selection. The test for positive selection has limited power (ANISIMOVA *et al.* 2001), hence the absence of signatures of positive selection does not exclude the possibility that positive selection played an important role. Moreover, positive selection can also target substitutions in non-coding regulatory regions of genes, for example by adaptive loss or gain of transcription factor binding sites, and this cannot be accounted for in a test for positive selection on amino acid content. It is also possible that many of these gene duplications brought about proximate fitness effects as a result of gene dose increase, for example through more efficient fructan synthesis or larger doses of IRIPs being exported to the apoplast.

Paleoclimate and evolution of Pooideae freezing tolerance

Paleoclimate and natural selection

Selection pressure from a cold climate is needed for adaptive evolution of molecular mechanisms involved in freezing tolerance to take place. Since the entire Pooideae lineage is dominated by cold tolerant species (GRASS PHYLOGENY WORKING GROUP 2001) it is likely that important molecular adaptations to freezing stress arose in an ancient Pooideae ancestor. Dramatic paleoclimatic changes have been implicated in adaptive gene content increase in many plants. The Cretaceous-Tertiary (K-T) boundary, so famously blamed for eradicating the dinosaurs, represents an exceptional rapid and radical change in paleoclimatic conditions. A study of the timing of whole genome duplications in major plant lineages demonstrated a remarkable correlation between the K-T boundary and paleopolyploidization events (FAWCETT *et al.* 2009). Fawcett and colleagues (2009) argued that plant lineages with double genomes had improved chances of surviving the K-T boundary, possibly due to improved stress tolerance as a consequence of polyploidization.

Another global scale climate change event, the Eocene-Oligocene (E-O) boundary climate cooling (~33.5-26 My ago) is coincident with the early evolution of the Pooideae lineage (COXALL and PEARSON 2007; ELDRETT *et al.* 2009; LIU *et al.* 2009; VINCENTINI *et al.* 2008; ZACHOS *et al.* 2001; ZANAZZI *et al.* 2007). One intriguing hypothesis is therefore that paleoclimatic conditions during the E-O boundary selected for improved freezing tolerance in a Pooideae ancestor essential for subsequent adaptation and ecological expansion into cold climate habitats.

If the hypothesis on the role of E-O cooling in Pooideae adaptation to cold climates is correct, we can expect Pooideae specific molecular evolution associated with the E-O cooling period of genes involved in freezing tolerance. This exact pattern was supported by the results in Paper IV, where we found that the timing of the gene duplication events giving birth to the Pooideae specific IRIP, CBF, and FST genes corresponded with the E-O cooling period. The importance of the E-O climate change in grass evolution and adaptation has previously been studied with a different focus. A drop in global carbon dioxide levels associated with the E-O

boundary has been linked to adaptive transitions from C3 to C4 photosynthesis in PACCAD grasses (VINCENTINI *et al.* 2008). This finding supports the importance of the paleoclimatic conditions at the E-O boundary in grass evolution.

Did the paleoclimate during the E-O boundary play dual roles?

It is evident that dramatic climate change represents a sudden increase in plant abiotic stress levels and will result in changes in the genetic constitution of a species as a consequence of altered selection pressure. But as mentioned in the introduction, abiotic stress can also trigger genome instability and generate novel DNA duplications through increased recombination rates (FILKOWSKI *et al.* 2004) and TE activity (BENNETZEN 2000; OLIVER and GREENE 2009; ZEH *et al.* 2009). In barley, for example, differences in TE activity resulting in divergence in genome size have been associated with adaptation to local microclimatic differences (KALENDAR *et al.* 2000). Although it is highly speculative, it is therefore possible that the E-O cooling played a dual role in the evolution of freezing tolerance and climate adaptation in Pooideae grasses; both triggering the creation of new genetic variation and at the same time selecting for those newly created and pre-existing genetic variants which conferred improved adaptation to the cooling climate.

Grass evolution time line; what we know and how we know it

Papers III and IV are dealing with estimating the absolute time of Pooideae specific gene duplication events putatively involved in evolution of freezing tolerance. Some of the results and conclusions are dependent on assumptions about the timing of major splits in the Poaceae species evolutionary tree. A discussion about what we know about the absolute time estimates of major transitions in the Poaceae evolutionary history is thus inevitable.

A central problem with reconstructing an absolute time scale of the Poaceae group evolution is the poor grass fossil record, and the ages of major evolutionary transitions are therefore still debated (PRASAD *et al.* 2005; VINCENTINI *et al.* 2008). The earliest Poaceae macro fossil record dates back 55 My (CREPET and FELDMAN 1991). Some of the floral characteristics of this fossil are common for members of the BEP (Bambusoideae, Ehrhartoideae, Pooideae),

and PACAAD clade (Panicoideae, Aristidoideae, Chloridoideae, Centothecoideae, Arundinoideae, Danthonioideae) (BREMER 2002). Hence, 55 My is considered a lower age boundary for the ancestor of most agricultural important grasses. In addition to this macro-fossil, two reports on grass paleo-pollen (HERENDEEN and CRANE 1995; JACOBS *et al.* 1999) and two reports of paleo-phytolith assemblages in different geological strata (PRASAD *et al.* 2005; STROMBERG 2005) shed light on the ancient evolutionary history of the Poaceae.

Because of the poor grass macro-fossil record, the age estimates of many important evolutionary transitions in the grass family are attributed to modern methods of molecular dating. These methods work by transforming evolutionary distances in a phylogenetic tree, originally inferred from differences in DNA sequences, into absolute time by (i) the use of one or more time constraint(s) on nodes in the tree based on fossil evidence, and (ii) using a statistical model for how evolutionary rates vary across the phylogeny. Several calibration methods assuming different evolutionary models are available. The most naive method is to enforce a strict molecular clock on the phylogeny, i.e. assuming that all branches in a phylogeny have identical evolutionary rates. However, this method is not particularly accurate since many gene phylogenies do not fit a global molecular clock when comparing different species. Newer and more sophisticated calibration methods take into account varying evolutionary rates in different parts of the phylogeny. In Paper IV, we used two methods to calibrate the phylogenetic trees; the Nonparametric Rate Smoothing (NPRS) method (SANDERSON 1997) which assumes different but auto-correlated evolutionary rates in all branches, and the recently published BEAST software (DRUMMOND and RAMBAUT 2007) that uses a relaxed molecular clock model but has no assumptions about the distribution of evolutionary rates in the phylogeny.

Two of the most comprehensive attempts to use molecular methods to date the ages of major evolutionary transition in the grasses are the studies of Bremer (2002) and Vicentini et al. (2008). The Bremer (2002) and Vicentini et al. (2008) studies used the NPRS and a Bayesian relaxed clock model, respectively, and reached very similar conclusions based on several independent fossil calibration points. The age of the stem node of the Poaceae was estimated to be 82 and 94 My, and the age of the most recent common ancestor of the BEP-PACCAD was estimated to be 50 and 52 My old, in the two studies, respectively. Furthermore, Vicentini

et al. (2008) estimated the age of the most recent common ancestor of the Pooideae sub-family, which contains the cereals and the forage grasses, to be approximately 43 My old.

One paper published in Science by Prasad et al. (2005) shed some doubts on the estimated ages of the major grass clades as provided by these molecular analyses. The authors studied what was reported to be fossilized dinosaur dung dated back to 65 My ago, and found that grass phytolith morphologies characteristic for extant Poaceae sub-families, including grasses from the BEP clade, were present at the time when dinosaurs dominated the earth (PRASAD *et al.* 2005). Using these novel findings in a molecular dating analysis assuming a strict molecular clock, the authors concluded that the age of the stem node of the Poaceae are approximately 85 My old, and that the most recent common ancestor of the BEP and PACCAD are much older than previously assumed, possibly as old as 80 My.

We can therefore divide the scientific debate on the ages of grasses into two competing hypothesis, the "Early BEP-PACCAD radiation" (EBPR) hypothesis supported by the Prasad (2005) study and the "Late BEP-PACCAD radiation" (LBPR) hypothesis supported by molecular dating using different dating methods, different DNA sequences, and various fossil calibration points (BREMER 2002; SANDVE and FJELLHEIM 2010; VINCENTINI *et al.* 2008). The competing hypotheses are presented in Figure 8. One key question is now; which version of the grass evolutionary history is the most correct one? In my opinion, the Vicentini et al. (2008) study makes a persuasive argument for the LBPR hypothesis, solely because of the amount of data pointing in the direction of the LBPR hypothesis. All fossil calibration schemes, except when using the Prasad (2005) data, produce extremely homogenous node ages estimates.

The EBPR versus LBPR debate have important implications for our conclusions in Paper IV. In this paper we assume that LBPR age estimates of the grass clades are correct and demonstrate that Pooideae specific gene duplications putatively important for evolution of freezing tolerance were clustered in the E-O cooling period. If we instead assume the EBPR hypothesis of Prasad et al. (2005), and set the rice-Pooideae divergence to ~75 My ago, then our age estimate of initial Pooideae specific duplication events are pushed back to ~62 My

ago. Most intriguingly, this estimate coincides with the even more dramatic climate change during the K-T boundary, a time where many other plant genomes are thought to have undergone adaptive increase in genome size (FAWCETT *et al.* 2009). Consequently, the interpretation of the Pooideae specific duplication events in the context of paleoclimatic change and adaptive evolution remains relevant, even if the LBPR hypothesis should turn out to be the correct model of Poaceae evolution.

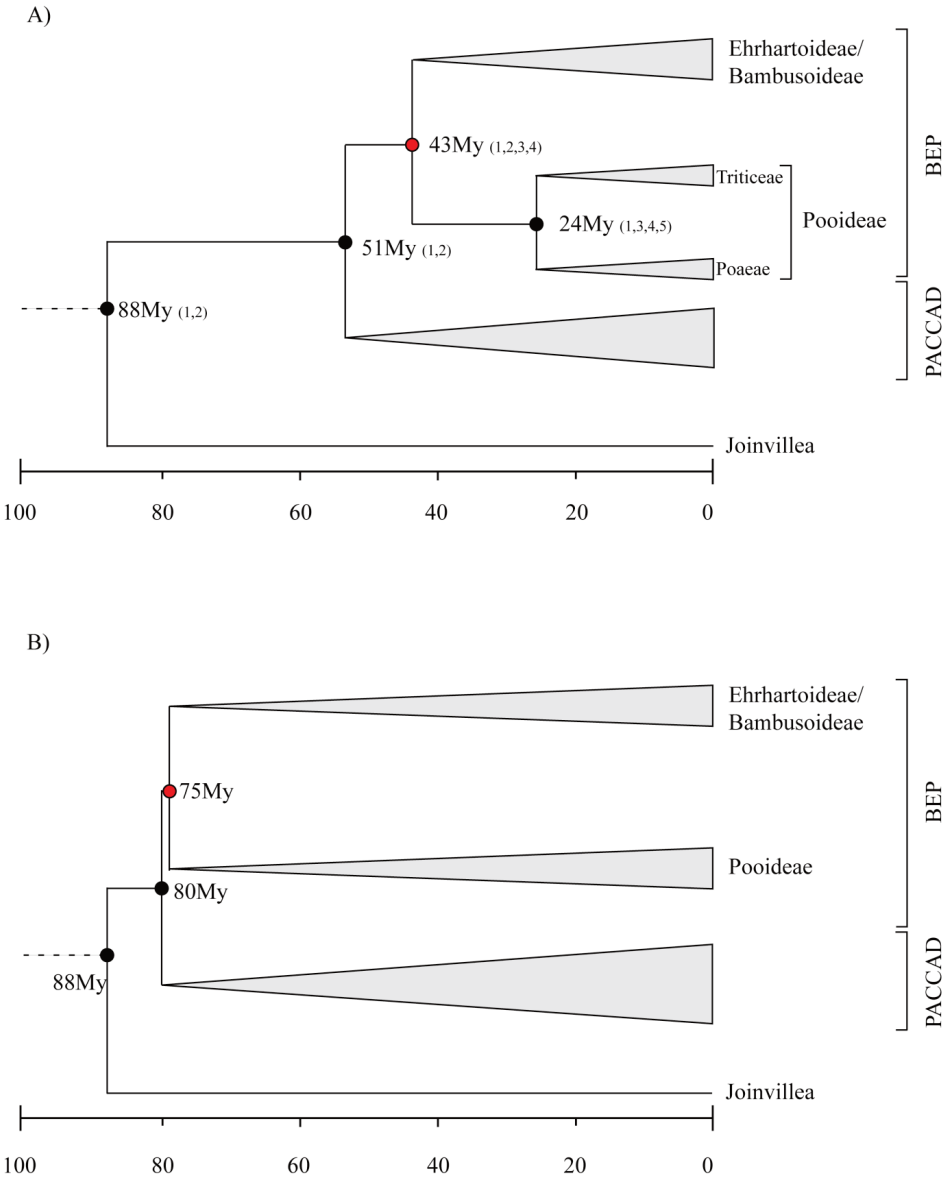


Figure 8. Competing hypotheses for the age of key nodes in the Poaceae evolutionary tree. A) The Late BEP-PACCAD radiation hypothesis with node ages supported by Vicentini et al. (2008) (1), Bremer (2002) (2), Gaut (2002) (3), Sandve and Fjellheim (2010) (4), and Sandve et al. (2008) (5). Node ages are presented as the mean of the estimates from these studies. B) Early BEP-PACCAD radiation hypothesis with node ages supported by Prasad et al. (2005). Grey triangles indicate further radiation within lineages, but do not represent a quantitative presentation of the timing, speed or extent of radiation.

Are molecular responses in laboratories relevant in nature?

Investigations into the molecular basis of CA and freezing tolerance have almost exclusively been done under controlled laboratory conditions. Such controlled conditions works to our advantage if our goal is to identify genes and molecular processes that are regulated by low temperature *per se*. However, it is seldom the case that laboratory conditions accurately reflect natural conditions. The conclusion from a recent study by Brachi et al. (2010) on genetic variation underlying variation in *A. thaliana* flowering time in nature provides a great example. In this study, the authors observed poor correlation between flowering time in the laboratory and in the field, and none of the well studied genes underlying flowering time variation in greenhouse studies were found to associate with flowering time in nature (BRACHI *et al.* 2010). This highlights the importance of taking gene-environment interactions into account when studying complex traits. If the ultimate goal is to understand the basis for freezing tolerance in nature we must therefore ask the question: what is the relevance of molecular responses under laboratory condition CA and freezing tolerance in natural conditions?

One of the biggest discrepancies between what goes on in nature compared to laboratory experiments on CA is related to the speed at which the temperature changes. In laboratory experiments plant are often moved directly from 20°C to low temperatures with constant light conditions, referred to as cold shock treatment (LEE *et al.* 2005; VOGEL *et al.* 2005). In natural conditions on the other hand, the temperature gradually decreases over the course of months with intermittent temperature fluctuations, and repeated shifts between warmer day time and cooler night time temperatures. A recent study on wheat under more natural CA conditions demonstrated that a gradual decrease in temperature produced novel and different patterns of transcriptional regulation compared to the commonly used laboratory conditions (WINFIELD *et al.* 2009). The CBF genes were for example regulated differently when the temperature changed gradually compared to if plants underwent cold shock (WINFIELD *et al.* 2010). Another discrepancy between laboratory and natural conditions is that in nature changes in day length, irradiance and light quality accompanies the decreasing temperature in natural systems. Interactions between light conditions and temperature have been demonstrated to modulate CA-associated gene expression (CROSATTI *et al.* 1999; GRAY *et al.* 1997), and are likely important in CA-responses in nature.

The debated role of ABA dependent transcriptional regulation in CA might also be linked to differences in laboratory versus natural conditions. Most experiments using cold shock do not support an important role of ABA signalling and ABA mediated transcriptional regulation early in the CA process (SHINOZAKI and YAMAGUCHI-SHINOZAKI 2000; THOMASHOW 1999). Because ABA synthesis during cold stress commences after the initial up-regulation of CBF pathways (THOMASHOW 1999), ABA is often excluded from the discussions of genetic pathways involved in CA. Although rarely mentioned in the literature, ABA has in fact been shown to increase the expression of CBF genes in *A. thaliana* (KNIGHT *et al.* 2004) and several other plant species (AGARWAL *et al.* 2006). Moreover, very little is known about the generation of ABA and the control of ABA-dependent gene expression during CA under natural conditions (GUSTA *et al.* 2005). It is therefore possible, and even likely, that ABA accumulation is more important for CA-transcriptional regulation in a natural CA-process compared to laboratory conditions.

It is certain that the widespread use of artificial laboratory conditions have biased our understanding of genes and genetic pathways involved in CA and freezing tolerance in nature. However, I must also stress that laboratory studies have been ever so central for our understanding of molecular mechanisms involved in freezing tolerance in nature, and the best example is the CBF pathway. Apart from the overwhelming numbers of experimental laboratory studies confirming the essential role of CBF genes for freezing tolerance, data from natural populations also supports the importance of CBF genes (e.g. ZHEN and UNGERER 2008).

Conclusions and future perspectives

Through the work presented in this thesis we have both confirmed previous findings and revealed novel insights into the genetics of CA and freezing tolerance of Pooideae grasses. Many of the genes identified to be involved in CA in *F. pratensis* confirmed previous findings in other plant species. Two of these genes, *FpTPT* and *FpQM,* are highly interesting candidate genes potentially underpinning phenotypic differences in freezing tolerance. The use of DArT

markers to associate genetic loci with difference in freezing tolerance identified and confirmed several chromosomal regions previously found to harbour QTLs for freezing tolerance and at the same time demonstrated the utility of the newly developed forage grass DArT resources. Patterns of molecular evolution in genes important for the CA process and acquirement of freezing tolerance suggest that lineage-specific gene duplications and evolution of a novel ice-binding protein domain in an early Pooideae ancestor could have been important for subsequent adaptation to cold climates. Our results also link the expansions of CBF, FST, and IRIP gene copy numbers to a global climate cooling during the E-O boundary ~34-26 My ago and we hypothesize that the E-O cooling was central in Pooideae evolution and adaptation to cold climates.

The novel sequences generated in our work and the rapidly increasing sequence databases of fully and partially sequenced Pooideae genomes will aid in future investigations into comparative- and functional genomics. Recently, the genomic sequence of the model Pooideae species *B. distachyon* was released (THE INTERNATIONAL BRACHYPODIUM INITIATIVE 2010), and in time an increasing numbers of T-DNA lines, transgenic resources would be available. This will enable further research and understanding of the molecular mechanisms important for Pooideae grass CA and freezing tolerance. Increased focus on importance of small RNA and epigenetic mechanism in the years to come will also likely broaden our perspectives on the complexity of the mechanisms involved in abiotic stress tolerance and climate adaptation.

Ongoing and future work will involve functional studies of the Pooideae specific IRIP genes including knock-down studies in transgenic *B. distachyon* plants, additional EST analysis of the Pooideae species CA-transcriptome using high throughput sequencing, and genome wide analysis of molecular evolutionary patterns of genes involved in CA. These (and other) future research efforts will contribute to an even better understanding of how, year after year, for millions of years, Pooideae grasses have come out victorious in the fierce battle against the northern winter climate.

Literature

ABE, H., T. URAO, T. ITO, M. SEKI, K. SHINOZAKI *et al.*, 2003 Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) Function as Transcriptional Activators in Abscisic Acid Signaling. Plant Cell **15:** 63-78.

AGARWAL, P., P. AGARWAL, M. REDDY and S. SOPORY, 2006 Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. Plant Cell Reports **25:** 1263-1274.

AGARWAL, P., and B. JHA, 2010 Transcription factors in plants and ABA dependent and independent abiotic stress signalling. Biologia Plantarum **54:** 201-212.

ALM, V., C. S. BUSSO, E. Å, H. RUDI, A. LARSEN *et al.*, submitted.

ALONSO, A., C. S. QUEIROZ and A. C. MAGALHÃES, 1997 Chilling stress leads to increased cell membrane rigidity in roots of coffee (*Coffea arabica* L.) seedlings. Biochimica et Biophysica Acta (BBA) - Biomembranes **1323:** 75-84.

AMME, S., A. MATROS, B. SCHLESIER and H. P. MOCK, 2006 Proteome analysis of cold stress response in *Arabidopsis thaliana* using DIGE-technology. Journal of Experimental Botany **57:** 1537-1546.

ANISIMOVA, M., J. P. BIELAWSKI and Z. YANG, 2001 Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. Molecular Biology and Evolution **18:** 1585-1592.

APEL, K., and H. HIRT, 2004 Reactive oxygen species: Metabolism, Oxidative Stress, and Signal Transduction. Annual Review of Plant Biology **55:** 373-399.

ARTUS, N. N., M. UEMURA, P. L. STEPONKUS, S. J. GILMOUR, C. LIN *et al.*, 1996 Constitutive expression of the cold-regulated *Arabidopsis thaliana* COR15a gene affects both chloroplast and protoplast freezing tolerance. Proceedings of the National Academy of Sciences of the United States of America **93:** 13404-13409.

BADAWI, M., J. DANYLUK, B. BOUCHO, M. HOUDE and F. SARHAN, 2007 The CBF gene family in hexaploid wheat and its relationship to the phylogenetic complexity of cereal CBFs. Molecular Genetics and Genomics **277:** 533-554.

BADAWI, M., Y. V. REDDY, Z. AGHARBAOUI, Y. TOMINAGA, J. DANYLUK *et al.*, 2008 Structure and functional analysis of wheat ICE (Inducer of CBF Expression) genes. Plant Cell and Physiology **49:** 1237-1249.

BAE, M. S., E. J. CHO, E. Y. CHOI and O. K. PARK, 2003 Analysis of the *Arabidopsis* nuclear proteome and its response to cold stress. Plant Journal **36:** 652-663.

BAJPAI, S., J. G. M. THEWISSEN and A. SAHNI, 2009 The origin and early evolution of whales: macroevolution documented on the Indian Subcontinent. Journal of Biosciences **34:** 673-686.

BECKER, B., and B. MARIN, 2009 Streptophyte algae and the origin of embryophytes. Annals of Botany **103:** 999-1004.

BENNETZEN, J. L., 2000 Transposable element contributions to plant gene and genome evolution. Plant Molecular Biology **42:** 251-269.

BHARDWAJ, P. K., AHUHA, P. S. AND KUMAR, S., 2010 Characterization of gene expression of *QM* from *Caragana jubata,* a plant species that grows under extreme cold. Molecular Biology Reports **37:** 1003-1010.

BOUCHÉ, N., A. SCHARLAT, W. SNEDDEN, D. BOUCHEZ and H. FROMM, 2002 A novel family of calmodulin-binding transcription activators in multicellular organisms. Journal of Biological Chemistry **277:** 21851-21861.

BOYKO, A., T. BLEVINS, Y. YAO, A. GOLUBOV, A. BILICHAK *et al.*, 2010 Transgenerational adaptation of *Arabidopsis* to stress requires DNA methylation and the function of Dicer-like proteins. PLoS ONE **5:** e9514.

BOYKO, A., and I. KOVALCHUK, 2008 Epigenetic control of plant stress response. Environmental and Molecular Mutagenesis **49:** 61-72.

BRACHI, B., N. FAURE, M. HORTON, E. FLAHAUW, A. VAZQUEZ *et al.*, 2010 Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. PLoS Genet **6:** e1000940.

BREMER, K., 2002 Gondwanaland evolution of the grass alliance of families (Poales) Evolution **56:** 1374-1387.

BROWSE, J., and Z. XIN, 2001 Temperature sensing and cold acclimation. Current Opinion in Plant Biology **4:** 241-246.

CATALÁN, P., P. TORRECILLA, J. Á. L. RODRIGUEZ and R. G. OLMSTEAD, 2004 Phylogeny of the festucoid grasses of subtribe Loliinae and allies (Poeae, Pooideae) inferred from ITS and trnL-F sequences. Molecular Phylogenetics and Evolution **31:** 517-541.

CATTIVELLI, L., P. BALDI, C. CROSATTI, N. DI FONZO, P. FACCIOLI *et al.*, 2002 Chromosome regions and stress-related sequences involved in resistance to abiotic stress in Triticeae. Plant Molecular Biology **48:** 649-665.

CHAKRABORTEE, S., C. BOSCHETTI, L. J. WALTON, S. SARKAR, D. C. RUBINSZTEIN *et al.*, 2007 Hydrophilic protein associated with desiccation tolerance exhibits broad protein stabilization function. Proceedings of the National Academy of Sciences **104:** 18073-18078.

CHANDRA, A., and A. DUBEY, 2010 Effect of ploidy levels on the activities of delta-1-pyrroline-5-carboxylate synthetase, superoxide dismutase and peroxidase in *Cenchrus* species grown under water stress. Plant Physiology and Biochemistry **48:** 27-34.

CHEN, C., WANDURAGALA, S., BECKER, D.F. AND DICKMAN, M.B., 2006 Tomato QM-like protein protects *Paccharomyces cerevisiae* cells against oxidative stress by regulating intracellular proline levels. Applied and Environmental Microbiology **72:** 4001–4006.

CHEN, H.-H., P. H. LI and M. L. BRENNER, 1983 Involvement of abscisic acid in potato cold acclimation. Plant Physiology **71:** 362-365.

CHINNUSAMY, V., M. OHTA, S. KANRAR, B.-H. LEE, X. HONG *et al.*, 2003 ICE1: a regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*. Genes & Development **17:** 1043-1054.

CHINNUSAMY, V., J. ZHU and J.-K. ZHU, 2007 Cold stress regulation of gene expression in plants. TRENDS in Plant Science **12:** 444-451.

CHOI, D. W., B. ZHU and T. J. CLOSE, 1999 The barley (*Hordeum vulgare* L.) dehydrin multigene family: sequences, allele types, chromosome assignments, and expression characteristics of 11 Dhn genes of cv Dicktoo. TAG Theoretical and Applied Genetics **98:** 1234-1247.

CHOI, H.-I., J.-H. HONG, J.-O. HA, J.-Y. KANG and S. Y. KIM, 2000 ABFs, a Family of ABA-responsive Element Binding Factors. Journal of Biological Chemistry **275:** 1723-1730.

COOK, D., S. FOWLER, O. FIEHN and M. F. THOMASHOW, 2004 A prominent role for the CBF cold response pathway in configuring the low-temperature metabolome of Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America **101:** 15243-15248.

COXALL, H. K., and P. N. PEARSON, 2007 The Eocene-Oligocene Transition, pp. 351-387 in *Deep-Time Perspectives on Climate Change: Marrying the Signal from Computer Models and biological Proxies*, edited by M. WILLIAMS, A. M. HAYWOOD, F. J. GREGORY and D. N. SCHMIDT. The Geological Society Bath.

CREPET, W. L., and G. D. FELDMAN, 1991 The earliest remains of grasses in the fossil record. American Journal of Botany **78:** 1010-1014.

CROSATTI, C., P. POLVERINO DE LAURETO, R. BASSI and L. CATTIVELLI, 1999 The interaction between cold and light controls the expression of the cold-regulated Barley gene cor14b and the accumulation of the corresponding protein. Plant Physiology **119:** 671-680.

CROW, K. D., and G. P. WAGNER, 2006 What is the role of genome duplication in the evolution of complexity and diversity? Molecular Biology and Evolution **23:** 887-892.

CUEVAS, J. C., R. LOPEZ-COBOLLO, R. ALCAZAR, X. ZARZA, C. KONCZ *et al.*, 2008 Putrescine is involved in *Arabidopsis* freezing tolerance and cold acclimation by regulating abscisic acid levels in response to low temperature. Plant Physiology **148:** 1094-1105.

DALL'OSTO, L., S. CAFFARRI and R. BASSI, 2005 A mechanism of nonphotochemical energy dissipation, independent from PsbS, revealed by a conformational change in the antenna protein CP26. Plant Cell **17:** 1217-1232.

DALLAIRE, S., M. HOUDE, Y. GAGNE, H. S. SAINI, S. BOILEAU *et al.*, 1994 ABA and low temperature induce freezing tolerance via distinct regulatory pathways in wheat. Plant Cell and Physiology **35:** 1-9.

DANYLUK, J., A. PERRON, M. HOUDE, A. LIMIN, B. FOWLER *et al.*, 1998 Accumulation of an acidic dehydrin in the vicinity of the plasma membrane during cold acclimation of wheat. Plant Cell **10:** 623-638.

DE WIT, E., S. E. V. LINSEN, E. CUPPEN and E. BEREZIKOV, 2009 Repertoire and evolution of miRNA genes in four divergent nematode species. Genome Research **19:** 2064-2074.

DEMEL, R. A., E. DORREPAAL, M. J. M. EBSKAMP, J. C. M. SMEEKENS and B. DE KRUIJFF, 1998 Fructans interact strongly with model membranes. Biochimica et Biophysica Acta (BBA) - Biomembranes **1375:** 36-42.

DES MARAIS, D. L., and M. D. RAUSHER, 2008 Escape from adaptive conflict after duplication in an anthocyanin pathway gene. Nature **454:** 762-765.

DHILLON, T., S. PEARCE, E. STOCKINGER, A. DISTELFELD, C. LI *et al.*, 2010 Regulation of freezing tolerance and flowering in temperate cereals: the VRN-1 connection. Plant Physiology**:** ahead of print.

DOENCH, J. G., C. P. PETERSEN and P. A. SHARP, 2003 siRNAs can function as miRNAs. Genes & Development **17:** 438-442.

DOHERTY, C. J., H. A. VAN BUSKIRK, S. J. MYERS and M. F. THOMASHOW, 2009 Roles for *Arabidopsis* CAMTA transcription factors in cold-regulated gene expression and freezing tolerance. Plant Cell **21:** 972-984.

DOMINGUEZ, T., M. L. HERNANDEZ, J. C. PENNYCOOKE, P. JIMENEZ, J. M. MARTINEZ-RIVAS *et al.*, 2010 Increasing omega-3 desaturase expression in tomato results in altered aroma profile and enhanced resistance to cold stress. Plant Physiology **153:** 655-665.

DONG, C.-H., M. AGARWAL, Y. ZHANG, Q. XIE and J.-K. ZHU, 2006a The negative regulator of plant cold responses, HOS1, is a RING E3 ligase that mediates the ubiquitination and degradation of ICE1. Proceedings of the National Academy of Sciences of the United States of America **103:** 8281-8286.

DONG, C.-H., X. HU, W. TANG, X. ZHENG, Y. S. KIM *et al.*, 2006b A putative *Arabidopsis* nucleoporin AtNUP160 is critical for RNA export and required for plant tolerance to cold stress. Mol and Cellular Biology **26:** 9533-9543.

DOWGERT, M. F., and P. L. STEPONKUS, 1984 Behavior of the plasma membrane of isolated protoplasts during a freeze-thaw cycle. Plant Physiology **75:** 1139-1151.

DOWGERT, M. F., J. WOLFE and P. L. STEPONKUS, 1987 The mechanics of injury to isolated protoplasts following osmotic contraction and expansion. Plant Physiology **83:** 1001-1007.

DRUMMOND, A., and A. RAMBAUT, 2007 BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology **7:** 214.

ELDRETT, J. S., D. R. GREENWOOD, I. C. HARDING and M. HUBER, 2009 Increased seasonality through the Eocene to Oligocene transition in northern high latitudes. Nature **459:** 969-973.

FARMER, A. A., T. M. LOFTUS, A. A. MILLS, K. Y. SATO, J. D. NEILL *et al.*, 1994 Extreme evolutionary conservation of QM, a novel c-Jun associated transcription factor. Human Molecular Genetics **3:** 723-728.

FAWCETT, J. A., S. MAERE and Y. VAN DE PEER, 2009 Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event. Proceedings of the National Academy of Sciences of the United States of America **106:** 5737-5742.

FILKOWSKI, J., A. YEOMAN, O. KOVALCHUK and I. KOVALCHUK, 2004 Systemic plant signal triggers genome instability. The Plant Journal **38:** 1-11.

FOWLER, S., and M. F. THOMASHOW, 2002 *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. Plant Cell **14:** 1675-1690.

FRICANO, A., F. RIZZA, P. FACCIOLI, D. PAGANI, P. PAVAN *et al.*, 2009 Genetic variants of HvCbf14 are statistically associated with frost tolerance in a European germplasm collection of Hordeum vulgare. TAG **119:** 1335-1348.

GALIBA, G., A. VÁGÚJFALVI, C. LI, A. SOLTÉSZ and J. DUBCOVSKY, 2009 Regulatory genes involved in the determination of frost tolerance in temperate cereals. Plant Science **176:** 12-19.

GALLOWAY, L. F., and J. R. ETTERSON, 2007 Transgenerational plasticity is adaptive in the wild. Science **318:** 1134-1136.

GAO, L., Y. GENG, B. LI, J. CHEN and J. YANG, 2010 Genome-wide DNA methylation alterations of *Alternanthera philoxeroides* in natural and manipulated habitats: implications for epigenetic regulation of rapid responses to environmental fluctuation and phenotypic variation. Plant, Cell & Environment**:** ahead of print.

GARAY-ARROYO, A., J. M. COLMENERO-FLORES, A. GARCIARRUBIO and A. A. COVARRUBIAS, 2000 Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. Journal of Biological Chemistry **275:** 5668-5674.

GAUDET, D. A., 1994 Progress towards understanding interactions between cold hardiness and snow mold resistance and development of resistant cultivars. Canadian Journal of Plant Pathology **16:** 241-246.

GONG, Z., C.-H. DONG, H. LEE, J. ZHU, L. XIONG *et al.*, 2005 A DEAD box RNA helicase is essential for mRNA export and important for development and stress responses in *Arabidopsis*. Plant Cell **17:** 256-267.

GONG, Z., H. LEE, L. XIONG, A. JAGENDORF, B. STEVENSON *et al.*, 2002 RNA helicase-like protein as an early regulator of transcription factors for plant chilling and freezing tolerance. Proceedings of the National Academy of Sciences of the United States of America **99:** 11507-11512.

GOULAS, E., M. SCHUBERT, T. KIESELBACH, L. A. KLECZKOWSKI, P. GARDESTROM *et al.*, 2006 The chloroplast lumen and stromal proteomes of Arabidopsis thaliana show differential sensitivity to short- and long-term exposure to low temperature. Plant Journal **47:** 720-734.

GRASS PHYLOGENY WORKING GROUP, 2001 Phylogeny and subfamilial classification of the grasses (Poaceae). Annals of the Missouri Botanical Garden **88:** 373-457.

GRAY, G. R., L. P. CHAUVIN, F. SARHAN and N. P. A. HUNER, 1997 Cold acclimation and freezing tolerance (a complex interaction of light and temperature). Plant Physiology **114:** 467-474.

GRIFFITH, M., P. ALA, D. S. C. YANG, W.-C. HON and B. A. MOFFATT, 1992 Antifreeze protein produced endogenously in winter rye leaves. Plant Physiology **100:** 593-596.

GRIFFITH, M., C. LUMB, S. B. WISEMAN, M. WISNIEWSKI, R. W. JOHNSON *et al.*, 2005 Antifreeze proteins modify the freezing process in planta. Plant Physiology **138:** 330-340.

GRIFFITH, M., and M. W. F. YAISH, 2004 Antifreeze proteins in overwintering plants: a tale of two activities. TRENDS in Plant Science **9:** 399-405.

GUSTA, L., R. TRISCHUK and C. J. WEISER, 2005 Plant cold acclimation: The role of abscisic acid. Journal of Plant Growth Regulation **24:** 308-318.

HAHN, M. W., 2009 Distinguishing among evolutionary models for the maintenance of gene duplicates. Journal of Heredity **100:** 605-617.

HANADA, K., C. ZOU, M. D. LEHTI-SHIU, K. SHINOZAKI and S.-H. SHIU, 2008 Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. Plant Physiology **148:** 993-1003.

HARTLEY, H., 1973 Studies on the origin, evolution, and distribution of the Gramineae. V. The subfamily Festucoideae Australian Journal of Botany **21:** 201-234.

HE, X., and J. ZHANG, 2005 Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. Genetics **169:** 1157-1164.

HEINEKE, D., A. KRUSE, U.-I. FLÜGGE, W. B. FROMMER, J. W. RIESMEIER *et al.*, 1994 Effect of antisense repression of the chloroplast triose-phosphate translocator on photosynthetic metabolism in transgenic potato plants. Planta **193:** 174-180.

HENDRY, G. A. F., 1993 Evolutionary origins and natural functions of fructans - a climatological, biogeographic and mechanistic appraisal. New Phytologist **123:** 3-14.

HERENDEEN, P. S., and P. R. CRANE, 1995 The fossil history of the monocotyledons, pp. 1–21 in *Monocotyledons: Systematics and Evolution*, edited by P. J. RUDALL, P. CRIBB, D. F. CUTLER and C. J. HUMPHRIES.

HERRERA, C. M., and P. BAZAGA, 2010 Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet *Viola cazorlensis*. New Phytologist **187:** 867-876.

HEY, S. J., E. BYRNE and N. G. HALFORD, 2010 The interface between metabolic and stress signalling. Annals of Botany **105:** 197-203.

HIGGINS, J. A., P. C. BAILEY and D. A. LAURIE, 2010 Comparative genomics of flowering time pathways using *Brachypodium distachyon* as a model for the temperate grasses. PLoS ONE **5:** e10065.

HINCHA, D. K., E. M. HELLWEGE, A. G. HEYER and J. H. CROWE, 2000 Plant fructans stabilize phosphatidylcholine liposomes during freeze-drying. European Journal of Biochemistry **267:** 535-540.

HISANO, H., A. KANAZAWA, A. KAWAKAMI, M. YOSHIDA, Y. SHIMAMOTO *et al.*, 2004 Transgenic perennial ryegrass plants expressing wheat fructosyltransferase genes accumulate increased amounts of fructan and acquire increased tolerance on a cellular level to freezing. Plant Science **167:** 861-868.

HISANO, H., A. KANAZAWA, M. YOSHIDA, M. O. HUMPHREYS, M. IIZUKA *et al.*, 2008 Coordinated expression of functionally diverse fructosyltransferase genes is associated with fructan accumulation in response to low temperature in perennial ryegrass. New Phytologist **178:** 766-780.

HITTINGER, C. T., and S. B. CARROLL, 2007 Gene duplication and the adaptive evolution of a classic genetic switch. Nature **449:** 677-681.

HOTH, S., M. MORGANTE, J.-P. SANCHEZ, M. K. HANAFEY, S. V. TINGEY *et al.*, 2002 Genome-wide gene expression profiling in *Arabidopsis thaliana* reveals new targets of abscisic acid and largely impaired gene regulation in the abi1-1 mutant. Journal of Cell Science **115:** 4891-4900.

HUMPHREYS, M., D. GASIOR, A. LESNIEWSKA-BOCIANOWSKA, Z. ZWIERZYKOWSKI and M. RAPACZ, 2007 Androgenesis as a means of dissecting complex genetic and physiological controls: selecting useful gene combinations for breeding freezing tolerant grasses. Euphytica **158:** 337-345.

HUNDERTMARK, M., and D. HINCHA, 2008 LEA (Late Embryogenesis Abundant) proteins and their encoding genes in *Arabidopsis thaliana*. BMC Genomics **9:** 118.

HUNER, N., G. ÖQUIST, V. HURRY, M. KROL, S. FALK *et al.*, 1993 Photosynthesis, photoinhibition and low temperature acclimation in cold tolerant plants. Photosynthesis Research **37:** 19-39.

HÄUSLER, R. E., N. H. SCHLIEBEN and U.-I. FLÜGGE, 2000 Control of carbon partitioning and photosynthesis by the triose phosphate/phosphate translocator in transgenic tobacco plants (*Nicotiana tabacum*). II. Assessment of control coefficients of the triose phosphate/phosphate translocator. Planta **210:** 383-390.

JABLONKA, E., and M. J. LAMB, 2005 *Evolution in Four Dimentions - Genetic, Epigenetic, Behavioural, and Symbolic Variation in the History of Life*. The MIT Press.

JACOBS, B. F., J. D. KINGSTON and L. L. JACOBS, 1999 The origin of grass-dominated ecosystems. Annals of the Missouri Botanical Garden **86:** 590-643.

JAGLO-OTTOSEN, K. R., S. J. GILMOUR, D. G. ZARKA, O. SCHABENBERGER and M. F. THOMASHOW, 1998 *Arabidopsis* CBF1 overexpression induces COR genes and enhances freezing tolerance. Science **280:** 104-106.

JI, X. M., W. VAN DEN ENDE, L. SCHROEVEN, S. CLERENS, K. GEUTEN *et al.*, 2007 The rice genome encodes two vacuolar invertases with fructan exohydrolase activity but lacks the related fructan biosynthesis genes of the Pooideae. New Phytologist **173:** 50-62.

JIAN, X., L. ZHANG, G. LI, L. ZHANG, X. WANG *et al.*, 2010 Identification of novel stress-regulated microRNAs from *Oryza sativa* L. Genomics **95:** 47-55.

JOHANSON, U., J. WEST, C. LISTER, S. MICHAELS, R. AMASINO *et al.*, 2000 Molecular analysis of FRIGIDA, a major determinant of natural variation in *Arabidopsis* flowering time. Science **290:** 344-347.

JOHANSSON, J., P. MANDIN, A. RENZONI, C. CHIARUTTINI, M. SPRINGER *et al.*, 2002 An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. Cell **110:** 551-561.

JOHN, U. P., R. M. POLOTNIANKA, K. A. SIVAKUMARAN, O. CHEW, L. MACKIN *et al.*, 2009 Ice recrystallization inhibition proteins (IRIPs) and freeze tolerance in the cryophilic Antarctic hair grass *Deschampsia antarctica* E. Desv. Plant, Cell & Environment **32:** 336-348.

JOHNSEN, Ø., F. C. G., N. NAGY, J. MØLMANN, O. G. DÆHLEN *et al.*, 2005 Climatic adaptation in *Picea abies* progenies is affected by the temperature during zygotic embryogenesis and seed maturation. Plant, Cell & Environment **28:** 1090-1102.

JOHNSON, E. S., 2004 Protein modification by SUMO. Annual Review of Biochemistry **73:** 355-382.

KALENDAR, R., J. TANSKANEN, S. IMMONEN, E. NEVO and A. H. SCHULMAN, 2000 Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. Proceedings of the National Academy of Sciences of the United States of America **97:** 6603-6607.

KARLSON, D., K. NAKAMINAMI, T. TOYOMASU and R. IMAI, 2002 A cold-regulated nucleic acid-binding protein of winter wheat shares a domain with bacterial cold shock proteins. Journal of Biological Chemistry **277:** 35248-35256.

KASHI, Y., and D. G. KING, 2006 Simple sequence repeats as advantageous mutators in evolution. TRENDS in Genetics **22:** 253-259.

KASUGA, M., Q. LIU, S. MIURA, K. YAMAGUCHI-SHINOZAKI and K. SHINOZAKI, 1999 Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Nat Biotech **17:** 287-291.

KAWAKAMI, A., Y. SATO and M. YOSHIDA, 2008 Genetic engineering of rice capable of synthesizing fructans and enhancing chilling tolerance. Journal of Experimental Botany **59:** 793-802.

KEREPESI, I., E. BANYAI-STEFANOVITS and G. GALIBA, 2004 Cold acclimation and abscisic acid induced alterations in carbohydrate content in calli of wheat genotypes differing in frost tolerance. Journal of Plant Physiology **161:** 131-133.

KHODAKOVSKAYA, M., R. MCAVOY, J. PETERS, H. WU and Y. LI, 2006 Enhanced cold tolerance in transgenic tobacco expressing a chloroplast ω-3 fatty acid desaturase gene under the control of a cold-inducible promoter. Planta **223:** 1090-1100.

KNIGHT, H., and M. R. KNIGHT, 2001 Abiotic stress signalling pathways: specificity and cross-talk. TRENDS in Plant Science **6:** 262-267.

KNIGHT, H., S. G. MUGFORD, B. ÜLKER, D. GAO, G. THORLBY *et al.*, 2009 Identification of SFR6, a key component in cold acclimation acting post-translationally on CBF function. The Plant Journal **58:** 97-108.

KNIGHT, H., E. L. VEALE, G. J. WARREN and M. R. KNIGHT, 1999 The sfr 6 mutation in *Arabidopsis* suppresses low-temperature induction of genes dependent on the CRT/DRE sequence motif. Plant Cell **11:** 875-886.

KNIGHT, H., D. G. ZARKA, H. OKAMOTO, M. F. THOMASHOW and M. R. KNIGHT, 2004 Abscisic acid induces CBF gene transcription and subsequent induction of cold-regulated genes via the CRT promoter element. Plant Physiology **135:** 1710-1717.

KNOX, A., T. DHILLON, H. CHENG, A. TONDELLI, N. PECCHIONI *et al.*, 2010 CBF gene copy number variation at Frost Resistance-2 is associated with levels of freezing tolerance in temperate-climate cereals. Theoretical and Applied Genetics **121:** 21-35.

KOBAYASHI, F., E. MAETA, A. TERASHIMA, K. KAWAURA, Y. OGIHARA *et al.*, 2008 Development of abiotic stress tolerance via bZIP-type transcription factor LIP19 in common wheat. Journal of Experimental Botany **59:** 891-905.

KOSMALA, A., A. BOCIAN, M. RAPACZ, B. JURCZYK and Z. ZWIERZYKOWSKI, 2009 Identification of leaf proteins differentially accumulated during cold acclimation between Festuca pratensis plants with distinct levels of frost tolerance. Journal of Experimental Botany **60:** 3595-3609.

KOSMALA, A., Z. ZWIERZYKOWSKI, D. GASIOR, M. RAPACZ, E. ZWIERZYKOWSKA *et al.*, 2006 GISH/FISH mapping of genes for freezing tolerance transferred from *Festuca pratensis* to *Lolium multiflorum*. Heredity **96:** 243-251.

KOSOVÁ, K., I. TOM PRÁSIL, P. PRÁSILOVÁ, P. VÍTÁMVÁS and J. CHRPOVÁ, 2010 The development of frost tolerance and DHN5 protein accumulation in barley (*Hordeum vulgare*) doubled haploid lines derived from Atlas 68×Igri cross during cold acclimation. Journal of Plant Physiology **167:** 343-350.

KRAUSE, G. H., 1988 Photoinhibition of photosynthesis. An evaluation of damaging and protective mechanisms. Physiologia Plantarum **74:** 566-574.

KROL, J., I. LOEDIGE and W. FILIPOWICZ, 2010 The widespread regulation of microRNA biogenesis, function and decay. Nature Review Genetics **11:** 597-610.

KUMBLE, K. D., J. DEMMER, S. FISH, C. HALL, S. CORRALES *et al.*, 2008 Characterization of a family of ice-active proteins from the Ryegrass, *Lolium perenne*. Cryobiology **57:** 263-268.

LANG-MLADEK, C., O. POPOVA, K. KIOK, M. BERLINGER, B. RAKIC *et al.*, 2010 Transgenerational inheritance and resetting of stress-induced loss of epigenetic gene silencing in *Arabidopsis*. Molecular Plant **3:** 594-602.

LEE, B.-H., D. A. HENDERSON and J.-K. ZHU, 2005 The *Arabidopsis* cold-responsive transcriptome and its regulation by ICE1. Plant Cell **17:** 3155-3175.

LEE, B.-H., H. LEE, L. XIONG and J.-K. ZHU, 2002 A mitochondrial complex I defect impairs cold-regulated nuclear gene expression. Plant Cell **14:** 1235-1251.

LESPINET, O., Y. I. WOLF, E. V. KOONIN and L. ARAVIND, 2002 The role of lineage-specific gene family expansion in the evolution of eukaryotes. Genome Research **12:** 1048-1059.

LI, H.-J., A.-F. YANG, X.-C. ZHANG, F. GAO and J.-R. ZHANG, 2007 Improving freezing tolerance of transgenic tobacco expressing sucrose: sucrose 1-fructosyltransferase gene from *Lactuca sativa*. Plant Cell, Tissue and Organ Culture **89:** 37-48.

LI, Y.-C., A. B. KOROL, T. FAHIMA and E. NEVO, 2004 Microsatellites within genes: Structure, function, and evolution. Molecular Biology and Evolution **21:** 991-1007.

LILLICO, S. G., J. C. MOTTRAM, N. B. MURPHY and S. C. WELBURN, 2002 Characterisation of the *QM* gene of *Trypanosoma brucei*. FEMS Microbiology Letters **211:** 123-128.

LIRA-MEDEIROS, C. F., C. PARISOD, R. A. FERNANDES, C. S. MATA, M. A. CARDOSO *et al.*, 2010 Epigenetic variation in mangrove plants occurring in contrasting natural environment. PLoS ONE **5:** e10326.

LIU, N., N.-Q. ZHONG, G.-L. WANG, L.-J. LI, X.-L. LIU *et al.*, 2007 Cloning and functional characterization of PpDBF1 gene encoding a DRE-binding transcription factor from *Physcomitrella patens*. Planta **226:** 827-838.

LIU, Q., M. KASUGA, Y. SAKUMA, H. ABE, S. MIURA *et al.*, 1998 Two Transcription Factors, DREB1 and DREB2, with an EREBP/AP2 DNA Binding Domain Separate Two Cellular Signal Transduction Pathways in Drought- and Low-Temperature-Responsive Gene Expression, Respectively, in Arabidopsis. Plant Cell **10:** 1391-1406.

LIU, Z., M. PAGANI, D. ZINNIKER, R. DECONTO, M. HUBER *et al.*, 2009 Global Cooling During the Eocene-Oligocene Climate Transition. Science **323:** 1187-1190.

LIVINGSTON, D., D. HINCHA and A. HEYER, 2009 Fructan and its relationship to abiotic stress tolerance in plants. Cellular and Molecular Life Sciences **66:** 2007-2023.

LIVINGSTON, D. P., III, and C. A. HENSON, 1998 Apoplastic sugars, fructans, fructan exohydrolase, and invertase in winter oat: Responses to second-phase cold hardening. Plant Physiology **116:** 403-408.

LLORENTE, F., J. C. OLIVEROS, J. M. MARTÍNEZ-ZAPATER and J. SALINAS, 2000 A freezing-sensitive mutant of *Arabidopsis*, frs1, is a new aba3 allele. Planta **211:** 648-655.

LOFTUS, T. M., Y. H. NGUYEN and E. J. STANBRIDGE, 1997 The QM protein associates with ribosomes in the rough endoplasmic reticulum. Biochemistry **36:** 8224-8230.

LUKENS, L. N., and S. ZHAN, 2007 The plant genome's methylation status and response to stress: implications for plant improvement. Current Opinion in Plant Biology **10:** 317-322.

LYNCH, M., and J. S. CONERY, 2000 The evolutionary fate and consequences of duplicate genes. Science **290:** 1151-1155.

LYNCH, M., and V. KATJU, 2004 The altered evolutionary trajectories of gene duplicates. TRENDS in Genetics **20:** 544-549.

MANTYLA, E., V. LANG and E. T. PALVA, 1995 Role of abscisic acid in drought-induced freezing tolerance, cold acclimation, and accumulation of LT178 and RAB18 proteins in *Arabidopsis thaliana*. Plant Physiology **107:** 141-148.

MARGUERON, R., and D. REINBERG, 2010 Chromatin structure and the inheritance of epigenetic information. Nat Rev Genet **11:** 285-296.

MARUYAMA, K., M. TAKEDA, S. KIDOKORO, K. YAMADA, Y. SAKUMA *et al.*, 2009 Metabolic pathways involved in cold acclimation identified by integrated analysis of metabolites and transcripts regulated by DREB1A and DREB2A. Plant Physiology **150:** 1972-1980.

MIDDLETON, A. J., A. M. BROWN, P. L. DAVIES and V. K. WALKER, 2009 Identification of the ice-binding face of a plant antifreeze protein. Febs Letters **583:** 815-819.

MIKKELSEN, M. D., and M. F. THOMASHOW, 2009 A role for circadian evening elements in cold-regulated gene expression in Arabidopsis. The Plant Journal **60:** 328-339.

MILLER, A., G. GALIBA and J. DUBCOVSKY, 2006 A cluster of 11 CBF transcription factors is located at the frost tolerance locus Fr-A m 2 in *Triticum monococcum*. Molecular Genetics and Genomics **275:** 193-203.

MIURA, K., J. B. JIN, J. LEE, C. Y. YOO, V. STIRM *et al.*, 2007 SIZ1-Mediated Sumoylation of ICE1 Controls CBF3/DREB1A Expression and Freezing Tolerance in *Arabidopsis*. Plant Cell **19:** 1403-1414.

MOLNAR, A., C. W. MELNYK, A. BASSETT, T. J. HARDCASTLE, R. DUNN *et al.*, 2010 Small silencing RNAs in plants are mobile and direct epigenetic modification in recipient cells. Science **328:** 872-875.

MONTECLARO, F. S., and P. K. VOGT, 1993 A Jun-binding protein related to a putative tumor suppressor. Proceedings of the National Academy of Sciences of the United States of America **90:** 6726-6730.

MOONEY, B. P., J. A. MIERNYK, C. M. GREENLIEF and J. J. THELEN, 2006 Using quantitative proteomics of Arabidopsis roots and leaves to predict metabolic activity. Physiologia Plantarum **128:** 237-250.

MOUILLON, J.-M., P. GUSTAFSSON and P. HARRYSON, 2006 Structural investigation of disordered stress proteins. Comparison of full-length dehydrins with isolated peptides of their conserved segments. Plant Physiology **141:** 638-650.

MULLER, P., X.-P. LI and K. K. NIYOGI, 2001 Non-Photochemical Quenching. A Response to Excess Light Energy. Plant Physiol. **125:** 1558-1566.

MURATA, N., S. TAKAHASHI, Y. NISHIYAMA and S. I. ALLAKHVERDIEV, 2007 Photoinhibition of photosystem II under environmental stress. Biochimica et Biophysica Acta (BBA) - Bioenergetics **1767:** 414-421.

NAKAMINAMI, K., D. T. KARLSON and R. IMAI, 2006 Functional conservation of cold shock domains in bacteria and higher plants. Proceedings of the National Academy of Sciences **103:** 10122-10127.

NISHIZAWA, A., Y. YABUTA and S. SHIGEOKA, 2008 Galactinol and raffinose constitute a novel function to protect plants from oxidative damage. Plant Physiology **147:** 1251-1263.

NOVILLO, F., J. M. ALONSO, J. R. ECKER and J. SALINAS, 2004 CBF2/DREB1C is a negative regulator of CBF1/DREB1B and CBF3/DREB1A expression and plays a central role in stress tolerance in *Arabidopsis*. Proceedings of the National Academy of Sciences of the United States of America **101:** 3985-3990.

NOVILLO, F., J. MEDINA and J. SALINAS, 2007 *Arabidopsis* CBF1 and CBF3 have a different function than CBF2 in cold acclimation and define different gene classes in the CBF regulon. Proceedings of the National Academy of Sciences of the United States of America **104:** 21002-21007.

OH, H. S., KWON, H., SUN, H.K. AND YANG, C.H., 2002 QM, a putative tumor suppressor, regulates proto-oncogene c-Yes. Journal Biological Chemistry **277:** 36489-36498.

OLIVER, K., R. , and W. K. GREENE, 2009 Transposable elements: powerful facilitators of evolution. Bioessays **31:** 703-714.

PARK, S. J., K. J. KWAK, H. J. JUNG, H. J. LEE and H. KANG, 2010 The C-terminal zinc finger domain of *Arabidopsis* cold shock domain proteins is important for RNA chaperone activity during cold adaptation. Phytochemistry **71:** 543-547.

PARVANOVA, D., A. POPOVA, I. ZAHARIEVA, P. LAMBREV, T. KONSTANTINOVA *et al.*, 2004 Low temperature tolerance of tobacco plants transformed to accumulate proline, fructans, or glycine betaine. Variable chlorophyll fluorescence evidence. Photosynthetica **42:** 179-185.

PEARCE, R. S., 2001 Plant freezing and damage. Annals of Botany **87:** 417-424.

PENNYCOOKE, J., H. CHENG, S. ROBERTS, Q. YANG, S. RHEE *et al.*, 2008 The low temperature-responsive, *Solanum* CBF1genes maintain high identity in their upstream regions in a genomic environment undergoing gene duplications, deletions, and rearrangements. Plant Molecular Biology **67:** 483-497.

PENNYCOOKE, J. C., M. L. JONES and C. STUSHNOFF, 2003 Down-regulating alpha-galactosidase enhances freezing tolerance in transgenic petunia. Plant Physiology **133:** 901-909.

POLLOCK, C. J., and T. JONES, 1979 Seasonal Patterns of Fructan Metabolism in Forage Grasses. New Phytologist **83:** 9-15.

POWLES, S. B., 1984 Photoinhibition of Photosynthesis Induced by Visible Light. Annual Review of Plant Physiology **35:** 15-44.

PRASAD, V., C. A. E. STROMBERG, H. ALIMOHAMMADIAN and A. SAHNI, 2005 Dinosaur coprolites and the early evolution of grasses and grazers. Science **310:** 1177-1180.

PRESTON, J. C., and E. A. KELLOGG, 2007 Conservation and divergence of *APETALA1/FRUITFULL*-like gene function in grasses: evidence from gene expression analyses. The Plant Journal **52:** 69-81.

PRITCHARD, J. K., J. K. PICKRELL and G. COOP, 2010 The genetics of human adaptation: Hard sweeps, soft sweeps, and polygenic adaptation. Current Biology **20:** R208-R215.

PUEBLA, A. F., G. L. SALERNO and H. G. PONTIS, 1997 Fructan metabolism in two species of *Bromus* subjected to chilling and water stress. New Phytologist **136:** 123-129.

PUHAKAINEN, T., M. W. HESS, P. MÄKELÄ, J. SVENSSON, P. HEINO *et al.*, 2004 Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in *Arabidopsis*. Plant Molecular Biology **54:** 743-753.

RANA, T. M., 2007 Illuminating the silence: understanding the structure and function of small RNAs. Nat Rev Mol Cell Biol **8:** 23-36.

RAPACZ, M., D. GASIOR, Z. ZWIERZYKOWSKI, A. LESNIEWSKA-BOCIANOWSKA, M. W. HUMPHREYS *et al.*, 2004 Changes in cold tolerance and the mechanisms of acclimation of photosystem II to cold hardening generated by anther culture of *Festuca pratensis* x *Lolium multiflorum* cultivars. New Phytologist **162:** 105-114.

RÉMUS-BOREL, W., Y. CASTONGUAY, J. CLOUTIER, R. MICHAUD, A. BERTRAND *et al.*, 2010 Dehydrin variants associated with superior freezing tolerance in alfalfa (*Medicago sativa* L.). Theoretical and Applied Genetics **120:** 1163-1174.

RENVOIZE, S. A., and W. D. CLAYTON, 1992 Classification and evolution of the grasses, pp. 3-37 in *Grass Evolution and Domestication*, edited by G. P. CHAPMAN. Cambridge University Press, Cambridge.

REYES, J. L., F. CAMPOS, H. WEI, R. ARORA, Y. YANG *et al.*, 2008 Functional dissection of Hydrophilins during *in vitro* freeze protection. Plant, Cell & Environment **31:** 1781-1790.

RIZZON, C., L. PONGER and B. S. GAUT, 2006 Striking similarities in the genomic distribution of tandemly arrayed genes in *Arabidopsis* and Rice. PLoS Comput Biol **2:** e115.

RORAT, T., 2006 Plant dehydrins — Tissue location, structure and function. Cellular & Molecular Biology Letters **11:** 536-556.

ROWAN, D. H. B., and D. SCHLUTER, 2008 Adaptation from standing genetic variation. TRENDS in Ecology and Evolution **23:** 38 - 44.

RUDI, H., S. R. SANDVE, L. OPSETH and O. A. ROGNLI, 2010 Identification of candidate genes important for frost tolerance in *Festuca pratensis* Huds. by transcriptional profiling. Plant Science**:** In press.

SAKUMA, Y., Q. LIU, J. G. DUBOUZET, H. ABE, K. SHINOZAKI *et al.*, 2002 DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. Biochemical and Biophysical Research Communications **290:** 998-1009.

SANDERSON, M. J., 1997 A nonparametric approach to estimating divergence times in the absence of rate constancy. Molecular Biology and Evolution **14:** 1218-1231.

SANDERSON, M. J., J. L. THORNE, N. WIKSTROM and K. BREMER, 2004 Molecular evidence on plant divergence times. Am. J. Bot. **91:** 1656-1665.

SANDVE, S., H. RUDI, T. ASP and O. ROGNLI, 2008 Tracking the evolution of a cold stress associated gene family in cold tolerant grasses. BMC Evolutionary Biology **8:** 245.

SANDVE, S. R., and S. FJELLHEIM, 2010 Did gene family expansions during the Eocene-Oligocene boundary climate cooling play a role in Pooideae adaptation to cool climates? Molecular Ecology **19:** 2075-2088.

SANGWAN, V., I. FOULDS, J. SINGH and R. S. DHINDSA, 2001 Cold-activation of *Brassica napus* BN115 promoter is mediated by structural changes in membranes and cytoskeleton, and requires Ca^{2+} influx. The Plant Journal **27:** 1-12.

SAUVAGE, C., N. DEROME, E. NORMANDEAU, J. ST.-CYR, C. AUDET *et al.*, 2010 Fast transcriptional responses to domestication in the brook charr salvelinus fontinalis. Genetics **185:** 105-112.

SAVOURÉ, A., X. J. HUA, N. BERTAUCHE, M. V. MONTAGU and N. VERBRUGGEN, 1997 Abscisic acid-independent and abscisic acid-dependent regulation of proline biosynthesis following cold and osmotic stresses in *Arabidopsis thaliana*. Molecular and General Genetics **254:** 104-109.

SCHROEVEN, L., W. LAMMENS, A. VAN LAERE and W. VAN DEN ENDE, 2008 Transforming wheat vacuolar invertase into a high affinity sucrose:sucrose 1-fructosyltransferase. New Phytologist **180:** 822-831.

SEKI, M., M. NARUSAKA, J. ISHIDA, T. NANJO, M. FUJITA *et al.*, 2002 Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. The Plant Journal **31:** 279-292.

SENDERSKA, E., J. MIGAS, A. WOŹNICZKA, W. KACZMAREK, A. KOZIEŁ *et al.*, 2005 Photosynthetic acclimation to cold in mountain Festuca ecotypes from Babia Góra (in Polish). in *Materiały Wydziałowej Sesji Koła Naukowego Rolników*, edited by M. KOPEĆ. UR, Wydział Rolniczo-Ekonomiczny, Kraków.

SHAKIROVA, F., C. ALLAGULOVA, M. BEZRUKOVA, A. AVAL'BAEV and F. GIMALOV, 2009 The role of endogenous ABA in cold-induced expression of the TADHN dehydrin gene in wheat seedlings. Russian Journal of Plant Physiology **56:** 720-723.

SHINOZAKI, K., and K. YAMAGUCHI-SHINOZAKI, 2000 Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. Current Opinion in Plant Biology **3:** 217-223.

SIDEBOTTOM, C., S. BUCKLEY, P. PUDNEY, S. TWIGG, C. JARMAN *et al.*, 2000 Phytochemistry: Heat-stable antifreeze protein from grass. Nature **406:** 256-256.

SKINNER, J., J. ZITZEWITZ, P. SZŰCS, L. MARQUEZ-CEDILLO, T. FILICHKIN *et al.*, 2005 Structural, functional, and phylogenetic characterization of a large CBF gene family in barley. Plant Molecular Biology **59:** 533-551.

SKØT, L., N. R. S. HAMILTON, S. MIZEN, K. H. CHORLTON and I. D. THOMAS, 2002 Molecular genecology of temperature response in *Lolium perenne*: 2. Association of AFLP markers with ecogeography. Molecular Ecology **11:** 1865-1876.

SMALLWOOD, M., and D. J. BOWLES, 2002 Plants in a cold climate. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **357:** 831-847.

SNEDDEN, W. A., and H. FROMM, 2001 Calmodulin as a versatile calcium signal transducer in plants. New Phytologist **151:** 35-66.

STEEMANS, P., A. L. HERISSE, J. MELVIN, M. A. MILLER, F. PARIS *et al.*, 2009 Origin and radiation of the earliest vascular land plants. Science **324:** 353-.

STEPONKUS, P. L., M. UEMURA, R. A. JOSEPH, S. J. GILMOUR and M. F. THOMASHOW, 1998 Mode of action of the COR15a gene on the freezing tolerance of *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America **95:** 14570-14575.

STRAND, Å., C. H. FOYER, P. GUSTAFFSON, P. GARDESTROM and V. HURRY, 2003 Altering flux through the sucrose biosynthesis pathway in transgenic *Arabidopsis thaliana* modifies photosynthetic acclimation at low temperatures and the development of freezing tolerance. Plant, Cell & Environment **26:** 523-535.

STRANGER, B. E., M. S. FORREST, M. DUNNING, C. E. INGLE, C. BEAZLEY *et al.*, 2007 Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science **315:** 848-853.

STROMBERG, C. A. E., 2005 Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the Cenozoic of North America. Proceedings of the National Academy of Sciences of the United States of America **102:** 11980-11984.

SUNKAR, R., V. CHINNUSAMY, J. ZHU and J.-K. ZHU, 2007 Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. TRENDS in Plant Science **12:** 301-309.

SUZUKI, I., D. A. LOS, Y. KANESAKI, K. MIKAMI and N. MURATA, 2000 The pathway for perception and transduction of low-temperature signals in *Synechocystis*. EMBO Journal **19:** 1327-1334.

SVENSSON, J. T., C. CROSATTI, C. CAMPOLI, R. BASSI, A. M. STANCA *et al.*, 2006 Transcriptome analysis of cold acclimation in barley albina and xantha mutants. Plant Physiology **141:** 257-270.

SZABADOS, L., and A. SAVOURÉ, 2010 Proline: a multifunctional amino acid. TRENDS in Plant Science **15:** 89-97.

SØRENSEN, J. G., M. M. NIELSEN and V. LOESCHCKE, 2007 Gene expression profile analysis of *Drosophila melanogaster* selected for resistance to environmental stressors. Journal of Evolutionary Biology **20:** 1624-1636.

TAJI, T., C. OHSUMI, S. IUCHI, M. SEKI, M. KASUGA *et al.*, 2002 Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in *Arabidopsis thaliana*. The Plant Journal **29:** 417-426.

TAKAHASHI, S., and N. MURATA, 2008 How do environmental stresses accelerate photoinhibition? TRENDS in Plant Science **13:** 178-182.

TEIXEIRA, F. K., and V. COLOT, 2009 Gene body DNA methylation in plants: a means to an end or an end to a means. EMBO Journal **28:** 997-998.

THE INTERNATIONAL BRACHYPODIUM INITIATIVE, 2010 Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature **463:** 763-768.

THOMASHOW, M. F., 1999 Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. Annual Review of Plant Physiology and Plant Molecular Biology **50:** 571-599.

TISHKOFF, S. A., F. A. REED, A. RANCIARO, B. F. VOIGHT, C. C. BABBITT *et al.*, 2007 Convergent adaptation of human lactase persistence in Africa and Europe. Nature Genetics **39:** 31-40.

TOMPA, P., and D. KOVACS, 2010 Intrinsically disordered chaperones in plants and animals. Biochemistry and Cell Biology **88:** 167-174.

TREMBLAY, K., F. OUELLET, J. FOURNIER, J. DANYLUK and F. SARHAN, 2005 Molecular characterization and origin of novel bipartite cold-regulated ice recrystallization inhibition proteins from cereals. Plant Cell and Physiology **46:** 884-891.

TREVASKIS, B., M. N. HEMMING, E. S. DENNIS and W. J. PEACOCK, 2007 The molecular basis of vernalization-induced flowering in cereals. TRENDS in Plant Science **12:** 352-357.

TUNNACLIFFE, A., and M. WISE, 2007 The continuing conundrum of the LEA proteins. Naturwissenschaften **94:** 791-812.

UEMURA, M., R. A. JOSEPH and P. L. STEPONKUS, 1995 Cold acclimation of *Arabidopsis thaliana* (Effect on plasma membrane lipid composition and freeze-induced lesions). Plant Physiology **109:** 15-30.

UEMURA, M., and P. L. STEPONKUS, 1999 Cold acclimation in plants: Relationship between the lipid composition and the cryostability of the plasma membrane. Journal of Plant Research **112:** 245-254.

VAZQUEZ, F., T. BLEVINS, J. AILHAS, T. BOLLER and F. MEINS, JR, 2008 Evolution of *Arabidopsis* MIR genes generates novel microRNA classes. Nucleic Acids Research **36:** 6429-6438.

VEREYKEN, I. J., V. CHUPIN, R. A. DEMEL, S. C. M. SMEEKENS and B. DE KRUIJFF, 2001 Fructans insert between the headgroups of phospholipids. Biochimica et Biophysica Acta (BBA) - Biomembranes **1510:** 307-320.

VIJN, I., and S. SMEEKENS, 1999 Fructan: More than a reserve carbohydrate? Plant Physiology **120:** 351-360.

VINCENTINI, A., J. C. BARBER, S. S. ALISCIONI, L. M. GIUSSANI and E. A. KELLOGG, 2008 The age of the grasses and clusters of origins of C_4 photosynthesis. Global Change Biology **14:** 2963-2977.

VOGEL, J. T., D. G. ZARKA, H. A. V. BUSKIRK, S. G. FOWLER and M. F. THOMASHOW, 2005 Roles of the CBF2 and ZAT12 transcription factors in configuring the low temperature transcriptome of Arabidopsis. The Plant Journal **41:** 195-211.

WALTERS, R. G., D. G. IBRAHIM, P. HORTON and N. J. KRUGER, 2004 A mutant of *Arabidopsis* lacking the triose-phosphate/phosphate translocator reveals metabolic regulation of starch breakdown in the light. Plant Physiology **135:** 891-906.

WANG, D., Y. GUO, C. WU, G. YANG, Y. LI *et al.*, 2008 Genome-wide analysis of CCCH zinc finger family in *Arabidopsis* and rice. BMC Genomics **9:** 44.

WARNER, J. R., and K. B. MCINTOSH, 2009 How common are extraribosomal functions of ribosomal proteins? Molecular Cell **34:** 3-11.

WEN, Y., J.-Z. SHAO, X.-X. PAN and L.-X. XIANG, 2005 Molecular cloning, characterization and expression analysis of QM gene from grass carp (*Ctenopharyngodon idellus*) homologous to Wilms' tumor suppressor. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology **141:** 356-365.

WINFIELD, M., C. LU, I. WILSON, J. COGHILL and K. EDWARDS, 2009 Cold- and light-induced changes in the transcriptome of wheat leading to phase transition from vegetative to reproductive growth. BMC Plant Biology **9:** 55.

WINFIELD, M. O., C. LU, I. D. WILSON, J. A. COGHILL and K. J. EDWARDS, 2010 Plant responses to cold: transcriptome analysis of wheat. Plant Biotechnology Journal**:** ahead of print.

XIO-CHUN, G. E., HUI, Z., SHU-XUAN, Z., JI-CHAO, C., CHONG-RONG, S. AND KAI-MING, C., 2002 Cloning and expression analysis of a human putative tumor suppressor gene homologue from rice. Acta Botanica Sinica **44:** 562-566.

XIONG, L., M. ISHITANI, H. LEE and J.-K. ZHU, 2001 The *Arabidopsis* LOS5/ABA3 locus encodes a molybdenum cofactor sulfurase and modulates cold stress- and osmotic stress-responsive gene expression. Plant Cell **13:** 2063-2083.

XIONG, Y., and S.-Z. FEI, 2006 Functional and phylogenetic analysis of a DREB/CBF-like gene in perennial ryegrass (*Lolium perenne* L.). Planta **224:** 878-888.

YAKOVLEV, I. A., C. G. FOSSDAL and Ø. JOHNSEN, 2010 MicroRNAs, the epigenetic memory and climatic adaptation in Norway spruce. New Phytologist**:** ahead of print.

YAMAGUCHI-SHINOZAKI, K., and K. SHINOZAKI, 1994 A novel cis-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. Plant Cell **6:** 251-264.

YANCEY, P., M. CLARK, S. HAND, R. BOWLUS and G. SOMERO, 1982 Living with water stress: evolution of osmolyte systems. Science **217:** 1214-1222.

YANG, J.-S., R. WANG, J.-J. MENG, Y.-P. BI, P.-L. XU *et al.*, 2010 Overexpression of *Arabidopsis* CBF1 gene in transgenic tobacco alleviates photoinhibition of PSII and PSI during chilling stress under low irradiance. Journal of Plant Physiology **167:** 534-539.

YOUNG, S. A., M. R. SALTZMAN, W. I. AUSICH, A. DESROCHERS and D. KALJO, 2010 Did changes in atmospheric CO2 coincide with latest Ordovician glacial-interglacial cycles? Palaeogeography, Palaeoclimatology, Palaeoecology**:** ahead of print.

ZACHOS, J., M. PAGANI, L. SLOAN, E. THOMAS and K. BILLUPS, 2001 Trends, rhythms, and aberrations in global climate 65 Ma to present. Science **292:** 686-693.

ZANAZZI, A., M. J. KOHN, B. J. MACFADDEN and D. O. TERRY, 2007 Large temperature drop across the Eocene-Oligocene transition in central North America. Nature **445:** 639-642.

ZARKA, D. G., J. T. VOGEL, D. COOK and M. F. THOMASHOW, 2003 Cold induction of *Arabidopsis* CBF genes involves multiple ICE (Inducer of CBF Expression) promoter elements and a cold-regulatory circuit that is desensitized by low temperature. Plant Physiology **133:** 910-918.

ZEH, D., W. , J. ZEH, A. and Y. ISHIDA, 2009 Transposable elements and an epigenetic basis for punctuated equilibria. Bioessays **31:** 715-726.

ZHANG, H.-Y., H. HE, L.-B. CHEN, L. LI, M.-Z. LIANG *et al.*, 2008 A genome-wide transcription analysis reveals a close correlation of promoter INDEL polymorphism and heterotic gene expression in rice hybrids. Molecular Plant **1:** 720-731.

ZHANG, J., Y. XU, Q. HUAN and K. CHONG, 2009 Deep sequencing of *Brachypodium* small RNAs at the global genome level identifies microRNAs involved in cold stress response. BMC Genomics **10:** 449.

ZHANG, X., S. G. FOWLER, H. CHENG, Y. LOU, S. Y. RHEE *et al.*, 2004 Freezing-sensitive tomato has a functional CBF cold response pathway, but a CBF regulon that differs from that of freezing-tolerant *Arabidopsis*. The Plant Journal **39:** 905-919.

ZHEN, Y., and M. C. UNGERER, 2008 Relaxed selection on the CBF/DREB1 regulatory genes and reduced freezing tolerance in the southern range of *Arabidopsis thaliana*. Molecular Biology and Evolution **25:** 2547-2555.

ZHOU, Z., M.-J. WANG, S.-T. ZHAO, J.-J. HU and M.-Z. LU, 2009 Changes in freezing tolerance in hybrid poplar caused by up- and down-regulation of PtFAD2 gene expression. Transgenic Research.

ZHOU, Z., G. ZHU, A. R. HARIRI, M.-A. ENOCH, D. SCOTT *et al.*, 2008 Genetic variation in human NPY expression affects stress response and emotion. Nature **452:** 997-1001.

ZHU, B., D. W. CHOI, R. FENTON and T. J. CLOSE, 2000 Expression of the barley dehydrin multigene family and the development of freezing tolerance. Molecular and General Genetics **264:** 145-153.

ZHU, J., C.-H. DONG and J.-K. ZHU, 2007 Interplay between cold-responsive gene regulation, metabolism and RNA processing during plant cold acclimation. Current Opinion in Plant Biology **10:** 290-295.

ZOU, C., M. D. LEHTI-SHIU, M. THOMASHOW and S.-H. SHIU, 2009 Evolution of stress-regulated gene expression in duplicate genes of *Arabidopsis thaliana*. PLoS Genet **5:** e1000581.

ZUTHER, E., K. BÜCHEL, M. HUNDERTMARK, M. STITT, D. K. HINCHA *et al.*, 2004 The role of raffinose in the cold acclimation response of *Arabidopsis thaliana*. FEBS Letters **576:** 169-173.

ÖRVAR, B. L., V. SANGWAN, F. OMANN and R. S. DHINDSA, 2000 Early steps in cold sensing by plant cells: the role of actin cytoskeleton and membrane fluidity. The Plant Journal **23:** 785-794.

Paper I

RUDI, H., S. R. SANDVE, L. OPSETH and O. A. ROGNLI, 2010 **Identification of candidate genes important for frost tolerance in Festuca pratensis Huds. by transcriptional profiling.** Plant Science In press

Plant Science xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Identification of candidate genes important for frost tolerance in *Festuca pratensis* Huds. by transcriptional profiling

Heidi Rudi[a,*], Simen R. Sandve[a], Lars M. Opseth[a], Arild Larsen[b], Odd Arne Rognli[a]

[a] Norwegian University of Life Sciences, Department of Plant and Environmental Sciences, N-1432 Ås, Norway
[b] Graminor AS, Torggården, N-8049 Bodø, Norway

ARTICLE INFO

Article history:
Received 5 March 2010
Received in revised form 12 July 2010
Accepted 15 July 2010
Available online xxx

Keywords:
Festuca pratensis
Frost tolerance
Cold acclimation
Winter survival
Perennial ryegrass
Expressed sequence tags

ABSTRACT

Studies of differential gene expression between cold acclimated (CA) and non-cold acclimated (NA) plants yield insight into how plants prepare for cold stress at the transcriptional level. Furthermore genes involved in the cold acclimation process are good candidate loci for genetic variation in frost tolerance and winter survival. In this study we combine different approaches to try to decode the genetics of cold acclimation and frost tolerance in meadow fescue (*Festuca pratensis* Huds). An EST library of cold acclimation responsive genes was established by suppression subtractive hybridization (SSH), and a microarray experiment was used to identify gene expression differences between high and low frost tolerance genotypes in response to cold acclimation. Many genes known to be involved in CA in other species were confirmed to be involved in CA in *F. pratensis*, however, 18% of the ESTs did not show significant homology to any database proteins. Seven genes were found to be differentially expressed (>2-fold) between high and low frost tolerance genotypes. Two of these genes, FpQM and FpTPT, represent interesting candidate genes for frost tolerance in perennial forage grasses.

1. Introduction

The *Pooideae* grasses (temperate grasses) is a large and economically important plant sub-family, including cereals (*Triticeae* tribe) and forage grasses (*Poaeae* tribe). Meadow fescue (*Festuca pratensis* Huds.) is the second most important forage grass species in Northern Europe. Species within the genera *Festuca* have a higher level of general stress tolerance compared with perennial ryegrass and is therefore excellent for studying plant adaptations to cold environment because of their adaptation to a life in northern most Europe.

Exposure to low, non-freezing temperatures induces a process in plants known as cold acclimation (CA). This process improves frost tolerance (FT) and survival of over-wintering plants through the expression of cold-responsive genes, which ultimately leads to altered physiological state of the plant [1]. The CA process is multigenic and drives major metabolic changes including changes in carbohydrates, proteins, nucleic acids, amino acids, growth regulators, phospholipids, and fatty acids [2]. Expression analyses using the *Arabidopsis* whole-genome array have shown that about 4–14% of the *Arabidopsis* genome is cold-responsive [3,4]. Similar results have been reported in *Pooideae* grasses [5–7].

Low temperature specific gene expression is mediated by different parallel signal transduction pathways. In *Arabidopsis* two main signaling pathways are known; one is depending on the involvement of the phytohormone abscisic acid (ABA) and one is not [8]. Both pathways trigger the expression of a range of transcription factors (TFs) that binds to CRT/DRE, ABRE, and MYCR/MYBR binding sites, amongst others, and regulates downstream transcription of CA and FT associated genes [9].

Many molecular processes which happen during CA are conserved between the dicot and monocots lineages [10]. However, most studies on CA transcriptional responses have been carried out on species that are not adapted to a perennial life in extreme winter climates such as the *Pooideae* forage grasses. This is important because adaptation to a perennial life history in harsh winter climates must have required changes at the genetic level which cannot be studied using an annual model species. If we constrict our research on cold and frost stress transcriptional responses to *Arabidopsis* or annual *Triticeae* species, we can only have limited insights into the genetic mechanisms underlying frost tolerance in important agricultural forage grass species. Some investigations into CA transcriptional responses in perennial forage grasses have been limited to *L. perenne* [11,12], a species which does not represent a typical frost tolerant forage grass.

Furthermore, identical CA conditions can result in large differences in interspecific FT levels. Such differences have been linked

* Corresponding author at: Norwegian University of Life Sciences, Department of Plant and Environmental Sciences, Box 5003, N-1432 Ås, Norway
Tel.: +47 91550493; fax: +47 64965901.
E-mail addresses: heidi.rudi@umb.no (H. Rudi), simen.sandve@umb.no (S.R. Sandve), lars.opseth@umb.no (L.M. Opseth), arild.larsen.@graminor.no (A. Larsen), odd-arne.rognli@umb.no (O.A. Rognli).

0168-9452/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.plantsci.2010.07.014

to several QTLs in *Pooideae* species [13,14], unpublished results including one QTL which co-locate with CBF transcription factors upstream in the regulation of CA-associated genes [15,16]. This might suggests that natural variation in regulation of downstream targets during CA may be a common cause for differences in FT levels. The relative importance of regulatory mutations versus structural mutations in adaptive evolution is a debated topic [17], however there are many examples of natural variation in gene regulation contributing to phenotypic evolution [18,19]. An attractive hypothesis is therefore that selection acts on variation in expression levels of important CA genes that are involved in minimizing damaging effects of cold stress.

The aim of this work was to identify genes differentially regulated in response to cold acclimation (CA) in the frost tolerant forage grass *F. pratensis* and to establish links between differences in CA induced gene expression and levels of frost tolerance (FT). This was achieved using a combination of experiments involving artificial selection for frost tolerance, phenotyping of FT using freezing tests, SSH-EST library construction, and microarray analysis.

2. Material and methods

2.1. Plant materials used to generate SSH cDNA libraries

Plant material used to construct the two SSH cDNA libraries were developed from a pair-cross between a genotype from a Yugoslavian cultivar (B14/16) and a genotype from a Norwegian cultivar (HF2/7). A total of 138 plants from the F_1 population were freeze tested by Alm et al. [20] and the two most extreme genotypes with the lowest and highest FT was selected. Five clonal ramets of the two genotypes were pre-grown at 17/12 °C day/night, 16/8 h light/darkness for 12 weeks. The light intensity was 135 μmol m^{-2} s^{-1} (400 W HQI-BT Osram bulbs). Crown tissues from the pre-grown non-acclimated (NA) clonal ramets were sampled at the same time as the pre-grown clonal ramets of the two genotypes were placed under cold acclimating conditions, i.e. 1 °C for 16 h photoperiod at 243 μmol m^{-2} s^{-1}. This was done in order to avoid effects of differences in developmental stage when comparing CA vs. NA plants, since the plants subjected to CA at 1 °C will stop growing at this temperature. Crown tissue of cold acclimated (CA) clonal ramets were sampled after 19 days of cold acclimation; a length of cold acclimation supposed to induce a state of maximum cold hardening under these conditions in meadow fescue. We ended up with two SSH cDNA libraries, one from the HF population and one from the LF population.

2.2. Divergent selections for freezing tolerance used to validate SSH and microarray results

In order to (i) validate the efficiency of the suppression subtractive hybridization (SSH) procedure in detecting candidate genes for frost tolerance and (ii) associate frost tolerance with gene expression levels, we developed two synthetic F_2 populations by divergent selections for frost tolerance. These selections were based on plant material from the 'B14/16 × HF2/7′ meadow fescue mapping family. One high frost tolerant (HF) and one low frost tolerant (LF) population was made by intercrossing the twelve plants with highest and lowest freezing tolerance, respectively. The freezing tolerance data of the 'B14/16 × HF2/7′ meadow fescue mapping family has been presented by Alm et al. (unpublished data). Crossings within each selected population were done by open pollination in isolation to get selected syn-1 populations. Twenty plants were randomly selected from $HF_{syn\text{-}1}$ and $LF_{syn\text{-}1}$, respectively, cloned and subjected to freezing tests as described by Larsen [21] using 12 replications. Frost tolerance was scored as regrowth after 10 and 20 days, according to a scale from 0 (dead) to 9 (without damage). Based on the average freezing scores the 10 most extreme high vs. low frost tolerance genotypes among the 20 HF_{syn1} and LF_{syn1} genotypes were selected for validation of the SSH procedure by hybridization of pooled mRNA extracted from crown tissues of non-cold acclimated (NA) and cold acclimated (CA) clonal ramets (see description below). The average FT of the plants from which the pooled mRNA samples were prepared was 6.49 ± 0.52 and 2.98 ± 0.51 for HF_{syn1} and LF_{syn1}, respectively.

2.3. Extraction of RNA from crown tissue

For use in microarray hybridization total RNA was isolated using a modified LiCl extraction protocol. Tissue was grinded to a fine powder with a prechilled mortar and 5 ml RNA extraction buffer (100 mM Tris pH 9.0, 1% SDS, 100 mM LiCl, 8 M LiCl and 10 mM EDTA) added per gram of plant material together with an equal volume of phenol (pH 8.0). The powder was quickly transferred into a 50 ml tube, vortexed and incubated at 60 °C in a water bath. After centrifugation (3000 rpm, 10 min) the supernatant was extracted with 1/2 volume of chloroform:isoamyl alcohol (24:1). The extraction was repeated once and one-third volume of 8 M LiCl was added to the supernatant, transferred to a 12 ml tube and incubated at 4 °C for 2 h. After centrifugation (10,000 rpm, 4 °C for 10 min), the pellet was washed (80% EtOH) dried and dissolved in 50 μl DEPC dH_2O per gram tissue. The RNA samples were DNase treated using Ambion DNA-free™ and purified using Qiagen RNeasy Mini Kit (Qiagen, Valencia, CA, USA).

2.4. Suppression subtractive hybridization

Subtracted cDNA libraries were constructed using the PCR-selected cDNA subtractive hybridization kit (SSH) according to manufacturer's protocol (Clontech, USA) and described by Diatchenko et al. [22]. PolyA$^+$ RNA was extracted from 300 mg crown tissue after 19 days of CA with oligo (dT) magnetic beads (DYNAL Dynabeads®mRNA DIRECT™Kit). The mRNA quality and quantity were measured on the Agilent 2100 Bioanalyzer.

Tester and driver cDNAs were prepared from 0.5 μg of Poly A$^+$ RNA extracted from CA and NA crown tissues, respectively. We ended up with two subtracted cDNA libraries, one from the HF genotypes and one from the LF genotypes, enriched for differentially expressed genes after CA. cDNA fragments were inserted into the pGEM®-T Easy Vector (Promega). Approximately 1000 clones were randomly selected for sequencing using M13 forward and reverse primers flanking the multiple cloning sites of the pGEM®-T Easy vector. DNA sequencing was performed using an automatic DNA sequencer (ABI Prism 3700) according to the manufacturer's instructions.

2.5. cDNA microarray

Randomly selected clones (559) among the 1000 sequenced were amplified by PCR using SP6 and T7 primers and the amplification products purified using the Montage® $PCR_{\mu 96}$ plate PCR Clean up kit (Millipore). EST's from the two SSH libraries were chosen based on quality and length of the sequences and printed together with 4161 *Lolium perenne* cDNAs originating from the other partners in the EU project GRASP [23] on a *Lolium/Festuca* cDNA microarray. Each PCR product was printed four times and the whole array was printed two times (two sub-arrays). Equal amounts of total RNA from crown tissues of non-acclimated (NA) and cold acclimated (CA) plants of the 10 extreme high frost tolerant $Hf_{syn\text{-}1}$ genotypes and the 10 extreme low frost tolerant $LF_{syn\text{-}1}$ genotypes were pooled separately generating four pools

Please cite this article in press as: H. Rudi, et al., Identification of candidate genes important for frost tolerance in *Festuca pratensis* Huds. by transcriptional profiling, Plant Sci. (2010), doi:10.1016/j.plantsci.2010.07.014

of RNA; HF_{syn1}-CA, HF_{syn1}-NA, LF_{syn1}-CA, LF_{syn1}-NA, and used for hybridization.

cDNA was obtained by reverse transcription from 40 μg and 20 μg total RNA and fluorescently indirect labeled with the fluorochromes cyanine-3 (Cy3) or cyanine-5 (Cy5), respectively. RNA was mixed with control luciferase polyA$^+$ mRNA (1 and 0.5 ng for Cy3- and Cy5-labeled samples, respectively) and first-strand synthesis was done in the presence of modified nucleotides (amino-allyl-dUTP) using random hexamer as primers (for details see [24]). Dyeswap with Cy5- and Cy3-labeled samples was carried out for each comparison. The slides were hybridized, washed and scanned for fluorescence emission using a ScanArray Express HT (Perkin Elmer, Boston, MA, USA).

2.6. Quantitative RT-PCR

Total RNA (100 mg) was extracted from each sample using RNeasy Plant Mini kit (Qiagen, Hilden, Germany). Frozen materials were placed in liquid nitrogen and grinded with a mortar and pestle. The tissue powder was treated as described in the manufacturer's protocol and we did DNase digestion on the column to eliminate genomic DNA contamination (RNeasy Mini Handbook, Qiagen). We then used RNeasy MinElute kit (Qiagen) to clean up RNA. RNA quality was analyzed using Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA) and RNA quantity was measured three times on a Nanodrop ND-1000 UV–Vis Spectrophotometer (Nanodrop Technologies, Wilmington, DE, USA) and averaged for real-time PCR quantification. 2.5 μg of total RNA was reversed transcribed using SuperScriptVilo cDNA synthesis kit (Invitrogen, Carlsbad, CA, USA). For quantitative RT-PCR we used the EXPRESS two-step qRT-PCR universal kit (Invitrogen) with the superscript VILO cDNA synthesis kit (Invitrogen). We used 2 μl of the cDNA synthesis reaction for each qPCR reaction in a total of 20 μl reaction volume. Primers were used at a concentration of 0.5 μM each and TaqMan probes at a concentration of 0.2 μM. Final ROX concentration was 0.05 μM. Transcript levels were analyzed using a ABI7500 real-time PCR machine (Applied Biosystems) according to the 7500 System software (Fast Cycling Program); 95 °C 20 s and 40 cycles of 95 °C 3 s, and 60 °C 30 s. As an endogenous control we used transcript analysis of *GADPH*. TaqMan MGB probes and primer sets were designed using Primer Express Software (Applied Biosystems). The primer and probe sequences are presented in Supplementary data, Table S1. Each RNA sample was assayed in duplicate and each experiment was replicated three times with independent biological materials. A standard curve was made for each primer and probe pairs and transcript levels were calculated relative to *GADPH* using the comparative threshold cycle method (ΔCT method). Fold changes in transcript levels over time and between treatments was calculated using the $\Delta\Delta C_t$ method.

2.7. Analyses of microarray data

The final expression level comparison was based on a single dyeswap comparison with eight technical replicates per microarray slide. The spot recognition software, ScanArray Express was used and global background correction (normalization based on all data) was performed. Expression data collected were further analyzed using the GeneSpring software (Silicon Genetics, Redwood City, CA, USA) using the cross-gene error model. Unfortunately, limitations in the production of the custom made microarrays in this study excluded the possibility of testing for significant expression changes; hence a 2-fold change threshold was applied to classify genes as having different expression profiles.

Table 1
Mean (±SD) of frost tolerance and heritability estimates of divergent selections based on a regrowth test. Phenotypic scale of freezing tolerance; 0 = no regrowth, dead to 9 = no visible damage, max regrowth.

Population	Base population	Selection HF	Selection LF
Base population (B14/16 × HF2/7)	4.48 ± 0.59	5.35 ± 0.16	3.61 ± 0.79
Syn-1		5.25 ± 1.34	4.21 ± 1.20
Selection differential (S)		0.87	0.87
Selection response (R)		0.77	0.27
Realized heritability (h^2)		0.89	0.31

2.8. Sequence analysis

EST sequences were run through a sequence quality check and assembly pipeline consisting of the Phred, Phrap and Consed programs [25–27]. Annotation of ESTs was done by selecting the best blastx search hit against Wheat TIGR10, Rice TIGR, Wheat Unigene, and UNIPROT databases.

3. Results

3.1. Divergent selection for frost tolerance

The results of the divergent selections for frost tolerance demonstrate good responses to selection especially for selection in direction high frost tolerance (Table 1). Additive genetic variation for freezing tolerance is thus large within the 'B14/16 × HF2/7' meadow fescue mapping population, illustrated by realized heritability estimates of h^2 = 0.89 and 0.39 for selection in direction high and low frost tolerance, respectively. The difference in realized heritability is explained by differences in genetic variance components in the populations selected for high and low frost tolerance, respectively. These selections should therefore be very suitable for testing differentially expression of genes involved in cold acclimation and freezing tolerance in meadow fescue.

3.2. EST sequencing and sequence analysis

The subtracted cDNA libraries contained DNA fragments ranging in size from 400 to approximately 1500 bp, with an average of 500 bp. From an initial 885 good sequences, we ended up with a total of 372 ESTs after sequence quality checks and contig assembly. An overview of the EST contig depth distribution is presented in Fig. 1. Blastx searches with an E-value cut off at <1E−4 yielded 322 ESTs homologous to putative proteins in the Uniprot, TIGR Wheat or TIGR Rice protein databases. Functional classification was done manually for all ESTs aided by the blastx searches, and the putative proteins were classified into 14 functional classes (Fig. 2). Proteins within "cell signaling", "cellular transport and membrane associated", "transcription" and "metabolism" are expected to be important for cold acclimation and these classes represent (21%) of the proteins. For genes known to be directly involved in the cold acclimation process we made a "cold stress" functional class that constituted 2% of the proteins. A large proportion (18%) of the ESTs could not be classified and were grouped into "unknown proteins". The three biggest classes, not considering "other" or "unknown" proteins, were "protein synthesis", "cell signaling" and "cellular transport and membrane associated" representing 21, 9 and 8% respectively.

3.3. Microarray analysis

The microarray analysis confirmed the enrichment of CA-associated genes in our EST collection. Among our 559 cDNAs

Please cite this article in press as: H. Rudi, et al., Identification of candidate genes important for frost tolerance in *Festuca pratensis* Huds. by transcriptional profiling, Plant Sci. (2010), doi:10.1016/j.plantsci.2010.07.014

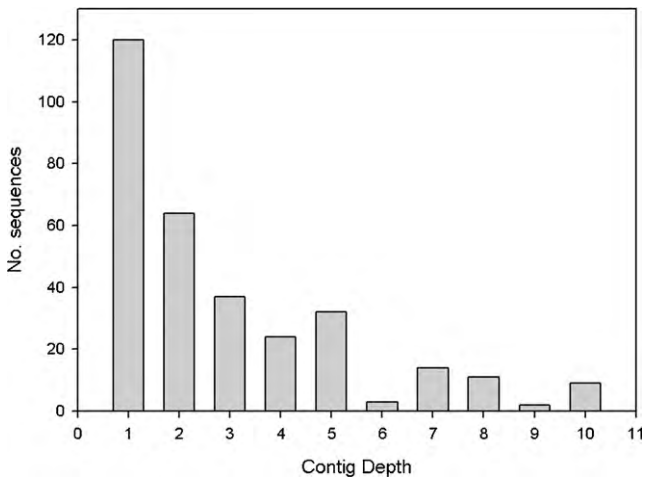


Fig. 1. Occurrence of ESTs in contigs. After sequence quality check and contig assembly we ended up with a total of 372 ESTs. A relatively high numbers were singletons in our subtracted cDNA libraries.

spotted on the array 147 (26%) genes were differentially expressed more than 2-fold, 111 upregulated and 36 downregulated when comparing HF_{syn1}-CA against HF_{syn1}-NA populations (data not shown). Several known candidate frost tolerance gene homologs were found among the upregulated genes (>2-fold), e.g. an ice recrystallization inhibition protein (IRIP), a 6-ft (fructan 6-fructosyltransferase) homolog, cr7 (low temperature membrane protein), and wcor80 (dehydrin) (Table 2). Analyses of differential expressions between HF_{syn1}-CA and LF_{syn1}-CA identified only 7 (1.3%) genes that displayed >2-fold difference, 4 upregulated and 3 downregulated (Table 3). Interestingly, two of these genes, the QM and TPT genes, have previously been associated with FT.

3.4. Relative expression of two candidate genes after cold acclimation

qRT-PCR were performed on two of these seven candidate genes for which other studies have indicated a role in plant FT, the tumor suppressor gene or QM gene [22] and a triose phosphate/phosphate translocator (TPT) [23]. The relative expression after cold acclimation was measured in the two parents in the *Festuca* mapping family (B14/16 and HF2/7) after 1, 2 and 3 weeks of CA and in three different tissues (crowns, stems and leaves). HF2/7 has higher slightly higher frost tolerance and much higher winter survival than B14/16 (unpublished data). The FpQM gene is upregulated after CA in stems and leaves (Fig. 3A). The relative expression is highest after 1 week CA and decreases in stems whereas in leaves it peaks at 2 weeks CA and then decreases. In the crown tissue the expression is lower compared to stems and leaves. The FpTPT gene, encoding a putative TPT, is upregulated after CA in crown tissue, stems and leaves in HF2/7 relative to the B14/16. As a summary, the expression peaks after 3 week CA in crows, 1 week in stems and 2 weeks in leaves of HF2/7 (Fig. 3B).

3.5. Characterization of a QM homologue

In our study we found that upregulation of a QM homologue were associated with differences in frost tolerance in *F. pratensis* genotypes. Based on the EST sequence of the partial FpQM gene we screened a *Festuca* BAC library [28] by PCR using gene-specific primers. The BAC clone was sequenced and Blast searches in the BrachyBase (http://www.brachybase.org/blast/) showed that *Brachypodium distachyon* contains at least two homologous genes on different chromosomes. A multiple sequence alignment of the deduced amino acid sequences of FpQM with other plant homologs showed very high conservation in the N-terminal and internal part of the protein (Supplementary data, Figure S1), with the presence of the ribosomal protein L10e signature (ADRLQTGMRGAFGKPQGVCARV) [29].

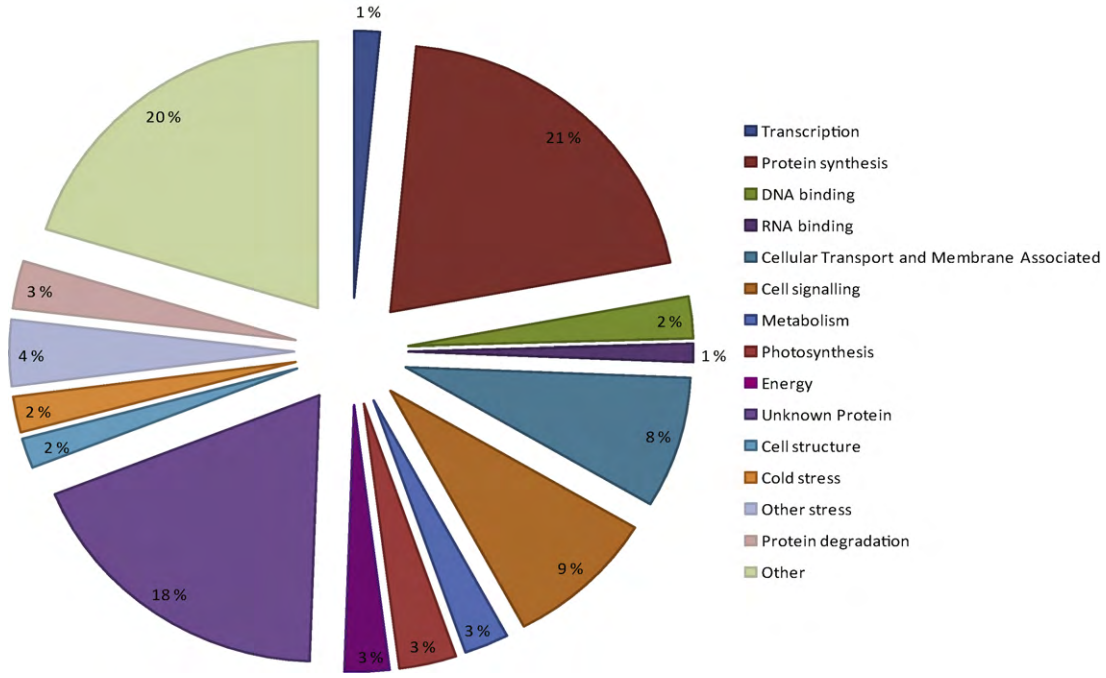


Fig. 2. Distribution of functional categories among EST clusters in meadow fescue. For 18% of the sequences no protein function could be found after blastx search against Wheat TIGR10, Rice TIGR, Wheat Unigene, and UNIPROT databases, with an E-value cut off at <1E−4. Functional classification was done manually for all ESTs aided by the blastx searches, and the putative proteins were classified into 14 functional classes. "Others" means proteins that have other functions than the ones described in the figure. The different classes are shown by different colors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Please cite this article in press as: H. Rudi, et al., Identification of candidate genes important for frost tolerance in *Festuca pratensis* Huds. by transcriptional profiling, Plant Sci. (2010), doi:10.1016/j.plantsci.2010.07.014

Table 2
Homologous meadow fescue genes associated with cold acclimation and frost tolerance in other grass species. Blastx search of ESTs isolated by subtractive cloning of cold acclimated (CA) low frost (LF) and high frost (HF)-tolerant meadow fescue. All have E-values ≤1E−05.

EST	Accession#	Homologous gene	Source
S2.59	X57554	Cold-regulated protein BLT14	*H. vulgare*
S2.59	L28093	LTS (low temperature specific) protein, TACR7	*T. aestivum*
ST3.70	TAE132439	mRNA for protein encoded by lt1.1 gene	*T. aestivum*
ST3.45	AF494041	Fructan 6-fructosyltransferase (6FT)	*L. perenne*
S5.10	U73212	Cold acclimation protein WCOR80	*T. aestivum*
S5.10	AF031250	Dehydrin-/LEA group 2-like protein	*T. elongatum*
S2.94	S47480	Chlorophyll a/b-binding protein type II, photosystem I	*L. temulentum*
S2.94	S52341	LHCI-680, photosystem I antenna protein	*H. vulgare*
ST3.40	NP_565489	LRR receptor protein kinase	*A. thaliana*
S2.5	AFO69331	BLTI-5 (low temperature induced)	*H. vulgare*
S2.5	AF271260	Cold-responsive protein (Wlt10)	*T. aestivum*
A2S.6	AF184280	Polyubiquitin (RUBQ2) gene	*O. sativa*
ST1.7	TAVDAC1	Voltage dependent, anion-selective channel (VDAC) protein	*T. aestivum*
ST3.135	CAB87814	Ice recrystallisation inhibition protein (AFP)	*L. perenne*
ST3.146	BLYINOPP	mRNA for vacuolar membrane proton-translocating inorganic pyrophosphatase	*H. vulgare*
ST1.45	JT0280	Sucrose synthase (EC 2.4.1.13)	*T. aestivum*
ST1.2	BA97282	Succinate dehydrogenase flavoprotein alpha subunit	*A. thaliana*
ST2.14	AJ311048	Alcohol dehydrogenase	*P. glaucum*
ST1.20	XP464488	Putative zinc finger and C2 domain protein	*O. sativa*
ST2.11	X57313	H2B histone	*Z. mays*
ST3.123	AJ496413	Manganese superoxide dismutase (SOD)	*A. camphorate*
S5.87	XP483065	Putative heat shock protein	*O. sativa*

4. Discussion

4.1. Functional groups involved in CA

In this study we have successfully combined the technologies of SSH and cDNA microarrays to identify genes differentially expressed between cold acclimated and non-acclimated plant tissues. In general, the results are comparable to other studies on CA as regards the distribution of functional classes in our EST libraries [30–32]. There was relatively high abundance of differentially expressed genes involved in cell signaling i.e. receptor-like protein kinases, calcium-modulated protein kinases, MAPK, casein kinases and serine/threonine protein kinases. Moreover, many genes involved in protein synthesis and metabolism (e.g. carbohydrates) were also differentially regulated between NA and CA plants.One important functional group commonly found to be regulated during CA is genes involved in photosynthesis [30]. In our study we used RNA from non-photosynthetic tissue from the crown, hence very few (3%) differentially expressed genes were classified as photosynthesis related. Recently Winfield et al. [7] demonstrated that the global transcriptional pattern in wheat during CA is very different between crown and leaf tissue. In fact, only 1–10% of the genes found to be regulated during CA were overlapping between the two different tissues [7], highlighting the importance of studying molecular processes in different tissues for a complete understanding of CA.

Few transcription factors were identified in our SSH library (1%), however this is likely due to the timing of RNA sampling during the CA process. RNA was sampled after 19 days of CA and the transcription factors involved in regulation of CA responses are most abundant in the very early stages of CA [33,34].

The co-location of QTLs for frost tolerance and map position of ESTs involved in CA could indicate the involvement of the EST-locus in frost tolerance. For one CA-associated EST, the position on the *F. pratensis* genetic map was known. This was a homolog of the ice recrystallization inhibition protein (IRIP) genes that has been shown to be *Pooideae* grass specific [35], bind to ice crystals [36] and enhance frost tolerance by reducing cell membrane destruction during freezing stress [37]. Interestingly the FpIRI1 gene mapped 1.5 cM distal to the confidence interval of a major QTL for frost tolerance in *F. pratensis* (unpublished results).

4.2. Expression differences in a few genes during CA is associated with divergent frost tolerance levels

Variation in gene expression during CA could represent candidate mechanisms underlying phenotypic differences in winter survival in natural populations. Mutants and transgenic plants with changes in the CA-pathways have been used to investigate the causative link between frost tolerance phenotypes and differences in gene expression [6,33,38]. In this study we used synthetically selected plants to link expression levels during CA with freezing tolerance phenotypes. Because abiotic stress tolerance is a highly quantitative trait, adaptive evolution from standing genetic variation are expected to happen through small changes in many genes, so-called soft sweeps [39]. This was demonstrated for gene expression levels in *Drosophila melanogaster*. Synthetic selection for divergent levels of abiotic stress tolerance have been shown to pro-

Table 3
tBlastX search of ESTs differentially expressed more than 2-fold between cold acclimated (CA) high frost (HF) and low frost (LF) genotypes after microarray hybridization. ND = not determined, because of no homology to any proteins in the tBlastX search.

EST	tBlastX	E-value
S2-82	Hypothetical protein	3E−14
S2-72	*Lolium perenne* clone LpSSR020 SSR marker sequence (AY919047)	3E−39
ST1-31	*Arabidopsis thaliana* putative tumor suppressor protein (At1g14320)	1E−72
A2S-95	Hypothetical protein	
ST3-96	ND	ND
ST2-22	*Zea mays* protein disulfide isomerase (PDIL2-3) (AY739290)	3E−50
ST3-127	*Zea mays* clone 285969 triose phosphate/phosphate translocator (EU965428)	6E−34

Please cite this article in press as: H. Rudi, et al., Identification of candidate genes important for frost tolerance in *Festuca pratensis* Huds. by transcriptional profiling, Plant Sci. (2010), doi:10.1016/j.plantsci.2010.07.014

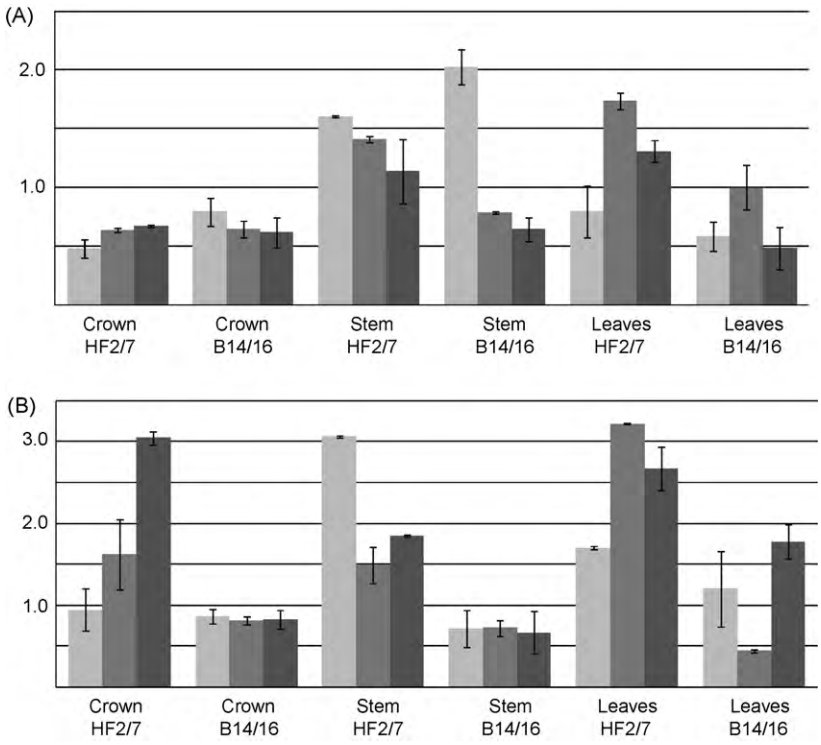


Fig. 3. Quantitative RT-PCR of the two candidate genes, FpTPT and FpQM in *F. pratensis* after cold acclimation (CA). Relative transcription was measured for the FpTPT (A) and FpQM (B) in crown, stem and leaves after 1 week (light grey), 2 weeks (grey) and 3 weeks (dark grey) of CA. The *X*-axis represents different length of CA treatments and tissues for two different frost tolerance genotypes (HF2/7 and B14/16). The *Y*-axis represents relative fold change in expression compared to non-acclimated (NA) controls.

duce small adaptive expression level changes in hundreds of genes [40]. Unfortunately, limitations in the production of the custom made microarrays in this study excluded the possibility of testing for significant expression changes in our study and a 2-fold change threshold was applied. With this criterion, only 1.3% of the genes experienced a change in expression level between plants with high and low frost tolerance. Furthermore, we analysed pooled RNA from multiple genotypes in high and low frost tolerance groups, respectively. In pooled RNA samples genotype specific expression differences not related to the frost tolerance *per se* is averaged out by the combination of many genotypes. However, this pooling strategy also eliminates genotype specific expression differences that truly are related to the selection for frost tolerance. Consequently, both the 2-fold filter and the RNA pooling strategy results in a highly conservative estimate of the actual numbers of genes differentially expressed between high and low frost tolerant genotypes. On the other hand, our conservative approach increases the chance that the expression differences associated with divergent frost tolerance are functionally linked to the phenotype.

4.3. Candidate genes linking gene expression and frost tolerance

Our results indicate that expression differences in only 7 genes are associated with differing levels of frost tolerance. Of these 7 genes, the QM-like and TPT-like genes have to some degree been functionally characterized in other plants and represent excellent candidates for being casually linked to the differences in frost tolerance. The QM gene was first isolated and identified as a putative tumor suppressor gene [41,42] and is highly conserved among eukaryotes [43]. As today the QM function is elusive. Functional studies have demonstrated that QM holds DNA-binding properties and regulate transcription [41,42]. Other studies have demonstrated ribosomal localization of the QM protein [44,45] and the conserved L10e ribosomal binding domain of QM is known to organize the architecture of the aminoacyl tRNA binding site of the ribosome [29]. QM homologs in both animals and plants have been implicated in abiotic and biotic stress responses. In carp fish, a QM homolog was upregulated subsequent to viral and bacterial infections [46]. In plants, upregulation of QM has been demonstrated during cold (and also other abiotic stress) in rice [47] and the temperate plant *Caragana jubata* that grows under extreme cold and high altitude in Himalaya [48]. Furthermore, the expression of a tomato QM gene in yeast protected against oxidative stress by increasing the accumulation of proline [49]. Because intercellular biochemical changes ultimately depend on mRNA-translation and protein synthesis, one hypothesis is that QM upregulation during cold stress (and also other biotic and abiotic stress) ensures that necessary intercellular biochemical adjustments are carried out efficiently to mitigate the adverse effects of the stress. However, ribosomal proteins are often implicated in extra-ribosomal functions [50], which also is the case for the QM gene [41], thus another hypothesis is that the QM protein are involved in transcriptional regulation of other genes involved in CA, e.g. genes in the proline synthesis pathway [49].

The TPT-like gene encodes a protein homologous to a triose phosphate/phosphate translocator (TPT). TPT proteins are located in the chloroplast membrane and are responsible for export of sucrose to the cytosol [51]. Accumulation of sucrose in the cytosol during CA is important for frost tolerance in *Arabidopsis* leaves [52,53] and in the Antarctic hair grass *Deschampsia antarctica* [54,55]. The causative link between sucrose accumulation and acquirement of frost tolerance could be related to stress alleviating functions of sucrose itself [56]. Sucrose has been demonstrated to make sugar–membrane interactions and stabilize membranes during freezing [57–59], and sucrose can also protect protein structure during freezing [60]. Moreover, the sucrose levels also affect the ability of the plant to produce other important CA-related compounds. In many plants [61–63], including *Pooideae* grasses [64],

Please cite this article in press as: H. Rudi, et al., Identification of candidate genes important for frost tolerance in *Festuca pratensis* Huds. by transcriptional profiling, Plant Sci. (2010), doi:10.1016/j.plantsci.2010.07.014

fructan accumulation have been demonstrated to improve cold stress tolerance and frost tolerance. The initial step in fructan synthesis is the conversion of two sucrose molecules to a fructan molecule by the fructocyl transferase enzyme 1-sucrose-sucrose-fructosyltransferase [65]. Hence, TPT expression levels could also affect frost tolerance by influencing the cytosolic fructan levels which ultimately depend on sucrose availability.

Acknowledgements

We thank Drs. Lidia Skuza and Magdalena Achrem for excellent help with the qRT-PCR and BAC sequencing work, and Øyvind Jørgensen and Britta Fromm for excellent technical assistance with the plant propagation and freezing tests. We are also indebted to Dr. Oene Dolstra and Carole Boucoiran, Plant Research International, Wageningen, The Netherlands, for performing the microarray hybridizations. We also thank Dr. Iain Donnison and Dr. Kerrie Farrar, IBERS, Aberystwyth University, UK, for the *Festuca* BAC library. This work was funded by the EU FP5 project GRASP (QLK5-CT-2002-00862) and project no. 173319/I10 funded by the Research Council of Norway and Graminor AS.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.plantsci.2010.07.014.

References

[1] M.F. Thomashow, Plant cold acclimation: freezing tolerance genes and regulatory mechanisms, Annu. Rev. Plant. Phys. 50 (1999) 571–599.
[2] P.H. Li, Subzero temperature stress physiology of herbaceous plants, Hort. Rev. 6 (1984) 373–416.
[3] S. Fowler, M.F. Thomashow, *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway, Plant Cell 14 (2002) 1675–1690.
[4] J.A. Kreps, Y. Wu, H.S. Chang, T. Zhu, X. Wang, J.F. Harper, Transcriptome changes for *Arabidopsis* in response to salt, osmotic, and cold stress, Plant Physiol. 130 (2002) 2129–2141.
[5] E.M Herman, K. Rotter, R. Premakumar, G. Elwinger, H. Bae, L. Ehler-King, S. Chen, D.P. Livingston 3rd, Additional freeze hardiness in wheat acquired by exposure to -3 degrees is associated with extensive physiological, morphological, and molecular changes, J. Exp. Bot. 57 (2006) 3601–3618.
[6] J.T. Svensson, C. Crosatti, C. Campoli, R. Bassi, A.M. Stanca, T.J. Close, L. Cattivelli, Transcriptome analysis of cold acclimation in barley albina and xantha mutants, Plant Physiol. 141 (2006) 257–270.
[7] M.O. Winfield, C. Lu, I.D. Wilson, J.A. Coghill, K.J. Edwards, Plant responses to cold: transcriptome analysis of wheat, Plant Biotechnol. J. (2010) 1–23.
[8] J.Z. Zhang, R.A. Creelman, J.K. Zhu, From laboratory to field. Using information from *Arabidopsis* to engineer salt, cold, and drought tolerance in crops, Plant Physiol. 135 (2004) 615–621.
[9] V. Chinnusamy, K. Schumaker, J.K. Zhu, Molecular genetic perspectives on cross-talk and specificity in abiotic stress signalling in plants, J. Exp. Bot. 55 (2004) 225–236.
[10] K. Nakashima, Y. Ito, K. Yamaguchi-Shinozaki, Transcriptional regulatory networks in response to abiotic stresses in *Arabidopsis* and grasses, Plant Physiol. 149 (2009) 88–95.
[11] T. Sawbridge, E.-K. Ong, C. Binnion, M. Emmerling, R. McInnes, K. Meath, N. Nguyen, M. O'Neill, F. O'Toole, C. Rhodes, J. Simmonds, P. Tian, K. Wearne, T. Webster, A. Winkworth, G. Spangenberg, Generation and analysis of expressed sequence tags in perennial ryegrass (*Lolium perenne* L.), Plant Sci. 165 (2003) 1089–1100.
[12] C. Zhang, S.Z. Fei, S. Warnke, L. Li, D. Hannapel, Identification of genes associated with cold acclimation in perennial ryegrass, J. Plant Physiol. 166 (2009) 1436–1445.
[13] Y.W. Xiong, S.Z. Fei, R. Arora, E.C. Brummer, R.E. Barker, G.W. Jung, S.E. Warnke, Identification of quantitative trait loci controlling winter hardiness in an annual x perennial ryegrass interspecific hybrid population, Mol. Breeding 19 (2007) 125–136.
[14] T. Yamada, E.S. Jones, N.O.I. Cogan, A.C. Vecchies, T. Nomura, H. Hisano, Y. Shimamoto, K.F. Smith, M.D. Hayward, J.W. Forster, QTL analysis of morphological, developmental, and winter hardiness-associated traits in perennial ryegrass, Crop Sci. 44 (2004) 925–935.
[15] A Fricano, F. Rizza, P. Faccioli, D. Pagani, P. Pavan, A. Stella, L. Rossini, P. Piffanelli, L. Cattivelli, Genetic variants of HvCbf14 are statistically associated with frost tolerance in a European germplasm collection of Hordeum vulgare, Theor. Appl. Genet. 119 (2009) 1335–1348.
[16] F. Sutton, D.G. Chen, X. Ge, D. Kenefick, Cbf genes of the Fr-A2 allele are differentially regulated between long-term cold acclimated crown tissue of freeze-resistant and–susceptible, winter wheat mutant lines, BMC Plant Biol. 9 (2009) 34.
[17] H.E Hoekstra, J.A. Coyne, The locus of evolution: evo devo and the genetics of adaptation, Evolution 61 (2007) 995–1016.
[18] M.D. Shapiro, M.E. Marks, C.L. Peichel, B.K. Blackman, K.S. Nereng, B. Jonsson, D. Schluter, D.M. Kingsley, Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks, Nature 428 (2004) 717–723.
[19] G.A. Wray, The evolutionary significance of cis-regulatory mutations, Nat. Rev. Genet. 8 (2007) 206–216.
[20] V. Alm, C. Fang, C.S. Busso, K.M. Devos, K. Vollan, Z. Grieg, O.A. Rognli, A linkage map of meadow fescue (*Festuca pratensis* Huds.) and comparative mapping with other Poaceae species, Theor. Appl. Genet. 108 (2003) 25–40.
[21] A. Larsen, Freezing tolerance in grasses. Methods for testing in controlled environments, Meld Norg Landbr Høgsk 57 (1978) 1–56.
[22] L. Diatchenko, Y-F.C. Lau, A.P. Campbell, A. Chenchik, F. Moqadam, B. Huang, S. Lukyanov, K. Lukyanov, N. Gurskaya, E.D. Sverdlov, P.D. Siebert, Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 6025–6030.
[23] A.B. Lübberstedt T, P.B. Holm, Development of ryegrass allele-specific (grasp) markers for sustainable grassland improvement—a new EU Framework V Project, Czech J. Genet. Plant Breed 39 (2003) 125–128.
[24] A. Aharoni, P.L.C. Keizer, H.J. Bouwmeester, Z. Sun, M. Alvarez-Huerta, H.A. Verhoeven, J. Blaas, A.M.M.L. van Houwelingen, R.C.H. De Vos, H. van der Voet, R.C. Jansen, M. Guis, J. Mol, R.W. Davis, M. Schena, A.J. van Tunen, A.P. O'Connell, Identification of the *SAAT* gene involved in strawberry flavor biogenesis by use of DNA microarrays, Plant Cell 12 (2000) 647–662.
[25] M. de la Bastide, W.R. McCombie, Assembling genomic DNA sequences with PHRAP, in: Current Protocols in Bioinformatics, John Wiley, 2007.
[26] B Ewing, L. Hillier, M.C. Wendl, P. Green, Base-calling of automated sequencer traces using phred. I. Accuracy assessment, Genome Res. 8 (1998) 175–185.
[27] D. Gordon, C. Abajian, P. Green, Consed: a graphical tool for sequence finishing, Genome Res. 8 (1998) 195–202.
[28] I.S. Donnison, D.M. O'Sullivan, A. Thomas, P. Canter, B. Moore, I. Armstead, H. Thomas, K.J. Edwards, I.P. King, Construction of a *Festuca pratensis* BAC library for map-based cloning in Festulolium substitution lines, Theor. Appl. Genet. 110 (2005) 846–851.
[29] M. Nishimura, T. Kaminishi, C. Takemoto, M. Kawazoe, T. Yoshida, A. Tanaka, S. Sugano, M. Shirouzu, T. Ohkubo, S. Yokoyama, Y. Kobayashi, Crystal structure of human ribosomal protein L10 core domain reveals eukaryote-specific motifs in addition to the conserved fold, J. Mol. Biol. 377 (2008) 421–430.
[30] M. Brautigam, A. Lindlof, S. Zakhrabekova, G. Gharti-Chhetri, B. Olsson, O. Olsson, Generation and analysis of 9792 EST sequences from cold acclimated oat, Avena sativa, BMC Plant Biol. 5 (2005), doi:10.1186/1471-2229-1185-1118.
[31] M. Houde, M. Belcaid, F. Ouellet, J. Danyluk, A.F. Monroy, A. Dryanova, P. Gulick, A. Bergeron, A. Laroche, M.G. Links, L. MacCarthy, W.L. Crosby, F. Sarhan, Wheat EST resources for functional genomics of abiotic stress, BMC Genomics 7/149 (2006), doi:10.1186/1471-2164-1187-1149.
[32] A.F. Monroy, A. Dryanova, B. Malette, D.H. Oren, M.R. Farajalla, W. Liu, J. Danyluk, L.W.C. Ubayasena, K. Kane, G.J. Scoles, F. Sarhan, P.J. Gulick, Regulatory gene candidates and gene expression analysis of cold acclimation in winter and spring wheat, Plant Mol. Biol. 64 (2007) 409–423.
[33] J.T. Vogel, D.G. Zarka, H.A. Van Buskirk, S.G. Fowler, M.F. Thomashow, Roles of the CBF2 and ZAT12 transcription factors in configuring the low temperature transcriptome of *Arabidopsis*, Plant J. 41 (2005) 195–211.
[34] B.H. Lee, D.A. Henderson, J.K. Zhu, The *Arabidopsis* cold-responsive transcriptome and its regulation by ICE1, Plant Cell 17 (2005) 3155–3175.
[35] S. Sandve, H. Rudi, T. Asp, O. Rognli, Tracking the evolution of a cold stress associated gene family in cold tolerant grasses, BMC Evol. Biol. 8 (2008) 245.
[36] C. Sidebottom, S. Buckley, P. Pudney, S. Twigg, C. Jarman, C. Holt, J. Telford, A. McArthur, D. Worrall, R. Hubbard, P. Lillford, Phytochemistry: heat-stable antifreeze protein from grass, Nature 406 (2000) 256.
[37] C. Zhang, S.-Z. Fei, R. Arora, D. Hannapel, Ice recrystallization inhibition proteins of perennial ryegrass enhance freezing tolerance, Planta 232 (2010) 155–164.
[38] H. Lee, L. Xiong, M. Ishitani, B. Stevenson, J.K. Zhu, Cold-regulated gene expression and freezing tolerance in an *Arabidopsis* thaliana mutant, Plant J. 17 (1999) 301–308.
[39] J.K. Pritchard, J.K. Pickrell, G. Coop, The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation, Curr. Biol. 20 (2010) R208–R215.
[40] J.G Sørensen, M.M. Nielsen, V. Loeschcke, Gene expression profile analysis of *Drosophila melanogaster* selected for resistance to environmental stressors, J. Evol. Biol. 20 (2007) 1624–1636.
[41] H.S. Oh, H. Kwon, H.K. Sun, C.H. Yang, QM, a putative tumor suppressor, regulates proto-oncogene c-Yes, J. Biol. Chem. 277 (2002) 36489–36498.
[42] F.S Monteclaro, P.K. Vogt, A Jun-Binding protein related to a putative tumor-suppressor, Proc. Natl. Acad. Sci. U.S.A. 90 (1993) 6726–6730.
[43] A.A. Farmer, T.M. Loftus, A.A. Mills, K.Y. Sato, J.D. Neill, T. Tron, M. Yang, B.L. Trumpower, E.J. Stanbridge, Extreme evolutionary conservation of QM, a novel c-Jun associated transcription factor, Hum. Mol. Genet. 3 (1994) 723–728.
[44] T.M. Loftus, Y.H. Nguyen, E.J. Stanbridge, The QM protein associates with ribosomes in the rough endoplasmic reticulum, Biochemistry 36 (1997) 8224–8230.

Please cite this article in press as: H. Rudi, et al., Identification of candidate genes important for frost tolerance in *Festuca pratensis* Huds. by transcriptional profiling, Plant Sci. (2010), doi:10.1016/j.plantsci.2010.07.014

[45] S.G. Lillico, J.C. Mottram, N.B. Murphy, S.C. Welburn, Characterisation of the QM gene of *Trypanosoma brucei*, FEMS Microbiol. Lett. 211 (2002) 123–128.
[46] Y. Wen, J.-Z. Shao, X.-X. Pan, L.-X. Xiang, Molecular cloning, characterization and expression analysis of QM gene from grass carp (Ctenopharyngodon idellus) homologous to Wilms' tumor suppressor, Comp. Biochem. Physiol. B: Biochem. Mol. Biol. 141 (2005) 356–365.
[47] G.E Xio-Chun, Z. Hui, Z. Shu-Xuan, C. Ji-Chao, S. Chong-Rong, C. Kai-Ming, Cloning and expression analysis of a human putative tumor suppressor gene homologue from rice, Acta. Bot. Sin. 44 (2002) 562–566.
[48] P.K. Bhardwaj, P.S. Ahuha, S. Kumar, Characterization of gene expression of QM from *Caragana jubata*, a plant species that grows under extreme cold, Mol. Biol. Rep. 37 (2010) 1003–1010.
[49] C. Chen, S. Wanduragala, D.F. Becker, M.B. Dickman, Tomato QM-like protein protects saccharomyces cerevisiae cells against oxidative stress by regulating intracellular proline levels, Appl. Environ. Microbiol. 72 (2006) 4001–4006.
[50] J.R. Warner, K.B. McIntosh, How common are extra-ribosomal functions of ribosomal proteins? Mol. Cell 34 (2009) 3–11.
[51] R.G. Walters, D.G. Ibrahim, P. Horton, N.J. Kruger, A mutant of arabidopsis lacking the triose-phosphate/phosphate translocator reveals metabolic regulation of starch breakdown in the light, Plant Physiol. 135 (2004) 891–906.
[52] J. Klotke, J. Kopka, N. Gatzke, A.G. Heyer, Impact of soluble sugar concentrations on the acquisition of freezing tolerance in accessions of *Arabidopsis* thaliana with contrasting cold adaptation—evidence for a role of raffinose in cold acclimation, Plant Cell Environ. 27 (2004) 1395–1404.
[53] A. Strand, C.H. Foyer, P. Gustafsson, P. Gardestrom, V. Hurry, Altering flux through the sucrose biosynthesis pathway in transgenic *Arabidopsis* thaliana modifies photosynthetic acclimation at low temperatures and the development of freezing tolerance, Plant Cell Environ. 26 (2003) 523–535.
[54] L.A. Bravo, L. Baschunan-Godoy, E. Perez-Torres, L.J. Corcuera, Cold hardiness in antarctic vascular plants in; Plant cold hardiness from laboratory to the field, ISBN-13: 978-1-84593-513-9 (2009) 198–213.
[55] J. Levitt, Responses of Plants to Environmental Stresses, 1, Academic Press, New York, 1980.
[56] K.A. Santarius, Freezing of isolated thylakoid membranes in complex media. VIII. Differential cryoprotection by sucrose, proline and glycerol, Physiol. Plant. 84 (1992) 87–93.
[57] J.H. Crowe, L.M. Crowe, J.F. Carpenter, A.S. Rudolph, C.A. Wistrom, B.J. Spargo, T.J. Anchordoguy, Interactions of sugars with membranes, BBA-Rev Biomembranes 947 (1988) 367–384.
[58] M. Caffrey, V. Fonseca, A.C. Leopold, Lipid-sugar interactions: relevance to anhydrous biology, Plant Physiol. 86 (1988) 754–758.
[59] T.J. Anchordoguy, A.S. Rudolph, J.F. Carpenter, J.H. Crowe, Modes of interaction of cryoprotectants with membrane phospholipids during freezing, Cryobiology 24 (1987) 324–331.
[60] J.F. Carpenter, J.H. Crowe, The mechanism of cryoprotection of proteins by solutes, Cryobiology 25 (1988) 244–255.
[61] D. Parvanova, A. Popova, I. Zaharieva, P. Lambrev, T. Konstantinova, S. Taneva, A. Atanassov, V. Goltsev, D. Djilianov, Low temperature tolerance of tobacco plants transformed to accumulate proline, fructans, or glycine betaine. Variable chlorophyll fluorescence evidence, Photosynthetica 42 (2004) 179–185.
[62] H.J. Li, A.F. Yang, X.C. Zhang, F. Gao, J.R. Zhang, Improving freezing tolerance of transgenic tobacco expressing-sucrose: sucrose 1-fructosyltransferase gene from Lactuca sativa, Plant Cell Tiss. Org. 89 (2007) 37–48.
[63] A. Kawakami, Y. Sato, M. Yoshida, Genetic engineering of rice capable of synthesizing fructans and enhancing chilling tolerance, J. Exp. Bot. 59 (2008) 793–802.
[64] H. Hisano, A. Kanazawa, A. Kawakami, M. Yoshida, Y. Shimamoto, T. Yamada, Transgenic perennial ryegrass plants expressing wheat fructosyltransferase genes accumulate increased amounts of fructan and acquire increased tolerance on a cellular level to freezing, Plant Sci. 167 (2004) 861–868.
[65] D. Livingston, D. Hincha, A. Heyer, Fructan and its relationship to abiotic stress tolerance in plants, Cell. Mol. Life Sci. 66 (2009) 2007–2023.

Please cite this article in press as: H. Rudi, et al., Identification of candidate genes important for frost tolerance in *Festuca pratensis* Huds. by transcriptional profiling, Plant Sci. (2010), doi:10.1016/j.plantsci.2010.07.014

Paper II

BARTOŠ, J., S.R. SANDVE, R. KÖLLIKER, D. KOPECKÝ, P. NĚMCOVÁ, Š. STOČES, L. ØSTREM, A. LARSEN, A. KILIAN, O. A. ROGNLI, J. DOLEŽEL. **Genetic mapping of DArT markers in the *Festuca-Lolium* complex and their use in marker-trait (freezing tolerance) association analysis**. Manuscript

Genetic mapping of DArT markers in the *Festuca-Lolium* complex and their use in marker-trait (freezing tolerance) association analysis

Jan Bartoš[1], Simen Rød Sandve[2], Roland Kölliker[3], David Kopecký[1], Pavla Němcová [1], Štěpán Stočes[1], Liv Østrem[4], Arild Larsen[5], Andrzej Kilian[6], Odd Arne Rognli[2], Jaroslav Doležel[1]

[1] CRH, Institute of Experimental Botany, Sokolovská 6, 77200 Olomouc, Czech Republic

[2] Department of Plant and Environmental Sciences, Norwegian University of Life Sciences, Ch.Magnus Falsens vei 1, 1432 Ås, Norway

[3] Agroscope Reckenholz-Tänikon Research Station ART, Reckenholzstrasse 191, CH-8046 Zurich, Switzerland

[4]Norwegian Institute for Agricultural and Environmental Research, Fureneset, N-6997 Hellevik i Fjaler, Norway

[5]Graminor AS, c/o Bioforsk Nord Bodø, Torggården, N-8049 Bodø, Norway

[4] Diversity Arrays Technology, 1 Wilf Crane Crescent, ACT 2600 Yarralumla, Australia

ABSTRACT

Species belonging to the *Festuca-Lolium* complex are important grassland, pasture and turf species and as such have been intensively studied. However, the limited availability of molecular markers and the out-crossing reproduction system of these species make genetic studies a challenging task. Here, we report on the saturation of *F. pratensis* and *L. multiflorum* genetic maps with diversity array technology (DArT) markers from the DArTFest array. In total, 658 DArT markers were placed on the genetic maps of the two species. Subsequently these mapped markers were sequenced and mapped *in silico* to the rice and *Brachypodium distachyon* genomes. The utility of the DArTFest array was then tested on a Festulolium population (FuRs0357) in an integrated analysis using the DArT marker map positions and sequence information to study associations between DArT markers and freezing tolerance. Ninety six markers were significantly associated with freezing tolerance and five of these markers were genetically mapped to chromosomes 2, 4 and 7. The three genomic loci associated with freezing tolerance in FuRs0357 co-localized with chromosome segments and QTLs previously implicated in freezing tolerance. The work presented clearly demonstrates the power of the DArTFest array in genetic studies of the *Festuca-Lolium* complex. The annotated DArTFest array resources could accelerate further studies and improvement of important traits in *Festuca-Lolium* species.

INTRODUCTION

The *Lolium–Festuca* species complex includes some of the world's most important forage grasses. Even though the forage grasses are closely related to the cereals and thus shares many characteristics with wheat, barley and other well studied crops, species of the *Lolium–Festuca* complex also possess several features unique to this group. This includes biological features, e.g. the perenniality of many forage grasses compared to the annuality of temperate cereals, and genomic features such as *Lolium–Festuca* specific chromosome rearrangements (ALM *et al.* 2003). Species within the *Lolium–Festuca* complex possess a range of complementary characteristics which are exploited in grass hybrid breeding.

Italian ryegrass (*Lolium multiflorum* Lam.) is one of the most important forage grasses of temperate regions. It is especially valued for its high dry matter yield and his excellent forage quality. Although primarily used for hay and silage production in temporary leys, it is also a substantial component of permanent grassland (PETER-SCHMID *et al.* 2008). However, *Lolium* species usually have poor persistence under abiotic stress. *Festuca* species on the other hand, generally have a better abiotic stress tolerance compared to *Lolium* species. Meadow fescue (*F. pratensis* Huds.) is a forage grass of high quality and yield potential which constitutes a significant component of species-rich permanent pastures and hay fields in alpine and eastern regions of Europe. In Scandinavia it is a major component also of intensively managed swards cut for silage (ROGNLI *et al.* 2010). Because of their close evolutionary relationships, introgression of *Festuca* genes into *Lolium* genomes (to make Festulolium hybrids) can be used to improve forage grass stress resistance (HUMPHREYS *et al.* 2005). In this respect *F. pratensis* is commonly used to improve freezing tolerance in such Lolium x Festuca hybrids (KOSMALA *et al.* 2006).

To better understand the links between the biological characteristics of the forage grasses and their genomic constitution and allelic variation, availability of genome wide molecular markers for high throughput characterization of *Lolium–Festuca* species is essential. **D**iversity **Ar**rays **T**echnology (DArT) is a high-throughput sequence independent genotyping method based on reduction of genome complexity and DNA microarray hybridization (JACCOUD *et al.* 2001). Through the past decade, it has become a valuable source of markers for genomes with limited sequence information. DArT arrays have been developed for crop species like banana, cassava, cereals as well as for model plants such as *Arabidopsis*, rice and sorghum (for

complete list see http://www.diversityarrays.com/publications.html). Although DArT markers were originally developed for diversity studies, they are frequently used for construction and saturation of genetic maps (TINKER *et al.* 2009; WENZL *et al.* 2006). DArT markers have also been used in the analysis of important agricultural traits (SINGH *et al.* 2010).

Recently, Kopecky et al. (2009) developed a DArTFest array containing 7680 probes for the *Festuca-Lolium* complex. Diverse accessions of three *Festuca* (*F. arundinacea, F. glaucescens and F. pratensis*) and two *Lolium* (*L. multiflorum* and *L. perenne*) species were used for array development. There are several reasons why the DArTFest array could become a valuable tool for genomic characterization of *Festuca-Lolium* species. Many of the DArT markers cross-hybridize to several different species which enables comparative genomics studies. DArT markers can also be sequenced and physically mapped to the genomes of fully sequenced model species to further understand the syntenic relationships between chromosomes of forage grasses and model plants. This may in turn lead to more efficient identification of genes underlying important agricultural traits.

According to breeding priorities, molecular dissection of agronomic traits in *L. multiflorum* has so far mainly been focused on resistance to diseases such as crown rust or bacterial wilt (e.g. STUDER *et al.* 2007; STUDER *et al.* 2006). In this context, a genetic linkage map based on 306 F_1 individuals has been established (*Xtg*-ART) and more recently been used for establishing a consensus linkage map in *Lolium* using EST-derived SSR markers (STUDER *et al.* 2010). Alm et al. (2003) established a *F. pratensis* mapping population (HF2/7 x B14/16) consisting of 138 F_1 individuals based on 446 RFLP, AFLP and SSR markers. This has been used for analysis of freezing and drought tolerance, winter hardiness, and vernalization sensitivity (ALM *et al.* submitted; ERGON *et al.* 2006). In this study we (i) saturate and improve these genetic maps of *F. pratensis* and *L. multiflorum* with DArTFest markers (ii) sequence the mapped markers to investigate the genomic origin of the DArT markers and the syntenic relationships to model genomes, and (iii) use these new DArTFest array resources in an integrated analysis of the agriculturally important trait freezing tolerance in a *L. perenne x F. pratensis* hybrid population.

MATERIAL AND METHODS

Plant material

The HF2/7 x B14/16 mapping family (ALM *et al.* 2003) consisting of 138 F1-offspring was used for the construction of the *F. pratensis* map, while the *Xtg*-ART mapping population (STUDER *et al.* 2006) consisting of 288 F1 individuals was used for the construction of the *L. multiflorum* map. In the association study the Festulolium population FuRs0357 was used. FuRs0357 (*L. perenne x F. pratensis*) originates from a wide genetic pool from several initial hybrids made from either *Festulolium* cv. Prior (LpFp, 4n) crossed with *L. perenne* (2x), or crosses between *L. perenne* (4n) and *F. pratensis* (2n). The initial hybrids were backcrossed twice onto diploid *L. perenne* to obtain BC_2 progenies and then put through two generations of seed propagation. Twenty one plants belonging to either a high freezing tolerance (HFT) group (11 plants) or a low freezing tolerance (LFT) group (10 plants) were used in the study.

DArT screening

Genomic DNA was isolated from individual plants of the mapping populations and FuRs0357 using Invisorb Spin Mini Kit (Invitek, Berlin, Germany). The previously developed DArTFest array (KOPECKY *et al.* 2009) was hybridized with fluorescently labelled genomic representations of individual plants which were prepared from genomic DNA by the same PstI/TaqI complexity reduction method as used for preparing the array (for details see AKBARI *et al.* 2006). Hybridization signals were converted into 0-1 scores using the DArTsoft software package developed at Diversity Arrays Technology Pty Ltd (DArT P/L, Yarralumla, Australia).

Genetic mapping

The *F. pratensis* and *L. multiflorum* maps were constructed in JoinMap 4.0 using the Kosambi mapping algorithm. Because DArT markers are dominant, marker files were coded as for double haploid populations (DH) and maps were calculated for each parent separately. Maps were then combined using bridging markers present in both parents. Inclusion of markers was decided by the following procedure. First, all markers were grouped into putative linkage groups based on likelihood ratio odds grouping. Next, initial maps were calculated and all markers with a chi-square value of >3 (highly distorted segregation) were removed. Subsequent recalculation of maps was performed until no markers had a chi-square score of >3. Finally, the two parental maps for each LG were combined into one. All final linkage maps were edited and finalized using the MapChart software (VOORRIPS 2002).

Association of DArT markers with freezing tolerance

Freezing tolerance was measured with the re-growth method (LARSEN 1978). In brief, 300 genotypes from FuRs0357 were put through controlled freezing stress which resulted in differential freezing induced damage and survival. Plants were randomly grouped in boxes using 6 replicates per genotype. The resulting HFT and LFT plants belonged to the 10% highest and lowest freezing tolerance phenotypes, respectively. The HFT group had an average freezing tolerance score of 7.06 (±0.22) and the LFT group had an average freezing tolerance score of 2.59 (±0.61) (range; 0 (dead) – 9 (no damage)). Associations between DArT marker genotypes and freezing tolerance groups were tested using Fisher's exact test. The null hypothesis is that the DArT marker genotypes are not associated with freezing tolerance; hence we expect random distribution of genotypes in HFT and LFT groups. We calculated the corresponding q-values from the Fisher Exact test p-values (STOREY 2002) to correct for multiple testing and false positives. A significance threshold of $q<0.05$ was used. Statistical analyses were carried out in R (R DEVELOPMENT CORE TEAM 2009) using the "fisher.test" and "q-value" functions in the base and qvalue packages, respectively.

Sequencing of selected clones/markers

All the DArT markers placed on the genetic maps of *F. pratensis* and *L. multiflorum* or significantly associated with freezing tolerance were selected for sequencing. The DNA was isolated according to standard DNA-isolation kit procedures. Reaction mix for cycle sequencing was prepared using standard BigDye chemistry (BigDye® Terminator v3.1 Cycle Sequencing Kit, Applied Biosystems). The dilution of reaction components was scaled to the final volume of 10 µl containing 3.2 pmol of universal M13 (forward or reverse) primer and 20 ng of sequence-ready (DArT clones) template. The reaction products were purified using the CleanSEQ kit (Agencourt Bioscience Corp., Beckman Coulter Comp., Beverly, MA, USA) and analyzed on an ABI 3730xl DNA analyzer (Applied Biosystems, Foster City, CA, USA). The obtained raw sequence data were assembled and edited using the DNA Baser software v.2 (HeracleSoftware, http://www.DnaBaser.com). Vector and adaptor sequences were removed prior to further analysis. All sequences were deposited in GenBank under accessions AAA – ZZZ.

Sequence analysis

All mapped markers were compared to each other to determine the extent of redundancy in DArT clones. Stand alone blastn software was used for the search with E-value set to e-10. Only reciprocal blast hits were taken into account. The map position of markers with significant hits was checked and markers which mapped to different map positions were not considered redundant. The sequences were further compared to known plant repeat sequences. To analyse repeat content TREP Release 10 (http://wheat.pw.usda.gov/ITMI/Repeats/) was joined with TIGR Plant Repeat Databases for Brassicaceae, Fabaceae, Poaceae and Solanaceae (OUYANG and BUELL 2004). This in-house build database contains 8,432 repetitive elements from 18 different genera. Repeat analysis was performed using RepeatMasker software (http://repeatmasker.org) with CrossMatch search engine (http://www.phrap.org/phredphrapconsed.html) and default settings. We compared DArT sequences with non-redundant protein sequences (nr) and non-human, non-mouse ESTs database (est_others) at GenBank in order to estimate the number of DArT markers derived from expressed loci. The protein search (blastx) was performed using blastcl3 with default settings and BLOSUM62 scoring matrix. Nucleotide search (blastn) was also performed by blastcl3 with default settings but reward for a nucleotide match was set to 2. Only the best blast hits (with lowest E-value) were taken into account.

Comparison of mapped DArTs to model genomes

The sequences of the mapped DArT markers were compared to the *Oryza sativa* ssp. *japonica* cultivar Nipponbare genome and the *Brachypodium distachyon* Bd21 genome using blastn with E-value set to e-10. Twelve pseudomolecules for the rice chromosomes (Build5) were downloaded from the IRGSP website (http://rgp.dna.affrc.go.jp/E/IRGSP/Build5/build5.html). The Bd21 genome was downloaded from Brachybase.org (http://files.brachypodium.org/). Only the best blast hits with alignment lengths of at least 50 bp were taken into account.

GISH analysis of Festulolium plants

Twelve individuals from FuRs0357 were selected randomly, six from HFT and six from LFT, for analysis of genomic constitution using Genomic *in* situ hybridization (GISH) as described by Kopecký et al. (2005). In brief, total genomic DNA of *F. pratensis* was labeled with digoxigenin using DIG-Nick Translation Kit (Roche Applied Science, Indianapolis, IN, USA) and used as a probe. Genomic DNA of *L. multiflorum* was sheared to ~200 bp fragments and

used as blocking DNA. Sites of probe hybridization were detected by anti-DIG-FITC conjugate (Roche). Chromosomes were counterstained with 1.5µg/ml 4',6-diamidino-2-phenylindole (DAPI) prepared in Vectashield antifade solution (Vector Laboratories, Burlingame, USA). Observations were made using an Olympus AX70 microscope equipped with epi-fluorescence and SensiCam B/W camera. ScionImage and Adobe Photoshop software were used for processing of color pictures.

RESULTS

Genetic mapping

As the DH mapping strategy was used for the construction of genetic maps, only markers segregating close to the expected 1:1 ratio (>0.4 and <0.6) were used. In total, 658 markers were placed on genetic maps, 148 markers in the *F. pratensis* population HF2xB14, and 529 markers in the *L. multiflorum* population *Xtg*-ART (Table 1, Fig. 1). Only 16 markers were mapped in both species.

Sequencing and sequence analysis of mapped markers

The 621 successfully sequenced DArT markers (out of 658 markers mapped) yielded 303.297 bp of sequence with an average length of 488.4 bp. A total of 399 DArT markers (64.3%) were found to represent unique genomic loci while the 222 remaining markers were redundant and assigned to 90 marker bins. The biggest bin consisted of six markers. Hence, 489 non-redundant DArT markers/loci were mapped when counting each bin as a unique locus (Table 2). Some markers belonging to a bin differ slightly in map position (see Fig. 1), which is most likely a result of genotyping errors or inaccuracies in the consensus map estimation.

To estimate the number of DArT markers derived from plant repetitive elements we performed blast search against an in-house built composite plant repeat database. Only 44 (7.1%) of the DArT markers contained repetitive elements supporting the notion that DArT markers represent hypo-methylated genomic regions. The number of retrotransposons (27) dominated over DNA-transposons (13) in DArT sequences. The remaining three hits in the plant repeat database were unclassified. To estimate the impact of repeat elements on marker redundancy we compared DArT markers derived from repetitive elements, with the marker bins. Interestingly, repetitive elements could only explain the presence of redundant markers for 7 (7.8%) of the marker bins.

In order to estimate the proportion of expressed sequences among our mapped DArT markers we performed a blastn search against GenBank "est_others" database. This identified 368 (59.3%) DArT markers with significant homology to expressed sequences. Blastx search against non-redundant protein sequences (nr) revealed 163 (26.2%) DArT markers with significant homology to known and hypothetical proteins, and of these 152 were also identified as expressed sequences in blastn search. In total, 379 DArT markers (293 non-redundant bins) were identified as potentially gene derived sequences. In addition, most of the repetitive element derived DArT markers (31 out of 44) had blast hits to ESTs and hypothetical proteins identified in data sets.

Comparison of mapped DArT markers to model genomes

The genetically mapped DArT markers were compared to the rice and *B. distachyon* genome in order to map the sequences *in silico* and investigate the synteny between model genomes and the *Festuca-Lolium* complex. Using the defined criteria for blast search, mapped DArT markers in *L. multiflorum* produced 228 (162 non-redundant bins) and 298 (225) hits against rice and *B. distachyon* genomes, respectively. Of the markers mapped in *F. pratensis,* 51 (40) and 71 (56) produced blast hits in rice and *B. distanchyon* genomes, respectively. Due to the low numbers of mapped *F. pratensis* markers, the analysis of syntenic relationships was only performed between *L. multiflorum* and rice (Table 3) and *L. multiflorum* and *B. distachyon* (Table 4). All significant homologies are listed in Supplementary table 1.

We identified 11 syntenic regions with more than 5 markers shared between particular *L. multiflorum* and rice chromosomes. The highest degree of synteny was found between *L. multiflorum* chromosome 3 (Lm3) and rice chromosome 1 (Os1), where 71.4% of Lm3 markers had a homologous sequence on Os1. In comparison, 12 syntenic regions with more than 5 shared markers were found between the genomes of *L. multiflorum* and *B. distachyon* (Bd). The most conserved syntenic relationships were found between Lm5-Bd4 and Lm3-Bd2, with 76.9% and 76.5% of the *L. multiflorum* markers having significant blast hits, respectively.

Analysis of markers associated with freezing tolerance

Out of 1868 DArT markers segregating in the FuRs0357 population, about 5% (96) had a significantly different distribution of genotype scores among HFT and LFT plants ($q<0.05$). Five of the significant markers (four non-redundant bins) had a genetic map position. These

five markers mapped to Fp4 (D556159), Fp7 (D560433), Lm2 (D560530), and Lm4 (D558701, D562536) (Table 5, Fig. 1). The FuRs0357 population was in fact derived from a cross between *L. perenne* and *F. pratensis.* Nevertheless, one can expect that map positions of DArT markers in *L. perenne* are highly co-linear with their position in *L. multiflorum* due to the close evolutionary relationship of their genomes (CATALÁN *et al.* 2004).

Seventy two (out of 96) of the markers associated with freezing tolerance were successfully sequenced. Their sequences were then mapped (by blast analysis) to the *B. distachyon* genome to further study the localization of association signals (Fig. 2).

GISH analysis of Festulolium plants

Microscopic analysis of mitotic metaphase plates of six LFT and six HFT plants revealed that FuRs0357 contained plants with different ploidy levels. Four of the LFT plants were tetraploid, while all six HFT plants were diploid. Such a large difference is not likely to occur by chance (Fisher Exact test; P=0.045), which indicates that the polyploidy levels are functionally linked to freezing tolerance. The higher ploidy level could affect the proportion of markers which cannot be called as "0" or "1" when converting fluorescence signal since it increases signal to noise ratio. Exactly this pattern was observed in the FuRs0357 population. The average numbers of not called DArT markers per plant were 14 in HFT and 127 in LFT populations (T-test; P=0.0003), reflecting the higher ploidy level in the LFT group.

GISH revealed *F. pratensis* chromatin only in a low proportion of plants, irrespective of their freezing tolerance phenotype. Two recombined chromosomes carrying terminal *F. pratensis* segments were found in one tetraploid LFT plant (Fig 3a), while two HFT plants had chromosomes carrying terminal *F. pratensis* segments (Fig 3b).

DISCUSSION

DArT marker mapping

The DArT technology has become a valuable resource for many plant species. More than thousand DArT markers have been genetically mapped in rye (BOLIBOK-BRAGOSZEWSKA *et al.* 2009) and oat (TINKER *et al.* 2009), and in barley more than two thousands have been mapped (WENZL *et al.* 2006). Here we have used the recently published DArTFest array (KOPECKY *et al.* 2009) and genetically mapped 658 DArT markers in the *Festuca-Lolium* complex. The mapped markers were not distributed equally between the two studied species;

and four times as many markers were mapped in *L. multiflorum* compared to *F. pratensis*. Among the accessions used for the DArTFest array development, Kopecky et al. (2009) found twice as many polymorphic loci in *L. multiflorum* compared to *F. pratensis* (*L. multiflorum*: 2148, *F. pratensis*: 1078), for the same number of plants screened. Based on this, we should expect about 50% fewer mapped DArT markers in *F. pratensis*. Some of the remaining 25% discrepancy in number of markers mapped in *L. multiflorum* compared to *F. pratensis* could be explained by a low genetic variation between the mapping parents of the *F. pratensis* mapping population. A higher level of polymorphism in *L. multiflorum* ecotypes compared to *F. pratensis* ecotypes was also observed when using co-dominant SSR markers (PETER-SCHMID *et al.* 2008). Based on studies of cpDNA variation Fjellheim et al. (2006) concluded that meadow fescue in Europe went through a bottleneck during or after the last glaciations.. Thus, the number of genetically mapped markers probably also reflect the overall diversity within the particular species.

Sequence analysis of DArT markers

Among the sequenced DArT markers, 64.3% were found to be unique. In similar studies (WITTENBERG *et al.* 2005) and (TINKER *et al.* 2009) found 56.3% and 48.1% unique markers among DArTs developed for *Arabidopsis* and oat, respectively. The levels of redundancy is strikingly similar despite the great difference in genome size, ranging from 157 Mbp for *A. thaliana* to 12,961 Mbp for oat (BENNETT and LEITCH 2005), and the contrasting repeat content of three species. In this study we found that only about 7% of DArT markers are derived from repetitive elements, hence the effect of inclusion of repetitive DNA on marker redundancy seems to be negligible. The origin of redundant DArT markers is therefore mostly a result of sampling multiple copies of the same probe locus in array development and possibly also sampling of probes deriving from tandem-duplicated genes which are frequent in plant genomes (HANADA *et al.* 2008; RIZZON *et al.* 2006).

DArT markers are derived from genomic representations prepared with methyl-sensitive restriction enzymes and hence should represent low-copy number genomic regions. As expected, a majority of the sequences were found to be derived from expressed parts of the genomes. About 60% of all markers had homology to the NCBI EST databases and one fourth had homology to the NCBI protein database. These groups are overlapping extensively; resulting in slightly more than 60% expressed non-redundant markers/bins. This is similar to the finding of (TINKER *et al.* 2009) in oat. Interestingly, most of the DArT markers with

repetitive origin had blast hits to ESTs and proteins. This further supports the tendency of DArT markers to originate from expressed genomic loci.

DArT map synteny with rice and rice and *B. distachyon*

Genome-wide comparisons of *L. multiflorum* to rice and *B. distachyon* based on DArT markers revealed common syntenic regions shared among the species. We found syntenic blocks shared between *L. multiflorum* and the model species for all chromosomes but Lm5. A few syntenic marker relationships between Lm5-Os3, Lm5-Os12, and Lm5-Bd1 were found as expected from other Pooideae species (BOLOT *et al.* 2009; THE INTERNATIONAL BRACHYPODIUM INITIATIVE 2010), although not many enough were identified to meet our criteria of a syntenic block on the chromosome scale (Table 3 and 4). The number of DArT markers used to infer syntenic regions in this paper is small, and many grass chromosomes have complex syntenic relationships (The International Brachypodium Initiative 2010). Thus the lack of expected syntenic blocks on Lm5 in our study could be related to uneven and low marker coverage of *L. multiflorum* chromosomes. We also found an additional syntenic relationship between Lm6 and Bd4, which have not been reported in cereals (BOLOT *et al.* 2009; THE INTERNATIONAL BRACHYPODIUM INITIATIVE 2010). However, in this case the number of markers shared is the lowest we have accepted; thus this syntenic relationship should be confirmed with additional markers.

The DArTFest resources and marker-trait analysis

As an example of how we can integrate the DArT marker map positions genetic maps and the DArT marker sequence information generated in this study, we performed a marker-trait association study in Festulolium populations with divergent freezing tolerance levels. It must be noted that the FuRs0357 population used in this study is anticipated to have high LD-levels due to the breeding scheme and selections; hence, any positive association signals are expected to only provide evidence of chromosome regions containing QTLs rather than identification of specific candidate genes.

We found DArT markers associated with the freezing tolerance phenotype on chromosome 2, 4 and 7 (Fig. 1, 2). Three significant markers, two mapped in *L. multiflorum* and one in *F. pratensis*, were mapped to the same central region of chromosome 4. This is in accordance with the findings by Kosmala et al. (2006) in a different Festulolium population. The same chromosomal region correspond to a QTL for freezing tolerance in *F. pratensis* (ALM *et al.*

submitted) (Fig. 2). This *F. pratensis* QTL is mapped to the same genomic region as the Fr-1 QTL for freezing tolerance in cereals, which is closlely linked to *Vrn1*, a key transcription factor involved in the transition from winter-growt to flowering (GALIBA *et al.* 2009; PRESTON and KELLOGG 2007). It has been debated whether the gene underlying the Fr-1 is directly involved in freezing tolerance, or if it simply is a pleiotropic effect of *Vrn1* function (CATTIVELLI *et al.* 2002; TREVASKIS *et al.* 2007). Recently, polymorphisms in *Vrn1* was shown to exhibit a pleiotropic effect on freezing tolerance by directly or indirectly affect expression of genes involved in the CA pathway (e.g. CBF genes) (DHILLON *et al.* 2010).

The existence of several *Vrn1*-like paralogs in grasses (PRESTON and KELLOGG 2007) complicates the comparative genomic analysis of syntenic relationships. In rice, two *Vrn1*-like paralogs (referred to as FUL1 and FUL2) are situated on Os3 and Os7, while in *B. distachyon* two *Vrn1*-like paralogs are found on Bd1 (HIGGINS *et al.* 2010). Syntenic relationships between forage grass chromosome 4 and both regions containing Vrn1-like gene copies on Bd1 were revealed by additional blast analysis of the significant DArT markers to Bd1 (Fig. 2). Furthermore, DArT markers significantly associated with freezing tolerance had blast hits flanking both FUL1 and FUL2 paralogs on Bd1 (Fig. 2). In conclusion the results indicate that the significant markers on Lm4 and Fp4 are (i) derived from the same genomic region, (ii) are likely to be in linkage disequilibrium with the same underlying QTL, and (iii) that the QTL could be functionally linked to variation in a *Vrn1* gene function and differences in the vernalization control rather than differences in the ability to withstand freezing.

The remaining significant DArT markers were mapped to the central part of Lm2 and the proximal part of Fp7. Chromosome 2 have been previously been implicated in freezing tolerance in Festulolium (KOSMALA *et al.* 2006) and *L. perenne* (SHINOZUKA *et al.* 2006). The significant marker on Fp7 mapped about 1.5 cM proximal to a freezing tolerance QTL in *F. pratensis* (ALM *et al.* submitted). However, no striking candidate genes are closely linked to the significant DArT markers on Lm2 and Fp7.

A common notion is that polyploid plants have improved stress tolerance compared to their diploid progenitors (JACKSON and CHEN 2010; ZHANG *et al.* 2010). However, the tolerance to physical stresses as freezing may decrease. A tetraploid C2 population of *F. pratensis* was shown to be less freezing tolerant and with lower genetic variation than its original diploid population (LARSEN 1978; LARSEN 1994). This is in agreement with our findings, where the

frequency of tetraploid plants was significantly higher in plants with low freezing tolerance. High freezing tolerance in Festulolium hybrids are sometimes explained by small (sometimes even undetectable by GISH) *F. pratensis* introgressions carrying desired genes and alleles into a *L. multiflorum* genomic background (KOSMALA *et al.* 2006). It is therefore possible that the low freezing tolerance in tetraploids can be explained by fewer *F. pratensis* introgressions as a result of their polyploid state or other genetic and/or epigenetic changes accompanying polyploidization event. The FuRs0357 Festulolium population was constructed by crossing 4x (autotetraploid) *L. perenne* with 2x *F. pratensis*, and backcrossing to 2x *L. perenne* for two generations. This should usually produce 2x hybrid plants with introgressed *F. pratensis* segments. Tetraploid hybrid plants among the BC2 progenies could have originated from unreduced gametes from the 4x *L. perenne* parent, and if this is the case they will contain nearly pure *Lolium* genomes.

LITTERATURE

AKBARI, M., P. WENZL, V. CAIG, J. CARLING, L. XIA *et al.*, 2006 Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. Theoretical and Applied Genetics **113:** 1409-1420.

ALM, V., C. S. BUSSO, E. Å, H. RUDI, A. LARSEN *et al.*, submitted.

ALM, V., C. FANG, C. S. BUSSO, K. M. DEVOS, K. VOLLAN *et al.*, 2003 A linkage map of meadow fescue (*Festuca pratensis* Huds.) and comparative mapping with other Poaceae species. Theoretical and Applied Genetics **108:** 25-40.

BENNETT, M. D., and I. J. LEITCH, 2005, http://www.kew.org/cvalues/.

BOLIBOK-BRAGOSZEWSKA, H., K. HELLER-USZYNSKA, P. WENZL, G. USZYNSKI, A. KILIAN *et al.*, 2009 DArT markers for the rye genome - genetic diversity and mapping. BMC Genomics **10:** 578.

BOLOT, S., M. ABROUK, U. MASOOD-QURAISHI, N. STEIN, J. MESSING *et al.*, 2009 The 'inner circle' of the cereal genomes. Current Opinion in Plant Biology **12:** 119-125.

CATALÁN, P., P. TORRECILLA, J. Á. L. RODRIGUEZ and R. G. OLMSTEAD, 2004 Phylogeny of the festucoid grasses of subtribe Loliinae and allies (Poeae, Pooideae) inferred from ITS and trnL-F sequences. Molecular Phylogenetics and Evolution **31:** 517-541.

CATTIVELLI, L., P. BALDI, C. CROSATTI, N. DI FONZO, P. FACCIOLI *et al.*, 2002 Chromosome regions and stress-related sequences involved in resistance to abiotic stress in Triticeae. Plant Molecular Biology **48:** 649-665.

DHILLON, T., S. PEARCE, E. STOCKINGER, A. DISTELFELD, C. LI *et al.*, 2010 Regulation of freezing tolerance and flowering in temperate cereals: the VRN-1 connection. Plant Physiology**:** ahead of print.

ERGON, Å., C. FANG, Ø. JØRGENSEN, T. S. AAMLID and O. A. ROGNLI, 2006 Quantitative trait loci controlling vernalisation requirement, heading time and number of panicles in meadow fescue (Festuca pratensis Huds.). Theoretical and Applied Genetics **112:** 232-242.

FJELLHEIM, S., O. A. ROGNLI, K. FOSNES and C. BROCHMANN, 2006 Phylogeographical history of the widespread meadow fescue (Festuca pratensis Huds.) inferred from chloroplast DNA sequences. Journal of Biogeography **33:** 1470-1478.

GALIBA, G., A. VÁGÚJFALVI, C. LI, A. SOLTÉSZ and J. DUBCOVSKY, 2009 Regulatory genes involved in the determination of frost tolerance in temperate cereals. Plant Science **176:** 12-19.

HANADA, K., C. ZOU, M. D. LEHTI-SHIU, K. SHINOZAKI and S.-H. SHIU, 2008 Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. Plant Physiology **148:** 993-1003.

HUMPHREYS, J., J. A. HARPER, I. P. ARMSTEAD and M. W. HUMPHREYS, 2005 Introgression-mapping of genes for drought resistance transferred from *Festuca arundinacea* var. glaucescens into *Lolium multiflorum*. Theoretical and Applied Genetics **110:** 579-587.

JACCOUD, D., K. PENG, D. FEINSTEIN and A. KILIAN, 2001 Diversity Arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Research **29:** e25-.

JACKSON, S., and Z. J. CHEN, 2010 Genomic and expression plasticity of polyploidy. Current Opinion in Plant Biology **13:** 153-159.

KOPECKY, D., J. BARTOS, A. LUKASZEWSKI, J. BAIRD, V. CERNOCH *et al.*, 2009 Development and mapping of DArT markers within the Festuca - Lolium complex. BMC Genomics **10:** 473.

KOPECKY, D., A. J. LUKASZEWSKI and J. DOLEŽEL, 2005 Genomic constitution of *Festulolium* cultivars released in the Czech Republic. Plant Breeding **124:** 454-458.

KOSMALA, A., Z. ZWIERZYKOWSKI, D. GASIOR, M. RAPACZ, E. ZWIERZYKOWSKA *et al.*, 2006 GISH/FISH mapping of genes for freezing tolerance transferred from *Festuca pratensis* to *Lolium multiflorum*. Heredity **96:** 243-251.

LARSEN, A., 1978 Freezing tolerance in grasses. Methods for testing in controlled environments. Meldinger fra Norges Landbrukshøgskole **57:** 1-56.

LARSEN, A., 1994 Breeding winter hardy grasses. Euphytica **77:** 231-237.

OUYANG, S., and C. R. BUELL, 2004 The TIGR Plant Repeat Databases: a collective resource for the identification of repetitive sequences in plants. Nucleic Acids Research **32:** 360-363.

PETER-SCHMID, M. K. I., B. BOLLER and R. KÖLLIKER, 2008 Habitat and management affect genetic structure of *Festuca pratensis* but not *Lolium multiflorum* ecotype populations. Plant Breeding **127:** 510-517.

PRESTON, J. C., and E. A. KELLOGG, 2007 Conservation and divergence of *APETALA1/FRUITFULL*-like gene function in grasses: evidence from gene expression analyses. The Plant Journal **52:** 69-81.

R DEVELOPMENT CORE TEAM, 2009 R: A language and environment for statistical computing, pp. R Foundation for Statistical Computing, Vienna.

RIZZON, C., L. PONGER and B. S. GAUT, 2006 Striking similarities in the genomic distribution of tandemly arrayed genes in *Arabidopsis* and Rice. PLoS Computational Biology **2:** e115.

ROGNLI, O. A., M. C. SAHA, S. BHAMIDIMARRI and S. VAN DER HEJDEN, 2010 Fescues, pp. pp 261-292 in *Fodder Crops and Amenity Grasses, Handbook of Plant Breeding*, edited by B. BOLLER, U. K. POSSELT and F. VERONESI. Springer Science+Business Media.

SHINOZUKA, H., H. HISANO, S. YONEYAMA, Y. SHIMAMOTO, E. JONES *et al.*, 2006 Gene expression and genetic mapping analyses of a perennial ryegrass glycine-rich RNA-binding protein gene suggest a role in cold adaptation. Molecular Genetics and Genomics **275:** 399-408.

SINGH, P., M. MERGOUM, T. ADHIKARI, T. SHAH, F. GHAVAMI *et al.*, 2010 Genetic and molecular analysis of wheat tan spot resistance effective against *Pyrenophora tritici-repentis* races 2 and 5. Molecular Breeding **25:** 369-379.

STOREY, J. D., 2002 A direct approach to false discovery rates. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **64:** 479-498.

STUDER, B., B. BOLLER, E. BAUER, U. POSSELT, F. WIDMER *et al.*, 2007 Consistent detection of QTLs for crown rust resistance in Italian ryegrass (*Lolium multiflorum* Lam.) across environments and phenotyping methods. Theoretical and Applied Genetics **115:** 9-17.

STUDER, B., B. BOLLER, D. HERRMANN, E. BAUER, U. POSSELT *et al.*, 2006 Genetic mapping reveals a single major QTL for bacterial wilt resistance in Italian ryegrass (*Lolium multiflorum* Lam.). Theoretical and Applied Genetics **113:** 661-671.

STUDER, B., R. KOLLIKER, H. MUYLLE, T. ASP, U. FREI *et al.*, 2010 EST-derived SSR markers used as anchor loci for the construction of a consensus linkage map in ryegrass (Lolium spp.). BMC Plant Biology **10:** 177.

THE INTERNATIONAL BRACHYPODIUM INITIATIVE, 2010 Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature **463:** 763-768.

TINKER, N., A. KILIAN, C. WIGHT, K. HELLER-USZYNSKA, P. WENZL *et al.*, 2009 New DArT markers for oat provide enhanced map coverage and global germplasm characterization. BMC Genomics **10:** 39.

TREVASKIS, B., M. N. HEMMING, E. S. DENNIS and W. J. PEACOCK, 2007 The molecular basis of vernalization-induced flowering in cereals. TRENDS in Plant Science **12:** 352-357.

VOORRIPS, R. E., 2002 MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. Journal of Heredity **93:** 77-78.

WENZL, P., H. LI, J. CARLING, M. ZHOU, H. RAMAN *et al.*, 2006 A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. BMC Genomics **7:** 206.

WITTENBERG, A., T. LEE, C. CAYLA, A. KILIAN, R. VISSER *et al.*, 2005 Validation of the high-throughput marker technology DArT using the model plant Arabidopsis thaliana. Molecular Genetics and Genomics **274:** 30-39.

ZHANG, X.-Y., C.-G. HU and J.-L. YAO, 2010 Tetraploidization of diploid Dioscorea results in activation of the antioxidant defense system and increased heat tolerance. Journal of Plant Physiology **167:** 88-94.

FIGURE LEGENDS

Figure 1
Genetic maps of seven *F. pratensis* and *L. multiflorum* chromosomes. DArTFest markers are red, and labeled with "D" followed by a clone number (six digits). Markers belonging to one redundant marker bin are indicated by the same superscript. Markers associated with freezing tolerance are labeled green.

Figure 2

Syntenic relationships between *F. pratensis*, *L. multiflorum*, and *B. distachyon* chromosome 4. Comparative map showing the syntenic relationships between the freezing tolerance QTL in *F. pratensis* and the DArT markers associated with freezing tolerance phenotype closely linked to VRN1 paralogs in the *B. distachyon* genome. The position of the DArT markers on Bd1 are extrapolated from the best blastn hit in the Bd genome. Red color text denotes bridging markers between maps. Green color text denotes DArT markers significantly associated with freezing tolerance phenotype flanking FUL1 and FUL2. For the Bd1 chromosome the left side numbers are physical position in base pair. Left side numbers for all other maps are genetic distance in cM. The rightmost map is reproduced from Alm *et al.* (submitted) where the black bar represent the 95% confidence interval for a freezing tolerance QTL.

TABLES

Table 1. Distribution of genetic markers among linkage groups

	Festuca pratensis		***Lolium multiflorum***	
	Non-DArT markers	DArT	Non-DArT markers	DArT
LG1	61	23 (27.4%)	63	67 (55.8%)
LG2	26	10 (27.8%)	43	73 (62.9%)
LG3	34	22 (39.3%)	49	73 (59.8%)
LG4	68	22 (24.4%)	61	106 (63.5)%
LG5	70	12 (14.6%)	43	71 (62.3%)
LG6	49	27 (35.5%)	35	63 (64.3%)
LG7	65	32 (33.0%)	68	76 (52.8%)
Total	**373**	**148 (28.4%)**	**352**	**529 (60.0%)**

Table 2. Redundancy of sequenced DArT markers

Extent of redundancy	**Distribution of DArT markers**	**Markers [%]**
unique	399	64.3
2 marker bins	64	20.6
3 marker bins	15	7.2
4 marker bins	8	5.2
5 marker bins	1	0.8
6 marker bins	2	1.9

Table 3. Synteny of *L. multiflorum* and rice as revealed by mapped DArT markers

	Lm1	**Lm2**	**Lm3**	**Lm4**	**Lm5**	**Lm6**	**Lm7**
Os01	1	1	15	1	1	1	-
Os02	2	-	1	3	2	16	-
Os03	-	1	-	18	-	-	2
Os04	2	10	2	2	2	1	-
Os05	7	-	1	1	-	-	-
Os06	-	-	-	2	-	-	15
Os07	-	7	-	1	-	2	-
Os08	-	2	-	-	-	-	7
Os09	-	-	-	-	7	1	-
Os10	7	-	-	-	1	1	-
Os11	1	-	1	5	1	2	1
Os12	-	-	1	-	3	-	1
Total	**20**	**21**	**21**	**33**	**17**	**24**	**26**

Note: Numbers correspond to non-redundant clusters mapped to each rice chromosome. Chromosome relationships with highest synteny are shaded.

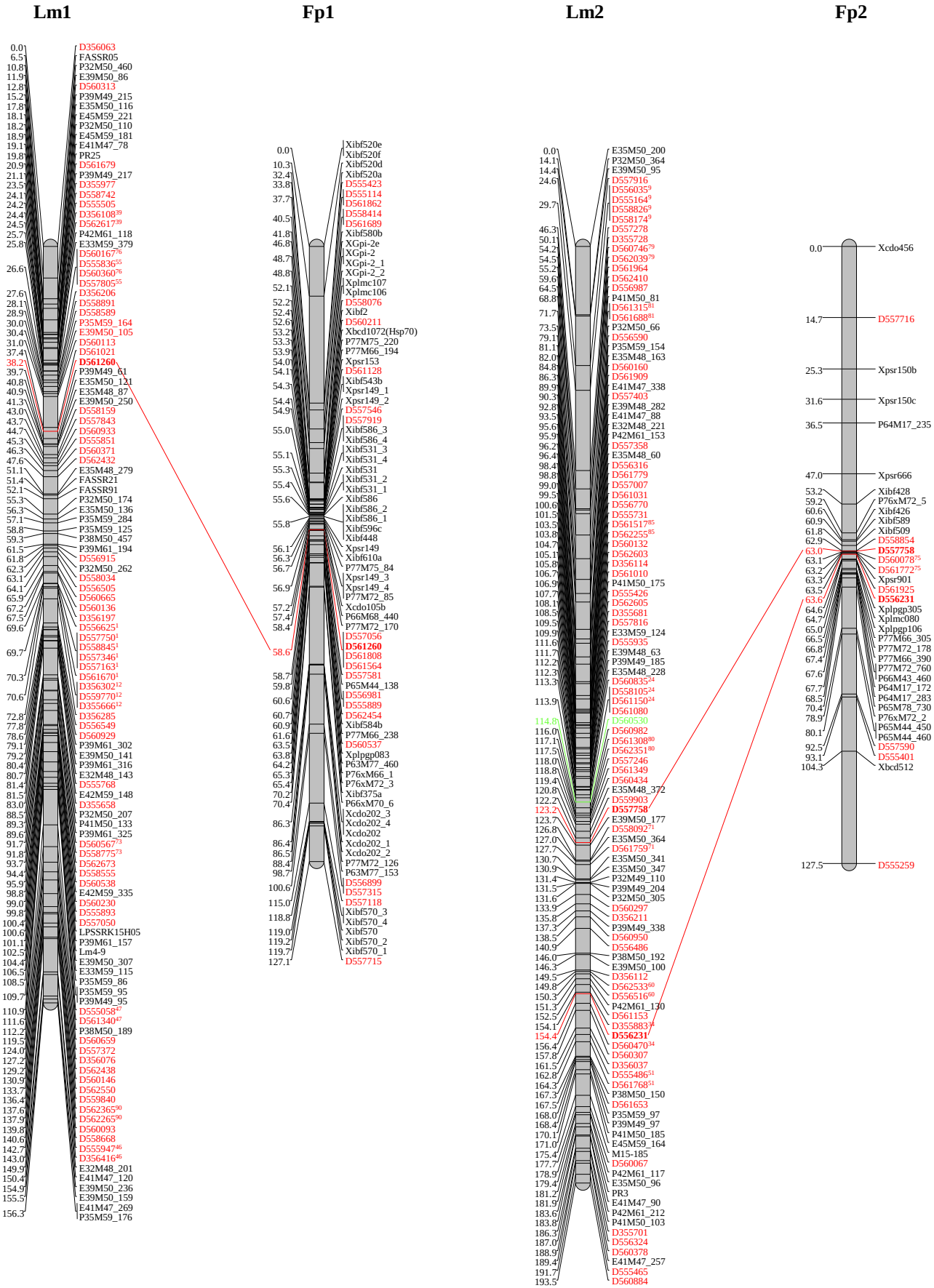
Table 4. Synteny of *L. multiflorum* and *B. distachyon* as revealed by mapped DArT markers

	Lm1	**Lm2**	**Lm3**	**Lm4**	**Lm5**	**Lm6**	**Lm7**
Bd01	1	11	3	30	1	3	22
Bd02	12	1	26	2	4	2	1
Bd03	10	2	1	4	1	20	8
Bd04	1	1	3	8	20	5	1
Bd05	1	15	1	1	1	1	2
Total	**25**	**30**	**34**	**45**	**26**	**31**	**34**

Note: Numbers correspond to non-redundant clusters mapped to each *B. distachyon* chromosome. Chromosome relationships with highest synteny are shaded.

Table 5. Mapped markers having significantly (Fisher Exact test) different distributions among HFT and LFT plants. "Presence" denotes if a hybridization signal is associated with the LFT or HFT group.

DArT ID	**q-value**	**Presence**	**LG**	**cM**
D560433	0,028	LFT	LG7 *F. pratensis*	67
D556159	0,042	HFT	LG4 *F. pratensis*	73
D558701	0,042	LFT	LG4 *L. multiflorum*	103
D562536	0,042	LFT	LG4 *L. multiflorum*	102
D560530	0,044	HFT	LG2 *L. multiflorum*	115


Lm1
Fp1
Lm2
Fp2
0.0 D356063
6.5 FASSR05
10.8 P32M50_460
11.9 E39M50_86
12.8 D560313
15.2 P39M49_215
17.8 E35M50_116
18.1 E45M59_221
18.2 P32M50_110
18.9 E45M59_181
19.1 E41M47_78
19.8 PR25
20.9 D561679
21.1 P39M49_217
23.5 D355977
24.1 D558742
24.2 D555505
24.4 D356108[39]
24.5 D562617[39]
25.7 P42M61_118
25.8 E33M59_379
26.6 D560167[76] D555836[55] D560360[76] D557805[55]
27.6 D356206
28.1 D558891
28.9 D558589
30.0 P35M59_164
30.4 E39M50_105
31.0 D560113
37.4 D561021
38.2 D561260
39.7 P39M49_61
40.8 E35M50_121
40.9 E35M48_87
41.3 E39M50_250
43.0 D558159
43.7 D557843
44.7 D560933
45.3 D555851
46.3 D560371
47.6 D562432
51.1 E35M48_279
51.4 FASSR21
52.1 FASSR91
55.3 P32M50_174
56.3 E35M50_136
57.1 P35M59_284
58.8 P35M59_125
59.3 P38M50_457
61.5 P39M61_194
61.8 D556915
62.3 P32M50_262
63.1 D558034
64.1 D556505
65.9 D560665
67.2 D560136
67.5 D356197
69.6 D556625[1] D557750[1] D558845[1] D557346[1] D557163[1]
70.3 D561670[1] D356302[12]
70.6 D559770[12] D355666[12]
72.8 D356285
77.8 D556549
78.6 D560929
79.1 P39M61_302
79.2 E39M50_141
80.4 P39M61_316
80.7 E32M48_143
81.4 D555768
81.5 E42M59_148
83.0 D355658
88.5 P32M50_207
89.3 P41M50_133
89.6 P39M61_325
91.7 D560567[73] D558775[73]
91.8 D562673
93.7 D558555
94.4 D560538
95.9 E42M59_335
98.8 D560230
99.0 D555893
99.8 D557050
100.4 LPSSRK15H05
100.6 P39M61_157
101.1 Lm4-9
102.5 E39M50_307
104.4 E33M59_115
106.5 P35M59_86
108.5 P35M59_95 P39M49_95
109.7 D555058[47]
110.9 D561340[47]
111.6 P38M50_189
112.2 D560659
119.5 D557372
124.0 D356076
127.2 D562438
129.2 D560146
130.9 D562550
133.7 D559840
136.4 D562365[90]
137.6 D562265[90]
137.9 D560093
139.8 D558668
140.6 D555947[46]
142.7 D356416[46]
143.0 E32M48_201
149.9 E41M47_120
150.4 E39M50_236
154.9 E39M50_159
155.5 E41M47_269
156.3 P35M59_176
0.0 Xibf520e
0.0 Xibf520f
10.3 Xibf520d
32.4 Xibf520a
33.8 D555423
37.7 D555114
40.5 D561862
41.8 D558414
46.8 D561689
48.7 Xibf580b
48.8 XGpi-2e
52.1 XGpi-2
52.2 XGpi-2_1
52.4 XGpi-2_2
52.6 Xplmc107
53.2 Xplmc106
53.3 D558076
53.9 Xibf2
54.0 D560211
54.1 Xbcd1072(Hsp70)
54.3 P77M75_220
54.4 P77M66_194
54.9 Xpsr153
55.0 D561128
55.1 Xibf543b
55.3 Xpsr149_1
55.4 Xpsr149_2
55.6 D557546
55.8 D557919
56.1 Xibf586_3
56.3 Xibf586_4
56.7 Xibf531_3
56.9 Xibf531_4
57.2 Xibf531
57.4 Xibf531_2
58.4 Xibf531_1
58.6 Xibf586
58.7 Xibf586_2
59.8 Xibf586_1
60.6 Xibf596c
60.7 Xibf448
60.9 Xpsr149
61.6 Xibf610a
63.5 P77M75_84
63.8 Xpsr149_3
64.2 Xpsr149_4
65.3 P77M72_85
65.4 Xcdo105b
70.2 P66M68_440
70.4 P77M72_170
86.3 D557056 D561260 D561808 D561564 D557581
86.4 P65M44_138
86.5 D556981
88.4 D555889
98.7 D562454
100.6 Xibf584b
115.0 P77M66_238
118.8 D560537
119.0 Xplpgp083
119.2 P63M77_460
119.7 P76xM66_1
127.1 P76xM72_3
Xibf375a
P66xM70_6
Xcdo202_3
Xcdo202_4
Xcdo202
Xcdo202_1
Xcdo202_2
P77M72_126
P63M77_153
D556899
D557315
D557118
Xibf570_3
Xibf570_4
Xibf570
Xibf570_2
Xibf570_1
D557715
0.0 E35M50_200
14.1 P32M50_364
14.4 E39M50_95
24.6 D557916
29.7 D556035[9] D555164[9] D558826[9] D558174[9]
46.3 D557278
50.1 D355728
54.2 D560746[79]
54.5 D562039[79]
55.2 D561964
59.6 D562410
64.5 D556987
68.8 P41M50_81
71.7 D561315[81] D561688[81]
73.5 P32M50_66
79.1 D556590
81.1 P35M59_154
82.0 E35M48_163
84.8 D560160
86.3 D561909
89.9 E41M47_338
90.3 D557403
92.8 E39M48_282
93.5 E41M47_88
95.6 E32M48_221
95.9 P42M61_153
96.2 D557358
96.4 E35M48_60
98.4 D556316
98.8 D561779
99.0 D557007
99.5 D561031
100.6 D556770
101.5 D555731
103.5 D561517[85]
103.8 D562255[85]
104.7 D560132
105.1 D562603
105.8 D356114
106.7 D561010
106.9 P41M50_175
107.7 D555426
108.1 D562605
108.5 D355681
109.5 D557816
109.9 E33M59_124
111.6 D555935
111.7 E39M48_63
112.2 P39M49_185
112.3 E35M48_228
113.3 D560835[24] D558105[24] D561150[24]
113.9 D561080
114.8 D560530
116.0 D560982
117.1 D561308[80]
117.5 D562351[80]
118.0 D557246
118.8 D561349
119.4 D560434
120.8 E35M48_372
122.2 D559903
123.2 D557758
123.7 E39M50_177
126.8 D558092[71]
127.0 E35M50_364
127.7 D561759[71]
130.7 E35M50_341
130.9 E35M50_347
131.4 P32M49_110
131.5 P39M49_204
131.6 P32M50_305
133.9 D560297
135.8 D356211
137.3 P39M49_338
138.5 D560950
140.9 D556486
146.0 P38M50_192
146.3 E39M50_100
149.5 D356112
149.8 D562533[60]
150.3 D556516[60]
151.3 P42M61_130
152.5 D561153
154.1 D355883[9]
154.4 D556231
156.4 D560470[34]
157.8 D560307
161.5 D356037
162.8 D555486[51]
164.3 D561768[51]
167.3 P38M50_150
167.5 D561653
168.0 P35M59_97
168.4 P39M49_97
170.1 P41M50_185
171.0 E45M59_164
175.4 M15-185
177.7 D560067
178.9 P42M61_117
179.4 E35M50_96
181.2 PR3
181.9 E41M47_90
183.6 P42M61_212
183.8 P41M50_103
186.3 D355701
187.0 D556324
188.9 D560378
189.4 E41M47_257
191.7 D555465
193.5 D560884
0.0 Xcdo456
14.7 D557716
25.3 Xpsr150b
31.6 Xpsr150c
36.5 P64M17_235
47.0 Xpsr666
53.2 Xibf428
59.2 P76xM72_5
60.6 Xibf426
60.9 Xibf589
61.8 Xibf509
62.9 D558854
63.0 D557758
63.1 D560078[75]
63.2 D561772[75]
63.3 Xpsr901
63.5 D561925
63.6 D556231
64.6 Xplpgp305
64.7 Xplmc080
65.0 Xplpgp106
66.5 P77M66_305
66.8 P77M72_178
67.4 P77M66_390
67.6 P77M72_760
67.7 P66M43_460
68.5 P64M17_172
70.4 P64M17_283
78.9 P65M78_730
80.1 P76xM72_2
92.5 P65M44_450
93.1 P65M44_460
104.3 D557590 D555401 Xbcd512
127.5 D555259

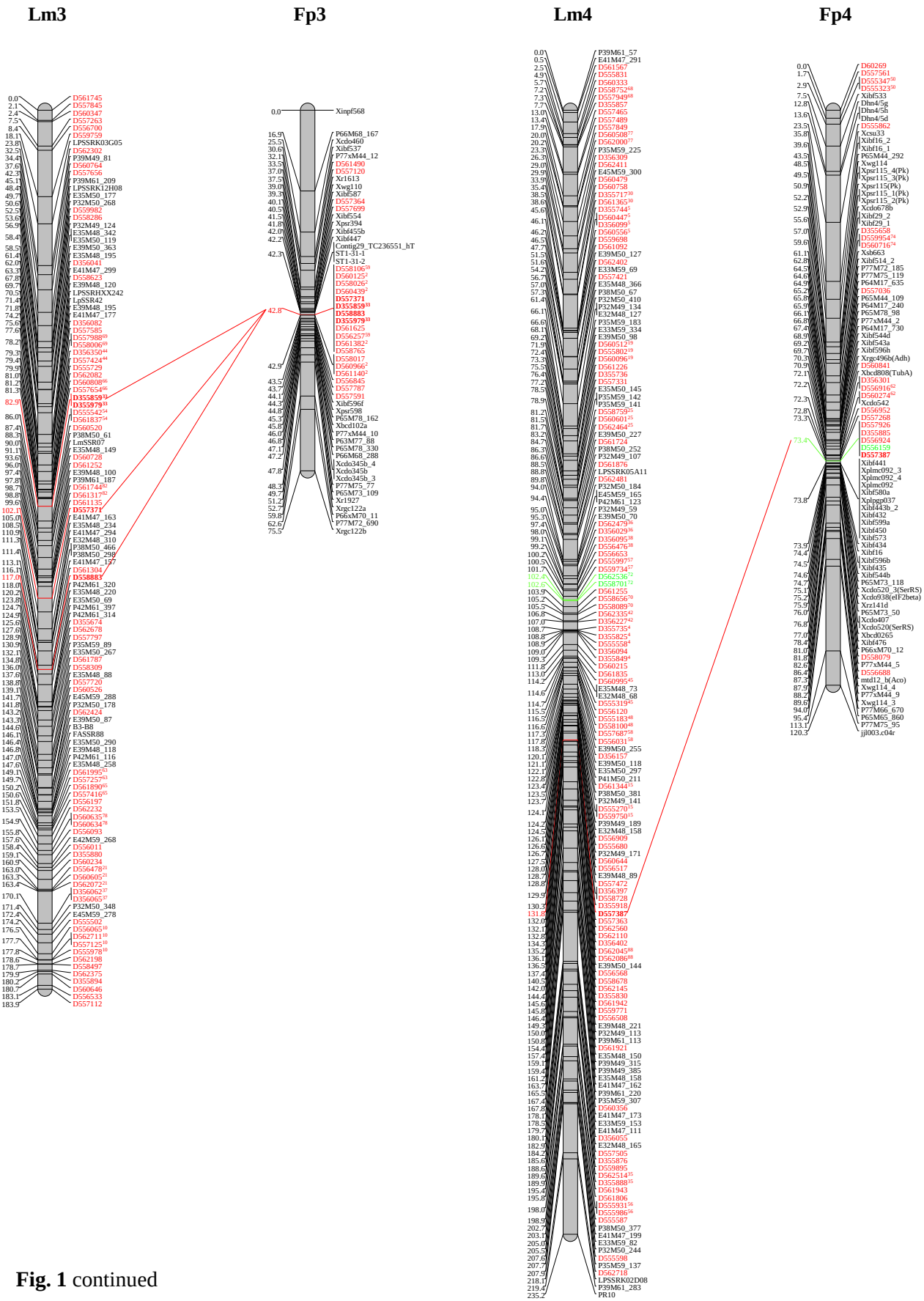


Fig. 1 continued

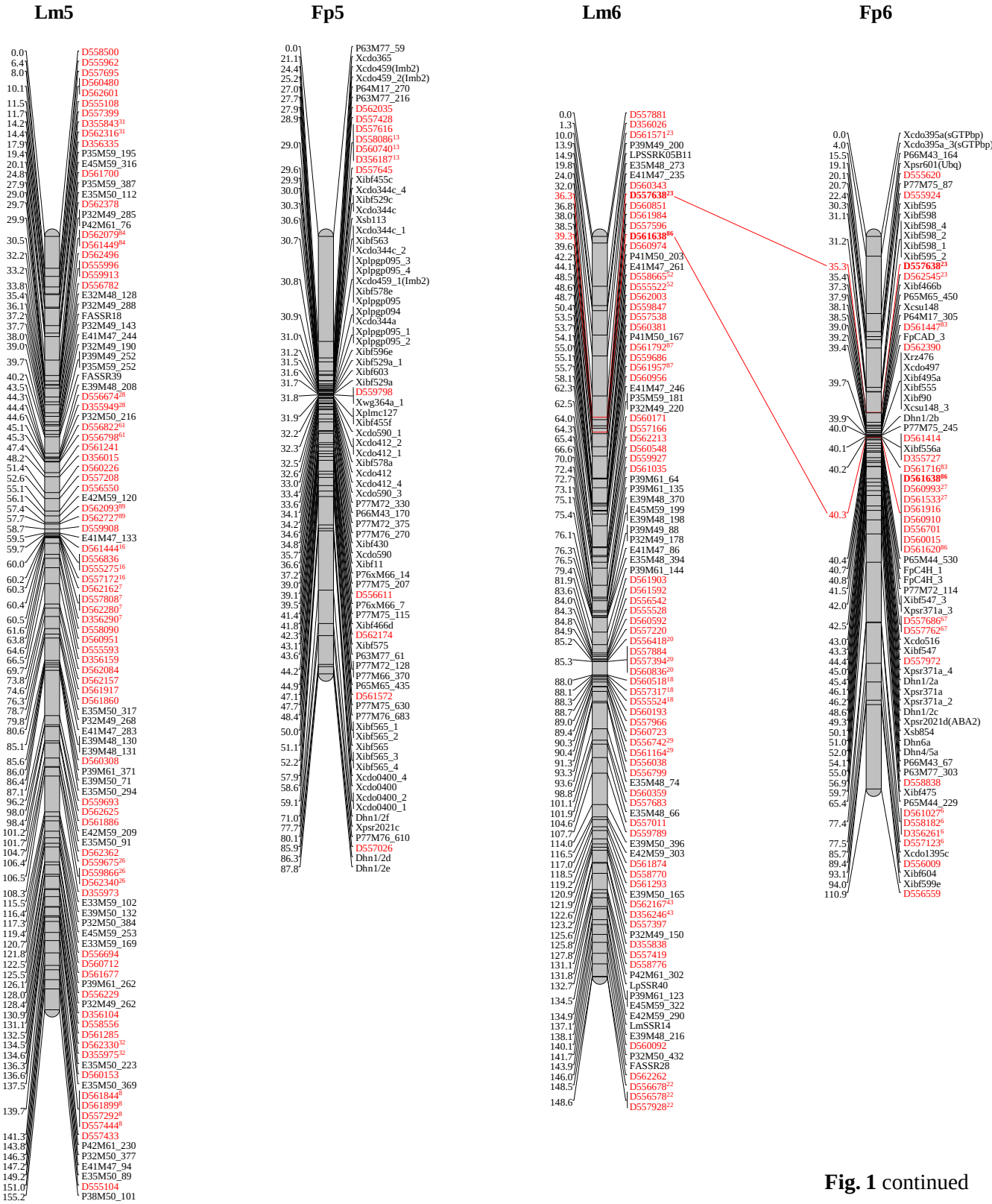


Fig. 1 continued

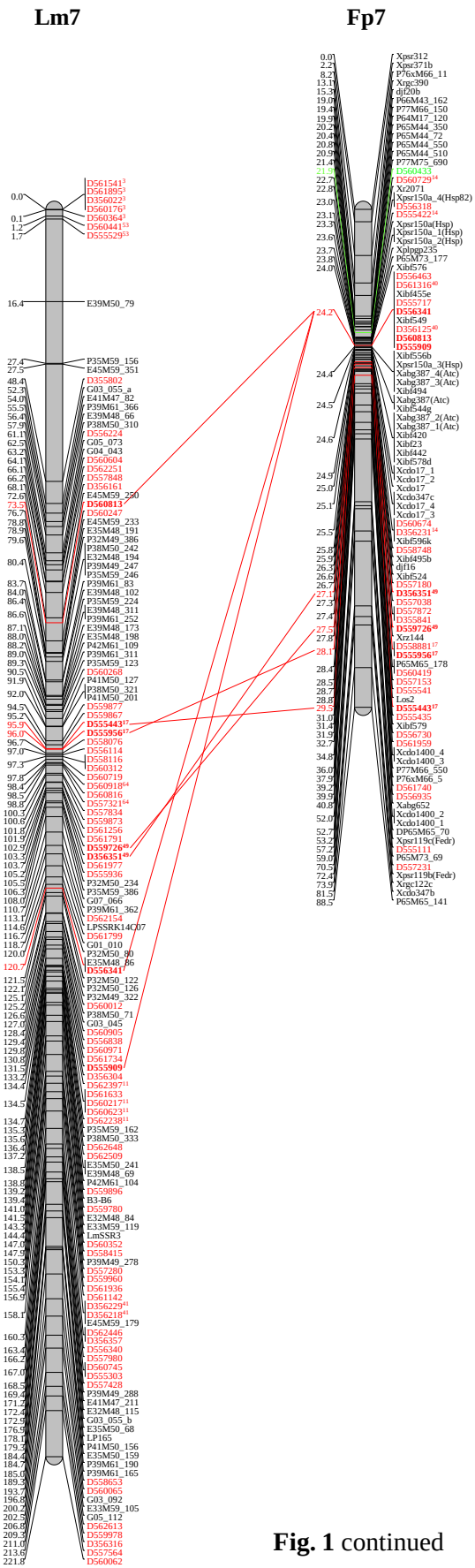
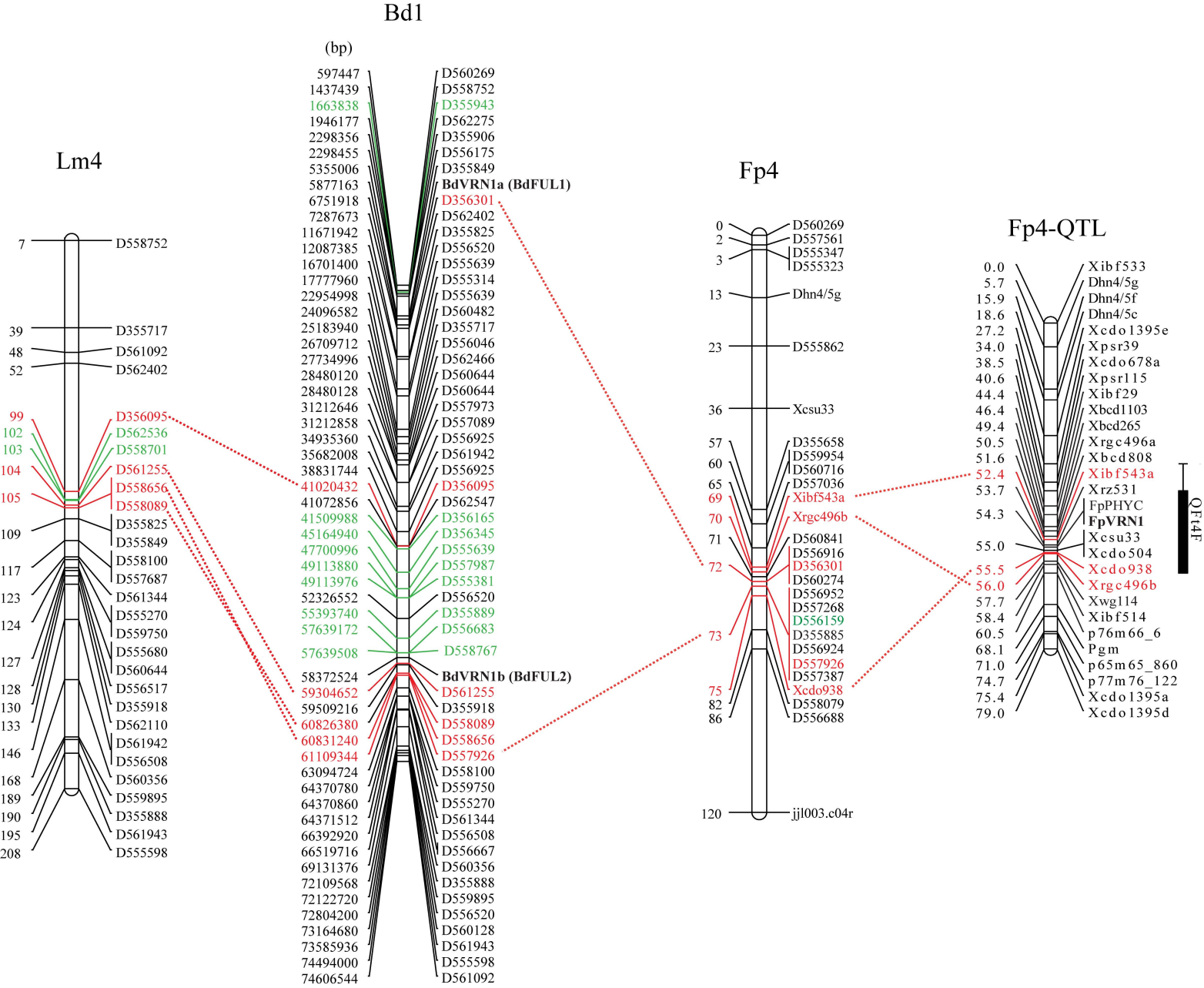


Fig. 1 continued


Lm4
Bd1
(bp)
Fp4
Fp4-QTL
7 D558752
39 D355717
48 D561092
52 D562402
99 D356095
102 D562536
103 D558701
104 D561255
105 D558656
D558089
109 D355825
D355849
117 D558100
D557687
123 D561344
124 D555270
D559750
127 D555680
D560644
128 D556517
130 D355918
133 D562110
146 D561942
D556508
168 D560356
189 D559895
190 D355888
195 D561943
208 D555598
597447 D560269
1437439 D558752
1663838 D355943
1946177 D562275
2298356 D355906
2298455 D556175
5355006 D355849
5877163 BdVRN1a (BdFUL1)
6751918 D356301
7287673 D562402
11671942 D355825
12087385 D556520
16701400 D555639
17777960 D555314
22954998 D555639
24096582 D560482
25183940 D355717
26709712 D556046
27734996 D562466
28480120 D560644
28480128 D560644
31212646 D557973
31212858 D557089
34935360 D556925
35682008 D561942
38831744 D556925
41020432 D356095
41072856 D562547
41509988 D356165
45164940 D356345
47700996 D555639
49113880 D557987
49113976 D555381
52326552 D556520
55393740 D355889
57639172 D556683
57639508 D558767
58372524 BdVRN1b (BdFUL2)
59304652 D561255
59509216 D355918
60826380 D558089
60831240 D558656
61109344 D557926
63094724 D558100
64370780 D559750
64370860 D555270
64371512 D561344
66392920 D556508
66519716 D556667
69131376 D560356
72109568 D355888
72122720 D559895
72804200 D556520
73164680 D560128
73585936 D561943
74494000 D555598
74606544 D561092
0 D560269
2 D557561
3 D555347
D555323
13 Dhn4/5g
23 D555862
36 Xcsu33
57 D355658
60 D559954
65 D560716
69 D557036
Xibf543a
70 Xrgc496b
71 D560841
72 D556916
D356301
D560274
D556952
D557268
D556159
73 D355885
D556924
D557926
D557387
75 Xcdo938
82 D558079
86 D556688
120 jjl003.c04r
0.0 Xibf533
5.7 Dhn4/5g
15.9 Dhn4/5f
18.6 Dhn4/5c
27.2 Xcdo1395e
34.0 Xpsr39
38.5 Xcdo678a
40.6 Xpsr115
44.4 Xibf29
46.4 Xbcd1103
49.4 Xbcd265
50.5 Xrgc496a
51.6 Xbcd808
52.4 Xibf543a
53.7 Xrz531
54.3 FpPHYC
FpVRN1
55.0 Xcsu33
Xcdo504
55.5 Xcdo938
56.0 Xrgc496b
57.7 Xwg114
58.4 Xibf514
60.5 p76m66_6
68.1 Pgm
71.0 p65m65_860
74.7 p77m76_122
75.4 Xcdo1395a
79.0 Xcdo1395d
QFt4F

Paper III

SANDVE, S., H. RUDI, T. ASP and O.A. ROGNLI, 2008. **Tracking the evolution of a cold stress associated gene family in cold tolerant grasses.** BMC Evolutionary Biology **8:** 245.

BMC Evolutionary Biology

BioMed Central

Research article

Tracking the evolution of a cold stress associated gene family in cold tolerant grasses

Simen R Sandve[1], Heidi Rudi[1], Torben Asp[2] and Odd Arne Rognli*[1]

Address: [1]Department of Plant and Environmental Sciences, University of Life Sciences, Ås, Norway and [2]Department of Genetics and Biotechnology, University of Aarhus, Forsøgsvej 1, 4200 Slagelse, Denmark

Email: Simen R Sandve - simen.sandve@umb.no; Heidi Rudi - heidi.rudi@umb.no; Torben Asp - Torben.Asp@agrsci.dk; Odd Arne Rognli* - odd-arne.rognli@umb.no

* Corresponding author

Published: 5 September 2008

BMC Evolutionary Biology 2008, **8**:245 doi:10.1186/1471-2148-8-245

Received: 30 April 2008
Accepted: 5 September 2008

This article is available from: http://www.biomedcentral.com/1471-2148/8/245

Abstract

Background: Grasses are adapted to a wide range of climatic conditions. Species of the subfamily Pooideae, which includes wheat, barley and important forage grasses, have evolved extreme frost tolerance. A class of ice binding proteins that inhibit ice re-crystallisation, specific to the Pooideae subfamily lineage, have been identified in perennial ryegrass and wheat, and these proteins are thought to have evolved from a leucine-rich repeat phytosulfokine receptor kinase (*LRR-PSR*)-like ancestor gene. Even though the ice re-crystallisation inhibition function of these proteins has been studied extensively *in vitro*, little is known about the evolution of these genes on the molecular level.

Results: We identified 15 putative novel ice re-crystallisation inhibition (IRI)-like protein coding genes in perennial ryegrass, barley, and wheat. Using synonymous divergence estimates we reconstructed the evolution of the IRI-like gene family. We also explored the hypothesis that the IRI-domain has evolved through repeated motif expansion and investigated the evolutionary relationship between a LRR-domain containing IRI coding gene in carrot and the Pooideae IRI-like genes. Our analysis showed that the main expansion of the IRI-gene family happened ~36 million years ago (Mya). In addition to IRI-like paralogs, wheat contained several sequences that likely were products of polyploidisation events (homoeologs). Through sequence analysis we identified two short motifs in the rice *LRR-PSR* gene highly similar to the repeat motifs of the IRI-domain in cold tolerant grasses. Finally we show that the LRR-domain of carrot and grass IRI proteins both share homology to an *Arabidopsis thaliana* LRR-trans membrane protein kinase (*LRR-TPK*).

Conclusion: The diverse IRI-like genes identified in this study tell a tale of a complex evolutionary history including birth of an ice binding domain, a burst of gene duplication events after cold tolerant grasses radiated from rice, protein domain structure differentiation between paralogs, and sub- and/or neofunctionalisation of IRI-like proteins. From our sequence analysis we provide evidence for IRI-domain evolution probably occurring through increased copy number of a repeated motif. Finally, we discuss the possibility of parallel evolution of LRR domain containing IRI proteins in carrot and grasses through two completely different molecular adaptations.

Background

The Poaceae family (grasses) contains some of the most economically important and well studied plant species, e.g. maize, wheat, barley, and rice. Generally speaking the Pooideae subfamily, which includes wheat, barley and forage grasses, are adapted to cold seasons. Many species in this subfamily can withstand temperatures far below freezing and intercellular ice formation [1,2]. Rice and maize on the other hand belongs to the subfamilies Ehrhartoideae and Panicoideae, respectively, and are adapted to warm and tropical climates. Pooideae lineage (from now on referred to as cold tolerant grasses) adaptation to cold climates makes grasses an interesting model system for studying climatic adaptation at the physiological and molecular level.

Frost tolerance adaptations are in many organisms associated with the evolution of antifreeze proteins (AFPs) [3]. AFPs can affect freezing- and ice crystallisation related stress via different mechanisms. Thermal hysteresis (TH) depresses the freezing point at which ice crystallisation initiates, which render it possible for organisms to survive under freezing temperatures. Ice re-crystallisation inhibition (IRI) on the other hand does not hinder ice crystallisation but manipulates the growth of the ice crystals such that small ice crystals grow at the expense of larger ice crystals, and this has been suggested to prevent or minimize the cellular damage in plants [4]. A third mode of AFP action is membrane stabilisation which has been reported for a fish AFP [5]. Animal AFPs generally possess high thermal hysteresis (TH) characteristics and lower ice crystallisation initiation temperature by 1–5°C [6,7]. Plant AFPs on the other hand have low TH-activity, but exhibits strong ice re-crystallisation inhibition (IRI) activity [6].

Genes encoding peptides with IRI capacity have evolved independently several times in different lineages of higher plants. These IRI peptides are homologous to diverse protein classes, e.g. thaumatin like proteins, endochitinases, endo-B-1,3-glucanase, and leucine rich repeat (LRR) containing proteins [6,8,9]. Three LRR-domain containing IRI proteins (LRR-IRI) have been identified in plants, one in carrot (DcAFP; accession number AAC6293) and two in wheat (TaIRI1 and TaIRI2 with accession numbers AAX81542 and AAX81543) [10,11]. DcAFP has been classified as a polygalacturonase-inhibiting protein (PGIP) but does not exhibit PGIP activity [12]. LRR motifs span across the entire processed DcAFP protein and form 10-loop beta-helix secondary structure with solvent exposed asparagine residues at putative ice binding sites [13]. *TaIRI1* and *TaIRI2* genes (accession numbers AY9968588 and AY968589) have been identified as homologous to the LRR-domain coding region of a rice phytosulfokine LRR receptor kinase (*OsLRR-PSR*: NP_001058711) and an *Arabidopsis* trans-membrane protein kinase (*AtLRR-TPK*: NP_200200). The wheat IRI peptides differ structurally from DcAFP in that the LRR-domain only comprises about half of the processed peptide [10].

In addition to the N-terminal LLR domain, wheat IRI proteins have a C-terminal repeat domain consisting of two similar A and B motifs, NxVxG and NxVxxG, respectively. This repeat domain has been reported to exhibit strong *in vitro* IRI capacity [14]. Interestingly, blast search yields no sequences with homology to the IRI-domain outside the subfamily of cold tolerant grasses [10]. Protein modelling has shown that the A and B repeated motifs of the IRI-domain folds into a B-roll with ice binding sites matching the prism face of ice [15]. Expression studies have shown that increased expression levels in wheat [10] and perennial ryegrass [Rudi et al, unpublished] are correlated to cold acclimation, but no *in vivo* studies to determine the localisation of these grass IRI peptides have been reported in the literature. However, *TaIRI1* and *TaIRI2* have been predicted to encode a N-terminal 20 amino acid signal peptide domain targeting the proteins to the secretory pathway, suggesting that the peptides could be located in the extracellular space [10].

While *DcAFP* is evolutionarily closely related to PGIPs; *TaIRI1* and *TaIRI2* genes are thought to have evolved from a LRR-PSR like ancestor gene. Furthermore, the evolutionary origin of the IRI-domain in grass IRI genes is much less obvious, because the IRI-domain is not homologous to any other sequences outside the cold tolerant grass lineage. Tremblay et al. [10] proposed a "TE-hypothesis" to explain this apparent lack of homologous coding regions; that the IRI-domain had arisen by a transposable element (TE) insertion. However, no TE signature sequence could be identified surrounding the IRI-domain [10], thus no empirical data supports the TE-hypothesis so far.

Here we report the identification and characterisation of novel LRR-IRI homologous genes in cold tolerant grass species. We perform a detailed study of the evolutionary relationships between *OsLRR-PSR* and IRI-like genes by analysing sequence divergence at synonymous sites. We also use synonymous site divergence to trace the evolutionary history of the IRI-like gene family with respect to gene duplication events. The evolution of gene families *per se* is in itself a much debated topic, and gene family expansion and subsequent functional diversification is thought to have been a significant factor contributing to adaptations to new environments [16,17]. The evolution of the IRI-like gene family of cold tolerant grasses is discussed in the context of the Duplication-Degeneration-Complementation (DDC) model [18]. Finally we address the unresolved matter of the evolutionary mechanism underlying the birth of the cold tolerant grass IRI-domain,

and propose a novel hypothesis on the evolution of this IRI-domain.

Results

Screening of perennial ryegrass BAC libraries

Initial screening of the perennial ryegrass (*Lolium perenne*) LTS18 BAC library with the *LpAFP* primer pair produced two hits, from which *LpIRI1* (EU680848) and *LpIRI2* (EU680849) were isolated. The NV#20F1-30 BAC library produced four hits with the LpIRIx primer pair, and three hits with the LpAFP primer pair. *LpIRI4* (EU680851) was subsequently isolated from one of the four positive LpIRIx hits and *LpIRI3* (EU680850) was isolated from one of three positive LpAFP hits. All genes isolated from perennial ryegrass were intronless and encoded putative peptides with high identity to the wheat *TaIRI1* and *TaIRI2* genes (blastp < 4e-10). *LpIRI1*, *LpIRI4*, and *LpIRI3* were similar in size and encoded peptides of 285, 242, and 254 amino acids, respectively. *LpIRI2* encoded a shorter ORF of 150 amino acids that was 94% identical to the LpIRI4 IRI-domain. The IRI-domain of LpIRI3 is identical to an earlier identified partial IRI peptide encoded by *LpAFP* (AJ277399).

Nucleotide alignments of the LpIRI-like genes showed that *LpIRI2* has undergone a deletion of almost the entire LRR-domain coding region, the only remains of it being a 102 base pair (bp) fragment upstream of the *LpIRI2* putative start codon. This could indicate that *LpIRI2* is a pseudogene or a non-functional allele. Non-functional sequences are expected to evolve under neutral expectation, which means that the rate of non-synonymous to synonymous substitutions (w) is expected to be 1. Average w between *LpIRI2* and the other perennial ryegrass sequences was estimated to be 0.56 which suggests that *LpIRI2* is under selective constraints despite the major deletion in the LRR-domain.

In silico *identification of IRI-like sequences*

The blastn EST search resulted in 189 wheat, 100 barley, 21 tall fescue (*F. arundinacea*), 5 Italian ryegrass (*L. multiflorum*) and 2 darnel ryegrass (*L. temulentum*) sequences. *In silico* mining produced from zero to eleven full length IRI-like sequences (i.e. with a start and stop codon) per species (Table 1), and sequences were annotated as follows; an initial two letters indicating latin species name, "C" indicating a contig of more than two ESTs, and lastly an identifier number. The number of ESTs in the contigs ranged from 2 to 40 (Table 2), with an average of 17 ESTs per contig. In addition we identified several partial IRI-like sequences; one partial wheat IRI-like contig of 10 ESTs, and three partial barley sequences of 2, 3, and 6 EST sequences. The partial sequences were not included in further analysis. A barley mRNA, AK249041, did not align to any barley contig but coded for a full length IRI-like ORF, hence we included this mRNA in our dataset. *In silico* mining with tall fescue, Italian ryegrass, and darnel ryegrass ESTs did not produce any full length IRI-like sequences. *In silico* EST mined sequences were considered validated if a core nucleotide sequence from NCBI had > 99% identity over > 200 bp. TaC3 was validated as being identical to *TaIRI1*, and HvC1 was validated by being identical to a full length mRNA sequence (AK252915). Accession numbers for ESTs belonging to full length IRI-like contigs acquired through *in silico* mining are included in additional material [see Additional file 1].

Table 1: All IRI-like sequences identified through EST *in silico* mining.

Species	No. EST sequences	IRI-like sequences	
		Full length	Partial
Wheat	189	11	1
Barley	100	3	3
Tall fescue	21	0	2
Italian ryegrass	5	0	1
Darnel ryegrass	2	0	1

IRI-like gene homologous ESTs were identified by a blastn search using *TaIRI1*.

Predicted protein structure characterisation

The *OsLRR-PSR* peptide sequence was used as a template for comparisons of the LRR-domains. The *OsLRR-PSR* is suitable for this purpose because it is a putative homologue to the LRR-domain of the IRI-like sequences. In the following peptide structure characterization we assume that the most recent common ancestor (MRCA) of all IRI-like grass genes encoded the same domain architecture as *OsLRR-PSR*. *OsLRR-PSR* was predicted to encode one LRR N-terminal domain (LRR-NT), 21 internal LRR homologous motifs (1–21), and a protein kinase (PK) C-terminal

Table 2: Full length IRI-like sequences identified through EST *in silico* mining and the number of ESTs per contig.

Species	Contig name	ESTs in contig
Wheat	TaC1	16
	TaC2	39
	TaC3	19
	TaC4	21
	TaC5	16
	TaC6	12
	TaC7	16
	TaC8	17
	TaC9	3
	TaC10	4
	TaC11	2
Barley	HvC1	23
	HvC2	40
	HvC3	15

domain. Only the LRR-NT and six internal LRR motifs were predicted with significant E-values by Pfam. Compared to *OsLRR-PSR* all predicted IRI-like peptides have reduced number of LRR motifs and lacks the PK-domain. Three internal LRR motifs did not align to any OsLRR-PSR LRR motifs. These have probably arisen through deletions in the LRR-domain causing two partial LRR motifs to merge into a novel LRR motif. Comparative sequence analysis suggests that deletions between LRR motifs 1–16, 2–16, and 12–16 of a *OsLRR-PSR*-like grass ancestor gene have resulted in three novel LRR-domains; LRR1b, 2b, and 16b (Fig. 1).

The IRI-domain also varies in size by number of repeated motifs (Fig. 1). About 60% of all sequences with an IRI-domain have 15 repeat motifs or more. Six sequences were detected to have a reduced number of repeat motifs, or had completely lost the IRI-domain. Analysis of codon based nucleotide alignments revealed that frameshift (FS) mutations could be identified in four of the IRI-like sequences (TaC3, TaC10, TaC11, and AK249041) that showed reduced IRI-domain size (data not shown). HvC3 is the only IRI-like sequence with a completely reduced IRI-domain. For all sequences in the HvC3 contig additional information on abiotic conditions under which the plants had been grown were included in the EST files. Without exception all ESTs originated from tissue sampled from etiolated barley seedlings, and not from cold acclimated tissue. This is congruent with HvC3 lacking the entire ice binding IRI-domain, suggesting that IRI-like paralogs are involved in several different stress responses.

Prediction of the subcellular location of the IRI-like peptides (see methods) predicted a signal peptide that targets the peptides to the secretory pathway present in all IRI-like peptides, except from LpIRI2. The lack of an LpIRI2 signal peptide, and the fact that LpIRI2 has undergone a deletion of almost the entire LRR-domain could suggest that *LpIRI2* is in fact a non-functional allele or pseudogene. However the results from the w estimates contradict the non-functionality hypothesis. Alternatively the lack of a signal peptide can be interpreted as that LpIRI2 simply has evolved a different function than the IRI-like peptides with a conserved signal peptide.

Phylogenetic analysis of IRI-like paralogs

OsLRR-PSR has highest homology to IRI-like genes outside the Pooideae subfamily (blastp E-value of $2e^{-46}$) and was therefore used as an out-group in the phylogenetic analysis. The paralog phylogenies of perennial ryegrass, barley, and wheat IRI-like peptide sequences are given in Figure 2A, 2B, and 2C, respectively. In the wheat phylogeny TaIRI2, the amino acid sequence with the highest number of conserved LRR-domains, diverge as a monophyletic branch, while all other wheat sequences with further reduced LRR-domain sizes form a second monophyletic group. This large monophyletic group can be subdivided into two smaller monophyletic clades (Fig. 2C). In group I a FS mutation in the IRI-domain of TaC3 shortens the ORF and separates TaC3 structurally from the other sequences. One might be tempted to speculate if this FS is due to alignment errors but the FS in the TaC3 contig have been validated independently by the full length mRNA *TaIRI1*. Group II contain IRI-like genes with high degree of divergent predicted peptide structure, i.e. putative peptides with strongly reduced and completely lost IRI-domain (TaC9 and TaC11) as well as predicted peptides with large IRI-domains. Alignments of predicted amino acid sequences with all FS mutations removed, that were used for the phylogenetic analysis, are presented in additional material [see Additional file 2, 3, and 4].

Estimation of synonymous divergence of IRI-like sequences

If we assume a molecular clock, synonymous substitution rates (dS) between two DNA sequences can be interpreted as a relative measurement of time since MRCA, thus for two paralogous genes dS can be interpreted as the time since gene duplication [19]. Without being able to account for all the IRI-like paralogs existing in a genome we cannot infer if two paralogs descend from a single duplication event (i.e. being true paralogs) or if they are products of two separate duplication events. We therefore restricted our initial analysis of IRI-like gene duplication events to only comprise the dS_{max} and dS_{min} for all pairwise comparisons. The dS_{max}-dS_{min} range can be interpreted as the evolutionary time span in which all duplications of IRI-like genes in our dataset have occurred.

The average dS between *OsLRR-PSR* and LRR-domains of all IRI-like sequences from cold tolerant grasses was estimated to 0.97 (standard deviation (SD) = 0.11). Hence if *OsLRR-PSR* is the true ortholog of IRI-like genes, the radiation of cold tolerant grasses from rice predates the initiation of IRI-like gene family expansion in our dataset (Fig. 3A). All paralog dS ranges for wheat, barley and perennial ryegrass overlapped, however there are large range differences. These range differences is caused by the low dS_{min} range of wheat and perennial ryegrass IRI-like sequences (Fig. 3A). For perennial ryegrass the dS between *LpIRI2* and *LpIRI4* lowered the dS_{min} estimate from 0.40 to 0.03. Dissection of wheat paralog dS estimates showed that the low wheat dS_{min} range could be traced back to low pairwise dS within the monophyletic clades I and II (Fig. 2C). If dS estimates from sequence pairs within clade I and II were removed the wheat dS_{min} shifted from 0.01 to 0.39.

Very low dS between paralog pairs might reflect the inclusion of highly diverged alleles in our paralog dataset. Sequencing errors in ESTs and inclusion of highly

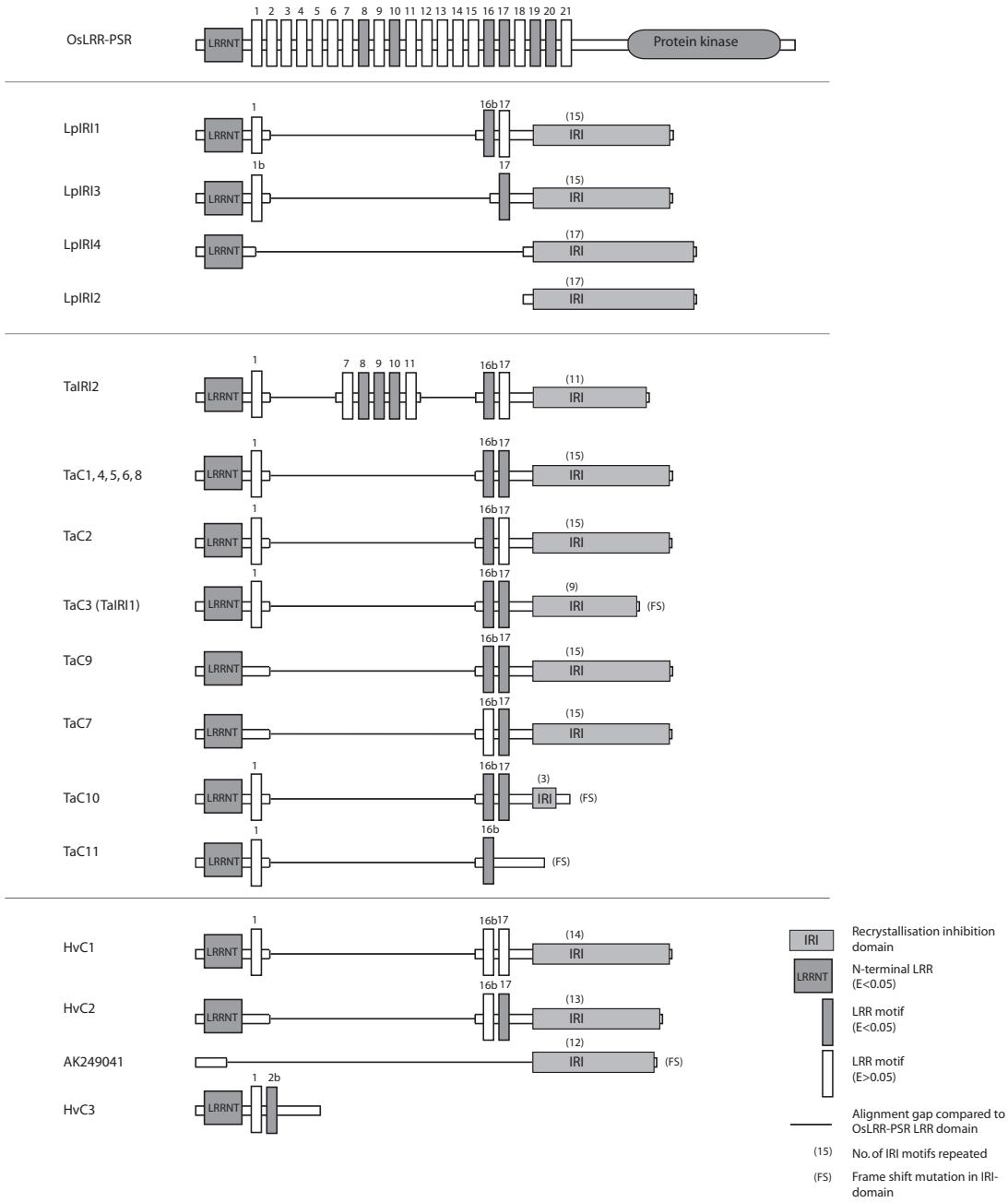


Figure 1

Predicted peptide structure of *OsLRR-PSR* and IRI-like sequences. Peptide structures of the LRR and IRI-domains of IRI-like sequences as predicted by Pfam and by visual characterisation, respectively. *OsLRR-PSR* was used as a template for peptide structure comparisons. Grey LRR-domains were predicted with significant P-values in Pfam ($P < 0.05$). Numbers of IRI motifs in the IRI-domains and whether frame shift mutations were identified are given in parentheses.

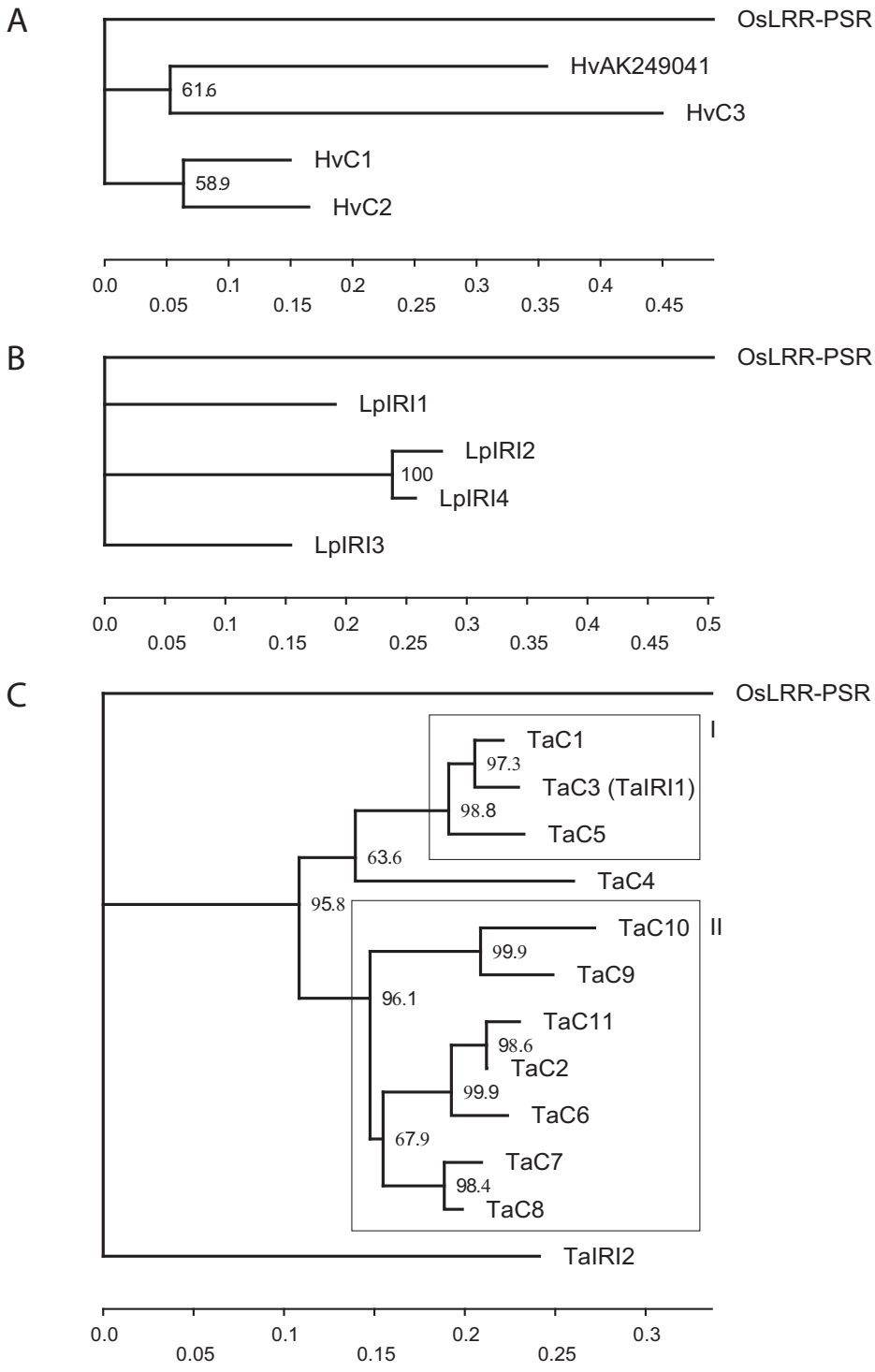


Figure 2

Maximum Likelihood paralog phylogenies of all full length IRI-like amino acid sequences. Bootstrap values from 1000 replicates are in percent at internal nodes. Bootstrap values < 50 are not shown. A) Phylogeny of barley IRI-like full length sequences. B) Phylogeny of perennial ryegrass IRI-like full length sequences. C) Phylogeny of wheat IRI-like full length sequences. Two monophyletic clades of very low within-clade synonymous divergence (I, II) are boxed.

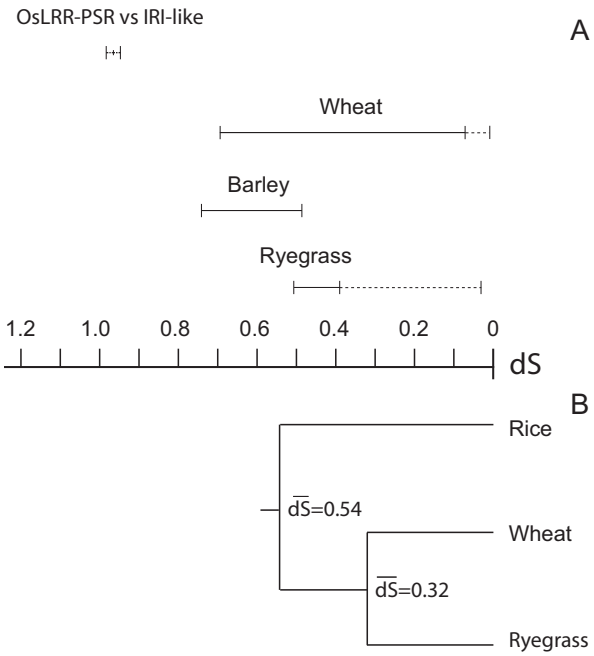


Figure 3

Synonymous divergence between paralogous and orthologous sequences. A) Maximum and minimum dS range between all IRI-like paralogous sequences of wheat, barley and perennial ryegrass and the average dS between *OsLRR-PSR* and all IRI-like sequences. Dotted lines for the divergence estimate between OsLRR-PSR and IRI-like sequences represent SD of the ortholog dS. Dotted lines for wheat and perennial ryegrass indicate changes in dS range if putative alleles are included. B) A tree indicating the average synonymous divergence, given in dS, estimated from ten control genes from rice, wheat, and perennial ryegrass.

diverged genotypes in our dataset could potentially give an inflated polymorphism level producing artificial contigs that are alleles rather than paralogs. To identify putative false paralogs we set an allelic dS threshold of dS < 0.03 (see methods section). Based on this definition we identified two putative allelic sequence pairs TaC2-TaC11 (dS = 0.01) and *LpIRI2-LpIRI4* (dS = 0.03).

Estimation of synonymous divergence between control genes

The ten control orthologs chosen for evolutionary rate control and their pairwise dS estimates are listed in Table 3. Based on these genes the average dS between rice and cold tolerant grasses were estimated to be 0.54 with an average SD between dS estimation methods of 0.08. The estimated divergence of rice and cold tolerant grasses based on the control genes (dS = 0.54) is substantially lower than the minimum equivalent estimate found by using *OsLRR-PSR* and IRI-like genes (dS = 0.72) (Fig. 3A and 3B). The average dS between perennial ryegrass and wheat genes were 0.32, with an average SD of 0.04 between the estimation methods.

Molecular analysis of LRR and IRI-domains

The alignment of the LRR-domain of *OsLRR-PSR* and the IRI-domains of a subset of the IRI-like peptide sequences shows two blocks of strikingly high conservation between the OsLRR-PSR and two IRI-domain motifs (Fig. 4). The two motifs have 2/3 of the motif sequence residues conserved between rice and cold tolerant grasses. The first motif has an Ile-Val substitution while the second motif has a Leu-Val substitution, neither of which are substitutions with large effect on hydrophilic properties.

Relatively low level of conservation between DcAFP and grass IRI-like peptide sequences was found (blastp 1e-4). All sequences in the LRR-domain alignment (Fig. 5) have a blast E-value of < 1e-18 to at least one other sequence in the alignment, but no larger blocks of conservation between DcAFP and any monocot LRR-domain can be identified. However we observe that several blocks of 2–4 conserved residues exist throughout the alignment of the LRR-domain between DcAFP and grass IRI-sequences.

Discussion

Until now only three cold tolerant grass IRI protein coding genes have been reported; a partial coding sequence of an IRI-domain from perennial ryegrass [14], and two highly identical full length mRNA paralogs from wheat [10]. Through *in silico* mining and BAC sequencing we have identified 15 full length and 8 partial novel IRI-like genes in cold tolerant grasses. In addition, we have obtained the complete sequence of *LpAFP*. The data accumulated leaves no doubt: cold tolerant grasses of the Pooideae subfamily have evolved a lineage specific family of IRI-like genes.

IRI-gene family radiation happened after the cold tolerant grass divergence

The prevailing hypothesis on the evolution of LRR-IRI-like genes belonging to cold tolerant grasses is that they are lineage specific and that an *OsLRR-PSR*-like gene is the MRCA [10]. This hypothesis was proposed based on sequence homology data only and we therefore re-examined this idea using more rigorous statistical methods by estimating synonymous divergence. When employing a commonly used mutation rate for grasses of $6.5*10^{-9}$ [20-22], estimated by Gaut et al. [23], the synonymous divergence level between *OsLRR-PSR* and IRI-like sequences suggested a MRCA about 75 Mya. This is slightly higher than upper thresholds of some published rice-Pooideae divergence estimates [22]. However, our estimate of divergence time between rice and cold tolerant grasses based on the control genes suggest a rice-Pooideae divergence

only 42 Mya. This is similar to divergence estimates published by Patterson et al. [24] and Salse et al. [20], dating back 41–47 and ~46 Mya, respectively. Our wheat and perennial ryegrass divergence estimates is dated ~10 million years prior to a previously published estimate of ~35 Mya [25].

The observed discrepancy between the two divergence estimates of rice and cold tolerant grasses in our study (Fig. 3A and 3B) can be interpreted in two different ways. The *OsLRR-PSR* gene is the true ortholog of IRI-like genes and the incongruent divergence time estimates are caused by differences in molecular clock rates. Or alternatively, if the molecular clock rate is similar, it follows that *OsLRR-PSR* diverged from IRI-like genes long before rice and cold tolerant grasses diverged (Fig. 3). The burst of IRI-like sequence duplications must then have occurred in the ancestor genome of rice, and this implies that the rice genome subsequently must have lost all genes belonging to the IRI-like gene family. Even though loss of genes and whole gene families is not an uncommon feature of plant genome evolution [16,24], elevated clock rate differences is a more parsimonious explanation to the divergence estimate differences seen in Figure 3A and 3B. Evolutionary rate differences are highly common among closely related species, different lineages of a species, and also within a genome [26,27]. When using a single gene family to estimate divergence between species, as with the IRI-like genes, deviation from the average genome clock rate would be expected. As an example the clock rate of the ten control genes varied from $4.1–7.1*10^{-9}$, with an average rate of $5.4*10^{-9}$, when using a divergence time between rice-cold tolerant grasses of 50 My.

Assuming true orthologous relationship between *OsLRR-PSR* and IRI-like genes we can calibrate an average molecular clock rate for IRI-like genes using the dS = 0.97 and assuming an absolute divergence time of 50 Mya (see methods). This gives us an estimate of an IRI-like gene family specific clock rate of $9.7*10^{-9}$. Employing this adjusted clock rate pushes the estimates for the initiation of IRI-like gene duplications forward to 36, 27, and 39 Mya for wheat, perennial ryegrass and barley, respectively. This is approximately 3–14 My after our estimate for divergence between rice and cold tolerant grasses based on the control genes.

Species specific differences in IRI-like sequence numbers

Twice as many IRI-like sequences, partial and full length, were identified in wheat compared to barley (Table 1). *In silico* mining is vulnerable to methodologically introduced uncertainties. For example, the fact that the wheat EST database at NCBI is more than twice as large (1.2 M ESTs) than for barley (500 K ESTs) could be a contributing factor to the differences in numbers of IRI-like mined sequences because we expect that the EST database size is positively correlated to transcriptome coverage of an organism. A separate effect of a larger EST database will be the inclusion of ESTs from an increased number of genotypes, which could be a source of introduction of allelic polymorphisms.

Even though methodological properties might elevate the number of wheat sequences identified to some extent, we believe that much of the difference in wheat and barley IRI-like sequence numbers are related to genomic ploidy level differences. Wheat (*Triticum aestivum*) is an allo-hexaploid originating about 8.000 years ago. It has three homoeologous genomes A, B, and D, which are estimated

Table 3: Evolutionary rate control genes and their pairwise synonymous distances.

Gene	Accession number			Synonymous distance (dS)	
	Ta	Lp	Os	Ta vs Lp dS	TaLp vs Os dS*
Cytosolic glyceraldehyde-3-phosphate dehydrogenase	EF592180	EF463063	NM_001059674	0.33 (0.04)	0.47 (0.06)
Actin	AB181991	AY014279	NM_001062196	0.30 (0.03)	0.49 (0.07)
Gibberellin 20-oxidase	Y14008	AY014281	NM_001058486	0.37 (0.02)	0.70 (0.13)
Phytochrome B	AF137331	AF137308	NM_001056445	0.39 (0.04)	0.57 (0.07)
Casein protein kinase 2 alpha subunit	AB052133	AB213317	NM_001065287	0.30 (0.05)	0.48 (0.09)
Na+/H+ antiporter precursor	AY461512	AY987047	NM_001074903	0.33 (0.04)	0.58 (0.09)
Putative plasma membrane Na+/H+ antiporter	AY326952	AY987046	CB634542	0.14 (0.02)	0.44 (0.06)
Myo-inositol phosphate synthase	AF542968	AY154382	NM_001055777	0.32 (0.04)	0.57 (0.07)
Cinnamoyl CoA reductase	CK161291	AF010290	NM_001052667	0.28 (0.01)	0.41 (0.03)
Fructan beta-(2,1) fructosidase	AJ564996	DQ016297	NM_001052039	0.39 (0.06)	0.72 (0.11)

Accession numbers are given for putative orthologous genes from wheat (Ta), perennial ryegrass (Lp) and rice (Os) used for evolutionary rate control. Their synonymous pairwise divergence was calculated as a mean of three estimates (see methods section), and the standard deviation of all three estimates is given in parenthesis. *denotes a mean synonymous distance of two pairwise comparisons between perennial ryegrass-rice and wheat-rice (referred to as TaLp-Os).

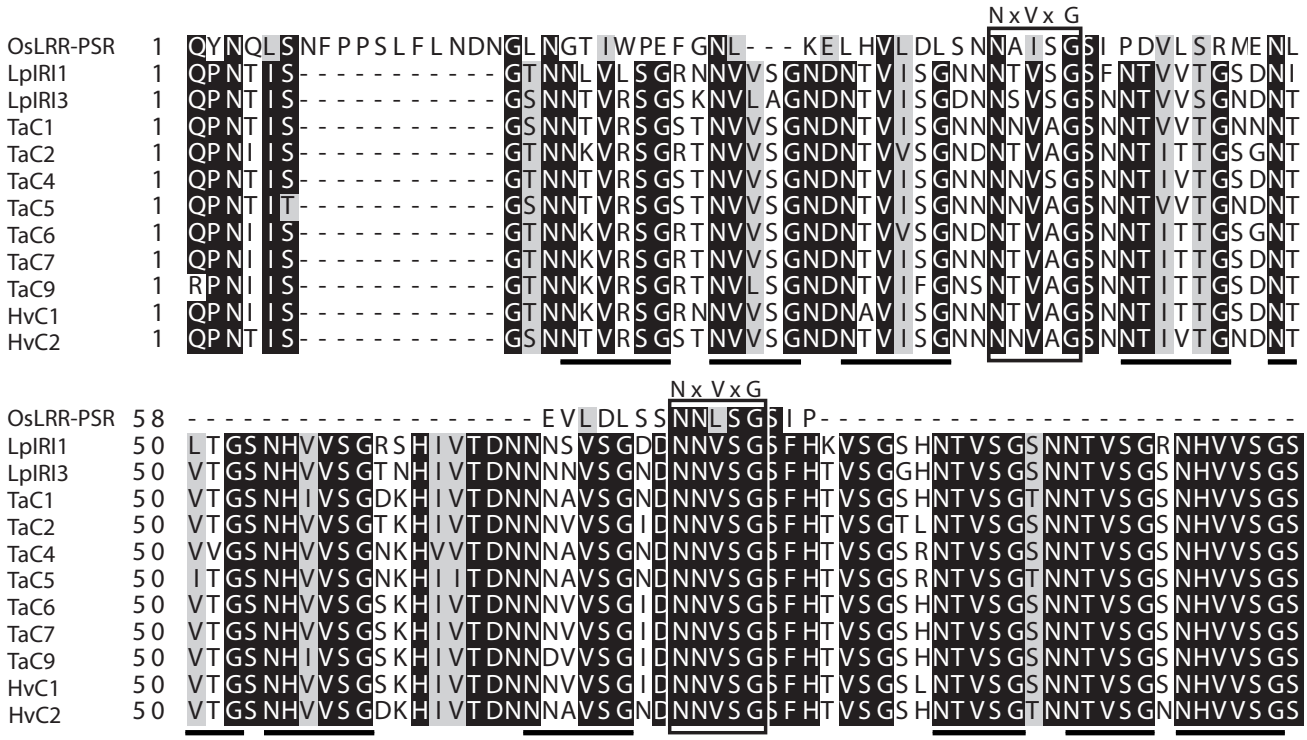


Figure 4

Amino acid alignment between IRI-like sequences and *OsLRR-PSR* at the initiation of the IRI-domain. Shaded residues are identical in > 90% of sequences. Black bars underlines IRI-domain A- and B motifs (NxVxG/NxVxxG). Boxed motifs are IRI-domain motifs shared between OsLRR-PSR and IRI-like sequences.

to have diverged 4.5–2.5 Mya [25]. Our results from the phylogenetic analyses, supported by the dS estimates, suggests that IRI-like sequences within monophyletic clade I and II (Fig. 2) could be homoeologous rather than paralogous. But low pairwise dS can alternatively reflect recent gene duplications. Consequently our inferences on the evolutionary relationship between putative wheat homoeologous sequences must be viewed in a critical manner.

A model of wheat IRI-like sequence evolution

As mentioned, wheat differs from the diploid grasses in our study in that it has many more and younger IRI-like paralogs, some of which probably can be accounted for by wheat polyploidisation. Using the IRI gene specific clock rate of $9.7*10^{-9}$ we built a model of the evolutionary history of wheat IRI-sequences using a phylogenetic approach (Fig. 6A). According to literature and the data acquired in this study wheat and barley specific duplications would be younger than ~35-25 My and wheat specific duplications would be younger than ~11 My [25]. The internal nodes of the phylogeny in Fig. 6 can therefore be interpreted as putative ancient duplication events happening before divergence within Pooideae (D_A), putative duplication events specific to wheat and barley (D_{WB}), putative speciation event between wheat and barley (S_{WB}), putative gene duplication events specific to wheat (D_W), speciation events of the A, B, and D genomes of wheat (S_{ABD}), and lastly allelic divergence (A).

The evolutionary model with adjusted clock rates strengthens our hypothesis on the homoeologous relationship of the wheat clade I and II sequences (Fig. 2). The divergence between TaC6 and TaC2/TaC11 are predicted slightly earlier than the polyploidisation event, thus this internal node could not be classified unambiguously. Sub-trees of clade I and II, in which we have included the barley sequence with lowest synonymous distance to each clade, are presented in Figure 6B and 6C, respectively. The sub-trees further support the hypothesis that clade I and II represents genes that mainly have arisen through polyploidisation events. In both sub-trees the closest related barley sequence is estimated to have diverged from the wheat clade I and II about 15 Mya, coinciding with wheat-barley divergence [25]. However without a complete knowledge of the orthologous relationships of the IRI-like sequences the inferences on evolutionary relationships are somewhat speculative.

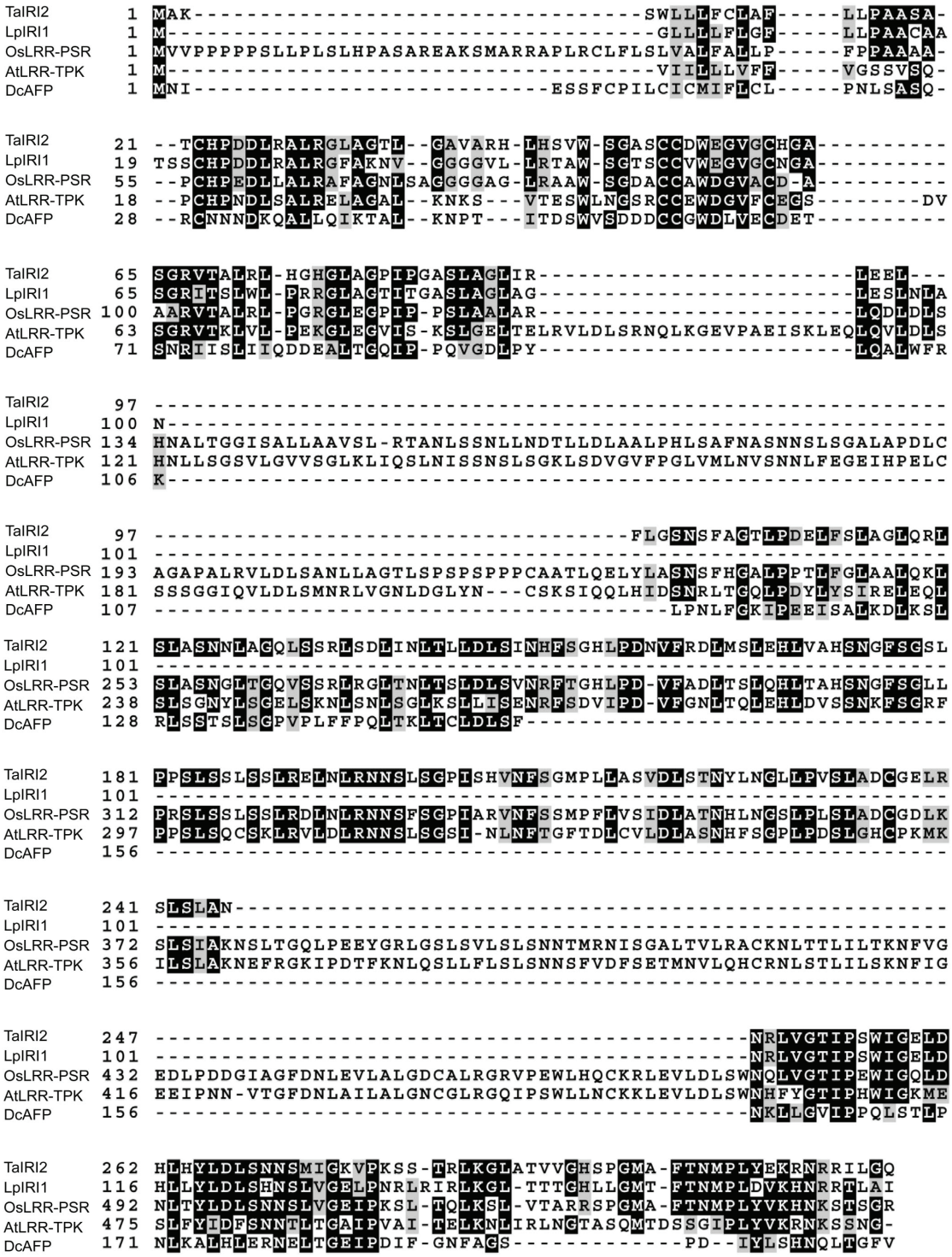


Figure 5

Amino acid alignment of homologous LRR-domains from monocot and dicot species. The alignment shows homologous LRR-domains between Arabidopsis (AtLRR-TPK), carrot (DcAFP), rice (OsLRR-PSR), wheat (TaIRI1) and perennial ryegrass (LpIRI1). All sequences in the alignment has a blastp E-value of > $1*10^{-18}$ to at least one other sequence in the alignment. Shaded residues are identical in > 60% of sequences.

Structural and functional diversification of the IRI-like gene family

A striking feature of the IRI-like gene family is the structural differentiation between paralogs (Fig. 1). Structural diversification of IRI-like genes, as seen in our sequence collection, would be expected to affect the spectrum of IRI-like peptide function, because both LRR and IRI-domains are known to be involved in substrate binding [15,28]. One interpretation of this pattern is that IRI-like sequences with complementary combinations of LRR motifs and IRI-domain sizes are selected for and retained in the genome, which is what we expect from the duplication-degeneration-complementation (DDC) model of paralog evolution [18]. DDC predicts that mutations in regulatory elements increase the probability of paralog retention because it leads to partitioning of ancestral functions (subfunctionalisation), and the model has proven to be an important contribution to understanding evolution of paralogous genes [29,30]. The DDC model has later been expanded to coding sequences [31,32], and recently a combination of regulatory and structural DDC has been demonstrated [33,34].

Regulatory subfunctionalisation in gene expression and tissue localisation has been demonstrated between *TaIRI1* and *TaIRI2* (TaC3) [10], two genes coding for highly divergent LRR-domain structure and length. In our study we have also found evidence that peptide structure divergence has led to sub- or neofunctionalisation. A barley IRI-like sequence contig (HvC3) with no IRI-domain still seems to play a functional role under etiolation. This suggests that the LRR-domain of IRI-like genes may play a functional role in multiple stress responses. Other LRR-domain containing genes in plants have also been shown to be involved in stress responses under drought stress and as a key membrane-bound regulator of absiscic acid signalling [35,36].

One interesting aspect of the structural divergence pattern is that all genes except *LpIRI2* are predicted to encode a conserved N-terminus signalling domain targeting the proteins to the secretory pathway. Secretion to the apoplast is expected for proteins with ice interacting functions. In the light of these data, an interpretation of the structural variability of LRR-domains, combined with the apparent conservation of the N-terminal signalling domain, is that IRI-like genes might be under selective pressure for a continuous ORF from the signalling domain across the LRR-domain and into the IRI-domain, conserving the crucial function of apoplast export of IRI peptides. The LRR-domain itself might not be under functional conservation. As an example: the full length sequenced mRNA AK249041 from barley has a N-terminal conserved predicted signal peptide motif, a completely reduced LRR-domain with no predicted LRR motifs, and an IRI-domain.

Less dramatic polymorphisms between paralogs, such as single amino acid substitutions or small motif number differences could potentially have a large effect on the functionality. Single amino acid substitutions have been shown to radically change AFP functionality in both plant and animal AFPs [13,37]. Chakrabartty and co-workers [38] showed that only small deletions in an AFP with repetitive structure from flounder altered the ice interacting properties dramatically. Thus, all the observed polymorphisms between IRI paralogs, even down to single amino acid substitutions, could potentially be of functional significance.

Birth of an IRI repeat domain

The molecular mechanisms underlying the metamorphosis from an *OsLRR-PSR*-like ancestor gene into the first bipartite IRI-like gene have been addressed by Trembley et al. [10]. They proposed the "transposable element hypothesis" (TE) suggesting that the IRI-domain is a TE insertion that has resulted in a FS mutation and caused the loss of the PK-domain. However no TE signatures were found flanking the IRI-domain [10]. Based on results from our sequence analysis (Fig. 4) we propose a competing hypothesis on the evolution of the IRI-domain, namely the repeated motif expansion (RME) hypothesis. It has been shown that expansions of domains by duplication of repeated motifs are common in genes of repetitive structure [39]. We suggest that IRI motifs have increased in copy number by a yet unknown mechanism, possibly illegitimate recombination, slippage, or uneven crossing over. Contrary to the TE-hypothesis the RME hypothesis can explain the evolution of the IRI-domain and at the same time account for the existence of two IRI motif-like blocks in *OsLRR-PSR* (Fig 4). Lastly, if the entire IRI-domain is a TE-insertion we would expect this TE sequence to be found at other loci in grasses. However no such reports of TE-like sequences homologous to the IRI-domain are known to our knowledge.

Convergent evolution of LRR containing AFPs

LRR-domain containing proteins are extremely abundant in plants. The largest LRR containing plant peptide group is LRR receptor kinases (LRR-RK), having more than 200 members in the *A. thaliana* genome [40]. Plant disease resistance associated genes (NBS-LRR) comprise another large LRR containing functional group [41]. Common for the function of LRR domains in any peptide is that they are associated with peptide-peptide recognition and binding interactions [42-44].

Through comparative protein domain analysis we have shown that LRR-domains of IRI-like genes are much less

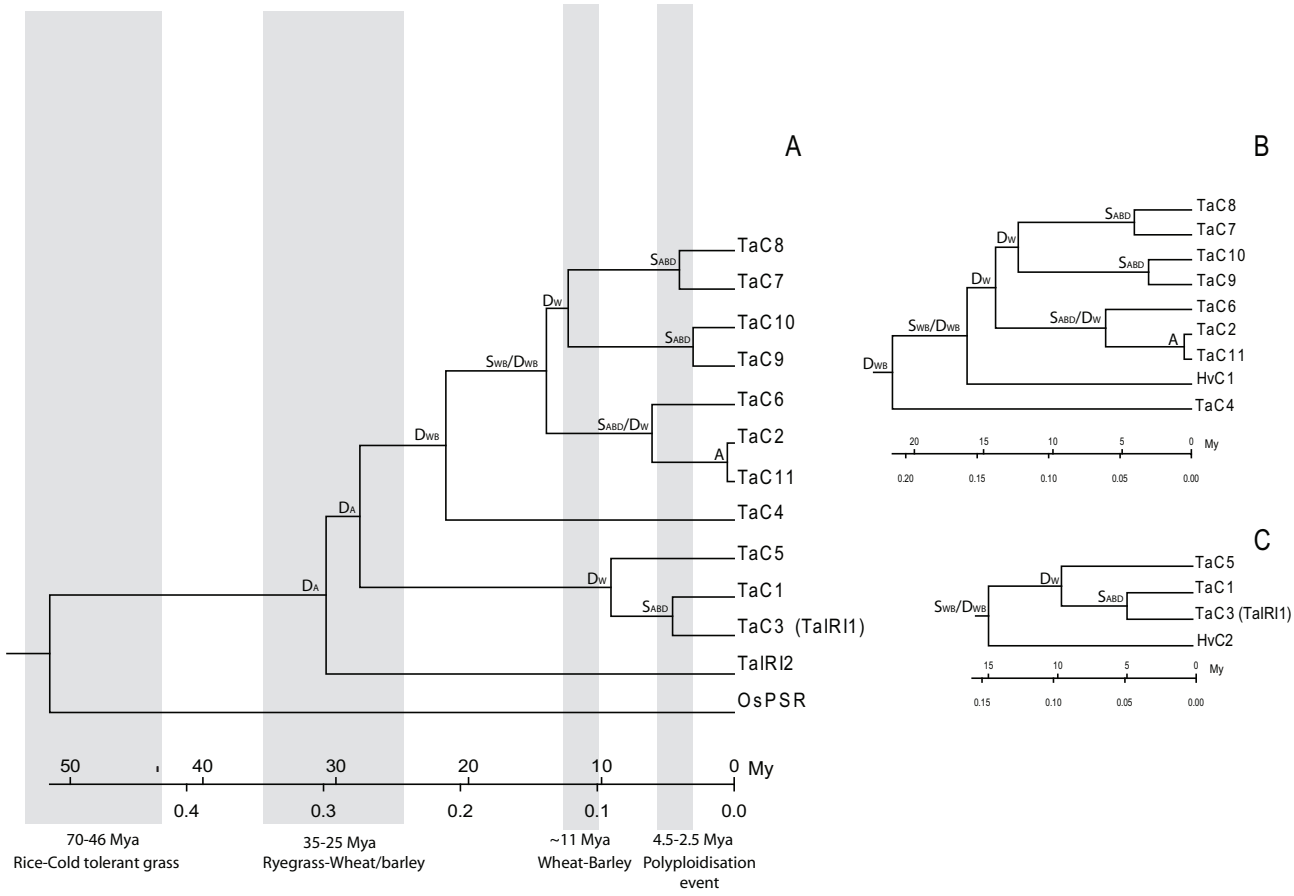


Figure 6

Evolutionary relationships between the wheat IRI-like sequences inferred with an IRI-like gene specific molecular clock rate. Phylogenetic trees based on pairwise synonymous distances made with Unweighted Pair Group Method with Arithmetic Mean (UPGMA). Distance scale shows synonymous site divergence and estimated absolute divergence time in million years. Internal nodes represent duplication, speciation, or allelic divergence events. D_A = ancient duplication events shared by wheat, barley, and perennial ryegrass; D_{WB} = duplication events shared by wheat and barley; D_W = duplication events exclusive for the wheat lineage; S_{WB} = speciation event between wheat and barley; S_{ABD} = speciation event for the A, B, and D genomes of wheat; A = allelic divergence event. A) Tree based on all wheat IRI-like sequences. B) Sub-tree of monophyletic clade II including the two most recently radiated IRI-like sequences in our dataset (HvC1 and TaC4). C) Sub-tree of monophyletic clade I including the single most recently radiated barley IRI-like sequence.

conserved compared to the predicted signal peptide motif flanking the N-terminus of the LRR-domains. We believe that this could be due to lack of selective constrains on the LRR-domain function itself, or perhaps selection for divergent LRR-domain functions as predicted by DDC. Whatever functional role today's IRI-like sequence LRR-domains might play; there is little doubt that the LRR-domain of IRI-like genes in cold tolerant grasses shares an ancient common ancestor with the LRR-domain of DcAFP (Fig. 5). However, while cold tolerant grass IRI-like proteins have evolved ice binding capacity through the evolution of an IRI-domain [14,15], DcAFP have evolved ice binding capacity through changes in the LRR-domain itself [13]. *DcAFP* and grass LRR-IRI genes are therefore intriguing examples of parallel evolution of function by two completely different molecular mechanisms; evolutionary alterations of a pre-existing LRR-domain and evolution of a novel repeat domain with ice binding properties.

Conclusion

The IRI-like genes identified by Sidebottom et al. [14] and Tremblay et al. [10], and in this study tell a tale of a complex evolutionary history that includes birth of an ice binding domain, a burst of gene duplication events after cold tolerant grasses radiated from rice, domain structure differentiation between paralogs, and sub- and/or neofunctionalisation of IRI-like proteins. Given more

detailed functional studies, the IRI-like gene family can provide a valuable example of how duplicated genes evolve novel functional spectres. The hypothesis that evolution of IRI-like genes has been important for Pooideae grass adaptation to cold climate [10] is strengthened by this study as we show that the evolution of the IRI-like gene family probably happened after the divergence from rice, and furthermore that the numbers of IRI-like genes are higher than earlier known.

Methods

In silico *IRI-like sequence mining*

A blastn search in the NCBI database was performed using *TaIRI1*. All sequences with blast E-value < $1*10^{-20}$ were downloaded from the EST and core nucleotide databases. Contigs were aligned with alignment parameters set to > 97% identity and > 40 nucleotides overlap using Sequencher (Gene Codes Corp., Ann Arbor, MI, USA). The 97% identity threshold was set to allow contig alignments to include different allelic forms and polymorphisms caused by EST sequencing errors. Non-coding nucleotides (i.e. promoter and 3'UTR) were removed after an initial prediction of open reading frame (ORF), and subsequently the sequences were realigned with identical parameters. All contigs were translated into their predicted amino acid sequence. Sequence contigs with lack of start and stop codon due to incomplete sequence coverage or putative sequence errors causing frame shift mutations were not included in the analysis. We validated the *in silico* mining method by aligning EST mined unigenes with four full length cDNA clones of grass IRI-like genes from the NCBI core nucleotide collection (barley; AK252915, AK249041/wheat; AY968588, AY968589).

BAC identification and sequencing

Two perennial ryegrass BAC libraries were used to identify novel IRI-like genes [45]. Primers for the initial identification of novel IRI-like genes were designed from coding sequences of *LpAFP* (AJ277399) and a partial sequence of a *Festuca pratensis* IRI-like homolog (EU684537). The LpAFP primer pair had forward primer 5'GATGAACAGCC GAATACGATTTCT3' and reverse primer 5'GCTTCCAGAT ACAACGTGGTTGCT3', denaturing at 94°C for 4 minutes and then 35 cycles of 94°C 30 s, 60°C 45 s, 72°C 45 s, and 72°C 10 min. Primer pairs designed from the *F. pratensis* sequence were forward primer 5'TGTCATATCGGGG AACAACA3' and reverse primer 5'ACATGGTTTCGTCCG GATAC3' denaturing at 94°C for 4 minutes and then 40 cycles of 94°C 10 s, 60°C 45 s, and 72°C 1.30 min, and 70°C for 10 min. We also designed a third primer pair, referred to as LpIRIx primer pair, with forward primer 5'GAATGCCGTATCTGGGGACC3' and reverse primer 5'GTGGTTCCCGGATACGGTATT3', based on multiple sequence data acquired from sequencing of the above mentioned genes. This primer pair was used under the same conditions as the LpAFP primers. DNA maxi-preps of the BAC-clones were preformed using the NucloBond BAC 100 Kit (MACHEREY-NAGEL, Düren, Germany). For BAC-sequencing 500 ng BAC-DNA was combined with 20 μM of primer, 8 μl BigDye 3.1 ready mix and dH_20, to 20 μl total volume. Following 5 min of denaturising at 95°C, 50 cycles were performed with 30 s at 95°C, 10 s 50°C, and 4 min 60°C. Subsequently, the sequencing reactions were precipitated and sequenced on an ABI PRISM 3100 (Applied Biosystems, Foster City, CA, USA).

Protein domain characterisation

Predicted peptide LRR domains were characterised using Pfam [46]. We verified the Pfam results by visual inspection of the sequences defining a LRR motif as LxxLxLxx, or variations of it where L is substituted with I, V, or A. To track the molecular evolution of the LRR-domain, *OsLRR-PSR* was used as a template for comparison to the predicted domains of IRI-like sequences. LRR motifs predicted by Pfam were considered significant if the Pfam E-value was lower than 0.05. IRI-like amino acid sequences were aligned with the LRR-domain of OsLRR-PSR and the LRR motifs in IRI-like sequences were named according to which of the LRR motif number in OsLRR-PSR they aligned to. IRI-domain characterisation was performed by visual scoring of the total number of repeat motifs (NxVxG/NxVxxG). IRI-domain repeat motifs were considered as "present", and counted, when they contained no more than one amino acid substitution compared to the consensus motifs. Signal peptide domains were predicted by TargetP [47].

Estimation of substitution rates

To estimate divergence times between putatively paralogous and orthologous sequences we used the average dS obtained from three different methods, Nei & Gojobori [48], Kumar [49], and the Li-Wu-Luo [50] method, in MEGA (4.0) [51]. As a control for evolutionary rates of IRI-like genes we calculated the average dS values of ten randomly selected orthologs from wheat, perennial ryegrass, and rice. Maximum likelihood estimation of non-synonymous to synonymous substitution ratios (w) was performed using Codeml in the PAML software package (v 3.15) [52]. The 3 × 4 codon substitution model was chosen for Codeml w estimations. PAL2NAL [53] was used to make codon based nucleotide alignments for the use in MEGA and PAML. The absolute time of divergence between orthologs and paralogs was estimated using a rate of $6.5*10^{-9}$ substitutions/synonymous site/year for grasses [23]. For estimation of mutation rates and absolute divergence times we used the relationship k = dS/2T, where k is the absolute rate of synonymous substitutions per year, dS is the synonymous substitution rate, and T is the absolute time since divergence. To identify putative alleles not grouped in the same contig due to methodo-

logical errors, a cut-off threshold of dS = 0.03 was used. This threshold was set on the basis of average inter-allelic dS for LRR-domains of 27 disease resistance like genes in *A. thaliana* [54], and inter-allelic dS_{max} of *LpIRI1* calculated from twelve European perennial ryegrass genotypes (dS_{max} = 0.015, data not published).

Molecular and phylogenetic analysis

All amino acid and nucleotide alignments were made by MAFFT [55] and manually edited in BioEdit [56], and the phylogenetic trees were constructed in Treefinder [57]. An AIC criteria test [58], implemented in the Modeltest option in Treefinder, was used to choose substitution model for the phylogenetic analysis. ML trees were bootstrapped with 1000 replicates. Synonymous distance based trees were inferred by UPGMA from a pairwise dS distance matrix in MEGA. Alignment figures were prepared by BoxShade http://www.ch.embnet.org/software/BOX_faq.html.

Authors' contributions

SRS conceived the study, carried out molecular genetics work and molecular analysis, and drafted the manuscript. HR carried out molecular genetics work, helped with data analysis, and participated in manuscript drafting. TA carried out molecular work and helped draft the manuscript. OAR participated in drafting the manuscript.

Additional material

Additional file 1

EST accession numbers. Table giving all EST accession numbers included in the full length IRI-like in silico *mined sequences.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-8-245-S1.pdf]

Additional file 2

Amino acid alignment of perennial ryegrass IRI-like sequences. Amino acid alignment of perennial ryegrass IRI-like sequences used for phylogenetic analysis.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-8-245-S2.pdf]

Additional file 3

Amino acid alignment of wheat IRI-like sequences. Amino acid alignment of wheat IRI-like sequences used for phylogenetic analysis.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-8-245-S3.pdf]

Additional file 4

Amino acid alignment of barley IRI-like sequences. Amino acid alignment of barley IRI-like sequences used for phylogenetic analysis.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-8-245-S4.pdf]

Acknowledgements

We thank Kjetil Fosnes for assistance with BAC sequencing and Dr. Magnus Dehli Vigeland, Dr. Gordon Allison, and anonymous reviewers for valuable comments on the manuscript. This study was funded through the KMB project *Festulolium* with Improved Forage Quality and Winter Survival for Norwegian Farming, project number 173319/110, funded by the Research Council of Norway and Graminor AS.

References

1. Pearce RS, Fuller MP: **Freezing of Barley Studied by Infrared Video Thermography.** *Plant Physiol* 2001, **125(1):**227-240.
2. Márquez E, Rada F, Fariñas M: **Freezing tolerance in grasses along an altitudinal gradient in the Venezuelan Andes.** *Oecologia* 2006, **150(3):**393-397.
3. Zachariassen KE, Kristiansen E: **Ice Nucleation and Antinucleation in Nature.** *Cryobiology* 2000, **41(4):**257-279.
4. Smallwood M, Bowles JD: **Plants in a cold climate.** *Philos Trans R Soc Lond B Biol Sci* 2002, **357(1423):**831-847.
5. Tomczak MM, Hincha DK, Estrada SD, Wolkers WF, Crowe LM, Feeney RE, Tablin F, Crowe JH: **A Mechanism for Stabilization of Membranes at Low Temperatures by an Antifreeze Protein.** *Biophys J* 2002, **82(2):**874-881.
6. Griffith M, Yaish MWF: **Antifreeze proteins in overwintering plants: a tale of two activities.** *Trends in Plant Science* 2004, **9(8):**399-405.
7. Barrett J: **Thermal hysteresis proteins.** *The International Journal of Biochemistry & Cell Biology* 2001, **33(2):**105-117.
8. Hon WC, Griffith M, Mlynarz A, Kwok YC, Yang DSC: **Antifreeze Proteins in Winter Rye Are Similar to Pathogenesis-Related Proteins.** *Plant Physiol* 1995, **109(3):**879-889.
9. Smallwood M, Worrall D, Byass L, Elias L, Ashford D, Doucet CJ, Holt C, Telford J, Lillford P, Bowles DJ: **Isolation and characterization of a novel antifreeze protein from carrot (Daucus carota).** *Biochem J* 1999, **340(2):**385-391.
10. Tremblay K, Ouellet F, Fournier J, Danyluk J, Sarhan F: **Molecular Characterization and Origin of Novel Bipartite Cold-regulated Ice Recrystallization Inhibition Proteins from Cereals.** *Plant Cell Physiol* 2005, **46(6):**884-891.
11. Worrall D, Elias L, Ashford D, Smallwood M, Sidebottom C, Lillford P, Telford J, Holt C, Bowles D: **A Carrot Leucine-Rich-Repeat Protein That Inhibits Ice Recrystallization.** *Science* 1998, **282(5386):**115-117.
12. Zhang D-Q, Wang H-B, Liu B, Feng D-R, He Y-M, Wang J-F: **Carrot Antifreeze Protein Does Not Exhibit the Polygalacturonase-inhibiting Activity of PGIP Family.** *Acta Genetica Sinica* 2006, **33(11):**1027-1036.
13. Zhang DQ, Liu B, Feng DR, He YM, Wang SQ, Wang HB, Wang JF: **Significance of conservative asparagine residues in the thermal hysteresis activity of carrot antifreeze protein.** *Biochemical Journal* 2004, **377:**589-595.
14. Sidebottom C, Buckley S, Pudney P, Twigg S, Jarman C, Holt C, Telford J, McArthur A, Worrall D, Hubbard R, *et al.*: **Phytochemistry: Heat-stable antifreeze protein from grass.** *Nature* 2000, **406(6793):**256-256.
15. Kuiper MJ, Davies PL, Walker VK: **A Theoretical Model of a Plant Antifreeze Protein from Lolium perenne.** *Biophys J* 2001, **81(6):**3560-3565.
16. Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud P-F, Lindquist EA, Kamisugi Y, *et al.*: **The Physcomitrella Genome Reveals Evolutionary Insights into the Conquest of Land by Plants.** *Science* 2008, **319(5859):**64-69.

17. Kondrashov FA, Kondrashov AS: **Role of selection in fixation of gene duplications.** *Journal of Theoretical Biology* 2006, **239(2):**141-151.
18. Force A, Lynch M, Pickett FB, Amores A, Yan Y-l, Postlethwait J: **Preservation of Duplicate Genes by Complementary, Degenerative Mutations.** *Genetics* 1999, **151(4):**1531-1545.
19. Blanc G, Hokamp K, Wolfe KH: **A Recent Polyploidy Superimposed on Older Large-Scale Duplications in the Arabidopsis Genome.** *Genome Res* 2003, **13(2):**137-144.
20. Salse J, Bolot S, Throude M, Jouffe V, Piegu B, Masood Quraishi U, Calcagno T, Cooke R, Delseny M, Feuillet C: **Identification and Characterization of Shared Duplications between Rice and Wheat Provide New Insight into Grass Genome Evolution.** *Plant Cell* 2008. tpc.107.056309
21. Ma J, Bennetzen JL: **Recombination, rearrangement, reshuffling, and divergence in a centromeric region of rice.** *Proceedings of the National Academy of Sciences of the United States of America* 2006, **103(2):**383-388.
22. Blanc G, Wolfe KH: **Widespread Paleopolyploidy in Model Plant Species Inferred from Age Distributions of Duplicate Genes.** *Plant Cell* 2004, **16(7):**1667-1678.
23. Gaut BS, Morton BR, McCaig BC, Clegg MT: **Substitution rate comparisons between grasses and palms: Synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene rbcL.** *Proceedings of the National Academy of Sciences* 1996, **93(19):**10274-10279.
24. Paterson AH, Bowers JE, Chapman BA: **Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics.** *Proceedings of the National Academy of Sciences* 2004, **101(26):**9903-9908.
25. Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R, Gornicki P: **Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the Triticum/Aegilops complex and the evolutionary history of polyploid wheat.** *Proceedings of the National Academy of Sciences* 2002, **99(12):**8133-8138.
26. Sloan DB, Barr CM, Olson MS, Keller SR, Taylor DR: **Evolutionary Rate Variation at Multiple Levels of Biological Organization in Plant Mitochondrial DNA.** *Mol Biol Evol* 2008, **25(2):**243-246.
27. DeRose-Wilson L, Gaut B: **Transcription-related mutations and GC content drive variation in nucleotide substitution rates across the genomes of Arabidopsis thaliana and Arabidopsis lyrata.** *BMC Evolutionary Biology* 2007, **7(1):**66.
28. Ellis J, Dodds P, Pryor T: **The generation of plant disease resistance gene specificities.** *Trends in Plant Science* 2000, **5(9):**373-379.
29. Gu Z, Rifkin SA, White KP, Li W-H: **Duplicate genes increase gene expression diversity within and between species.** *Nat Genet* 2004, **36(6):**577-579.
30. Hughes AL, Friedman R: **Expression Patterns of Duplicate Genes in the Developing Root in Arabidopsis thaliana.** *Journal of Molecular Evolution* 2005, **60(2):**247-256.
31. Tocchini-Valentini GD, Fruscoloni P, Tocchini-Valentini GP: **From the Cover: Structure, function, and evolution of the tRNA endonucleases of Archaea: An example of subfunctionalization.** *Proceedings of the National Academy of Sciences* 2005, **102(25):**8933-8938.
32. Cusack BP, Wolfe KH: **When gene marriages don't work out: divorce by subfunctionalization.** *Trends in Genetics* 2007, **23(6):**270-272.
33. Babbitt CC, Haygood R, Wray GA: **When Two Is Better Than One.** *Cell* 2007, **131(2):**225-227.
34. Hittinger CT, Carroll SB: **Gene duplication and the adaptive evolution of a classic genetic switch.** *Nature* 2007, **449(7163):**677-681.
35. Osakabe Y, Maruyama K, Seki M, Satou M, Shinozaki K, Yamaguchi-Shinozaki K: **Leucine-Rich Repeat Receptor-Like Kinase1 Is a Key Membrane-Bound Regulator of Abscisic Acid Early Signaling in Arabidopsis.** *Plant Cell* 2005, **17(4):**1105-1119.
36. Chini A, Grant JJ, Seki M, Shinozaki K, Loake GJ: **Drought tolerance established by enhanced expression of the CC-NBS-LRR gene, ADR1, requires salicylic acid, EDS1 and ABI1.** *The Plant Journal* 2004, **38(5):**810-822.
37. Haymet ADJ, Ward LG, Harding MM, Knight CA: **Valine substituted winter flounder "antifreeze": preservation of ice growth hysteresis.** *FEBS letters* 1998, **430:**301-306.
38. Chakrabartty A, Yang DS, Hew CL: **Structure-function relationship in a winter flounder antifreeze polypeptide. II. Alteration of the component growth rates of ice by synthetic antifreeze polypeptides.** *J Biol Chem* 1989, **264(19):**11313-11316.
39. Wicker T, Yahiaoui N, Keller B: **Illegitimate recombination is a major evolutionary mechanism for initiating size variation in plant resistance genes.** *The Plant Journal* 2007, **51(4):**631-641.
40. Shiu S-H, Bleecker AB: **Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases.** *Proceedings of the National Academy of Sciences* 2001, **98(19):**10763-10768.
41. Belkhadir Y, Subramaniam R, Dangl JL: **Plant disease resistance protein signaling: NBS-LRR proteins and their partners.** *Current Opinion in Plant Biology* 2004, **7(4):**391-399.
42. Ogawa M, Shinohara H, Sakagami Y, Matsubayashi Y: **Arabidopsis CLV3 Peptide Directly Binds CLV1 Ectodomain.** *Science* 2008, **319(5861):**294.
43. Thomas CM, Jones DA, Parniske M, Harrison K, Balint-Kurti PJ, Hatzixanthis K, Jones JDG: **Characterization of the Tomato Cf-4 Gene for Resistance to Cladosporium fulvum Identifies Sequences That Determine Recognitional Specificity in Cf-4 and Cf-9.** *Plant Cell* 1997, **9(12):**2209-2224.
44. Meyers BC, Shen KA, Rohani P, Gaut BS, Michelmore RW: **Receptor-like Genes in the Major Resistance Locus of Lettuce Are Subject to Divergent Selection.** *Plant Cell* 1998, **10(11):**1833-1846.
45. Farrar K, Asp T, Lubberstedt T, Xu ML, Thomas AM, Christiansen C, Humphreys MO, Donnison IS: **Construction of two Lolium perenne BAC libraries and identification of BACs containing candidate genes for disease resistance and forage quality.** *Molecular Breeding* 2007, **19(1):**15-23.
46. Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R, *et al.*: **Pfam: clans, web tools and services.** *Nucl Acids Res* 2006, **34(suppl_1):**D247-251.
47. Nielsen H, Engelbrecht J, Brunak S, von Heijne G: **Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites.** *Protein Eng* 1997, **10(1):**1-6.
48. Nei M, Gojobori T: **Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions.** *Mol Biol Evol* 1986, **3(5):**418-426.
49. Nei M, Kumar S: **Molecular Evolution and Phylogenetics.** Oxford University Press; 2000.
50. Li WH, Wu CI, Luo CC: **A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes.** *Mol Biol Evol* 1985, **2(2):**150-174.
51. Tamura K, Dudley J, Nei M, Kumar S: **MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0.** *Mol Biol Evol* 2007, **24(8):**1596-1599.
52. Yang Z: **PAML: a program package for phylogenetic analysis by maximum likelihood.** *Computational Applied Bioscience* 1997, **13(5):**555-556.
53. Suyama M, Torrents D, Bork P: **PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments.** *Nucl Acids Res* 2006, **34(suppl_2):**W609-612.
54. Bakker EG, Toomajian C, Kreitman M, Bergelson J: **A Genome-Wide Survey of R Gene Polymorphisms in Arabidopsis.** *Plant Cell* 2006, **18(8):**1803-1818.
55. Katoh K, Kuma K-i, Toh H, Miyata T: **MAFFT version 5: improvement in accuracy of multiple sequence alignment.** *Nucl Acids Res* 2005, **33(2):**511-518.
56. Hall TA: **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** *Nucleic Acids Symposium Series* 1999, **41:**95-98.
57. Jobb G, von Haeseler A, Strimmer K: **TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics.** *BMC Evolutionary Biology* 2004, **4(1):**18.
58. Aikake H: **A new look at the statistical model identification.** *IEEE Transactions on Automatic Control* 1974, **19(3):**716-723.

Paper IV

SANDVE, S. R., and S. FJELLHEIM, 2010. **Did gene family expansions during the Eocene-Oligocene boundary climate cooling play a role in Pooideae adaptation to cool climates?** Molecular Ecology **19:** 2075-2088.

MOLECULAR ECOLOGY

Molecular Ecology (2010) **19**, 2075–2088 doi: 10.1111/j.1365-294X.2010.04629.x

Did gene family expansions during the Eocene–Oligocene boundary climate cooling play a role in Pooideae adaptation to cool climates?

SIMEN RØD SANDVE and SIRI FJELLHEIM

Department of Plant and Environmental Sciences, Norwegian University of Life Sciences, PO Box 5003, 1431 Aas, Norway

Abstract

Adaptation to cool environments is a common feature in the core group of the grass subfamily Pooideae (Triticeae and Poeae). This suggest an ancient evolutionary origin of low temperature stress tolerance dating back prior to the initiation of taxonomic divergence of core Pooideae species. Viewing the Pooideae evolution in a palaeo-climatic perspective reveals that taxonomic divergence of the core Pooideae group initiated shortly after a global super-cooling period at the Eocene–Oligocene boundary (~33.5–26 Ma). This global climate cooling altered distributions of plants and animals and must have imposed selection pressure for improved low temperature stress responses. Lineage-specific gene family expansions are known to be involved in adaptation to new environmental stresses. In Pooideae, two gene families involved in low temperature stress response, the C-repeat binding factor (*CBF*) and fructosyl transferase (*FT*) gene families, has undergone lineage-specific expansions. We investigated the timing of these gene family expansions by molecular dating and found that Pooideae-specific expansion events in *CBF* and *FT* gene families took place during Eocene–Oligocene super-cooling period. We hypothesize that the E–O super-cooling exerted selection pressure for improved low temperature stress response and frost tolerance in a core Pooideae ancestor, and that those individuals with multiple copies of *CBF* and *FT* genes were favoured.

Keywords: climate change, evolution, gene family evolution, Pooideae

Received 22 December 2009; revision received 2 March 2010; accepted 7 March 2010

Introduction

The grass family Poaceae is thought to have originated ~70–80 million years ago (Ma) (Crepet & Feldman 1991; Janssen & Bremer 2004; Prasad *et al*. 2005) and is one of the largest plant families with ~10 000 species, many of which have immense ecological, cultural, and economical value. From a human perspective the sub-family Pooideae (herein also referred to as pooids), which includes the cereals and forage grasses, is one of the most important plant groups. Thus the pooids have been the subject of intense research, including great effort to understand the ancient evolutionary history of the group.

Pooideae is thought to have originated from a common ancestor shared with bamboo (Bambusoideae sub-family) and rice (Ehrhartoideae sub-family) (Barker *et al*. 2001) about 46–41.4 Ma (Gaut 2002; Stromberg 2005; Sandve *et al*. 2008). Adaptation to cool environments is a common feature among the core pooid species (Triticeae and Poeae tribes) (Barker *et al*. 2001), suggesting a common evolutionary origin for low temperature tolerance. The present natural distribution of pooids is skewed towards cooler climatic zones in the northern and southern hemisphere, the northern parts of Canada and Greenland and Antarctica being the extreme latitudinal limits (Hartley 1973). Any adaptations to low temperature stress common to all core pooids must date back prior to the most recent common ancestor (MRCA) of the core pooids, which is thought to be ~25 million years (Myr) old (Gaut 2002; Sandve *et al*. 2008).

Coinciding with early pooid evolution (45–35 Ma) was a global climate transition towards a cooler climate,

Correspondence: Simen Rød Sandve, Fax: +47 64965901;
E-mail: simen.sandve@umb.no

often referred to as the Cenozoic 'greenhouse to ice-house transition' (Zachos *et al.* 2001, 2003). Ice-encapsulation of the Arctic and Antarctic initiated ~45 Ma (Lear *et al.* 2000; Moran *et al.* 2006; Eldrett *et al.* 2007) and there is evidence for establishment of sea ice in the Arctic ~47.5 Ma (Stickley *et al.* 2009). Cenozoic cooling rate climaxed in a super-cooling event at the Eocene–Oligocene (E–O) boundary that lasted from 33.5 to ~26 Ma (Zachos *et al.* 2001; Liu *et al.* 2009). Mean annual continental temperature at present semi-high latitudes (~44°N) is estimated to have dropped as much as 8°C (Zanazzi *et al.* 2007) and was probably accompanied by increased seasonality (Eldrett *et al.* 2009). This climate cooling is associated with large scale extinctions and shifts in global distributions of both fauna and flora (Ivany *et al.* 2000; Coxall & Pearson 2007; Seiffert 2007).

A recent study of a large collection of plant lineages, including the grasses, shows that plants that had undergone whole genome duplications were more likely to survive the Cretaceous–Tertiary (K–T) boundary mass extinction (Fawcett *et al.* 2009). The evolutionary advantage of increased gene content under extreme environmental disturbance is thought to be related to increased rate of evolution and adaptation due to functional redundancy (Crow & Wagner 2006). Similar to the K–T boundary, but much less extreme, the E–O boundary represents a global sudden shift in climate driven selection pressure. However, no whole genome duplications in core pooids are associated with the E–O boundary (Blanc & Wolfe 2004; Fawcett *et al.* 2009). On the other hand, adaptive increase in gene content is not restricted to whole genome duplications. Selection for lineage specific expansion of individual gene families can also play an important role in adaptive evolution (Kondrashov & Kondrashov 2006; Chen *et al.* 2008; Zou *et al.* 2009). Thus it is possible that the E–O climate cooling selected for individuals with redundancy of genes involved in low temperature stress response which then facilitated rapid adaptation.

Two well studied gene families known to be important in cold stress response in the pooids are the fructosyl transferase (*FT*) and the C-repeat binding factor (*CBF*) gene families (Note: The *FT* abbreviation used herein must not be confused with Flowering Locus T). Interestingly, it is shown that both of these gene families have undergone Pooideae-specific gene family expansion events (Badawi *et al.* 2007; Hisano *et al.* 2008). However the timing of these expansion events has not yet been studied. Fructosyl transferase genes encode fructan-sugar synthesizing enzymes. Fructan accumulation in grasses initiates at exposure to low temperatures and is thought to play two roles in adaptation to cool environments by (i) acting as storage nutrients in overwintering plants and (ii) stabilizing cell membranes under freezing stress (reviewed in Livingston *et al.* 2009). Recently a causal link between fructan accumulation and low temperatures stress tolerance has been confirmed through transgenic studies in rice (Kawakami *et al.* 2008), tobacco (Konstantinova *et al.* 2002; Parvanova *et al.* 2004; Li *et al.* 2007) and pooids (Hisano *et al.* 2004). *CBF*s are transcription factors that regulate downstream genes under abiotic stress, including a large proportion of the cold responsive genes (Agarwal *et al.* 2006). Two *CBF* sub-groups, the *CBFIIIc/d* and *CBFIV*, has been identified as Pooideae specific and contain genes that are expressed under cold treatment (Badawi *et al.* 2007). Moreover, several studies have linked polymorphisms in *CBFIIIc/d* and *CBFIV* genes to phenotypic differences in frost tolerance levels of wheat and barley (Knox *et al.* 2008; Fricano *et al.* 2009; Sutton *et al.* 2009).

The acquirement of Pooideae-specific *CBFIIIc/d*, *CBFIV* and *FT* gene families are good candidate events to represent ancient adaptive gene family expansion events. One interesting question is therefore; could *CBFIIIc/d*, *CBFIV* and *FT* expansions be associated to the E–O cooling period? The aim of this study was therefore to analyse the timing of the initial expansions of the Pooideae-specific *CBFIIIc/d*, *CBFIV* and *FT* gene families and assess if these expansion events were associated with the E–O climate cooling period. Our results indicate that expansion events in all three gene families are likely to have happened in a core pooid ancestor within the E–O cooling period. The possible implications of Pooideae-specific evolution of gene families involved in cold and frost stress for the Pooideae-lineage adaptation to cold climates is discussed.

Materials and methods

Sequences and alignments

We retrieved pooid and rice *FT* and *CBF* sequences identified by Hisano *et al.* (2008) and Badawi *et al.* (2007), respectively, from the NCBI database. Our initial sequence collection included eight species of core pooids, five species from the Triticeae tribe (*Triticum aestivum*, *T. monococcum*, *Hordeum vulgare*, *H. brevisubulatum* and *Secale cereale*) and three Poeae tribe species (*Festuca arundinacea*, *Avena sativa* and *Lolium perenne*). We chose to include *A. sativa* in the Poeae tribe based on recommendations in recently published treatments (Quintanar *et al.* 2007; Soreng *et al.* 2007). In addition we included 10 rice *CBF*-homologs and one *Arabidopsis thaliana FT*-homolog. To minimize alignment uncertainties only coding sequences were used in the analyses. Protein alignments for the corresponding nucleotide sequences

were first made in MAFFT (Katoh *et al.* 2005) and then protein alignments were used to guide nucleotide alignments with Pal2Nal (Suyama *et al.* 2006). Due to substantial variation in the degree of conservation across the gene length, all alignments were manually checked and trimmed using BioEdit (Hall 1999) to contain only highly conserved domains. The *Arabidopsis thaliana FT*-homolog (NP564798) and a rice *FT*-homolog (NP001052748) was retrieved from the NCBI database to be used as outgroups in *FT* phylogenies.

Phylogenetic analysis

Reliability of phylogeny-based molecular dating depends on alignment quality, phylogenetic analysis, calibration methods, and the use of correct calibration time (Fawcett *et al.* 2009). We used multiple phylogenetic methods and calibration methods to ensure that the node age estimates were not biased by a specific method. The phylogenetic tree estimations were done using three different methods, all with differing underlying assumptions. Two methods were Bayesian methods, MRBAYES (Huelsenbeck & Ronquist 2001) and BEAST (Drummond & Rambaut 2007) while one method were based on Maximum Likelihood (ML) and was implemented in Treefinder (Jobb *et al.* 2004). Further, the methods differed in their assumption of a molecular clock. MRBAYES and Treefinder have no assumption of a molecular clock, while BEAST assumes a relaxed molecular clock. In this study we used the uncorrelated lognormal relaxed clock to allow local rate variability. The evolutionary models used in the ML phylogenetic analysis were chosen based on the Akaike Information Criteria (AIC) model test implemented in Treefinder. The AIC model test selected J2, a sub-model of TIM (transitional model), and TVM (transversional model) with gamma distributed substitution rate heterogeneity (five classes) for the *CBF* and *FT* datasets, respectively. Because these models are not included in MRBAYES or BEAST, we used the GTR model (general time reversible) with gamma distributed rate heterogeneity for the Bayesian phylogenetic inference. The ML phylogenies were 50% consensus trees made from 100 bootstrap replicates. MRBAYES estimations were performed by running two MCMC chains for 100 000–500 000 generations with nst = 6 (GTR model) and rates = gamma. After each run the convergence diagnostics were inspected. In addition to visual convergence inspection, we controlled that the split frequencies between the MCMC chains were <0.01 in all analysis and that the convergence diagnostic (PSRF) were ~1. Fifty percent consensus trees with a burn-in equal to 25–50% of the total number of trees were subsequently constructed (>5000 trees sampled). For BEAST phylogenies a single chain was run for 10 000 000 generations under a Yule tree model using GTR substitution model with five categories of gamma distributed rates. Trees were sampled every 1000 generations. Burnin was set to 50% of chain length (sampling 5000 trees) after inspecting convergence and sampling statistics. The effective sample size (ESS) for each run was checked in Tracer software (v1.5). The ESS of the posterior distribution were >500 in all analysis, indicating that the MCMC sampling were well mixed.

Initially, two consensus trees were made from all *FT* and *CBF* sequences, respectively. The two initial trees were used to confirm that *CBFIIIc/d*, *CBFIV* and *FT* were indeed Pooideae-specific as reported in Hisano *et al.* (2008) and Badawi *et al.* (2007). Next, new reduced alignments were made containing only *CBFIIIc/d* and *CBFIV* sequences with one or two outgroup sequences. From these reduced alignments calibration-phylogenies were estimated for the use in node age estimation. The *CBFIIIc/d* and *CBFIV* alignments were 382 and 530 nucleotides long, respectively. The average pairwise similarity between sequences was 0.80 (±0.05) in the *CBFIIIc/d* alignment and 0.83 (±0.05) in the *CBFIV* alignment. The calibration-phylogenies were not allowed to contain polytomies because this was observed to severely affect node age estimates. Hence, the phylogenies were trimmed to only contain well resolved clades (no bootstrap values or posterior probabilities smaller than 0.8). For the *FT* analysis slightly different alignments were used in MRBAYES/ML compared to the BEAST analysis. Instead of the divergent *A. thaliana FT*-homolog outgroup, a more closely related rice homolog (NP_001052748) was used in BEAST analysis. This also affected the alignment length used in analysis. The *FT* alignment in MRBAYES and ML analysis were 522 nucleotides long and in the BEAST analysis 1701 nucleotides long, with average sequence similarity of 0.77 (±0.08) and 0.79 (±0.07), respectively. All Gen Bank accession numbers for sequences that were used in the final analyses can be found in the respective phylogenetic trees in the result section. Alignments used in the final calibration analysis can be found in the Supporting Information.

Calibrations of phylogenetic trees

We can divide the calibration methods (i.e. transformation of the phylogenies from relative to absolute time) used in this article into two types. The first type is the calibration methods which were used on phylogenetic trees inferred with MRBAYES and ML methods. In these calibrations the 50% consensus tree topology and branch lengths are assumed to be 'true' and no phylogenetic uncertainties are incorporated into the calibrations.

Two such methods were used, the Local Rates Minimum Deformation (LRMD) and Non Parametric Rate Smoothing (NPRS) (Sanderson 1997), implemented in the Treefinder software. Both LRMD and NPRS assume that rates are more similar in neighbouring branches and try to minimize the rate change along the lineages. Such autocorrelated rates in a tree is a fair assumption for a gene family derived phylogeny. As novel paralogs evolve, rates will change in the tree due to gene redundancy and relaxed selection. Two distantly related paralog genes in the tree would be expected to have less similar rates compared to close orthologs in the tree. The difference between LRMD and NPRS methods lies in their cost function and how these rate changes are estimated. The second type of calibration method that was used was integrated in the BEAST analysis. Phylogenetic reconstruction in BEAST assumes a relaxed molecular clock which enables simultaneous inference of phylogeny and absolute time calibration. In addition BEAST incorporates phylogenetic uncertainties (i.e. variation in branch length and topology) into the confidence intervals of the node age estimates while the only uncertainty incorporated in the LRMD and NPRS calibrations in this study is uncertainty of the age of the calibration point used.

As calibration point for estimation of absolute time we used the Pooideae-Rice divergence, i.e. the node which defined the split between the closest rice homolog and the Pooideae-specific clades. Three recent publications using different dating techniques have estimated the rice-pooid split to 46 Ma (Gaut 2002), 42 Ma (Sandve *et al.* 2008) and 41.4 Ma (Stromberg 2005). See the discussion for more details on the reliability of these rice-pooid divergence dates. To reflect the uncertainties linked to the rice-pooid divergence time, a time window between 46 and 41.4 Ma was used as the rice-pooid split in the calibration procedure. The MRCA between Triticeae and Poeae orthologous sequences represent the divergence between Triticeae and Poeae lineages, which is estimated to be ~25 Myr old (Gaut 2002; Sandve *et al.* 2008). We used these nodes (in which we have a prior expectation of age) to assess the reliability of our calibration results. Extreme deviations from the expected Triticeae–Poeae MRCA age would indicate potential errors in our node age estimates.

For the LRMD and NPRS calibrations a point estimate of the node ages and 95% confidence interval (CI) were estimated by repeating the calibration procedure 1000 times, each time with a random calibration time point drawn from a uniform distribution in the defined calibration time window (46–41.4 Ma). LRMD and NPRS calibrations could only use the one distance from one Pooideae-specific sequence to the rice homolog at one time. To control for the possibility that calibrations with different Pooideae sequences produce different results, we calibrated every tree multiple times with several Pooideae sequences. As parameters in the BEAST calibration we used a uniformly distributed prior probability on the rice-pooid split similar to the LRMD and NPRS calibrations (46–41.4 Ma). Furthermore, we used a normally distributed prior on one of the Triticeae–Poeae splits in each phylogeny. This prior was set as the average of all Triticeae–Poeae split estimates (see 'Results' section) from the LRMD and NPRS calibrations with its associated standard deviation.

Randomization test

Clustering of randomly sampled events in time can occur simply by chance. To assess if the initial gene family expansion events in the *CBFIIIc/d*, *CBFIV* and *FT* clades cluster closer to the E–O cooling period than expected by chance, we used a randomization test (Manly 2006). The null hypothesis is that the initial expansion events in *CBFIIIc/d*, *CBFIV* and *FT* clades are not clustered closer to the E–O cooling period than expected by the null distribution of three random events. We calculated the null distribution by drawing all possible sets of three expansion events in a time window spanning from Pooideae-origin (set to 43.7 Ma) to the divergence of Triticea–Poaea lineages (set to 25 Ma) (Gaut 2002; Sandve *et al.* 2008) using the statistical environment R (R Development Core Team 2009). We performed this test on the results from all three calibration methods separately and on the average node age estimates across all calibration methods. Divergence from the null hypothesis was evaluated using the standard Euclidean distance as test statistic; i.e. the test statistic is defined as the sum of the squared distances from all initial gene duplication events to the middle of the cool period following the E–O boundary (29.75 Ma). The exact *P*-value was calculated as the proportion of all possible draws of sets of three events that produced a smaller or equally small test statistic when compared to our data. Small values of the test statistic led to rejection of the null hypothesis in favour of the alternative hypothesis that the clustering is not a chance event. The reason we use the distance to the middle of the cool period instead of the distance to the E–O boundary itself, is because selection pressure for improved low temperature stress response must have been an ongoing process during the whole super-cooling period. R-script for the randomization test can be found in Supporting Information.

Test for positive selection

One type of adaptive evolution is favourable changes of one or more amino acids in a protein, also referred to

as positive selection. This type of evolution is expected to leave the signature at the molecular level of a non-synonymous substitution rate (d*N*) larger than synonymous substitution rate (d*S*) (d*N*/d*S* = ω > 1). We used pylogenetic analysis with maximum likelihood (PAML v4.4) (Yang *et al.* 2000) to test if the pooid-specific *CBFIIIc/d*, *CBFIV*, or *FT* genes showed any signatures of such positive selection (ω > 1). The same alignments as used in the phylogenetic analysis was used in the PAML analysis, however all outgroup sequences were removed. As suggested in Yang *et al.* (2000), model 0, 1, 2, 3, 7, and 8 were fitted to the data and likelihood ratio tests (LRT) were performed to evaluate if models which allow for positively selected sites (model 2 and model 8) had a better fit compared to models that did not allow for positive selection (model 1 and model 7). The comparison of model 0 and model 3 was done to test if ω was constant or varied among codons. The LRT was performed by comparing two times the log-likelihood difference between two nested models under a χ^2 distribution (Yang *et al.* 2000). See Yang *et al.* (2000) for a detailed description of PAML models.

Results

Phylogenetic estimations

In the initial phylogeny of all *CBF* sequences *CBFIV* grouped in a monophyletic clade with *OsCBFIV-1.B.1* as the closest related rice homolog (Supporting Information). This is in accordance with Badawi *et al.* (2007). The *CBFIIIc/d* topology obtained by Badawi *et al.* (2007) could not be recreated by MRBAYES or ML estimations when all *CBF* sequences were included in the phylogenetic analysis (Supporting Information). The *OsCBFIIIb-1H* was identified as the closest rice homolog to the Pooideae *CBFIIIc/d* sequences, but the *CBFIIIc/d* sequences did not form a monophyletic clade (Supporting Information). In the MRBAYES phylogeny, one sequence (*TmCBFIIId-17*) clustered outside the entire Pooideae *CBFIIIc/d* clade and the *OsCBFIIIb-1H* formed a polytomy within the assumed Pooideae-specific *CBFIIIc/d* clade. The ML phylogeny showed similar tendencies, but in the ML phylogeny all Pooideae-specific *CBFIIIc/d* sequences were clustered in one monophyletic clade which also included the *OsCBFIIIb-1H*. Subsequently we re-estimated MRBAYES and ML phylogenies using a reduced number of divergent outgroup sequences to increase the length of the alignment with high conservation and thus improve parameter estimations. We were now able to recreate the monophyletic relationship of Pooideae-specific *CBFIIIc/d* sequences as presented in Badawi *et al.* (2007). The Bayesian *FT* phylogeny supported the monophyletic origin of the Pooideae-specific *FT* sequences as presented by Hisano *et al.* (2008) with rice vacuolar invertase 3 (AF276704) as the closest related rice sequence (data not shown).

After removing the sequences which created polytomies all the phylogenetic methods produced identical topologies. The final calibration tree for *CBFIIIc/d* contained 14 genes; three were rice homologs (AY327040, AF300970, AP008215), three belonged to the Poeae and eight belonged to Triticeae. The *CBFIIIc/d* phylogeny contained two nodes representing divergence between Triticeae and Poeae, however in the initial phylogenetic estimations with MRBAYES the sub-clade containing one of these Triticeae–Poeae splits had a very low Bayesian support (0.62). This led us to remove *TmCBFIII-16* and re-estimate the *CBFIIc/d* phylogeny as this sequence have a high likelihood of being incorrectly placed in the phylogeny. The *CBFIV* calibration phylogeny consisted of 17 genes; two rice homologs (AY327040, AY785894), one sequence belonging to Poeae and 15 belonging to Triticeae. Two different *FT* phylogenies both including 17 genes were used. For the LRMD and NPRS analyses one *A. thaliana* homolog, one rice homolog, seven Poeae sequences, and nine Triticeae sequences was included. In the BEAST analysis the *A. thaliana* homolog was swapped with a distant rice homolog (*Os04g0692800*). The accession numbers for all sequences included in the final calibrated trees can be found in Figs 1–4.

Tree calibration and node age estimation

Table 1 contains point estimates and 95% confidence limits for the ages of the initial expansion events in the three Pooideae-specific clades. Switching the Pooideae calibration sequences within the same phylogeny in LRMD and NPRS calibrations did not affect the node age estimates more than ±250 000 years, hence only results from one calibration run is presented. When averaging over all phylogenetic estimation methods and calibration methods, the initial expansion events were estimated to have happened within the E–O cooling period; at 33.4 (±1.93), 27.1 (±1.46), and 31.1 (±1.63) Ma in *CBFIIIc/d*, *CBFIV* and *FT* clades respectively. Overall, BEAST produced the oldest and LRMD produced the youngest estimates for the initial expansion events. BEAST also produced the largest CIs, reflecting that BEAST takes into account uncertainties linked to branch lengths and topology when estimating the 95% CI of node ages. All node age estimates for the initial expansions in *CBFIV* and *FT* clades were placed inside the boundary of the E–O cooling period. Moreover, 95% CI from the BEAST analysis does not extend much beyond the E–O cooling period (Figs 1 and 2). Our results therefore strongly suggest that the initial expansion events in

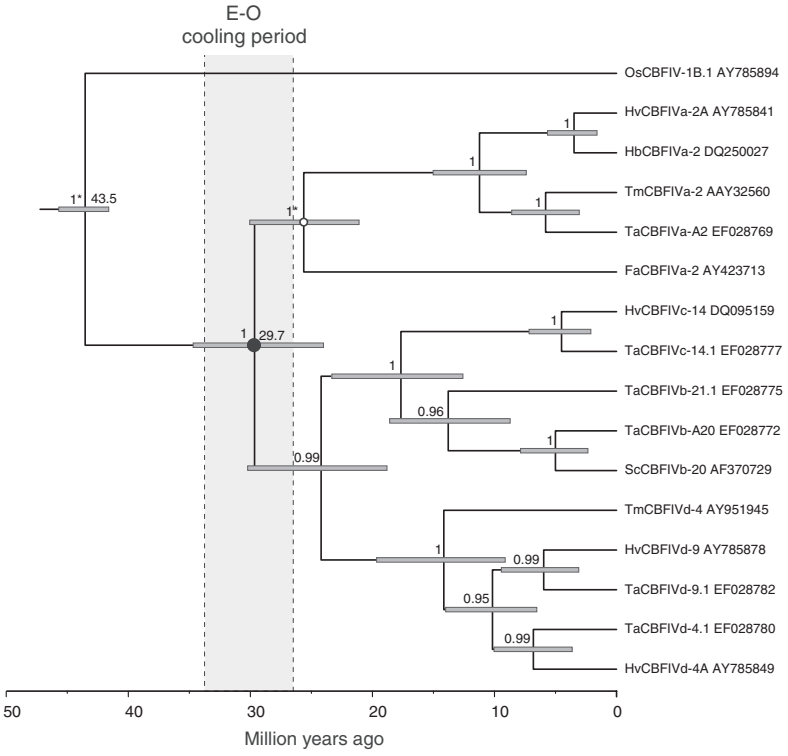


Fig. 1 Calibrated *CBFIV* phylogeny from BEAST. Calibrated phylogeny from BEAST. Outgroup sequence (AY327040) is removed from the figure. Grey bars represent 95% CI, numbers to the left of each node represents Bayesian posterior probabilities, and numbers to the right of the nodes are the average node age estimates. Black circle represents the initial gene family expansion event, grey circle represents secondary expansion event ancestral to the core pooids, white smaller circles represents the putative Triticeae–Poeae divergence node. Star denotes the calibration points at which prior probabilities were defined. Gene names in the phylogenies consist of species abbreviation and gene classification according to Badawi *et al.* (2007) and Gen Bank accession numbers. Species abbreviations are as follows: Ta=*T. aestivum*. Tm = *T. monococcum*. Hv = *H. vulgare*. As = *A. sativa*. Os = *O. sativa*.

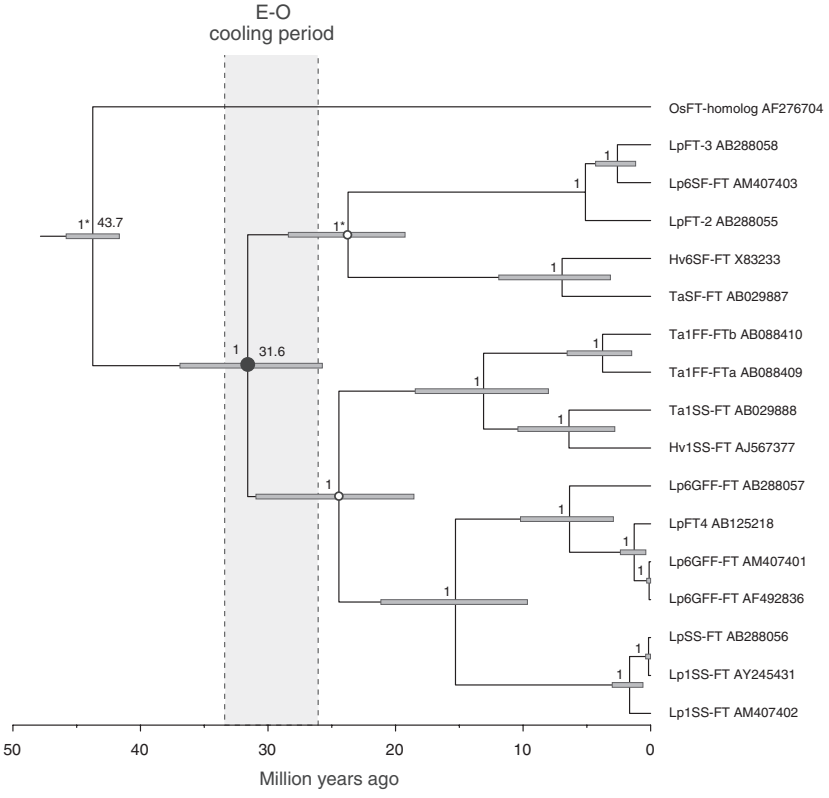


Fig. 2 Calibrated *FT* phylogeny from BEAST. Calibrated phylogeny from BEAST. Outgroup sequence (Os04g0692800) is removed from the figure. Grey bars represent 95% CI, numbers to the left of each node represents Bayesian posterior probabilities, and numbers to the right of the nodes are the average node age estimates. Black circle represents the initial gene family expansion event and white smaller circles represent the putative Triticeae–Poeae divergence nodes. Star denotes the calibration points at which prior probabilities were defined. Gene names in the phylogenies consist of species abbreviation, putative gene function followed by the Gen Bank accession number. Species abbreviations are as follows: Ta = *T. aestivum*. Hv = *H. vulgare*. Lp = *L. perenne*. Os = *O. sativa*.

CBFIV and *FT* clades happened within the E–O cooling period. The results from the *CBFIIIc/d* clade differ from the *CBFIV* and *FT* clades in two ways. First, only the LRMD calibration results places the initial expansion event within the E–O cooling period (Table 1; Fig. 4). Second, the *CBFIIIc/d* clade has a second expansion

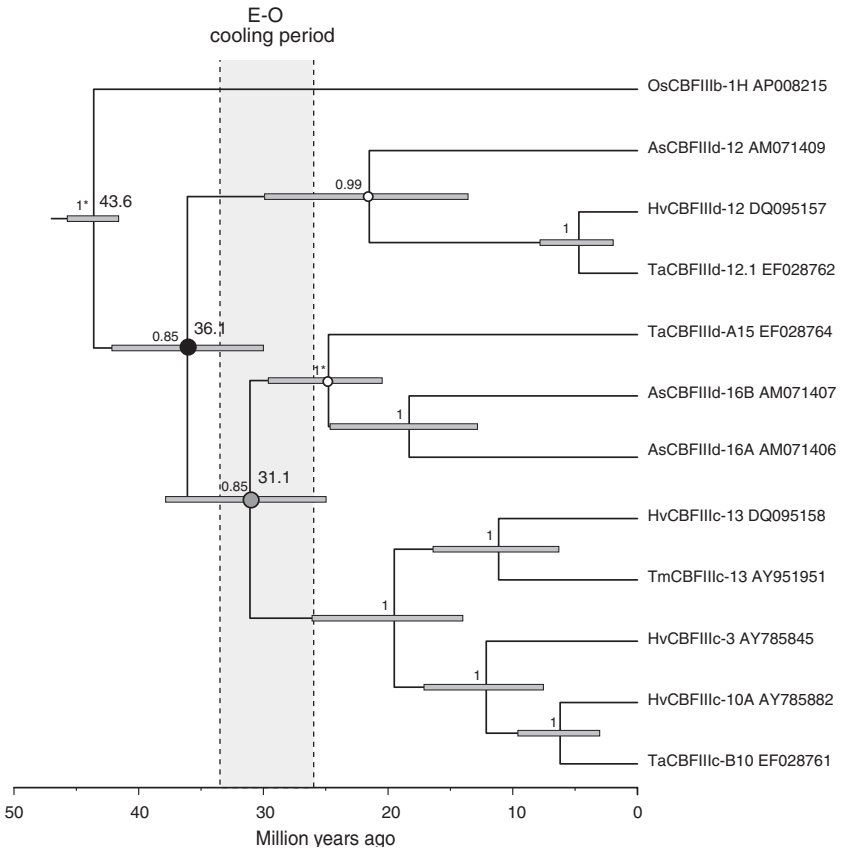


Fig. 3 Calibrated *CBFIIIc/d* phylogeny from BEAST. Calibrated phylogeny from BEAST. Outgroup sequences (AY327040, AF300970) are removed from the figure. Grey bars represent 95% CI, numbers to the left of each node represents Bayesian posterior probabilities, and numbers to the right of the nodes are the average node age estimates. Black circle represents the initial gene family expansion event and white smaller circles represents the putative Triticeae–Poeae divergence node. Star denotes the calibration points at which prior probabilities were defined. Gene names in the phylogenies consist of species abbreviation and gene classification according to Badawi *et al.* (2007) and Gen Bank accession numbers. Species abbreviations are as follows: Ta = *T. aestivum*. Tm = *T. monococcum*. Hv = *H. vulgare*. As = *A. sativa*. Os = *O. sativa*.

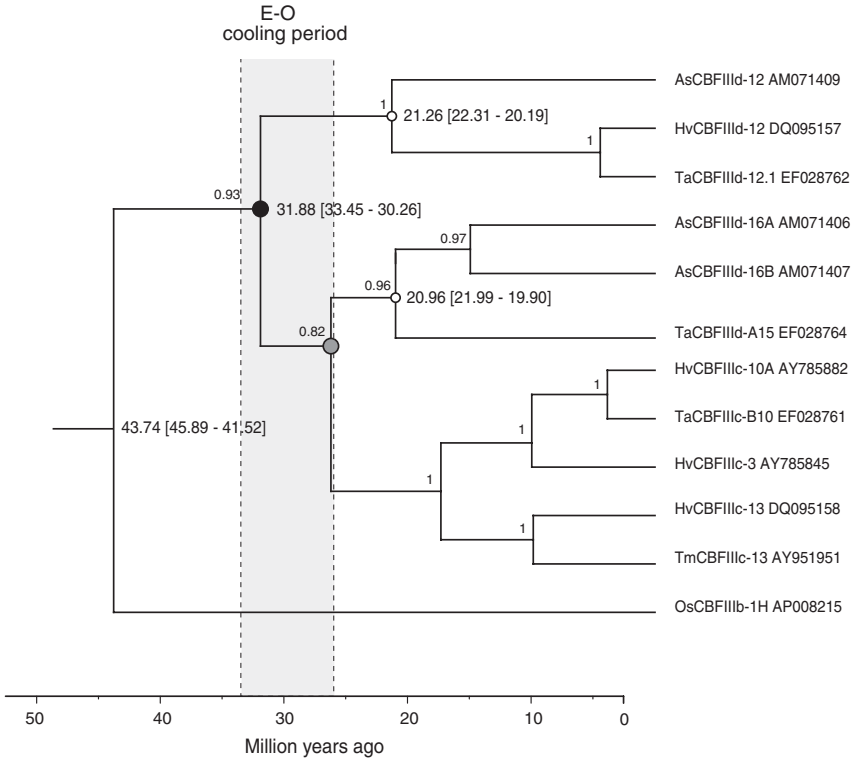


Fig. 4 LRMD calibrated *CBFIIIc/d* phylogeny from MRBAYES. Calibrated MRBAYES phylogeny using the Local Rate Minimum Deformation method. Outgroup sequences (AY327040, AF300970) are removed from the figure. Black circle represents the initial gene family expansion event, grey circle represents secondary expansion event ancestral to the core pooids, and white smaller circles represent the putative Triticeae–Poeae divergence nodes. Gene names in the phylogenies consist of species abbreviation and gene classification according to Badawi *et al.* (2007) and Gen Bank accession numbers. Species abbreviations are as follows: Ta = *T. aestivum*. Tm = *T. monococcum*. Hv = *H. vulgare*. As = *A. sativa*. Os = *O. sativa*.

Table 1 Estimated age (million years) of the initial expansion of Pooideae-specific *FT* and *CBF* gene families

Calibration method	Software	Statistical framework	Gene family	Initial expansion	95% CI
LRMD	MRBAYES	*Bayesian, no clock*	*CBFIIIc/d*	31.8	33.5–30.3
			CBFIV	26.7	28.0–25.3
			FT	28.8	30.2–27.4
	Treefinder	*ML, no clock*	*CBFIIIc/d*	31.3	32.8–29.7
			CBFIV	26.1	27.4–24.8
			FT	30.1	31.6–28.6
NPRS	MRBAYES	*Bayesian, no clock*	*CBFIIIc/d*	33.6	34.9–31.6
			CBFIV	26.2	27.5–24.9
			FT	31.8	33.5–30.3
	Treefinder	*ML, no clock*	*CBFIIIc/d*	34.2	35.9–32.5
			CBFIV	26.7	28.0–25.3
			FT	33.0	34.6–31.3
Uncorrelated log-normal	BEAST	Bayesian, relaxed clock	*CBFIIIc/d*	36.1	29.9–42.1
			CBFIV	29.6	24.0–34.7
			FT	31.6	25.7–36.9

Point estimates from Local Rates Minimum Deformation, Near Parametric Rate Smoothing, and BEAST calibration methods are presented with their estimated 95% CI.

event dating back prior to the Triticea–Poaea split. In the calibrations that estimated the initial expansion to have happened prior to the E–O cooling, this second expansion was placed inside the E–O cooling period (compare Figs 3 and 4). Estimated node age for this secondary expansion event in the *CBFIIIc/d* clade ranged from 25.8 to 31.1 Ma between all calibrated phylogenies in our study. Consequently, even though the age of the initial *CBFIIIc/d* expansion event is uncertain, it seems likely that there was at least one *CBFIIIc/d* expansion event inside the E–O cooling period in a core pooid ancestor.

In the *CBFIIIc/d*, *CBFIV* and *FT* phylogeny two, one and two nodes, respectively, represent a putative MRCA for the Triticeae and Poeae lineages. In the LRMD and NPRS calibrations the youngest and oldest Triticeae–Poeae divergence node were dated to 21.0 and 30.1 Ma, with an average age of the Triticeae–Poeae divergence of 24.5 Ma (±2.93). The average age and its standard deviation were used in subsequent BEAST calibrations as a prior on one of the Triticeae–Poeae divergence nodes in each tree. The age estimates of the Triticeae–Poeae splits in the *FT* and *CBFIV* clades were distributed on both sides (older and younger) of the average age, depending on phylogenetic method and calibration method (data not shown). However, in the *CBFIIIc/d* calibrations the ages of the Triticeae–Poeae divergence nodes were younger than 24.5 Myr in all analyses. The average age of the Triticeae–Poeae divergence in MRBAYES phylogenies calibrated with LRMD and NPRS were 21.25 (±0.2) Myr. The same trend is seen in the un-constrained Triticeae–Poeae split node in the BEAST analysis (21.5 Myr) (Fig. 3).

Randomization test for random clustering of gene duplications events

We performed a randomization test to assess if the clustering of *CBFIIIc/d*, *CBFIV*, and *FT* initial expansion events inside the E–O cooling period is likely to occur by chance. This was done by investigating the probability of acquiring similar clustering when any three random events were drawn from the same defined time interval. Not all calibration methods produced a clustering which gave a significant test. The randomization test was significant for the LRMD calibration results ($P = 0.034$ for both ML and Bayesian phylogenies) but not for the NPRS calibration results (ML; $P = 0.14$, Bayesian; $P = 0.10$) or BEAST results ($P = 0.16$). If we used average node ages from all calibration results (across all phylogenetic methods and calibration methods) the test was significant ($P = 0.04$). The overall low *P*-values reflect that is unlikely that a clustering of expansion events in the E–O cooling period is due to chance.

Analysis of substitution rates

Model 3 with variable ω had a significantly better fit compared to the one-ratio ω model for all gene families ($P < 0.001$, d.f. = 4), but only *CBFIIIc/d* family had a ω-category with $\omega > 1$. Tests for positive selection were significant in the *CBFIIIc/d* clade but only with the model 7– model 8 comparison ($P < 0.01$, d.f. = 2). The Bayes Empirical Bayes (BEB) (Yang *et al.* 2005) analysis implemented in the PAML analysis identified amino acids number 64 (A), 71 (R), and 100 (Q) with a posterior

probability greater than 0.54, 0.64, and 0.89 for being positively selected sites in the *CBFIIIc/d* alignment. The *CBFIV* and *FT* families had no signatures of positive selection. The ω-categories estimated under model 3 was <0.6 in both gene families and neither model 2 or model 8 had a significantly better fit compared to model 1 or model 7, respectively. All parameter estimates from the PAML analyses can be found in Supporting Information.

Discussion

Node ages and estimate reliability

The reliability of phylogeny based molecular dating depends on alignment quality, phylogenetic analysis, calibration methods, and the use of correct calibration time and point (Fawcett *et al.* 2009). Since we only use one calibration point in our study the reliability of our results depends mainly on the use of correct calibration time and the identification of the true MRCA node between rice and the pooid sequences. In our analysis we have assumed that rice and Pooideae shared a MRCA ~46–41.4 Ma. This time window was based on results from three recent publications which used three different methods of phylogenetic dating. It is important to be aware that even though the three estimates were derived from different methods, two of them were (directly and indirectly) based on an identical underlying assumption. Sandve *et al.* (2008) used a synonymous substitution rate to date the MRCA of rice and pooids. The synonymous substitution rate was originally estimated by Gaut *et al.* (1996) based on the assumption that the MRCA of rice-maize are 50 Myr old, an age which was inferred from fossil evidence discussed in a paper by Stebbins (1981). In Gaut (2002) the same assumption is used (MRCA of rice-maize 50 Myr - old), but in this study the dating was done by NPRS calibration of a multigene phylogeny. Although this is a weakness one of the most comprehensive studies of the grass family evolution done by Bremer (2002) supports the use of 50 Myr as the rice-maize divergence. In his study Bremer used four fossil based calibration points spanning the entire depth of the grass family phylogeny. Moreover, Stromberg (2005) used a very different approach to acquire her rice-pooid split estimate, using a combination of phytolith assemblages and grass macrofossils (see citations in Stromberg 2005) as calibration points. In conclusion, up to date literature concerning the age of Pooideae-rice split support the choice of calibration time interval used in this study.

The second assumption our calibration results depend on is that we have identified the true MRCA node between rice and the Pooideae-specific clades. A way of assessing this is to examine the age estimates of other nodes in the tree where an expected age can be defined. One such node is the putative MRCA of the Triticeae–Poeae lineages which is estimated to be ~25 Myr old in two other studies (Gaut 2002; Sandve *et al.* 2008). The average age of the putative Triticeae–Poeae split across all LRMD and NPRS calibrations were 24.5 Myr (±2.93 Myr), indicating that our calibration results seems dependable. However, there were deviations in the *CBFIIIc/d* clade from the expected Triticeae–Poeae divergence age which deserve some attention. In all calibrations the Triticeae–Poeae divergence in *CBFIIIc/d* were found to be younger than the average age, and younger than any of the Triticeae–Poeae divergence nodes in *FT* and *CBFIV* clades. A reason for this could be the second expansion event in *CBFIIIc/d* as calibration results are known to be increasingly unreliable with increasing distance to the calibration point in the phylogeny (Battistuzzi *et al.* 2010).

In the *CBFIV* and *FT* clades the initial expansion event is consistently placed within the E–O transition super-cooling period, regardless of the method. The initial expansion event in *CBFIIIc/d* was on the other hand not unambiguously placed within the E–O cooling period (Figs 3 and 4). The incongruence in *CBFIIIc/d* calibration results with initial expansion event in the *CBFIIIc/d* clade ranging from 36.1 to 31.3 Ma raises an important question. Which of the *CBFIIIc/d* calibrated trees most correctly reflect the true evolutionary history? This question is unfortunately hard to answer. There are no evidence to suggest that a relaxed molecular clock method (e.g. BEAST) produces any superior phylogenetic reconstruction than a non-clock based method (e.g. MRBAYES and ML) (Wertheim *et al.* 2010). Neither can we imagine any good reason to strongly favour the results from NPRS over LRMD, or vice versa. However it is important to note, for the evolutionary interpretation of our results, that *CBFIIIc/d* has two expansion events prior to the core pooid group divergence ~25 Ma. Thus at least one ancient expansion event in *CBFIIIc/d* was placed within the E–O cooling period in all calibrations (compare Figs 3 and 4). Any ancient expansions that are common to all core pooids, even secondary ones, could be adaptive.

The evolutionary impact of gene family expansions

Increase in gene content by gene family expansions is thought to be important for adaptive evolution (Lespinet *et al.* 2002; Chen *et al.* 2008; Flagel & Wendel 2009). In humans for example there are evidence for adaptive copy number increase of *AMY1* for starch digestion (Perry *et al.* 2007), and in many genes related to immunity, olfactory function and protein secretion (Nguyen

et al. 2006). In insects adaptive copy number increase of enzymes that break down toxic compounds have been demonstrated in evolution of insecticide resistance (Ferrarri & Georghiou 1991). Interestingly, a recent study by Chen *et al.* (2008) on an Antarctic fish species provides insights into the genetic basis of adaptation to colder environments in animals. An increase in copy numbers relative to closely related warm water adapted species was found for >100 genes potentially important for physiological fitness under cold stress (Chen *et al.* 2008). One of these proteins was a homologue to the fish type II antifreeze proteins. Similarly, lineage specific expansion of the ice interacting ice re-crystallization inhibition proteins (IRIPs) have previously been shown to have occurred in cold tolerant Pooideae grasses (Sidebottom *et al.* 2000; Sandve *et al.* 2008).

In this study we suggest that copy number increase of *FT*, *CBFIIIc/d*, and *CBFIV* clades could have been involved in adaptation of pooids to cold stress, and subsequent ecological expansion into cooler habitat types. One interesting question that arises is therefore; in which way did an increased number of *FT* and *CBF* genes provide better fitness in cold environments? Evolutionary models for maintenance of gene duplicates predict several mechanisms in which a gene duplication event might be adaptive. First, a proximate fitness advantage of a gene duplication event is either an effect of increased gene dose (Ferrarri & Georghiou 1991) or functional redundancy (Gu 2003; Moore *et al.* 2005). *FT* genes encode enzymes involved in fructan sugar synthesis. Doubling of the *FT* enzyme dose could have resulted in a better sugar metabolism efficiency and thus more rapid fructan accumulation, or perhaps elevated maximum fructan levels (or a combination). Fitness effect of a higher dose of *CBFIIIc/d* and *CBFIV* genes could similarly have led to increased expression of transcription factors under abiotic stress. This might have (i) led to a quicker stress response due to more rapid accumulation of downstream gene products, or (ii) provided the plant with a longer and stronger response to cold stress if higher *CBF* transcript levels also lead to *CBF* mRNA being present in the cell for a longer time. If *CBF* and *FT* gene duplications did not provide any fitness enhancing gene dose effects in a pooid ancestor, these expansions would still have provided functional redundancy and genetic buffering. Genetic buffering protects against deleterious effects of null mutations (Gu *et al.* 2003) and could be advantageous in environments where for example just a slight decrease in frost tolerance (weak deleterious mutations) would be detrimental to the plants. However, the adaptive advantage (strength of selection) for maintaining duplicates purely as genetic buffering has been questioned by many researchers because of the extremely low rates of null mutations (Hahn 2009). Both the dose effect and genetic buffering models would predict patterns of nucleotide variation reflecting purifying selection, which is consistent with substitution patterns observed in the *FT* and *CBFIV* clades.

Second, other models of adaptive increase in gene copy numbers are gene specialization (sub-functionalization), and neofunctionalization (Ohno 1970; Gibson & Goldberg 2009; Hahn 2009). These models assume positive selection on post-duplication mutations rather than an adaptive consequence of the duplication event itself (Hahn 2009). Evolution of lineage specific gene expression patterns is involved in adaptive evolution (Ferea *et al.* 1999; Rifkin *et al.* 2003; Nuzhdin *et al.* 2004), thus gene specialization or neofunctionalization in *CBFIIIc/d* and *CBFIV* genes might have resulted in favourable changes in abiotic stress signal networks leading to novel gene expression patterns and improved cold stress response. The significant test for positive selection in the *CBFIIIc/d* clade ($P < 0.01$) could be interpreted as a support for such adaptive changes in target site specificity. Extant *FT* paralogs have undergone evolution of different substrate specificities (Livingston *et al.* 2009). This could indicate that gene specialization of redundant *FT* enzymes in a pooid ancestor could have played a role in evolution of low temperature stress tolerance. Even though we found no signatures for positive selection in the *FT* gene family we cannot fully exclude positive selection in the *FT* clade due to the relatively low power of such tests (Hahn 2009).

Genetic basis for Pooideae-lineage adaptation to cooler climates

Functional studies provide compelling evidence that evolution of *CBFIIIc/d*, *CBFIV*, and *FT* genes are important for present day low temperature stress response in Pooideae (see 'Introduction'). Hence the rise of *FT*, *CBFIIIc/d*, and *CBFIV* clades in a pooid ancestor did likely affect the evolution of cold stress tolerance. On the other hand, a number of *FT* and *CBF* genes are also known to be involved in drought stress (Xue *et al.* 2008; Livingston *et al.* 2009; Xu *et al.* 2009) and some authors argue that evolution of drought stress was important in early Pooideae evolution (Schardl *et al.* 2008). If we assume that the E–O climate change influenced aridity levels, the rise of *FT*, *CBFIIIc/d* and *CBFIV* clades could have played dual roles in adaptation to abiotic stress. Unfortunately very little data on global aridity level changes across the E–O boundary is available. One recent palaeo-climate study from terrestrial systems in central North America could not identify resolvable

aridity changes across the E–O boundary (Zanazzi *et al.* 2007), whereas the climate cooling at the E–O boundary is evident from many studies (see 'Introduction'). Hence, the evidence from both functional gene studies and palaeo-climate data are in support for a link between adaptation to cold climate and the expansion of the *FT*, *CBFIIIc/d*, and *CBFIV* gene families.

Two additional Pooideae-specific evolutionary innovations have also been suggested to be important for adaptation to cold climates: (i) evolution of vernalization responsiveness related to the *VRN1* gene (Preston & Kellogg 2008) and (ii) evolution of the Pooideae-specific ice re-crystallization inhibition protein (*IRIP*) gene family (Sidebottom *et al.* 2000; Tremblay *et al.* 2005). The timing of *VRN1* recruitment in vernalization responsiveness is unknown, thus any link to the E–O climate cooling remains undeterminable. *IRIP* genes are thought to play an important role in frost stress; the *IRIP* gene expression is regulated by exposure to cold temperature (Tremblay *et al.* 2005) and the IRI-domain of *IRIP* proteins binds to ice crystals and modify their size and shape *in vitro* (Sidebottom *et al.* 2000). Furthermore, the *IRIP* genes have undergone multiple rounds of gene duplication (Sandve *et al.* 2008). Average synonymous substitutions per synonymous site (*d*S) is assumed to be unaffected by natural selection, thus represents a measurement of relative time. The *d*S between *TaIRI2* (AAX81543) and other wheat *IRIP* paralogs in Sandve *et al.* (2008) can therefore be used as an estimate for time since initial expansion of the wheat IRIP gene family. The average dS between *TaIRI2* and all other wheat paralogs were 0.69 (±0.1), and between all wheat *IRIP* paralogs and the rice homolog 0.99. An *ad hoc* dating of the initial expansion event in the *IRIP* gene family can then be done by using the relationship of yearly synonymous substitution rate = *d*S/2T, where *T* is time since divergence from MRCA. By this method, the initial gene family expansion is dated to 26.4 (±2.64) Ma, if we assume the rice-pooid divergence at 43.7 Ma. This strongly suggest that a fourth expansion event in a gene family involved in low temperature stress might be associated with the E–O super-cooling period.

Conclusions

We have identified three Pooideae-specific gene families, *CBFIIIc/d*, *CBFIV*, and *FT*, which have undergone expansion events during the E–O cooling in a pooid ancestor. A fourth gene expansion event in the Pooideae-specific IRIP gene family has also likely occurred during the E–O cooling period. Moreover, the clustering of the initial expansion events close to, or inside E–O cooling period is not likely to be a chance events (*P*-values ranging from 0.034 to 0.14, depending on the calibration results used). Our results shed light on how Pooideae species might have evolved into a group of species well adapted to cold environments. We suggest a model of cold stress tolerance evolution in Pooideae where strong selection pressure on a pooid ancestor during the E–O cooling led to adaptive evolution of better cold tolerance through expansions of *CBFIIIc/d*, *CBFIV*, *FT* and *IRIP* gene familes (Fig. 5). This adaptive increase in gene content could have been linked to proximate gene dose effects or perhaps subsequent gene specialization, or neofunctionalization.

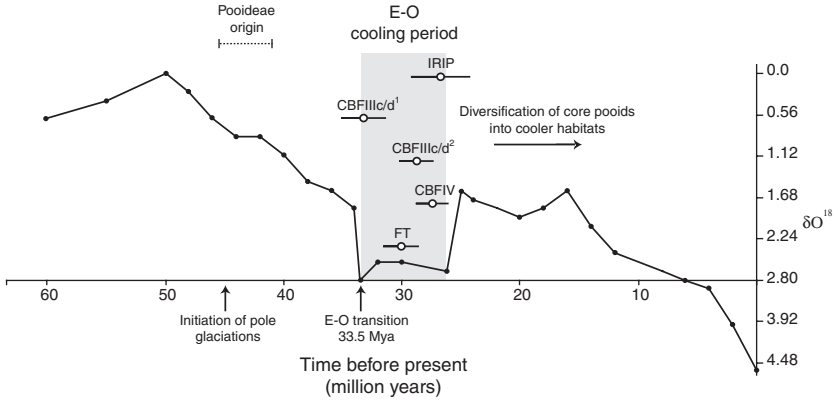


Fig. 5 Rise of gene families involved in low temperature stress tolerance in a palaeo-climatic perspective. White dots represent the average age of all estimates for *CBFIIIc/d*, *CBFIV*, *FT* and *IRIP* expansions common for the Pooideae lineage. CBFIIIc/d^1 and CBFIIIc/d^2 denotes the first and the second expansion event in the *CBFIIIc/d* clade, respectively. Black bars are standard deviations of the average age between all calibration methods. The estimate of initial IRIP expansion is calculated based on the synonymous distance between wheat IRIP paralogs (data collected from Sandve *et al.* 2008). Scale on *y*-axis is average global $\delta^{18}O‰$ values (a proxy for global temperature) and is extrapolated from Zachos *et al.* (2001).

Stromberg (2005) concluded from studying phytolith records in North America that the ecological expansion of pooids happened >10 Myr after the onset of E–O boundary. However, these data are geographically constricted to mid-latitude (~40–45°N) and thus cannot be used to draw conclusions about the timing of ecological expansion of pooids in more fringe localities. It is therefore possible that novel cool-climate adaptations allowed for ecological range expansion in fringe habitats which in turn led to taxonomic radiation of core pooids in cold environments. Further research into early palaeo-distribution of pooid species (40–20 Ma) in more extreme latitudes would therefore be of great value for the understanding of Pooideae evolution.

Acknowledgements

We would like to thank Magnus Dehli Vigeland, Lynn Huynh, and Caroline Stromberg, and three anonymous reviewers for valuable comments on the manuscript, and Christian Ritz for help and input on statistical issues.

References

Agarwal P, Agarwal P, Reddy M, Sopory S (2006) Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. *Plant Cell Reports*, **25**, 1263–1274.

Badawi M, Danyluk J, Boucho B, Houde M, Sarhan F (2007) The *CBF* gene family in hexaploid wheat and its relationship to the phylogenetic complexity of cereal *CBF*s. *Molecular Genetics and Genomics*, **277**, 533–554.

Barker N, Clark L, Davis J *et al.* (2001) Phylogeny and subfamilial classification of the grasses (Poaceae). *Annals of the Missouri Botanical Garden*, **88**, 373–457.

Battistuzzi FU, Filipski A, Hedges SB, Kumar S (2010) Performance of relaxed clock methods in estimating evolutionary divergence times and their credibility intervals. *Molecular Biology and Evolution*, doi: 10.1093/molbev/msq014.

Blanc G, Wolfe KH (2004) Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *The Plant Cell*, **16**, 1667–1678.

Bremer K (2002) Gondwanaland evolution of the grass alliance of families (Poales). *Evolution*, **56**, 1374–1387.

Chen Z, Cheng CHC, Zhang J *et al.* (2008) Transcriptomic and genomic evolution under constant cold in Antarctic notothenioid fish. *Proceedings of the National Academy of Sciences, USA*, **105**, 12944–12949.

Coxall HK, Pearson PN (2007) The Eocene–Oligocene transition. In: *Deep-Time Perspectives on Climate Change: Marrying the Signal from Computer Models and Biological Proxies* (eds Williams M, Haywood AM, Gregory FJ, Schmidt DN), pp. 351–387. The Geological Society, Bath.

Crepet WL, Feldman GD (1991) The earliest remains of grasses in the fossil record. *American Journal of Botany*, **78**, 1010–1014.

Crow KD, Wagner GP (2006) What is the role of genome duplication in the evolution of complexity and diversity? *Molecular Biology and Evolution*, **23**, 887–892.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **8**, 214.

Eldrett JS, Harding IC, Wilson PA, Butler E, Roberts AP (2007) Continental ice in Greenland during the Eocene and Oligocene. *Nature*, **446**, 176–179.

Eldrett JS, Greenwood DR, Harding IC, Huber M (2009) Increased seasonality through the Eocene to Oligocene transition in northern high latitudes. *Nature*, **459**, 969–973.

Fawcett JA, Maere S, Van de Peer Y (2009) Plants with double genomes might have had a better chance to survive the Cretaceous–Tertiary extinction event. *Proceedings of the National Academy of Sciences, USA*, **106**, 5737–5742.

Ferea TL, Botstein D, Brown PO, Rosenzweig RF (1999) Systematic changes in gene expression patterns following adaptive evolution in yeast. *Proceedings of the National Academy of Sciences, USA*, **96**, 9721–9726.

Ferrarri JA, Georghiou GP (1991) Quantitative genetic variation of esterase activity associated with a gene amplification in *Culex quinquefasciatus*. *Heredity*, **66**, 265–272.

Flagel LE, Wendel JF (2009) Gene duplication and evolutionary novelty in plants. *New Phytologist*, **183**, 557–564.

Fricano A, Rizza F, Faccioli P *et al.* (2009) Genetic variants of *HvCbf14* are statistically associated with frost tolerance in a European germplasm collection of *Hordeum vulgare*. *Theoretical and Applied Genetics*, **119**, 1335–1348.

Gaut B (2002) Evolutionary dynamics of grass genomes. *New Phytologist*, **154**, 15–28.

Gaut BS, Morton BR, McCaig BC, Clegg MT (1996) Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene rbcL. *Proceedings of the National Academy of Sciences, USA*, **71**, 10274–10279.

Gibson TA, Goldberg DS (2009) Questioning the ubiquity of neounctionalization. *Plos Computational Biology*, **5**, 1.

Gu X (2003) Evolution of duplicate genes versus genetic robustness against null mutations. *Trends in Genetics*, **19**, 354–356.

Gu Z, Steinmetz LM, Gu X, Scharfe C, Davis RW, Li W-H (2003) Role of duplicate genes in genetic robustness against null mutations. *Nature*, **421**, 63–66.

Hahn MW (2009) Distinguishing among evolutionary models for the maintenance of gene duplicates. *Journal of Heredity*, **100**, 605–617.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.

Hartley H (1973) Studies on the origin, evolution, and distribution of the Gramineae. V. The subfamily Festucoideae. *Australian Journal of Botany*, **21**, 201–234.

Hisano H, Kanazawa A, Kawakami A *et al.* (2004) Transgenic perennial ryegrass plants expressing wheat fructosyltransferase genes accumulate increased amounts of fructan and acquire increased tolerance on a cellular level to freezing. *Plant Science*, **167**, 861–868.

Hisano H, Kanazawa A, Yoshida M *et al.* (2008) Coordinated expression of functionally diverse fructosyltransferase genes is associated with fructan accumulation in response to low temperature in perennial ryegrass. *New Phytologist*, **178**, 766–780.

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754–755.

Ivany LC, Patterson WP, Lohmann KC (2000) Cooler winters as a possible cause of mass extinctions at the Eocene/Oligocene boundary. *Nature*, **407**, 887–890.

Janssen T, Bremer K (2004) The age of major monocot groups inferred from 800 *rbcL* sequences. *Botanical Journal of the Linnean Society*, **146**, 385–398.

Jobb G, von Haeseler A, Strimmer K (2004) TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evolutionary Biology*, **4**, 18.

Katoh K, Kuma K-i, Toh H, Miyata T (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, **33**, 511–518.

Kawakami A, Sato Y, Yoshida M (2008) Genetic engineering of rice capable of synthesizing fructans and enhancing chilling tolerance. *Journal of Experimental Botany*, **59**, 793–802.

Knox A, Li C, Vágújfalvi A *et al.* (2008) Identification of candidate *CBF* genes for the frost tolerance locus *Fr-A^{m}2* in *Triticum monococcum*. *Plant Molecular Biology*, **67**, 257–270.

Kondrashov FA, Kondrashov AS (2006) Role of selection in fixation of gene duplications. *Journal of Theoretical Biology*, **239**, 141–151.

Konstantinova T, Parvanova D, Atanassov A, Djilianov D (2002) Freezing tolerant tobacco, transformed to accumulate osmoprotectants. *Plant Science*, **163**, 157–164.

Lear CH, Elderfield H, Wilson PA (2000) Cenozoic deep-sea temperatures and global ice volumes from Mg/Ca in benthic foraminiferal calcite. *Science*, **287**, 269–272.

Lespinet O, Wolf YI, Koonin EV, Aravind L (2002) The role of lineage-specific gene family expansion in the evolution of eukaryotes. *Genome Research*, **12**, 1048–1059.

Li H-J, Yang A-F, Zhang X-C, Gao F, Zhang J-R (2007) Improving freezing tolerance of transgenic tobacco expressing sucrose: sucrose 1-fructosyltransferase gene from *Lactuca sativa*. *Plant Cell, Tissue and Organ Culture*, **89**, 37–48.

Liu Z, Pagani M, Zinniker D *et al.* (2009) Global cooling during the Eocene–Oligocene climate transition. *Science*, **323**, 1187–1190.

Livingston D, Hincha D, Heyer A (2009) Fructan and its relationship to abiotic stress tolerance in plants. *Cellular and Molecular Life Sciences*, **66**, 2007–2023.

Manly BFJ (2006) *Randomization, Bootstrap and Monte Carlo Methods in Biology*, 3rd edn. Chapman & Hall, London.

Moore RC, Grant SR, Purugganan MD (2005) Molecular population genetics of redundant floral-regulatory genes in *Arabidopsis thaliana*. *Molecular Biology and Evolution*, **22**, 91–103.

Moran K, Backman J, Brinkhuis H *et al.* (2006) The Cenozoic palaeoenvironment of the Arctic Ocean. *Nature*, **441**, 601–605.

Nguyen D-Q, Webber C, Ponting CP (2006) Bias of selection on human copy-number variants. *Plos Genetics*, **2**, 2.

Nuzhdin SV, Wayne ML, Harmon KL, McIntyre LM (2004) Common pattern of evolution of gene expressoin level and protein sequence in *Drosophila*. *Molecular Biology and Evolution*, **21**, 1308–1317.

Ohno S (1970) *Evolution by Gene Duplication*. Springer, New York.

Parvanova D, Ivanov S, Konstantinova T *et al.* (2004) Transgenic tobacco plants accumulating osmolytes show reduced oxidative damage under freezing stress. *Plant Physiology and Biochemistry*, **42**, 57–63.

Perry GH, Dominy NJ, Claw KG *et al.* (2007) Diet and the evolution of human amylase gene copy number variation. *Nature Genetics*, **39**, 1256–1260.

Prasad V, Stromberg CAE, Alimohammadian H, Sahni A (2005) Dinosaur coprolites and the early evolution of grasses and grazers. *Science*, **310**, 1177–1180.

Preston JC, Kellogg EA (2008) Discrete developmental roles for temperate cereal grass *VERNALIZATION1/FRUITFULL*-like genes in flowering competency and the transition to flowering. *Plant Physiology*, **146**, 265–276.

Quintanar A, Castroviejo S, Catalan P (2007) Phylogeny of the tribe Aveneae (Pooideae, Poaceae) inferred from plastid *trnT-F* and nuclear ITS sequences. *American Journal of Botany*, **94**, 1554–1569.

R Development Core Team (2009) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rifkin SA, Kim J, White KP (2003) Evolution of gene expression in the *Drosophila melanogaster* subgroup. *Nature Genetics*, **33**, 138–144.

Sanderson MJ (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution*, **14**, 1218–1231.

Sandve S, Rudi H, Asp T, Rognli O (2008) Tracking the evolution of a cold stress associated gene family in cold tolerant grasses. *BMC Evolutionary Biology*, **8**, 245.

Schardl CL, Craven KD, Speakman S *et al.* (2008) A novel test for host-symbiont codivergence indicates ancient origin of fungal endophytes in grasses. *Systematic Biology*, **57**, 483–498.

Seiffert ER (2007) Evolution and extinction of Afro-Arabian primates near the Eocene–Oligocene boundary. *Folia Primatologica*, **78**, 314–327.

Sidebottom C, Buckley S, Pudney P *et al.* (2000) Phytochemistry: heat-stable antifreeze protein from grass. *Nature*, **406**, 256–256.

Soreng RJ, Davis JI, Voionmaa MA (2007) A phylogenetic analysis of the Poaceae tribe Poeae *sensu lato* based on morphological characters and sequence data from three plastid-encoded genes: evidence for reticulation, and a new classification for the tribe. *Kew Bulletin*, **62**, 425–454.

Stebbins GL (1981) Coevolution of grasses and herbivores. *Annals of the Missouri Botanical Garden*, **68**, 75–86.

Stickley CE, St John K, Koc N *et al.* (2009) Evidence for middle Eocene Arctic sea ice from diatoms and ice-rafted debris. *Nature*, **460**, 376–379.

Stromberg CAE (2005) Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the Cenozoic of North America. *Proceedings of the National Academy of Sciences, USA*, **102**, 11980–11984.

Sutton F, Chen D-G, Ge X, Kenefick D (2009) *CBF* genes of the Fr-A2 allele are differentially regulated between long-term cold acclimated crown tissue of freeze-resistant and -susceptible, winter wheat mutant lines. *BMC Plant Biology*, **9**, 34.

Suyama M, Torrents D, Bork P (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, **34**, W609–612.

Tremblay K, Ouellet F, Fournier J, Danyluk J, Sarhan F (2005) Molecular characterization and origin of novel bipartite

cold-regulated ice recrystallization inhibition proteins from cereals. *Plant and Cell Physiology*, **46**, 884–891.

Wertheim JO, Sanderson MJ, Worobey M, Bjork A (2010) Relaxed molecular clocks, the bias-variance trade-off, and the quality of phylogenetic inference. *Systematic Biology*, **59**, 1–8.

Xu Z-S, Ni Z-Y, Li Z-Y *et al.* (2009) Isolation and functional characterization of *HvDREB1*—a gene encoding a dehydration-responsive element binding protein in *Hordeum vulgare*. *Journal of Plant Research*, **122**, 121–130.

Xue G-P, McIntyre C, Glassop D, Shorter R (2008) Use of expression analysis to dissect alterations in carbohydrate metabolism in wheat leaves during drought stress. *Plant Molecular Biology*, **67**, 197–214.

Yang Z, Nielsen R, Goldman N, Pedersen A-MK (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics*, **155**, 431–449.

Yang Z, Wong WSW, Nielsen R (2005) Bayes empirical inference of amino acid sites under positive selection. *Molecular Biology and Evolution*, **22**, 1107–1118.

Zachos J, Pagani M, Sloan L, Thomas E, Billups K (2001) Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, **292**, 686–693.

Zachos JC, Wara MW, Bohaty S *et al.* (2003) A transient rise in tropical sea surface temperature during the Paleocene–Eocene thermal maximum. *Science*, **302**, 1551–1554.

Zanazzi A, Kohn MJ, MacFadden BJ, Terry DO (2007) Large temperature drop across the Eocene–Oligocene transition in central North America. *Nature*, **445**, 639–642.

Zou C, Lehti-Shiu MD, Thomashow M, Shiu S-H (2009) Evolution of stress-regulated gene expression in duplicate genes of *Arabidopsis thaliana*. *PLoS Genetics*, **5**, e1000581.

S.R.S.'s research interests include molecular evolution, genomics, and association genetics. Lately his research has focused on different aspects of adaptations to cold climates in the grass sub-family Pooideae. S.F. is working on plant evolution and phylogeography as well as with genetic diversity and conservation of genetic resources.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1. Bayesian phylogeny including all CBF sequences included in the study

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Associated papers

Mol Breeding
DOI 10.1007/s11032-010-9479-0

SHORT COMMUNICATION

High-throughput genotyping of unknown genomic terrain in complex plant genomes: lessons from a case study

Simen Rød Sandve · Heidi Rudi · Guro Dørum · Paul Ragnar Berg · Odd Arne Rognli

Received: 21 August 2009 / Accepted: 25 June 2010

Abstract Novel high-throughput genotyping technologies have facilitated rapid genotyping of single nucleotide polymorphisms in non-model organisms. Most plant species have complex genomes with a large proportion of their genes having one or more paralogous copies due to single gene duplications and ancient or recent polyploidization events. These paralogous gene copies are potential sources of genotyping errors, and hence genotyping of plant genomes is inherently difficult. Here we present a case study that exemplifies paralog-related problems in high-throughput genotyping of plant genomes. We used the MassARRAY genotyping platform to genotype the *LpIRI* locus in *L. perenne* populations; this gene is thought to be involved in low-temperature stress tolerance. The dissection of the molecular genetics underlying the genotyping results provides a good example of how unknown paralogs can mask the true genotype of the locus, instructive to the non-specialist plant researcher and breeder.

Keywords Genotyping error · Paralogs · Plant · MassARRAY

Electronic supplementary material The online version of this article (doi:10.1007/s11032-010-9479-0) contains supplementary material, which is available to authorized users.

S. R. Sandve · H. Rudi · O. A. Rognli (✉)
Department of Plant and Environmental Sciences, Norwegian University of Life Sciences, Ås, Norway
e-mail: odd-arne.rognli@umb.no

G. Dørum
Department of Chemistry, Biotechnology, and Food Science, Norwegian University of Life Sciences, Ås, Norway

P. R. Berg
Centre for Integrative Genetics (CIGENE), Norwegian University of Life Sciences, Ås, Norway

Results

A genotyping case study

High-throughput genotyping (HTG) technologies like the Sequenom MA platform, and Illumina's GoldenGate and Infinium platforms, permit custom-designed HTG of single nucleotide polymorphism (SNPs) (Bray et al. 2001; Butler and Ragoussis 2008) suitable for non-model organisms with few genomic resources. However, using custom-designed HTG SNP assays for non-model species also comes with a cost: the uncertainty related to genotyping errors caused by unknown paralogs. The validity of the HTG results from "unknown genomic terrain" is sometimes difficult to assess without prior expectations of sequence variation, or without time-consuming and expensive sequencing validations of the genotyping results. The case study presented herein highlights this problem.

Published online: 11 July 2010

In the quest for novel genes associated with frost tolerance, we investigated allele frequencies of candidate genes in three experimental populations (high frost tolerance, low frost tolerance, and unselected) of the diploid forage grass species *Lolium perenne* L. The experimental populations originated from crosses between five unrelated parental genotypes from a European collection of *L. perenne* germplasm. High and low frost tolerance populations had been through two generations of divergent selection for frost tolerance and each selection round was followed by a single round of open pollination (random matings). The unselected population had been through two cycles of random mating. Thus all populations were expected to be in Hardy–Weinberg equilibrium (HWE). Thirty individuals from each population were genotyped.

MassARRAY genotyping of *LpIRI1*

The genotyping was conducted on the Sequenom mass spectrometry-based MassARRAY (MA) genotyping platform (Bray et al. 2001). One of the loci under investigation was *LpIRI1* (EU680848), a gene previously sequenced from a BAC clone. The allelic variation at the *LpIRI1* loci was investigated in all five parental genotypes prior to MA genotyping by direct sequencing of *LpIRI1* PCR products (primers can be found in Electronic Supplementary Material). Sequencing revealed SNP heterozygosity in all parental genotypes; hence we expected a maximum of ten haplotypes in the experimental populations. The primers used in the MA assay can be found in Electronic Supplementary Material. A selection of five SNPs was chosen for genotyping.

In the initial attempt to genotype *LpIRI1* SNPs, genomic DNA was used as a template. The genotype score plots showed clustering of genotype classes for most SNP loci (see example in Fig. 1a), as expected for a successful genotyping. However, there was evidence that the genotyping results did not reflect the true genotypes at the SNP loci. First, even at SNP loci with relatively well-clustered genotypes, some of the homozygous genotypes (GG cluster in Fig. 1a) tended to be located closer to the heterozygote than expected. The expected spatial distribution of genotypes in the MA genotype score plots is along the x and y axes for the homozygous genotypes and along the diagonal for the heterozygous genotypes (see Fig. 1c). Second, two out of the five SNP loci deviated from HWE ($P < 0.0001$), SNP1 having an excess of heterozygote genotypes while SNP2 had no heterozygote genotypes at all. The other three SNP loci had a small excess of heterozygote genotypes compared to the Hardy–Weinberg expectations but did not deviate significantly from expected

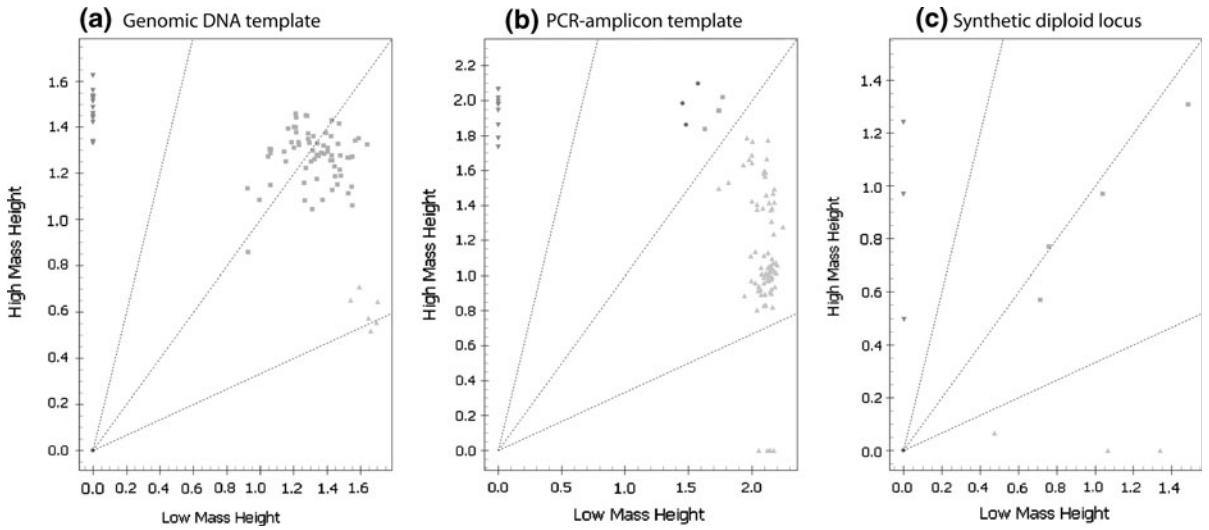


Fig. 1 MassARRAY (MA) genotyping results for *LpIRI1* SNP1. *Filled up triangle* = homozygous G-genotypes, *filled down triangle* = homozygous C-genotypes, *filled square* = heterozygous genotypes, and *filled circle* = no call (failed reaction). MA genotype scores from **a** assay performed on genomic template, **b** assay performed on PCR-amplicons, and **c** assay performed on a synthetic diploid locus made by combining different cloned haplotypes

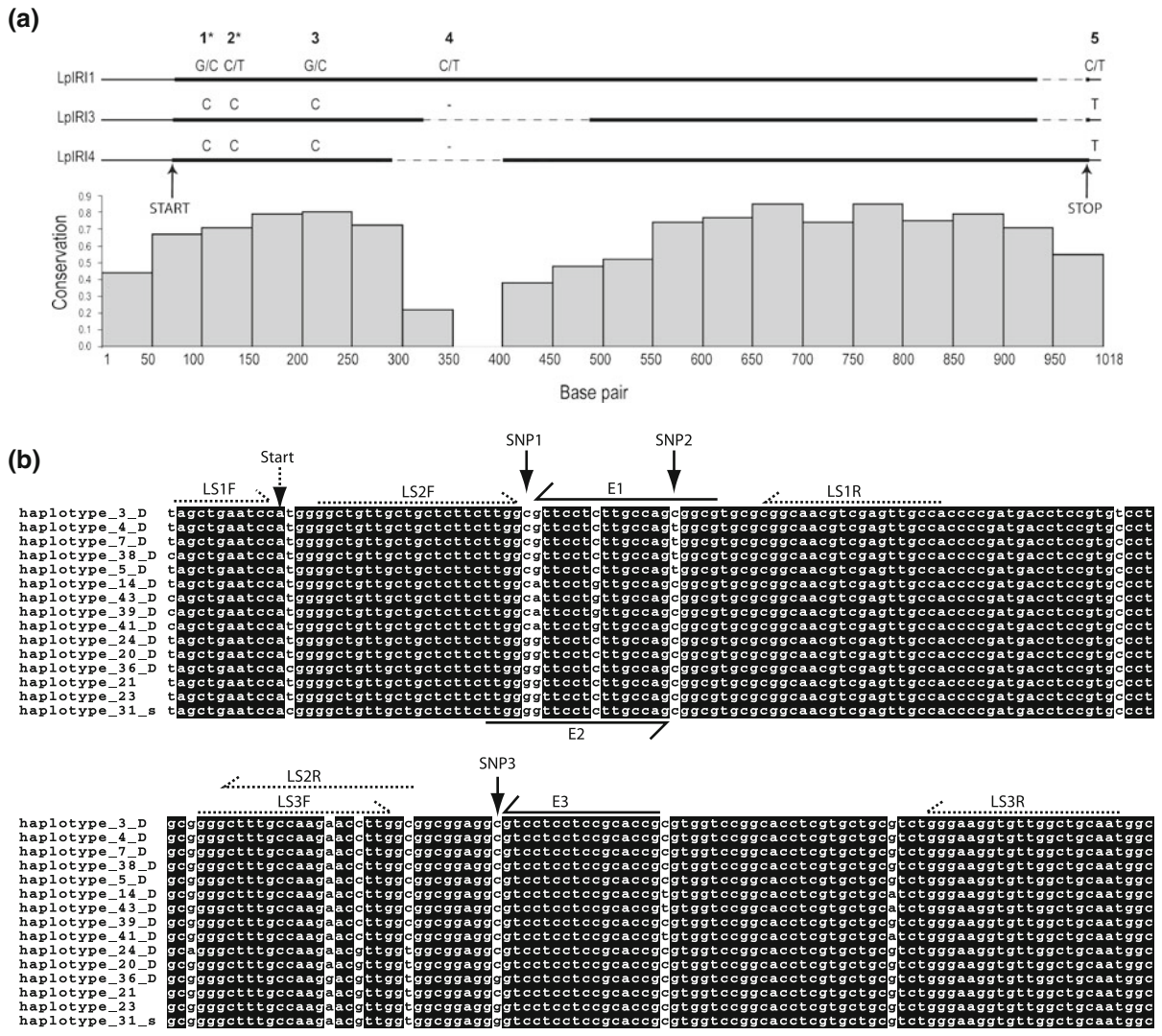


Fig. 2 Polymorphic sites and sequence conservation between *LpIRI* paralogs. **a** Degree of conservation between three distantly related *LpIRI*-paralogs. Conservation across the coding sequence is indicated by *grey bars*. Each *bar* represents the average degree of conservation across 50 base pairs. * Indicates loci that significantly deviated ($P < 0.0001$) from HWE. *Dotted lines* indicate large InDels (>50 bp). **b** Degree of conservation between the 15 nearly identical *LpIRI1*-like paralogs. *Black background* indicates 100% similarity. *Dotted black arrows* indicate the alignment of the SNP locus-specific (LS) primers for SNP 1–3. *Black arrows* indicate alignment of extension (E) primers for SNP 1–3

proportions. As a footnote, all SNPs in two additional genes (*LpPhyC* and *Lp6FT*) in the same population were in HWE, as expected from our experimental design (data not shown). Deviations from the neutral expectations for SNPs in *LpIRI1* but not in other genes indicate genotyping errors in the dataset.

Extensive BAC library exploration subsequently revealed two additional full-length paralogs, *LpIRI3* and *LpIRI4* (Sandve et al. 2008). An analysis of the homology between *LpIRI1*, *LpIRI3*, and *LpIRI4* revealed relatively high homology across the coding region, only interrupted by some InDels (Fig. 2a). None of the primers used in the MA assay aligned perfectly with *LpIRI3* or *LpIRI4* but for SNP1, for example, none of the mismatches between MA primers and *LpIRI3* was closer than four nucleotides from the 3′ end of the primers, i.e., where mismatches have the strongest preventive effect on unspecific amplification. Hence paralog mixing of *LpIRI3* in our MA assay could not be ruled out.

To circumvent any paralog-caused genotyping errors due to *LpIRI3*, *LpIRI4*, or any additional unknown paralogs, we then attempted to genotype the *LpIRI1*-SNPs using *LpIRI1* PCR amplicons as template for MA (PCR-MA). The forward primer was designed in the 5′ flanking region of *LpIRI1* which has no homology to either *LpIRI3* or *LpIRI4* (primers can be found in Electronic Supplementary Material). The amplicon was designed to be 550 bp long for sequencing purposes; hence only SNP loci 1–4 were included in the PCR-MA analysis. Surprisingly, the PCR-MA did not produce well-separated genotype classes (Fig. 1b). At SNP1, only four of the GG genotypes are classified as "perfect" GG genotypes (tight clustering along the *x*-axis), while all the other GG genotypes are poorly clustered and lie closer to the heterozygous genotypes in the plot. This pattern could theoretically be caused by (1) unequal allele amplification, i.e., less amplification of C alleles compared to G alleles, (2) a mix of G and C allele signals from unknown paralogs, or (3) DNA contamination and PCR amplification of the *LpIRI1* locus from several genotypes simultaneously. It is possible that unequal allele amplification contributed in some part to the poor clustering in the MA plot (Fig. 1a, b), but probably not to a large extent. If for example unequal allele amplification were highly influential in Fig. 1b, the GG genotypes close to the diagonal must in reality be GC genotypes. This is highly unlikely because the populations would then have a large excess of heterozygous genotypes compared to the Hardy–Weinberg expectations. Unspecific amplification of unknown paralog loci with C alleles would explain our results without having to assume that most GG genotypes are true heterozygotes.

Identification of *LpIRI1*-like nearly identical paralogs

In an attempt to determine the cause of the PCR-MA genotyping failure, we sub-cloned and sequenced one *LpIRI1*-specific PCR product per plant from an independent subsample of 40 genotypes in the experimental populations. Five clones from each genotype were randomly picked and sequenced and only haplotypes that occurred in a minimum of two independent PCR products were used in the analysis. Fifteen *LpIRI1*-like nearly identical haplotypes were identified (GenBank accession numbers GU201907–GU201921) (Fig. 2b). Many genotypes had more than two *LpIRI1* haplotypes and as many as five different *LpIRI1*-like haplotypes were found in one individual. Because we only sequenced five clones from each PCR product, the total number of *LpIRI1*-like haplotypes in our population is probably larger. The average identity between all nearly identical paralog (NIP) sequences was 98.6% (±0.06%) calculated as the average pairwise sequence identity across the entire alignment (455 nucleotides, not including a six-codon gap).

Maximum likelihood and Bayesian phylogenies estimated with Treefinder (Jobb et al. 2004) and MrBayes (Huelsenbeck and Ronquist 2001) supported a NIP haplotype phylogeny with strong support for two main haplotype groups (Fig. 3). Analyses of synonymous substitution divergence in MEGA (Tamura et al. 2007), using an IRIP-specific molecular clock (Sandve et al. 2008), suggest these groups diverged ~2 million years ago (Fig. 3). With two exceptions (haplotypes 20 and 36), the two NIP groups reflect the presence or absence of a three-codon deletion (Fig. 3). An evolutionary scenario where identical deletions arise independently in two loci within ~2 million years is highly unlikely. A better explanation to the incongruous relationship between topology and deletion is gene conversion or non-allelic homologous recombination, which is not uncommon between tightly linked paralogs (Lam and Jeffreys 2006; Storz et al. 2007). Tests for recombination were carried out in DNAsp (Rozas et al. 2003) and revealed that within haplotype group I one recombination event was predicted if haplotypes with the deletions are included in the analysis. However, no recombination events were predicted when the haplotypes with deletions were omitted from the analysis. Even though we can conclude that *LpIRI1*-like haplotypes belong to at least two loci, it is hard to determine the true number of *LpIRI1*-like NIPs without larger sequencing efforts.

MassARRAY on synthetic diploid genotypes

Next, we performed PCR-MA using the *LpIRI1*-like clones as template (1) to assess how genotype score plots of a bona fide diploid *LpIRI1* locus would appear, and (2) to confirm that the use of PCR amplicons per se (Fig. 1b) was not interfering with the genotyping. Ten combinations of four different cloned *LpIRI1*-like

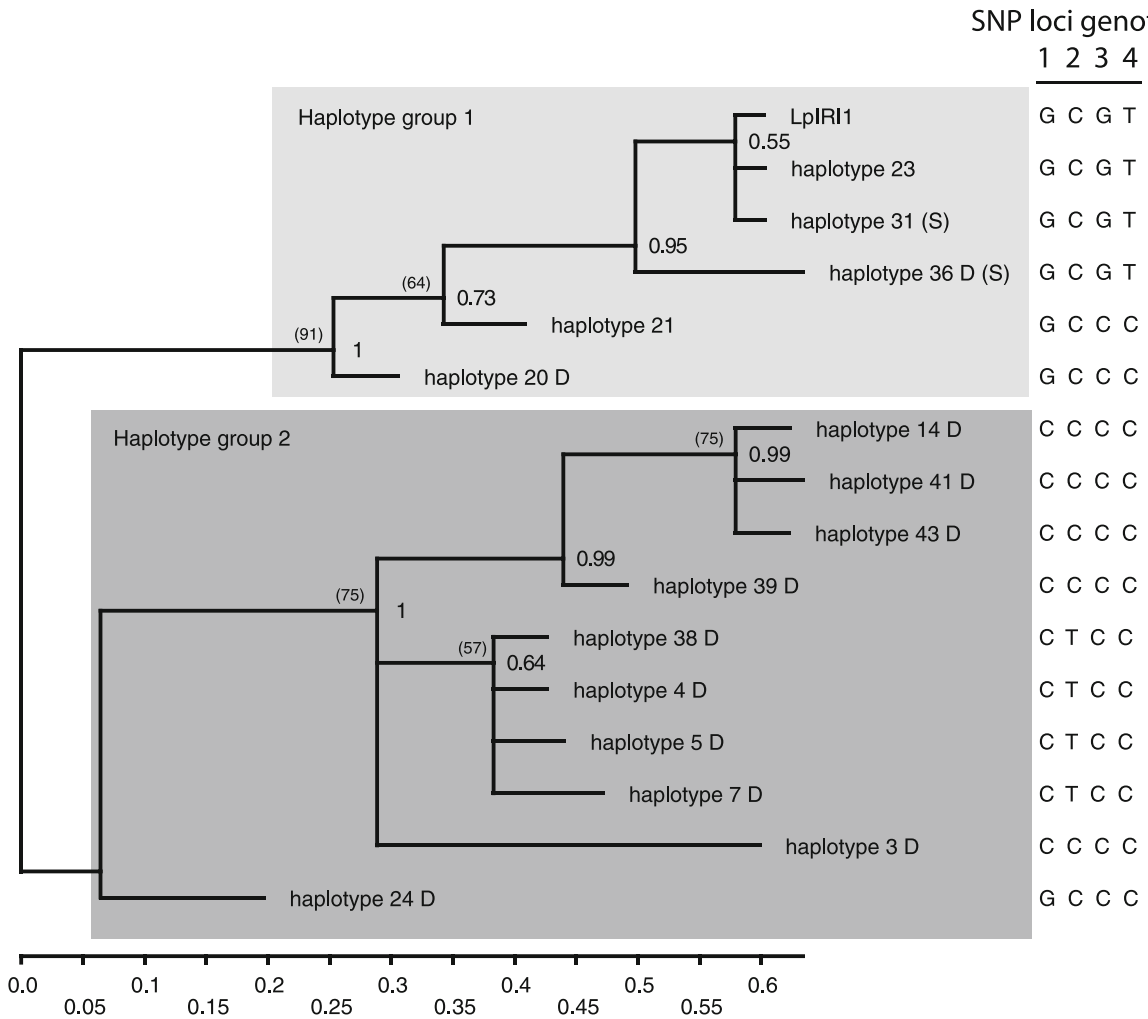


Fig. 3 Phylogeny of the *LpIRI1*-like nearly identical haplotypes. Bayesian phylogeny of *LpIRI1*-like haplotypes. *Numbers without brackets* at internal nodes represent Bayesian posterior probabilities, while *numbers in brackets* are the corresponding bootstrap values in percent (1,000 replicates) from the maximum likelihood phylogeny. D = haplotypes with a three-codon deletion, S = haplotypes with a premature STOP codon

haplotypes (haplotypes 3, 4, 21, and 24) were used. All synthetic diploid genotypes were correctly genotyped by PCR-MA, confirming that failure of the initial PCR-MA was not related to the use of PCR products per se. Furthermore, the synthetic diploids had genotype scores distributed perfectly along the *x*, *y*, or diagonal axes (Fig. 1c) and no tendency of "cloudy" clustering of heterozygote genotypes or shifts in the spatial location of homozygote genotypes, as seen in genomic-MA results (Fig. 1a). Mismatches between primer and template as close as four bases from the 3′ end did not influence genotyping results. It should be noted that we did not test all haplotypes in this experiment and it is possible that mismatches closer to the primer 3′ end can have added to the complexity of the poor MA-plot clustering (Fig. 1a, b). Lastly, genotyping data from the synthetic diploids provide indirect proof that genotypes in the populations carried *LpIRI1*-like haplotype combinations that must have affect the genotyping results. To mention one example, genotype 36 from the low frost tolerance population contained at least three *LpIRI1*-like haplotypes (3, 23, and 24). This combination would give a 2:1 ratio of G:C and C:G at SNP1 and SNP3, respectively, and thus result in a MA-plot genotype that lies in between the homozygous and heterozygous.

Discussion

What was to be a simple genotyping experiment turned into an extensive study that included BAC library screening, BAC sequencing, clone-based sequencing, and multiple rounds of HTG, before we were able to resolve the molecular genetics underlying the SNP genotyping results. Our example might be an extreme one in the sense that the *LpIRI1*-like paralogs are almost identical. However, we do think it is important to stress that unknown paralogs can pose a serious threat to genotyping accuracy in non-sequenced genomes.

Minimizing genotyping errors in HTG

Minimizing the probability of genotyping errors caused by polymorphisms between paralogous loci is crucial. This encompasses careful experimental planning which involves selection of SNP loci to be genotyped, thorough bioinformatics analysis, and post-analysis data clean-up. When dealing with non-sequenced genomes, a major issue of HTG is lack of knowledge of existing paralogs. To maximize the information basis of the HTG experiment, a BLAST search against all fully sequenced plant genomes and expressed sequence tag (EST) databases of closely related species should be done to determine whether the locus of interest belongs to a gene family. The size of gene families varies between species (Rensing et al. 2008); therefore the existence of paralogs in one genome does not accurately predict the presence of paralogs in another. Even so, a BLAST identification of paralogs in related plant species can help flag loci that can potentially contribute to genotyping errors.

Figure 4 shows the expected paralog sequence divergence for different evolutionary scenarios using a discrete version of the codon-based Markov model of Yang et al. (2000) and the synonymous substitution rate of 6.5×10^{-9} (Gaut et al. 1996). The results shed light on the importance of which SNPs to choose to minimize the probability of unspecific paralog genotyping signals. In our case study we attempted to genotype an intronless gene belonging to a gene family. If we assume that all codons on average are under purifying selection, the entire gene-length of intronless genes is diverging at a very slow rate (Fig. 4). Thus, genotyping of SNPs in coding region of paralogs are much more likely to result in

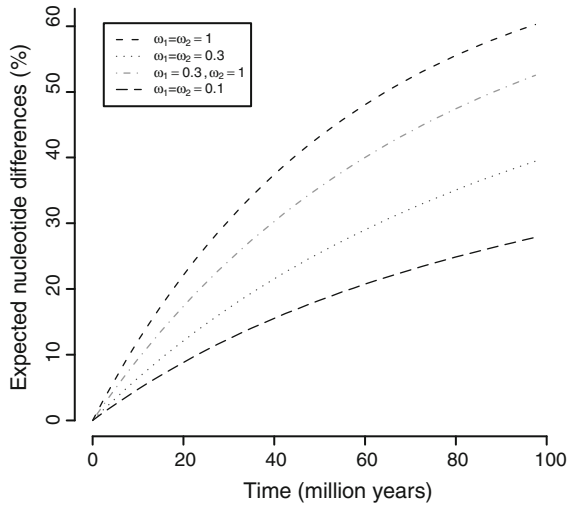


Fig. 4 Expected nucleotide differences accumulated between two paralogs over time. Different evolutionary scenarios are represented, starting from the *top curve*: Scenario 1, Two non-functional paralog sequences (e.g., intron sequences), $\omega_1 = \omega_2 = 1$; Scenario 2, intermediate purifying selection for both paralogs, $\omega_1 = \omega_2 = 0.3$; Scenario 3, one gene under intermediate purifying selection while the other evolves as a pseudogene, $\omega_1 = 0.3$, $\omega_2 = 1$; Scenario 4, both paralogs under strong purifying selection, $\omega_1 = \omega_2 = 0.1$. ω is the ratio of non-synonymous to synonymous substitutions in a gene

genotyping errors. In hindsight, genotyping SNPs in flanking non-coding regions of *LpIRI1* could have been a better strategy for avoiding paralog-related genotyping errors. Although genotyping SNPs in flanking regions might provide better locus specificity, one should be aware that this simultaneously increases the chance of genotyping errors caused by unequal allele amplification due to higher allelic variation.

A study of paralog age distribution in the genomes of maize, cotton, wheat, lettuce, sunflower, barley, ice plant, potato, soybean, rice, and *Arabidopsis thaliana* shows that the age distribution has an exponential decay, with most paralogs being younger than 60 million years (Blanc and Wolfe 2004). Based on our paralog divergence model, 60 million years translates to an expected divergence level of $\sim$50% for non-coding freely evolving sequences (scenario 1, Fig. 4). However, many genes belonging to gene families are not evolving as neutral sequences but are under selective constraints, and thus the average expected divergence level for two paralogs in a plant genome could be significantly lower than 50%

(Fig. 4). A potentially much bigger problem arises when HTG is put to the test in recent polyploid species. For example, in hexaploid wheat most genes have three homoeologous copies which diverged about 2.5–4.5 million years ago (Huang et al. 2002). Even if two out of three homoeologs had evolved as neutral loci, the expected level of differences between the homoeologs would still be very low.

Furthermore, if the locus of interest is known to be part of a gene family, PCR-amplicon genotyping could provide a solution. But as our study showed, if nearly identical paralogs exist, even a PCR-MA could be unreliable. Thus, a thorough validation of the specificity of the PCR through cloning and sequencing of PCR products is recommended before any HTG is carried out on amplicons from genes belonging to a gene family.

Lastly, data quality control is important before allowing genotyping results to be included in any final analysis. Testing for deviations from expected segregation can reveal SNP loci with genotyping errors caused by paralogs (or other genotyping error sources), even when raw data from HTG appear normal (Fig. 1a). On the other hand, exclusion of SNPs that deviate from HWE should not be carried out without a critical examination of the SNP. This could lead to unfortunate rejections of biologically significant SNP loci that deviate from HWE because of selection. In this case study, however, the populations that were used had undergone one generation of random mating and were expected to be in HWE.

Does the choice of HTG platform matter?

One very interesting question that arises is: does the choice of HTG platform matter with regard to paralog related genotyping errors? SNP genotyping in soybean has provided some insights into the performance of MA and GoldenGate platforms in complex genomes. In soybean, which has recently undergone a whole genome duplication (Blanc and Wolfe 2004), the paralog-related genotyping failure rate was estimated to be 23% for PCR-based STS genotyping targeting the 3′ region of ESTs (Choi et al. 2007). The same authors later designed a high-throughput MA assay for SNPs located in those EST sequences that produced reliable STS genotypes (unpublished data referred to in Hyten et al. 2008) and 30% of these SNP assays did not produce a reliable assay, most likely due to paralog mixing. Hyten et al. (2008) genotyped 384 soybean SNPs using the GoldenGate assay and reported that only 10% of the SNP assays failed. However, because the genotyped plants were from recombinant inbred line mapping populations, the authors were able to successfully genotype more SNPs than if the mapping populations had been F2 or backcross populations (Hyten et al. 2008). In heterozygous out-crossing populations the frequency of SNPs which would have been classified as failed assays due to poor genotype clustering would have increased substantially. Identical conclusions have been drawn in recent studies on the performance of HTG in polyploid wheat genomes. Bérard et al. (2009) and Akhunov et al. (2009) showed that SNPlex, MassARRAY, and GoldenGate could be used with great success in inbred wheat lines, but they also concluded that HTG in heterozygous genomes would be extremely challenging, and much less successful, due to interference from homoeologs.

Different genotyping technologies vary in the way SNPs are detected and scored, hence these differences might affect the propensity to paralog-related genotyping errors. The two major differences between the two platforms discussed above, GoldenGate and MA, are the locus-specific amplification step and the different genotype detection technologies implemented. In the GoldenGate assay the locus-specific PCR step is performed with universal primers, while the MA assay uses custom-made locus-specific primers. The result of this could be a higher chance of unequal allele amplification in the MA assay, but this ultimately depends on the prior knowledge of sequence variation at that locus in the population under investigation. The second big difference between MA and GoldenGate is how they translate molecular signals (SNP genotypes) into genotype calls; GoldenGate uses fluorescence signals while MA uses the mass of a single base extension reaction to score the genotypes. This step, and downstream software handling of these genotype signals, could play a role. A side-by-side comparison of the performance of HTG platforms in paralog-rich genomes would therefore be a valuable study that could provide the plant research community with guidelines for choosing the HTG technology best suited for their purpose.

Acknowledgments We thank Kjetil Fosnes for technical assistance, and Odd-Arne Olsen, Torben Asp, and reviewers for valuable comments on the manuscript. We thank Magnus Dehli Vigeland for help with statistical modelling. The Bayesian phylogenetic analysis was carried out through the Bioportal server at University of Oslo. This work was funded through the KMB project *Festulolium* with Improved Forage Quality and Winter Survival for Norwegian Farming, project number 173319/I10, funded by the Research Council of Norway and Graminor AS.

References

Akhunov E, Nicolet C, Dvorak J (2009) Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. Theor Appl Genet 119: 507–517

Bérard A, Le Paslier MC, Dardevet M, Exbrayat-Vinson F, Bonnin I, Cenci A, Haudry A, Brunel D, Ravel C (2009) High-throughput single nucleotide polymorphism genotyping in wheat (*Triticum* spp.). Plant Biotechnol J 7:364–374

Blanc G, Wolfe KH (2004) Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. Plant Cell 16:1667–1678

Bray MS, Boerwinkle E, Doris PA (2001) High-throughput multiplex SNP genotyping with MALDI-TOF mass spectrometry: practice, problems and promise. Hum Mutat 17:296–304

Butler H, Ragoussis J (2008) BeadArray-based genotyping. In: Starkey J, Elaswarapu R (eds) Methods in molecular biology, vol 439: Genomics protocols, 2nd edn. Humana Press Inc., Totowa, NJ, pp 53–74

Choi I-Y, Hyten DL, Matukumalli LK, Song Q, Chaky JM, Quigley CV, Chase K, Lark KG, Reiter RS, Yoon M-S, Hwang E-Y, Yi S-I, Young ND, Shoemaker RC, van Tassell CP, Specht JE, Cregan PB (2007) A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. Genetics 176:685–696

Gaut BS, Morton BR, McCaig BC, Clegg MT (1996) Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene rbcL. Proc Natl Acad Sci USA 93:10274–10279

Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R, Gornicki P (2002) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. Proc Natl Acad Sci USA 99: 8133–8138

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755

Hyten D, Song Q, Choi I-Y, Yoon M-S, Specht J, Matukumalli L, Nelson R, Shoemaker R, Young N, Cregan P (2008) High-throughput genotyping with the GoldenGate assay in the complex genome of soybean. Theor Appl Genet 116:945–952

Jobb G, von Haeseler A, Strimmer K (2004) TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. BMC Evol Biol 4:18

Lam K-WG, Jeffreys AJ (2006) Processes of copy-number change in human DNA: the dynamics of α-globin gene deletion. Proc Natl Acad Sci USA 103:8921–8927

Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud P-F, Lindquist EA, Kamisugi Y, Tanahashi T, Sakakibara K, Fujita T, Oishi K, Shin-I T, Kuroki Y, Toyoda A, Suzuki Y, Hashimoto S-I, Yamaguchi K, Sugano S, Kohara Y, Fujiyama A, Anterola A, Aoki S, Ashton N, Barbazuk WB, Barker E, Bennetzen JL, Blankenship R, Cho SH, Dutcher SK, Estelle M, Fawcett JA, Gundlach H, Hanada K, Heyl A, Hicks KA, Hughes J, Lohr M, Mayer K, Melkozernov A, Murata T, Nelson DR, Pils B, Prigge M, Reiss B, Renner T, Rombauts S, Rushton PJ, Sanderfoot A, Schween G, Shiu S-H, Stueber K, Theodoulou FL, Tu H, Van de Peer Y, Verrier PJ, Waters E, Wood A, Yang L, Cove D, Cuming AC, Hasebe M, Lucas S, Mishler BD, Reski R, Grigoriev IV, Quatrano RS, Boore JL (2008) The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. Science 319:64–69

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496–2497

Sandve SR, Rudi H, Asp T, Rognli OA (2008) Tracking the evolution of a cold stress associated gene family in cold tolerant grasses. BMC Evol Biol 8:245

Storz JF, Baze M, Waite JL, Hoffmann FG, Opazo JC, Hayes JP (2007) Complex signatures of selection and gene conversion in the duplicated globin genes of house mice. Genetics 177:481–500

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Yang Z, Nielsen R, Goldman N, Pedersen A-MK (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155:431–449

Plant Science xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Review

Molecular mechanisms underlying frost tolerance in perennial grasses adapted to cold climates

Simen R. Sandve[a], Arkadiusz Kosmala[b], Heidi Rudi[a], Siri Fjellheim[a], Marcin Rapacz[c], Toshihiko Yamada[d], Odd Arne Rognli[a,*]

[a] Department of Plant and Environmental Sciences, Norwegian University of Life Sciences, N-1432 Ås, Norway
[b] Institute of Plant Genetics Polish Academy of Sciences, Strzeszynska 34, 60-479 Poznan, Poland
[c] Department of Plant Physiology, Agricultural University in Krakow, Podluzna 3, 30-239 Krakow, Poland
[d] Field Science Center for Northern Biosphere Hokkaido University Kita 11, Nishi 10, Kita-ku, Sapporo 060-0811, Japan

ARTICLE INFO

Article history:
Received 10 March 2010
Received in revised form 13 July 2010
Accepted 17 July 2010
Available online xxx

Keywords:
Cold acclimation
Frost tolerance
CBF
Photosynthetic acclimation
AFP
Fructans

ABSTRACT

We review recent progress in understanding cold and freezing stress responses in forage grass species, notably *Lolium* and *Festuca* species. The chromosomal positions of important frost tolerance and winter survival QTLs on *Festuca* and *Lolium* chromosomes 4 and 5 are most likely orthologs of QTLs on *Triticeae* chromosome 5 which correspond to a cluster of CBF-genes and the major vernalization gene. Gene expression and protein accumulation analyses after cold acclimation shed light on general responses to cold stress. These responses involve modulation of transcription levels of genes encoding proteins involved in cell signalling, cellular transport and proteins associated with the cell membrane. Also, abundance levels of proteins directly involved in photosynthesis were found to be different between genotypes of differing frost tolerance levels, stressing the importance of the link between the function of the photosynthetic apparatus under cold stress and frost tolerance levels. The significance of the ability to undergo photosynthetic acclimation and avoid photoinhibition is also evident from numerous studies in forage grasses. Other interesting candidate mechanisms for freezing tolerance in forage grasses are molecular responses to cold stress which have evolved after the divergence of temperate grasses. This includes metabolic machinery for synthesis of fructans and novel ice-binding proteins.

Contents

1. Introduction

Plants in cold climates are frequently exposed to sub-zero temperatures in the autumn, winter, and spring seasons. Exposure to sub-zero temperatures requires a battery of molecular and physiological adaptations to minimize frost related injuries which potentially can be fatal. Cold acclimation (CA) is a process whereby plants in response to low but non-freezing temperatures undergo

* Corresponding author at: Department of Plant and Environmental Sciences, Norwegian University of Life Sciences, P.O. Box 5003, N-1432 Ĺs, Norway.
E-mail addresses: simen.sandve@umb.no (S.R. Sandve), akos@igr.poznan.pl (A. Kosmala), heidi.rudi@umb.no (H. Rudi), siri.fjellheim@umb.no (S. Fjellheim), rrrapacz@cyfronet.pl (M. Rapacz), yamada@fsc.hokudai.ac.jp (T. Yamada), odd-arne.rognli@umb.no (O.A. Rognli).

0168-9452/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.plantsci.2010.07.011

a range of biological changes in order to increase their frost tolerance (FT) and prepare for the winter season [1]. The process of CA encompasses biological modifications on many levels, e.g. modulation of gene expression levels [2], accumulation and degradation of proteins [3–5] changes in sugar content [6], and changes in the photosynthetic machinery [7].

A large part of the research concerning plant cold stress response and FT has been carried out using the model species, *Arabidopsis thaliana* (named herein as *Arabidopsis*) and *Oryza sativa* (rice) [1]. As a consequence, the research on low-temperature stress responses in non-model species has focussed on genetic mechanisms, which originally were identified in the model species and have been conserved between plant lineages over hundreds of millions of years. In many cases a direct inference of gene function based on homology between model dicot plants and agriculturally important species is elusive. Moreover, neither *Arabidopsis* nor rice is adapted to a perennial life in extreme winter climates. This is important because adaptation to a perennial life history in harsh winter climates must have required changes at the genetic level which cannot be studied using an annual model species. Hence if we only use model plant species to investigate the molecular basis of cold and frost stress response this might provide limited insights into the genetic mechanisms underlying these traits in important agricultural species.

The *Pooideae* grasses (temperate grasses) is a large and economically important sub-family including cereals (*Triticeae* tribe) and forage grasses (*Poaeae* tribe). Divergence of temperate grasses from the most recent common ancestor shared with rice is thought to have happened ~46–42 million years ago (Mya) [8,9]. Parallel to the origin and early evolution of the *Pooideae* group the global climate became gradually cooler [10]. As opposed to rice, which is adapted to warm and humid environments, *Pooideae* grasses radiated in cooler environments [11]. This is reflected by the present distribution of *Pooideae* species which is extremely skewed towards cooler environments [12]. Thus evolution of cold and frost stress responses, either through fine tuning of ancient abiotic stress responses or evolution of novel adaptations to cold environments must have been central for the *Pooideae* sub-family.

The evolutionary history of temperate grasses makes them an excellent model system for studying adaptations to cold and frost stress. During the last decade several research groups have focussed their research on understanding the cold and freezing stress responses in forage grass species (*Poeae* tribe), mainly *Lolium* and *Festuca* species, and recently also *Phleum pratense* L. Species of the *Poeae* tribe are excellent models for plant adaptations to cold environment because of their adaptation to habitats in the northernmost part of the Northern hemisphere, i.e. the circumpolar arctic region. The research on the *Poeae* species includes mapping of quantitative trait loci (QTL), transcriptomics, proteomics, and functional gene studies, and plant physiology research of low-temperature (CA) response and FT. We do not intend for this review to serve as an elaborate review of every aspect of CA and FT but aim to summarize the progress in CA and FT research on forage grasses from the past decade. We will start by highlighting results from studies at the -omics level, and then focus on three specific mechanisms involved in CA and FT where research performed on forage grasses has made significant contributions; (1) photoinhibition, (2) ice interacting proteins, and (3) fructan synthesis.

2. QTL mapping and genomics

Winter survival (WS) is a very complex trait determined by combinations of frost, desiccation, water logging, ice-encasement, anoxia, and snow cover. However, FT is the single component that generally explains most of the variation in WS [13]. There are few reports of Quantitative Trait Locus (QTL) mapping of FT and WS in forage grass species. Mapping of frost and drought tolerance QTLs using regrowth tests, and WS QTLs based on field survival, in the 'B14/16 × HF2/7' full-sib family of meadow fescue (*Festuca pratensis* Huds.) was reported by Alm [14] and Alm et al. unpublished results. Major QTLs for FT/WS were located on chromosomes 1F, 2F, 5F and 6F, and for drought tolerance traits on chromosomes 1F, 3F, 4F, and 5F. In many cases the QTLs co-located with genes involved in abiotic stress tolerance, e.g. orthologs of barley *Dhn* (dehydrin) genes. QTLs for several stress tolerance traits mapped to the same regions on *Festuca* chromosomes 1F, 4F and 5F. Frost and drought both induce cell dehydration which induces many stress-responsive genes via the DREB1A/CBF-transcription factors [1,15] (Fig. 1). Co-location of WS QTLs with drought tolerance QTLs indicates that dehydration is the major factor behind these QTLs, while co-location with FT QTLs indicate frost tolerance as the main factor. WS QTLs not co-located with any component stress factors are most probably caused by genes affecting seasonal adaptation, e.g. photoperiodic sensitivity. Coincidence of QTLs provided thus information about the roles of frost and dehydration in the genetic control of WS, and illustrates the advantage of mapping several stress tolerance traits in the same family. Two FT/WS QTLs on chromosome 5F most likely correspond to *Fr-A1* [16] and *Fr-A2* [17] on wheat homoeologous group 5A, while a small QTL for FT on chromosome 4F, located at the position of *FpVRN1*, an ortholog of the wheat *VRN1* gene [18], most likely is a pleiotropic effects of vernalization and/or a photoperiodic genes as in wheat [19,20].

A single QTL for electrical conductivity (i.e. an indirect measurement for FT) was detected on linkage group (LG) 4 in the p150/112 mapping population of perennial ryegrass (*Lolium perenne*) [21]. Comparison with QTLs on *Triticeae* homoeologous chromosomes 4 was rather inconclusive due to lack of common markers, but the authors discussed whether the location corresponded to the vernalization/FT (*VRN1/Fr1*) region on homoeologous chromosomes 5 of wheat [16]. No QTL for WS was detected in this family. Identification of QTLs controlling FT and WS in an annual (*Lolium multiflorum*) × perennial (*L. perenne*) ryegrass interspecific hybrid population was reported by Xiong et al. [22]. They measured FT by ion leakage and electrical conductivity following natural cold acclimation in the field in two consecutive years. The two methods gave very similar results. WS was recorded by visual scoring in the field in two springs. Two QTLs associated with WS on linkage groups 4 and 5, and one QTL for FT on LG 5 were consistently detected over several years and different maps (male and female). They proposed that the QTL for WS on LG 4 could correspond to the QTL for electrical conductivity reported by Yamada et al. [21] and might be caused by the action of an ortholog of the vernalization gene of wheat and barley. The other consistent and overlapping QTLs detected for WS and FT on LG 5 by Xiong et al. [22] were proposed to be orthologous to the QTL for FT detected in a homoeologous position in wheat and barley chromosomes 5, similar to the situation in *Festuca*.

Twenty CBF-genes have been identified in barley (*Hordeum vulgare*), of which 11 are found in two tight tandem clusters on the long arm of chromosome 5H in the same region as the *Fr-H2* frost resistance locus [23,24]. An orthologous genomic region in *Triticum monococcum* contains similar CBF gene clusters located at the *Fr-A^{m}2* frost resistance QTL [17,25]. In *Lolium perenne* five *LpCBF* genes were identified of which four mapped tightly on LG5 in a position orthologous to the CBF-clusters in wheat and barley [26]. One CBF gene (*FpCBF6*), which is induced and has peak expression 2 h after start of cold acclimation, has been identified in meadow fescue [Alm *et al.*, unpublished results]. Based on phylogenetic analysis, *FpCBF6* was placed in the subclade CBF3-3a [27] together with *OsDREB1A/CBF3, HvCBF6, TaCBF6, LpCBF3 and FaDREB1A.* It maps on meadow fescue chromosome 5F and co-locate within the major FT/WS QTL that most likely is orthologous to the *Fr-H2* and *Fr-A^{m}2* frost resistance QTLs in barley and wheat, respectively [Alm *et al.*, unpublished results].

Please cite this article in press as: S.R. Sandve, et al., Molecular mechanisms underlying frost tolerance in perennial grasses adapted to cold climates, Plant Sci. (2010), doi:10.1016/j.plantsci.2010.07.011

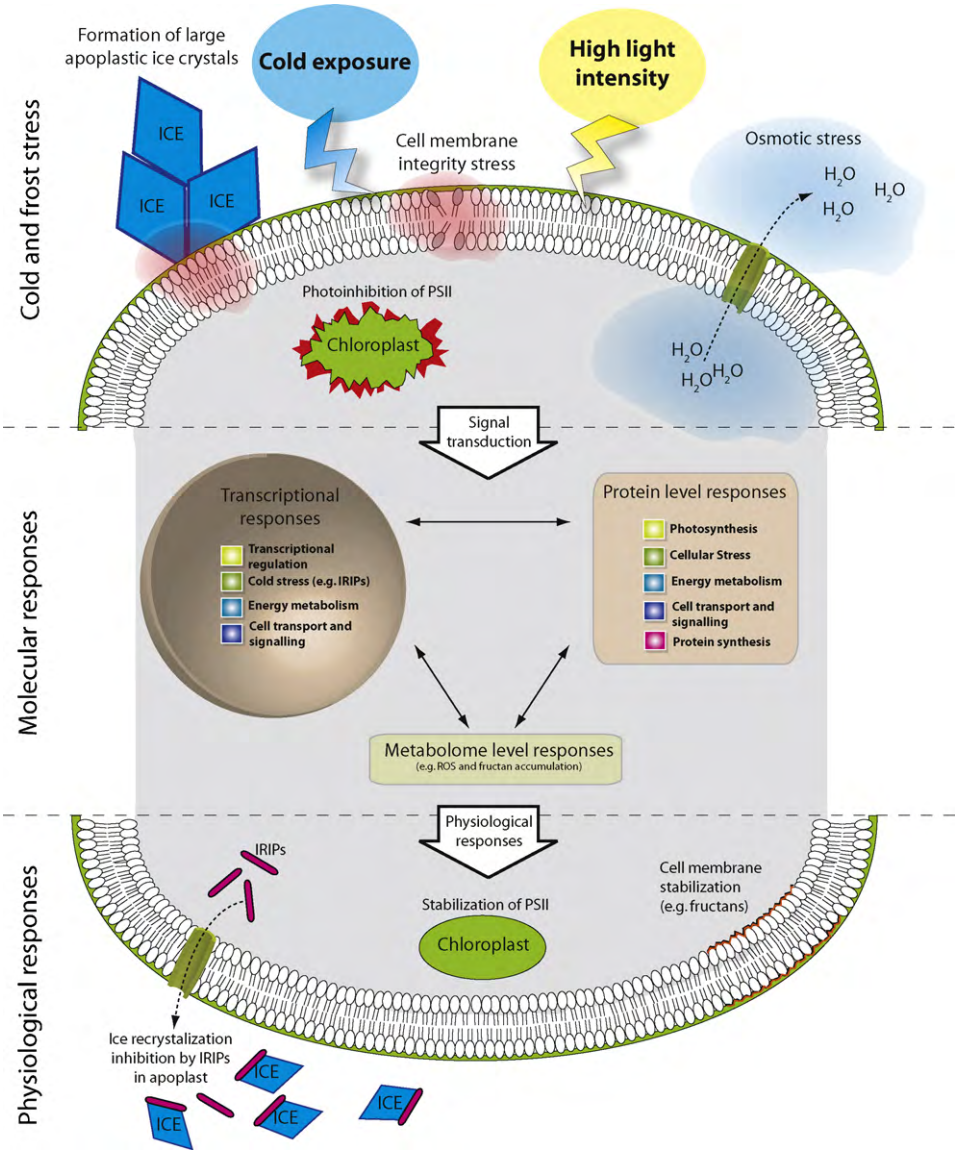


Fig. 1. Summary model of molecular mechanisms important for frost tolerance in perennial grasses. The figure does not depict all mechanisms involved in frost tolerance but those discussed in this review. (1) *Top panel:* cold temperatures in combination with high light intensity result in various stresses and physiological responses. Ice formation in apoplast can lead to direct physical stress on the cell membrane but also osmotic stress due to incorporation of water molecules into growing ice crystals. High light intensity combined with low temperature during autumn and early winter trigger photoinhibition of the chloroplast. (2) *Middle panel:* molecular responses to cold stress during autumn and early winter encompass large scale modulations of gene expression levels and abundance of proteins responsible for further transcriptional regulation and signalling (e.g. CBFs), energy metabolism and modification of the photosynthetic machinery, and proteins or enzymes that directly or indirectly help stabilize cell membrane integrity (e.g. IRIPs and fructosyl transferases). (3) *Lower panel:* known and presumed physiological responses to avoid cellular destruction during cold and freezing stress, and secure survival of perennial grasses through the winter.

The coincident location of several of the QTLs in *Festuca* and *Lolium* with QTLs and genes in *Triticeae* species indicate the action of structural or regulatory genes that are conserved across evolutionarily distant species [28]. In this respect CBF-transcription factors and dehydrin genes regulating the expression of cold and drought regulated genes, and the vernalization response genes appear to play decisive roles. The major structural difference between the *Festuca/Lolium* and *Triticeae* homoeologous chromosomes 4 and 5 [29,30] is especially interesting in comparative genomics of frost tolerance. The major vernalization gene *VRN1* and the FT QTLs/*CBF*-gene clusters are both located on homoeologous chromosomes 5 in *Triticeae*, while in *Festuca/Lolium* they are on different chromosomes, i.e. (*VRN1*) on chromosome 4 and (FT QTLs/*CBF*-genes) on chromosome 5. This makes it feasible to separate the indirect effects of the major vernalization gene on CA and FT from the effects of genes on chromosome 5 directly involved in developing frost tolerance.

3. Transcriptomics of cold acclimation and frost tolerance

Comprehensive research on transcriptional modulation during CA has been carried out in *Arabidopsis* using different microarray technologies and different statistical criteria. These studies have estimated the number of *Arabidopsis* genes being regulated by CA to be in the order of 2–13% [31–34]. Recently, CA transcriptional responses have also been investigated in cold tolerant Pooideae species such as winter and spring wheat (*Triticum aestivum*) cultivars [35,36], barley (*H. vulgare*) [37], *L. perenne* [38], and *F. pratensis* [39]. Only two of these studies [36,37] surveyed the entire genome and the results showed that Pooideae grasses have a comparable

Please cite this article in press as: S.R. Sandve, et al., Molecular mechanisms underlying frost tolerance in perennial grasses adapted to cold climates, Plant Sci. (2010), doi:10.1016/j.plantsci.2010.07.011

number (~9–12%) of cold regulated genes to that of *Arabidopsis*. However, these two studies were carried out on cereal species (wheat and barley) hence we can only assume that this global picture of the amount of transcriptional changes during CA is similar in forage grasses. The overall picture of the type of transcriptional changes occurring during CA in Pooideae grasses is also very similar to what has been found in *Arabidopsis* (see above mentioned references). Genes involved in signaling, regulation of transcription, cellular transport, cell membrane, and genes with putative protective roles during freezing-related stress are up-regulated, while genes involved in metabolism, respiration and photosynthesis are down-regulated [35–39].

More interestingly is therefore what we have learned from studying CA transcriptional responses in Pooideae grasses that is different or novel compared to the *Arabidopsis* studies. One aspects of CA transcriptional response which cannot be studied in *Arabidopsis* is the difference in cold regulated genes between over-wintering perennial crown tissue and the annual leaf tissue. In wheat strikingly little overlap was observed (1–10%) between genes which underwent transcriptional changes in leaf compared to crown tissue during CA [36]. A similar trend is also evident in forage grasses. For example, genes involved in photosynthesis is common in CA-EST libraries from leaves [38] but almost absent in crown tissue [39]. This does not imply that CA processes in leaves are not important for winter survival of perennial tissues, in fact leaves are very important for production of CA-associated metabolites such as sugars; it simply means that understanding biologically relevant transcriptional responses to cold stress in agriculturally important Pooideae grasses is difficult if the focus is on annual non-grass model plants only. Another important lesson from studying CA transcriptomes of Pooideae grasses is the abundance of Pooideae-specific cold regulated genes. This includes for example genes encoding novel ice interacting proteins and novel enzymes involved in sugar metabolism [37–39] (treated in greater detail in the last part of this review).

In *Arabidopsis* the genetic basis for differences in frost tolerance has been investigated by studying variation in transcriptional response during CA using inbred mutant lines [33]. This type of experiment is unfortunately not feasible in Pooideae forage grasses due to their self-incompatibility and outbreeding nature. However in Pooideae cereals, where inbreeding is possible, such experiments have been carried out. chloroplast development pathway mutants, Svensson et al. demonstrated that only 11% of all cold responsive genes maintained normal regulation during CA in plants with non-functioning chloroplasts [37]. This emphasizes the importance of photosynthesis-related molecular mechanisms during CA and acquirement of frost tolerance in Pooideae grasses.

An alternative to experiments with inbred mutant genotypes is to study genetic differences between genotypes selected for divergent frost tolerance, and this has been carried out in forage grasses. By analyzing contrasting gene expression patterns after long-term CA (19 days) between high frost tolerant (HFT) and low-frost tolerant (LFT) genotypes of *F. pratensis*, Rudi et al. [39] identified potential candidate genes underlying the differences in frost tolerance. Surprisingly only 7 (1.3%) of the genes studied were >2-fold differentially expressed between the HFT and LFT genotypes. This result could imply that the difference in FT levels is not causally linked to large global variation in transcriptional responses during CA but instead determined by either (i) transcriptional differences in a few important genes or (ii) post-transcriptional and post-translational mechanisms. However we must stress that Rudi et al. [39] studied transcription differences after 19 days of CA; hence variation in early transient transcriptional regulation would not be detected in this study.

Three out of the seven genes found to differ in expression between high- and low-frost tolerant genotypes had significant blast results to a tumor-suppressor protein, a phosphate/phosphate translocator, and a protein disulfide isomerise [39]. Based only on homology these three genes are putatively involved in regulation of translation, oxidative protein folding and metabolic regulation, respectively. There are evidence from other studies to support the involvement of the tumor-repressor homolog and the phosphate/phosphate translocator homolog in freezing stress. Recently, the tumor-suppressor protein homolog has been cloned in *Caragana jubata*, a temperate plant that grows under extreme cold and at higher altitudes in Himalaya, and this gene was shown to be up-regulated after low-temperature treatments [40]. Further, sucrose flux through sucrose biosynthesis has been shown to modify development of freezing tolerance in *Arabidopsis* [41]. Because phosphate/phosphate translocators play a vital role in the sucrose biosynthesis pathway, increased expression of this gene would be expected to influence sucrose metabolism and therefore possibly also frost tolerance. This assumes of course that the same or a similar molecular mechanism is conserved in monocot grasses.

4. Frost tolerance and proteomics

Even though there is a hierarchical, and to some extent deterministic, relationship between the transcriptome and the proteome response, gene expression levels and protein levels are in many cases not strictly correlated (e.g. [42]). Hence, genome-wide transcription analyses do not offer the complete picture of plant molecular responses during CA and cold stress. It is therefore important to complement gene expression studies with studies of proteomic changes under CA and investigate how changes in protein abundance during CA correlate with differences in FT.

Many CA-proteomic studies have been carried out in *Arabidopsis* [43–46] and rice [47–51]. Even though rice belongs to the grass family the results from the studies in rice are only partly transferable to Pooideae grasses since rice cannot cold acclimate and develop tolerance to frost. Some investigations have been conducted on cereal species with CA-ability but these studies are mainly restricted to studying a few genes, e.g. *Wsc120* in *T. aestivum* [52] and *Cor14b* in *H. vulgare* [7,53]. Only one comprehensive study of proteome responses to CA has been carried out in Pooideae forage grass [4]. In this study *F. pratensis* genotypes with distinct levels of FT were selected for comparative analyses of leaf protein accumulation before and after 2, 8, 26 h, and 3, 5, 7, 14 and 21 days of CA. High-throughput two-dimensional electrophoresis (2-DE) in combination with electrospray ionization mass spectrometry was used to study how the proteome of the different genotypes changed in response to CA. Comparisons between HFT and LFT plants revealed a total of 41 (5.1%) proteins which showed a minimum of 1.5-fold difference in abundance during the CA process. The largest differences in protein abundance (28.1%) appeared relatively early, most often on the 2nd day of CA. At the time point when maximal levels of FT and maximal difference in FT between HFT and LFT individuals were reached (the 21st day of CA), 10 out of the 41 proteins (24%) had >1.5 fold differences in abundance between HFT and LFT.

The majority of differentially accumulated *F. pratensis* proteins were proteins which are directly involved in photosynthesis, stressing the importance of the link between the function of the photosynthetic apparatus under cold stress and FT levels (Fig. 1). Several of the *F. pratensis* proteins identified had not been reported before to possibly be involved in the development of CA and FT, even in model species. This group includes chloroplast-localized Ptr ToxA binding protein 1, globulin 2, 50S ribosomal protein L10 from chloroplasts, 30S ribosomal protein S10 from chloroplasts, ADP (adenosine diphosphate) – glucose pyrophosphatase, and ADP-ribosylation factor 1 [4]. A direct comparison of the results obtained for *F. pratensis* with those described by others, e.g. for *Arabidopsis*

Please cite this article in press as: S.R. Sandve, et al., Molecular mechanisms underlying frost tolerance in perennial grasses adapted to cold climates, Plant Sci. (2010), doi:10.1016/j.plantsci.2010.07.011

[43–46] is unfortunately not possible due to different methodological approaches of the studies. In the case of *F. pratensis* the protein abundance of HFT and LFT plants were compared before CA and at eight different time points of CA while in the *Arabidopsis* studies the abundance of particular proteins of non-acclimated and cold acclimated plants was mainly analyzed (e.g. [43]). Also an extract of total leaf proteins from *F. pratensis* was used, whereas in *Arabidopsis* protein extracts derived from different cell compartments, including the nucleus [44], plasma membranes [46] and chloroplasts [45], were applied.

Kosmala et al. [4] reported that degradation of the oxygen-evolving enhancer protein 1 was observed during CA of *F. pratensis*. Studies in rice have shown degradation of the Rubisco large subunit, Rubisco activase, sedoheptulose-1, 7-bisphosphatase, PS II oxygen-evolving complex protein 2, ATP (adenosine triphosphate) synthase alpha chain, and ATP synthase CF1 beta chain during low-temperature treatment e. It is well known that the ability of cells to adapt to new environmental conditions requires a rapid reconstruction of existing regulatory pathways. Thus, protein degradation is one of the most essential components of plant responses to environmental stimuli [54].

In another recent study Kosmala et al. [5] analyzed the protease activity after 3 weeks of CA in two *F. pratensis* genotypes with high and low T_{EL50} values (T_{EL50} = temperature of 50% survival determined by electrical conductivity). A differences in the protease activity of HFT and LFT genotypes appeared after 2 h of cold acclimation and the activity of the proteases was significantly higher in the LFT genotype than the HFT genotype at almost all the studied time points during 21 days of CA. After 3 weeks of CA the protease activities increased 3-fold in LFT and 2-fold in the HFT plant compared to the activities detected before CA [5]. However, at this time point it would be difficult to determine which of the identified proteolytic activities actually was associated with cold acclimation and differentially expressed FT among the *F. pratensis* genotypes.

The proteomic approach presented herein for *F. pratensis* should be further extended to cover a wider part of the proteome. Furthermore, integration of plant responses to cold and frost stress on the transcriptome, proteome, and metabolome levels, similar to that described for *Arabidopsis* (e.g. [55–57]) is necessary to fully understand the mechanisms underlying genetic differences in FT among genotypes of forage grass species.

5. Photoinhibition avoidance in cold and freezing tolerance

Photoinhibition is the process whereby light energy absorbed in the photosynthetic light processes exceeds energy demand of the dark processes which leads to PSII over-reduction and subsequent inhibition of the photosynthetic capacity (Fig. 1). This can result not only in destruction of the photosynthetic apparatus but also to damage of whole cells due to production of reactive oxygen species accompanying PSII over-reduction [58]. Under low temperatures in winter and spring, the photosynthetic fixation of CO_2 is very limited, a condition which can cause photoinhibition even under relatively low irradiance [59]. The tolerance to this cold-induced photoinhibition seems to be closely related to freezing tolerance, a relationship which is partially due to common mechanisms of acclimation to both stresses as demonstrated in winter wheat and rye (*Secale cereale* L.) [60]. PSII over-reduction may also act as one of the signals triggering gene expression of cold regulated genes involved in CA as shown in rye [61] (Fig. 1). Reduction of PSII capacity must be carefully controlled by the plant to avoid processes that destroy photosystems and cells due to the overproduction of reactive oxygen species under photoinhibition [62–64]. As a consequence photosynthetic acclimation to photoinhibitory conditions are also induced under cold [60].

Two main strategies of acclimation of the photosynthetic apparatus are found in higher plants during cold hardening [60,65]. The first strategy, called the photochemical mechanism, depends on increasing the energy demand by increasing carbon assimilation and carbon metabolism [60]. This strategy is observed especially in over-wintering herbaceous plants which are not able to accumulate sufficient amounts of photoassimilates during the summer. Accumulation of high amounts of photoassimilates in autumn and early winter acts as an energy source for the cold acclimation process [60,66]. The second strategy relies on an intensification of protective non-photochemical mechanisms that harmlessly dissipate excess excitation energy as heat. The non-photochemical mechanism is the most important in equilibrating between absorption and utilization of light energy [63] and has been shown to be important also in herbaceous plants when low temperatures generate a strong photoinhibitory pressure [67].

Perennial forage grasses are very interesting model systems for studying mechanisms of photosynthetic acclimation to cold. Firstly, acclimation of the photosynthetic apparatus to high light-low temperature conditions is essential for perennial forage grasses to develop winter hardiness and FT. Laboratory experiments and field experiment over multiple winters has showed that the capacity for photosynthetic acclimation was correlated with genotypic differences in WS [68]. Furthermore, genotypes with impaired photosynthetic acclimation were shown to be unable to undergo proper CA and increase their freezing tolerance [68]. However the correlation between winter hardiness (or freezing tolerance) and the ability of photosynthetic acclimation was ~0.7, indicating that in the case of some plants, freezing tolerance and tolerance to cold-induced photoinhibition are at least partially independent mechanisms [68]. Secondly, there seems to be a link between FT levels and the type of photosynthetic acclimation which grasses undergo. Studies on androgenic genotypes of allotetraploid *Festulolium* cultivars (*F. pratensis* (4×) × *L. multiflorum* (4×)) demonstrated that plants with higher winter hardiness in the field, and higher FT in the laboratory were also more tolerant to cold-induced photoinactivation of PSII. The higher tolerance to photoinhibition was mainly due to increased non-photochemical dissipation of excess energy, however in one of the androgenic genotypes reduced non-photochemical dissipation was compensated by an increased electron transport (i.e. photochemical mechanism) [68]. Further studies have shown that perennial *F. pratensis* genotypes with high FT has increased dissipation capacity compared to biennial *L. multiflorum* genotypes in which only increased quenching of photochemical energy was observed [69].

Experiments with *F. pratensis* × *L. multiflorum* hybrids illustrate that transfer of the non-photochemical mechanism from *Festuca* to *Lolium* results in increasing frost tolerance [68,69]. More detailed studies of the genomic composition of *Festuca* × *Lolium* hybrids have given some indications of which chromosomal regions that harbour the underlying genes. Hybrid genotypes with a *L. multiflorum* genetic background and a chromosome 4 introgression from *F. pratensis* were more frost tolerant than pure *Lolium* genotypes [69] and this is probably due to changes in the non-photochemical mechanism of photosynthetic acclimation. In a chromosome substitution population in which each *L. perenne* chromosome has been replaced by its *F. pratensis* homologue [70], the chromosome 4-substitution line has both higher expression of non-photochemical mechanism of the photosynthetic acclimation and the highest freezing tolerance [69].

Some evidence also exists for the importance of adaptation to photoinhibition under low temperatures in natural populations of *Poeae* species. Photosynthetic acclimation to cold conditions is thought to play a role in the altitudinal range distribution of *Festuca* species in Babia Gora Mountain in Poland. In Babia Gora *Festuca rubra* distribution is limited to lower altitude habitats compared

Please cite this article in press as: S.R. Sandve, et al., Molecular mechanisms underlying frost tolerance in perennial grasses adapted to cold climates, Plant Sci. (2010), doi:10.1016/j.plantsci.2010.07.011

to species of mountainous fescues (*Festuca versicolor* and *Festuca supina*) which have more efficient non-photochemical mechanisms of photosynthetic acclimation to low temperatures than *F. rubra* [71].

6. *Pooideae* lineage specific genes: ice interacting proteins and fructan metabolism

Several gene families, or sub-classes of gene families, are known to be lineage specific for the *Pooideae* grasses, for example sub-clades of the CBF-transcription factor family [72], the fructocyl transferase (FST) gene family [73], and one *Pooideae*-specific ice re-crystalization inhibition protein coding gene family [9]. It has been speculated if these *Pooideae*-specific gene family expansions have been important for adaptation of a common *Pooideae* ancestor to cold climates and subsequently ecological expansion of *Pooideae* into cooler habitats [74]. In the following sections we will focus on IRIP and FTS genes and their importance for FT in *Pooideae* grasses. Since relatively little research has been done on CBFs in forage grasses we recommend a recent review on grass CBFs in cereals [75] for further discussion on this topic.

6.1. Ice interacting proteins

Anti-freeze proteins (AFP), i.e. proteins that modify the water freezing process, have evolved independently in many organisms [76] and are thought to be an important adaptation to a life in frost-exposed environments. AFPs can be classified into two groups according to their function; (1) AFPs that affect the temperature at which water freezes (thermal hysteresis), and (2) AFPs that are ice re-crystallization inhibition proteins (IRIP). Only the latter type is thought to be important in plant AFP action [77]. Ice re-crystallization is the process where larger ice crystals grow at the expense of smaller crystals. However, when IRIPs are present under freezing conditions ice re-crystallization is inhibited, forcing new small crystals to form rather than an increase in the crystal size. It is thought that controlled ice crystal growth provides protection from physical damage caused by ice formation in the apoplastic space (Fig. 1).

The first report of ice-binding proteins in plants came from experiments with rye (*S. cereale*) [78]. Subsequently many plant AFPs of different evolutionary origin have been discovered [79,80]. In forage grasses the first report of an ice interacting protein were in a study by Sidebottom et al. [80]. The authors reported the discovery of a cold-induced mRNA transcript encoding a partial ice re-crystallization inhibition protein (IRIP) isolated from *L. perenne* (LpAFP, AJ277399). Several IRIP homologs have subsequently been isolated from forage grass species and it is now clear that they belong to a Pooideae-specific multi-gene family [9,81–83]. Protein modelling and functional protein studies have shown that NxVxG/NxVxxG repeat motifs forms a beta-roll with a beta-sheet ice-binding face that can bind to ice and is responsible for ice re-crystallization inhibitory action *in vitro* [84,85]. The grass IRIP genes encode a signal peptide that targets the proteins to the apoplast [9], and apoplastic export of IRIPs have been shown to happen *in vivo* in transgenic *Arabidopsis* plants [82].

Interestingly the ice interacting domain of IRIPs is not homologous to proteins in rice, *Arabidopsis*, or any other plant outside the *Pooideae* sub-family but have probably evolved through repeat-motif expansion events after the *Pooideae* divergence [9]. Synthetic AFPs have been shown to enhance FT, but the importance of IRIPs in determining FT in forage grasses have yet to be demonstrated. However, recently *Arabidopsis* plants expressing *L. perenne* IRIP genes were shown to display improved cell membrane stability during freezing and increased FT [86]. This strongly supports the notion that IRIP genes are important for survival in frost-exposed environments.

6.2. Fructan synthesis

Fructans are soluble fructosyloligosaccharides derived from sucrose, which are used for carbohydrate storage in some (~15%) flowering plants, including the cereals and forage grasses [87]. It has long been recognized that temperate grasses accumulate fructans in response to seasonal drop in temperature in the field and under exposure to cold temperatures in the laboratory [65,73,88–92]. Yet historically, the direct functional significance of fructan synthesis under low-temperature stress and the role of fructan synthesis in adaptation to cold climates have been controversial [93]. Research on forage grasses from the past decade have provided evidence that fructan molecules themselves have a direct effect on freezing and low-temperature stress tolerance [94,95]. The question today is not *if* but *how* fructans help increase the FT.

The study of fructan biosynthesis under cold temperatures has been important to understand the role of fructans in cold stress. Fructan sugars are synthesised by four fructosyl transferase (FST) enzymes [94]. The initial step in fructan synthesis is the conversion of two sucrose molecules to a fructan molecule by the 1-Sucrose-sucrose-fructosyltransferases (1-SST). Further polymerization are then carried out by other FSTs, i.e. Fructan:fructan 1-fructosyltransferases (1-FFT), Fructan:fructan 6G-fructosyltransferases (6G-FFT), and Sucrose:fructan 6-fructosyltransferases (6-SFT), resulting in five possible types of fructan sugars [see [94] for review]. Several studies on cereals and forage grasses have demonstrated that FST gene expression levels are regulated by exposure to low temperatures [89,96]. However, the most elaborate study on expression of FST genes under CA was done by Hisano et al. [73] in *L. perenne*. The authors investigated the expression patterns of multiple FST genes during the CA process and found that it was consistent with the proposed model of fructan synthesis pathway. Expression levels of genes involved in the initial steps of fructan synthesis (1-SST and 6G-FFT) increased shortly after transferring the plants to cold conditions, while genes encoding enzymes that elongate the fructan chains (putative 6-SFT) increased their expression levels throughout the cold acclimation process. The correlation between FT gene expression and fructan sugar accumulation under CA has been demonstrated through several experiments [73,97]. Tamura et al. [98] reported that the PpFT1 gene from timothy (*P. pratense*) is a novel gene with unique enzymatic properties that differ from previously cloned plant 6-SFTs, and is involved in the synthesis of highly polymerized levans in timothy. Accumulation of large amounts of carbohydrates as fructans with high degree of polymerization by *PpFT1* and other enzymes might be involved in the superior winter hardiness of timothy.

The direct functional role of fructans on cold stress tolerance is thought to be linked to increased cell membrane stabilization (Fig. 1). Fructans have been shown to interact with model cell membranes [99,100] and change their stability under freezing [101]. A pioneering study using transgenic plants have provided evidence that fructans stabilize cell membranes *in vivo* under freezing. Hisano et al. [6] made transgenic *L. perenne* plants that overexpressed wheat 1-SST and 6-SFT genes and showed that increased fructan synthesis reduced electrolyte leakage (i.e. reduced cell membrane damage) during freezing. Similar results have later been replicated in transgenic tobacco expressing a lettuce 1-SST gene [102]. Two additional studies using transgenic plants with increased fructan content have showed increased freezing tolerance in tobacco [103] and increased chilling tolerance in rice [104], however, these studies did not measure cell membrane damage.

Please cite this article in press as: S.R. Sandve, et al., Molecular mechanisms underlying frost tolerance in perennial grasses adapted to cold climates, Plant Sci. (2010), doi:10.1016/j.plantsci.2010.07.011

7. Concluding remarks

Perennial forage grasses are essential elements of sustainable and multifunctional farming systems providing both feed for ruminants and ecosystem services, e.g. carbon sequestration, soil formation and protection, nutrient cycling, and aesthetic landscape values. This is especially the case in high-latitude and high-altitude temperate regions where perennial grasses make up a very large part of the agricultural land. Climate changes are predicted to give more unpredictable and unstable winter climates with more cycles of freezing-thawing and less long-lasting snow cover. These climates pose new combinations of winter stresses, and in many cases increasing stress, for perennial plants in many regions.

In this review we have summarized present knowledge on important molecular mechanisms underlying cold and freezing stress survival in perennial forage grasses (Fig. 1). Even though the last decade of research have disclosed important molecular systems involved in winter survival and adaptation to cool environments we have still a long way to go to understand frost tolerance at the whole plant level. Important topics to address in future research include: how does plants sense temperature changes – are there one universal or several different termosensor systems? We need to understand more about the detailed role of carbohydrate composition and turnover at the whole plant level during different types of winter climates; how do plants regulate respiration during winter, and how is photosynthetic activity controlled at low light intensity during warmer periods of the winter at higher latitudes, or at low light intensity under snow cover? We also lack understanding of the reversion from a cold acclimated state to a growth state, the de-hardening process, and how this is regulated at the molecular level. Repeated cycles of hardening/de-hardening/re-hardening during winters are likely outcomes of the predicted climate changes affecting the northern hemisphere. A better understanding of these processes is vital in order to predict and mitigate the effects of climate change and develop forage grass cultivars with superior combinations of winter survival, dry matter production and nutritive quality.

References

[1] M.F. Thomashow, Plant cold acclimation: freezing tolerance genes and regulatory mechanisms, Annu. Rev. Plant Physiol. Plant Mol. Biol. 50 (1999) 571–599.
[2] S.J. Gilmour, D.G. Zarka, E.J. Stockinger, M.P. Salazar, J.M. Houghton, M.F. Thomashow, Low temperature regulation of the *Arabidopsis* CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression, Plant J. 16 (1998) 433–442.
[3] C.L. Guy, Cold acclimation and freezing stress tolerance: role of protein metabolism, Annu. Rev. Plant Physiol. Plant Mol. Biol. 41 (1990) 187–223.
[4] A. Kosmala, A. Bocian, M. Rapacz, B. Jurczyk, Z. Zwierzykowski, Identification of leaf proteins differentially accumulated during cold acclimation between *Festuca pratensis* plants with distinct levels of frost tolerance, J. Exp. Bot. 60 (2009) 3595–3609.
[5] A. Kosmala, A. Bocian, Z. Zwierzykowski, Activities of proteolytic enzymes during cold acclimation of *Festuca pratensis*, in: P. Krajewski, P. Kachlicki, B. Naganowska (Eds.) Genetics and Genomics in Crop Improvement – From a Model Plant to a New Cultivar, Treatises and Monographs IPG PAS, Poznań, 2009.
[6] H. Hisano, A. Kanazawa, A. Kawakami, M. Yoshida, Y. Shimamoto, T. Yamada, Transgenic perennial ryegrass plants expressing wheat fructosyltransferase genes accumulate increased amounts of fructan and acquire increased tolerance on a cellular level to freezing, Plant Sci. 167 (2004) 861–868.
[7] M. Rapacz, B. Wolanin, K. Hura, M. Tyrka, The effects of cold acclimation on photosynthetic apparatus and the expression of COR14b in four genotypes of barley (*Hordeum vulgare*) contrasting in their tolerance to freezing and high-light treatment in cold conditions, Ann. Bot. 101 (2008) 689–699.
[8] B. Gaut, Evolutionary dynamics of grass genomes, New Phytol. 154 (2002) 15–28.
[9] S. Sandve, H. Rudi, T. Asp, O. Rognli, Tracking the evolution of a cold stress associated gene family in cold tolerant grasses, BMC Evol. Biol. 8 (2008) 245.
[10] J. Zachos, M. Pagani, L. Sloan, E. Thomas, K. Billups, Trends, rhythms, and aberrations in global climate 65 Ma to present, Science 292 (2001) 686–693.
[11] N. Barker, L. Clark, J. Davis, M. Duvall, G. Guala, C. Hsiao, E. Kellogg, H. Linder, R. Mason-Gamer, S. Mathews, M. Simmons, R. Soreng, R. Spangler, Phylogeny and subfamilial classification of the grasses (Poaceae), Ann. Mo. Bot. Gard. 88 (2001) 373–457.
[12] H. Hartley, Studies on the origin, evolution, and distribution of the gramineae. V. The subfamily Festucoideae, Aust. J. Bot. 21 (1973) 201–234.
[13] S. Pulli, K. Hjortsholm, A. Larsen, B. Gudleifsson, S. Larsson, B. Kristiansson, L. Hömmö, A.M. Tronsmo, P. Ruuth, C. Kristensson, Development and Evaluation of laboratory Testing Methods for Winterhardiness Breeding, Publications – Nordic Gene Bank, 1996, pp. 1–68.
[14] V. Alm, Comparative genome analysis of meadow fescue (*Festuca pratensis* Huds.): genetic linkage mapping and QTL analyses of frost and drought tolerance. Agricultural University of Norway, 2001.
[15] K. Shinozaki, K. Yamaguchi-Shinozaki, Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways, Curr. Opin. Plant Biol. 3 (2000) 217–223.
[16] G. Galiba, S.A. Quarrie, J. Sutka, A. Morgounov, J.W. Snape, RFLP mapping of the vernalization (Vrn1) and frost resistance (Fr1) genes on chromosome 5A of wheat, TAG Theor. Appl. Genet. 90 (1995) 1174–1179.
[17] A. Vágújfalvi, G. Galiba, L. Cattivelli, J. Dubcovsky, The cold-regulated transcriptional activator Cbf3 is linked to the frost-tolerance locus Fr-A2 on wheat chromosome 5A, Mol. Genet. Genom. 269 (2003) 60–67.
[18] Å. Ergon, C. Fang, Ø. Jørgensen, T.S. Aamlid, O.A. Rognli, Quantitative trait loci controlling vernalisation requirement, heading time and number of panicles in meadow fescue (*Festuca pratensis* Huds.), TAG Theor. Appl. Genet. 112 (2006) 232–242.
[19] A.E. Limin, D.B. Fowler, Developmental traits affecting low-temperature tolerance response in near-isogenic lines for the vernalization locus Vrn-A1 in wheat (*Triticum aestivum* L. em Thell), Ann. Bot. 89 (2002) 579–585.
[20] T. Dhillon, S. Pearce, E. Stockinger, A. Distelfeld, C. Li, A.K. Knox, I. Vashegyi, A. Vagujfalvi, G. Galiba, J. Dubcovsky, Regulation of freezing tolerance and flowering in temperate cereals: the VRN-1 connection, Plant Physiol. (2010), 110.159079.
[21] T. Yamada, E.S. Jones, N.O.I. Cogan, A.C. Vecchies, T. Nomura, H. Hisano, Y. Shimamoto, K.F. Smith, M.D. Hayward, J.W. Forster, QTL analysis of morphological, developmental, and winter hardiness-associated traits in perennial ryegrass, Crop Sci. 44 (2004) 925–935.
[22] Y. Xiong, S. Fei, R. Arora, E. Brummer, R. Barker, G. Jung, S. Warnke, Identification of quantitative trait loci controlling winter hardiness in an annual × perennial ryegrass interspesific hybrid population, Mol. Breed. 19 (2007) 125–136.
[23] E. Francia, D. Barabaschi, A. Tondelli, G. Laidò, F. Rizza, A. Stanca, M. Busconi, C. Fogher, E. Stockinger, N. Pecchioni, Fine mapping of a Hv CBF gene cluster at the frost resistance locus Fr-H2 in barley, TAG Theor. Appl. Genet. 115 (2007) 1083–1091.
[24] J. Skinner, P. Szűcs, J. von Zitzewitz, L. Marquez-Cedillo, T. Filichkin, E. Stockinger, M. Thomashow, T. Chen, P. Hayes, Mapping of barley homologs to genes that regulate low temperature tolerance in *Arabidopsis*, TAG Theor. Appl. Genet. 112 (2006) 832–842.
[25] A. Miller, G. Galiba, J. Dubcovsky, A cluster of 11 CBF transcription factors is located at the frost tolerance locus Fr-A m 2 in *Triticum monococcum*, Mol. Genet. Genom. 275 (2006) 193–203.
[26] K. Tamura, T. Yamada, A perennial ryegrass CBF gene cluster is located in a region predicted by conserved synteny between Poaceae species, TAG Theor. Appl. Genet. 114 (2007) 273–283.
[27] Y. Xiong, S.-Z. Fei, Functional and phylogenetic analysis of a DREB/CBF-like gene in perennial ryegrass (*Lolium perenne* L.), Planta 224 (2006) 878–888.
[28] K.M. Devos, Updating the 'crop circles', Curr. Opin. Plant Biol. 8 (2005) 155–162.
[29] V. Alm, C. Fang, C.S. Busso, K.M. Devos, K. Vollan, Z. Grieg, O.A. Rognli, A linkage map of meadow fescue (*Festuca pratensis* Huds.) and comparative mapping with other Poaceae species, TAG Theor. Appl. Genet. 108 (2003) 25–40.
[30] S. Sim, T. Chang, J. Curley, S.E. Warnke, R.E. Barker, G. Jung, Chromosomal rearrangements differentiating the ryegrass genome from the *Triticeae*, oat, and rice genomes using common heterologous RFLP probes, TAG Theor. Appl. Genet. 110 (2005) 1011–1019.
[31] S. Fowler, M.F. Thomashow, *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway, Plant Cell 14 (2002) 1675–1690.
[32] M.A. Hannah, A.G. Heyer, D.K. Hincha, A global survey of gene regulation during cold acclimation in *Arabidopsis thaliana*, PLoS Genet. 1 (2005) e26.
[33] J.T. Vogel, D.G. Zarka, H.A. Van Buskirk, S.G. Fowler, M.F. Thomashow, Roles of the CBF2 and ZAT12 transcription factors in configuring the low temperature transcriptome of *Arabidopsis*, Plant J. 41 (2005) 195–211.
[34] B.-H. Lee, D.A. Henderson, J.-K. Zhu, The *Arabidopsis* cold-responsive transcriptome and its regulation by ICE1, Plant Cell 17 (2005) 3155–3175.
[35] A. Monroy, A. Dryanova, B. Malette, D. Oren, M. Ridha Farajalla, W. Liu, J. Danyluk, L. Ubayasena, K. Kane, G. Scoles, F. Sarhan, P. Gulick, Regulatory gene candidates and gene expression analysis of cold acclimation in winter and spring wheat, Plant Mol. Biol. 64 (2007) 409–423.
[36] M.O. Winfield, C. Lu, I.D. Wilson, J.A. Coghill, K.J. Edwards, Plant responses to cold: transcriptome analysis of wheat, Plant Biotechnol. J. (2010), doi:10.1111/j.1467-7652.2010.00536.x.
[37] J.T. Svensson, C. Crosatti, C. Campoli, R. Bassi, A.M. Stanca, T.J. Close, L. Cattivelli, Transcriptome analysis of cold acclimation in barley albina and xantha mutants, Plant Physiol. 141 (2006) 257–270.

Please cite this article in press as: S.R. Sandve, et al., Molecular mechanisms underlying frost tolerance in perennial grasses adapted to cold climates, Plant Sci. (2010), doi:10.1016/j.plantsci.2010.07.011

[38] C. Zhang, S.-Z. Fei, S. Warnke, L. Li, D. Hannapel, Identification of genes associated with cold acclimation in perennial ryegrass, J. Plant Physiol. 166 (2009) 1436–1445.
[39] H. Rudi, S. Sandve, L.M. Opseth, A. Larsen, O.A. Rognli, Identification of candidate genes important for frost tolerance differences in *Festuca pratensis* Huds. by transcriptional profiling, Unpublished results, 2010.
[40] P. Bhardwaj, P. Ahuja, S. Kumar, Characterization of gene expression of QM from *Caragana jubata*, a plant species that grows under extreme cold, Mol. Biol. Rep. 37 (2010) 1003–1010.
[41] Å. Strand, C.H. Foyer, P. Gustafsson, P. Gardestrom, V. Hurry, Altering flux through the sucrose biosynthesis pathway in transgenic *Arabidopsis thaliana* modifies photosynthetic acclimation at low temperatures and the development of freezing tolerance, Plant Cell Environ. 26 (2003) 523–535.
[42] B.P. Mooney, J.A. Miernyk, C.M. Greenlief, J.J. Thelen, Using quantitative proteomics of *Arabidopsis* roots and leaves to predict metabolic activity, Physiol. Plant. 128 (2006) 237–250.
[43] S. Amme, A. Matros, B. Schlesier, H.P. Mock, Proteome analysis of cold stress response in *Arabidopsis* thaliana using DIGE-technology, J. Exp. Bot. 57 (2006) 1537–1546.
[44] M.S. Bae, E.J. Cho, E.Y. Choi, O.K. Park, Analysis of the *Arabidopsis* nuclear proteome and its response to cold stress, Plant J. 36 (2003) 652–663.
[45] E. Goulas, M. Schubert, T. Kieselbach, L.A. Kleczkowski, P. Gardestrom, W. Schroder, V. Hurry, The chloroplast lumen and stromal proteomes of *Arabidopsis thaliana* show differential sensitivity to short- and long-term exposure to low temperature, Plant J. 47 (2006) 720–734.
[46] Y. Kawamura, M. Uemura, Mass spectrometric approach for identifying putative plasma membrane proteins of *Arabidopsis* leaves associated with cold acclimation, Plant J. 36 (2003) 141–154.
[47] S.X. Cui, F. Huang, J. Wang, X. Ma, Y.S. Cheng, J.Y. Liu, A proteomic analysis of cold stress responses in rice seedlings, Proteomics 5 (2005) 3162–3172.
[48] M. Hashimoto, S. Komatsu, Proteomics analysis of rice seedlings during cold stress, Proteomics 7 (2007) 1293–1302.
[49] N. Imin, T. Kerim, B.G. Rolfe, J.J. Weinman, Effect of early cold stress on the maturation of rice anthers, Proteomics 4 (2004) 1873–1882.
[50] D.-G. Lee, N. Ahsan, S.-H. Lee, J.J. Lee, J.D. Bahk, K.Y. Kang, B.-H. Lee, Chilling stress-induced proteomic changes in rice roots, J. Plant Physiol. 166 (2009) 1–11.
[51] S.P. Yan, Q.Y. Zhang, Z.C. Tang, W.A. Su, W.N. Sun, Comparative proteomic analysis provides new insights into chilling stress responses in rice, Mol. Cell. Proteom. 5 (2006) 484–496.
[52] P. Vitamvas, G. Saalbach, I.T. Prasil, V. Capkova, J. Opatrna, J. Ahmed, WCS120 protein family and proteins soluble upon boiling in cold-acclimated winter wheat, J. Plant Physiol. 164 (2007) 1197–1207.
[53] C. Dal Bosco, M. Busconi, C. Govoni, P. Baldi, A.M. Stanca, C. Crosatti, R. Bassi, L. Cattivelli, cor gene expression in barley mutants affected in chloroplast development and photosynthetic electron transport, Plant Physiol. 131 (2003) 793–802.
[54] R.D. Vierstra, Proteolysis in plants: mechanisms and functions, Plant Mol. Biol. 32 (1996) 275–302.
[55] D. Cook, S. Fowler, O. Fiehn, M.F. Thomashow, A prominent role for the CBF cold response pathway in configuring the low-temperature metabolome of *Arabidopsis*, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 15243–15248.
[56] G.R. Gray, D. Heath, A global reorganization of the metabolome in *Arabidopsis* during cold acclimation is revealed by metabolic fingerprinting, Physiol. Plant. 124 (2005) 236–248.
[57] F. Kaplan, J. Kopka, D.W. Haskell, W. Zhao, K.C. Schiller, N. Gatzke, D.Y. Sung, C.L. Guy, Exploring the temperature-stress metabolome of *Arabidopsis*, Plant Physiol. 136 (2004) 4159–4168.
[58] G.H. Krause, Photoinhibition of photosynthesis. An evaluation of damaging and protective mechanisms, Physiol. Plant. 74 (1988) 566–574.
[59] S. Takahashi, N. Murata, How do environmental stresses accelerate photoinhibition? Trends Plant Sci. 13 (2008) 178–182.
[60] N. Huner, G. Öquist, V. Hurry, M. Krol, S. Falk, M. Griffith, Photosynthesis, photoinhibition and low temperature acclimation in cold tolerant plants, Photosyn. Res. 37 (1993) 19–39.
[61] C. Ndong, J. Danyluk, N.P.A. Huner, F. Sarhan, Survey of gene expression in winter rye during changes in growth temperature, irradiance or excitation pressure, Plant Mol. Biol. 45 (2001) 691–703.
[62] G.R. Gray, L.P. Chauvin, F. Sarhan, N.P.A. Huner, Cold acclimation and freezing tolerance (a complex interaction of light and temperature), Plant Physiol. 114 (1997) 467–474.
[63] P. Muller, X.-P. Li, K.K. Niyogi, Non-photochemical quenching. a response to excess light energy, Plant Physiol. 125 (2001) 1558–1566.
[64] S.B. Powles, Photoinhibition of photosynthesis induced by visible light, Annu. Rev. Plant Physiol. 35 (1984) 15–44.
[65] C.J. Pollock, T. Jones, Seasonal patterns of fructan metabolism in forage grasses, New Phytol. 83 (1979) 9–15.
[66] M. Rapacz, F. Janowiak, Physiological effects of winter rape (*Brassica napus* var. oleifera) prehardening to frost. I. Frost resistance and photosynthesis during cold acclimation, J. Agron. Crop Sci. 181 (1998) 13–20.
[67] W.W. Adams III, B. Demmig-Adams, T.N. Rosenstiel, A.K. Brightwell, V. Ebbert, Photosynthesis and photoprotection in overwintering plants, Plant Biol. (Stuttg) 4 (2002) 545–557.
[68] M. Rapacz, D. Gasior, Z. Zwierzykowski, A. Lesniewska-Bocianowska, M.W. Humphreys, A.P. Gay, Changes in cold tolerance and the mechanisms of acclimation of photosystem II to cold hardening generated by anther culture of *Festuca pratensis* × *Lolium multiflorum* cultivars, New Phytol. 162 (2004) 105–114.
[69] M. Humphreys, D. Gasior, A. Lesniewska-Bocianowska, Z. Zwierzykowski, M. Rapacz, Androgenesis as a means of dissecting complex genetic and physiological controls: selecting useful gene combinations for breeding freezing tolerant grasses, Euphytica 158 (2007) 337–345.
[70] J. King, I.P. Armstead, I.S. Donnison, H.M. Thomas, R.N. Jones, M.J. Kearsey, L.A. Roberts, A. Thomas, W.G. Morgan, I.P. King, Physical and genetic mapping in the grasses lolium perenne and *Festuca pratensis*, Genetics 161 (2002) 315–324.
[71] E. Senderska, J. Migas, A. Woźniczka, W. Kaczmarek, A. Kozieł, M. Rapacz, Photosynthetic acclimation to cold in mountain Festuca ecotypes from Babia Góra (in Polish), in: M. Kopeć (Ed.) Materiały Wydziałowej Sesji Koła Naukowego Rolników, UR, Wydział Rolniczo-Ekonomiczny, Kraków, 2005.
[72] M. Badawi, J. Danyluk, B. Boucho, M. Houde, F. Sarhan, The CBF gene family in hexaploid wheat and its relationship to the phylogenetic complexity of cereal CBFs, Mol. Genet. Genom. 277 (2007) 533–554.
[73] H. Hisano, A. Kanazawa, M. Yoshida, M.O. Humphreys, M. Iizuka, K. Kitamura, T. Yamada, Coordinated expression of functionally diverse fructosyltransferase genes is associated with fructan accumulation in response to low temperature in perennial ryegrass, New Phytol. 178 (2008) 766–780.
[74] S.R. Sandve, S. Fjellheim, Did gene family expansions during the eocene–oligocene boundary climate cooling play a role in Pooideae adaptation to cool climates? Mol. Ecol. 19 (2010) 2075–2088.
[75] G. Galiba, A. Vágújfalvi, C. Li, A. Soltész, J. Dubcovsky, Regulatory genes involved in the determination of frost tolerance in temperate cereals, Plant Sci. 176 (2009) 12–19.
[76] P.L. Davies, J. Baardsnes, M.J. Kuiper, V.K. Walker, Structure and function of antifreeze proteins, Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci. 357 (2002) 927–933.
[77] M. Griffith, C. Lumb, S.B. Wiseman, M. Wisniewski, R.W. Johnson, A.G. Marangoni, Antifreeze proteins modify the freezing process in planta, Plant Physiol. 138 (2005) 330–340.
[78] M. Griffith, P. Ala, D.S.C. Yang, W.-C. Hon, B.A. Moffatt, Antifreeze protein produced endogenously in winter rye leaves, Plant Physiol. 100 (1992) 593–596.
[79] M. Griffith, M.W.F. Yaish, Antifreeze proteins in overwintering plants: a tale of two activities, Trends Plant Sci. 9 (2004) 399–405.
[80] C. Sidebottom, S. Buckley, P. Pudney, S. Twigg, C. Jarman, C. Holt, J. Telford, A. McArthur, D. Worrall, R. Hubbard, P. Lillford, Phytochemistry: heat-stable antifreeze protein from grass, Nature 406 (2000) 256–1256.
[81] K.D. Kumble, J. Demmer, S. Fish, S. Hall, S. Corrales, A. DeAth, C. Elton, R. Prestidge, S. Luxmanan, C.J. Marshall, D.A. Wharton, Characterization of a family of ice-active proteins from the ryegrass, *Lolium perenne*, Cryobiology 57 (2008) 263–268.
[82] U.P. John, R.M. Polotnianka, K.A. Sivakumaran, O. Chew, L. Mackin, M.J. Kuiper, J.P. Talbot, G.D. Nugent, J. Mautord, G.E. Schrauf, G.C. Spangenberg, Ice recrystallisation inhibition proteins (IRIPs) and freeze tolerance in the cryophilic Antarctic hairgrass *Deschampsia antarctica* E. Desv, Plant Cell Environ. 32 (2009) 336–348.
[83] K. Tremblay, F. Ouellet, J. Fournier, J. Danyluk, F. Sarhan, Molecular characterization and origin of novel bipartite cold-regulated ice recrystallization inhibition proteins from cereals, Plant Cell Physiol. 46 (2005) 884–891.
[84] M.J. Kuiper, P.L. Davies, V.K. Walker, A theoretical model of a plant antifreeze protein from lolium perenne, Biophys. J. 81 (2001) 3560–3565.
[85] A.J. Middleton, A.M. Brown, P.L. Davies, V.K. Walker, Identification of the ice-binding face of a plant antifreeze protein, FEBS Lett. 583 (2009) 815–819.
[86] C. Zhang, S.-Z. Fei, R. Arora, D. Hannapel, Ice recrystallization inhibition proteins of perennial ryegrass enhance freezing tolerance, Planta (2010).
[87] G. Hendry, The ecological significance of fructan in a contemporary flora, New Phytol. 106 (1987) 201–216.
[88] A. De Cugnac, Les glucides des Graminees. Importance de fructoholosides, Bull. Soc. Chim. Biol. 13 (1931) 125–132.
[89] F. del Viso, A.F. Puebla, C.M. Fusari, A.C. Casabuono, A.S. Couto, H.G. Pontis, H.E. Hopp, R.A. Heinz, Molecular characterization of a putative sucrose: fructan 6-fructosyltransferase (6-SFT) of the cold-resistant patagonian grass bromus pictus associated with fructan accumulation under low temperatures, Plant Cell Physiol. 50 (2009) 489–503.
[90] B. Gonzalez, J. Boucaud, J. Salette, J. Langlois, Fructan and cryoprotection in ryegrass (*Lolium perenne* L.), New Phytol. 115 (1990) 319–323.
[91] D. Smith, R.D. Groteluesche, Carbohydrates in grasses. I. Sugar and fructosan composition of the stem bases of several northern-adapted grasses at seed maturity, Crop Sci. 6 (1966) 263–266.
[92] T. Yukawa, M. Kobayashi, Y. Watanabe, Studies on fructan accumulation in wheat (*Triticum aestivum* L.), Jpn. J. Crop Sci. 64 (1995) 801–806.
[93] G.A.F. Hendry, Evolutionary origins and natural functions of fructans – a climatological, biogeographic and mechanistic appraisal, New Phytol. 123 (1993) 3–14.
[94] D. Livingston, D. Hincha, A. Heyer, Fructan and its relationship to abiotic stress tolerance in plants, Cell. Mol. Life Sci. 66 (2009) 2007–2023.
[95] I. Vijn, S. Smeekens, Fructan: more than a reserve carbohydrate? Plant Physiol. 120 (1999) 351–360.
[96] A. Kawakami, M. Yoshida, Molecular characterization of sucrose: sucrose 1-fructosyltransferase and sucrose: fructan 6-fructosyltransferase associated with fructan accumulation in winter wheat during cold hardening, Biosci. Biotechnol. Biochem. 66 (2002) 2297–2305.

Please cite this article in press as: S.R. Sandve, et al., Molecular mechanisms underlying frost tolerance in perennial grasses adapted to cold climates, Plant Sci. (2010), doi:10.1016/j.plantsci.2010.07.011

[97] J. Chalmers, A. Lidgett, N. Cummings, Y.Y. Cao, J. Forster, G. Spangenberg, Molecular genetics of fructan metabolism in perennial ryegrass, Plant Biotechnol. J. 3 (2005) 459–474.
[98] K.-I. Tamura, A. Kawakami, Y. Sanada, K. Tase, T. Komatsu, M. Yoshida, Cloning and functional analysis of a fructosyltransferase cDNA for synthesis of highly polymerized levans in timothy (*Phleum pratense* L.), J. Exp. Bot. 60 (2009) 893–905.
[99] R.A. Demel, E. Dorrepaal, M.J.M. Ebskamp, J.C.M. Smeekens, B. de Kruijff, Fructans interact strongly with model membranes, Biochim. Biophys. Acta (BBA) – Biomemb. 1375 (1998) 36–42.
[100] I.J. Vereyken, V. Chupin, R.A. Demel, S.C.M. Smeekens, B. De Kruijff, Fructans insert between the headgroups of phospholipids, Biochim. Biophys. Acta (BBA) – Biomemb. 1510 (2001) 307–320.
[101] D.K. Hincha, E.M. Hellwege, A.G. Heyer, J.H. Crowe, Plant fructans stabilize phosphatidylcholine liposomes during freeze-drying, Eur. J. Biochem. 267 (2000) 535–540.
[102] H.-J. Li, A.-F. Yang, X.-C. Zhang, F. Gao, J.-R. Zhang, Improving freezing tolerance of transgenic tobacco expressing sucrose: sucrose 1-fructosyltransferase gene from *Lactuca sativa*, Plant Cell Tissue Organ Cult. 89 (2007) 37–48.
[103] D. Parvanova, A. Popova, I. Zaharieva, P. Lambrev, T. Konstantinova, S. Taneva, A. Atanassov, V. Goltsev, D. Djilianov, Low temperature tolerance of tobacco plants transformed to accumulate proline, fructans, or glycine betaine. Variable chlorophyll fluorescence evidence, Photosynthetica 42 (2004) 179–185.
[104] A. Kawakami, Y. Sato, M. Yoshida, Genetic engineering of rice capable of synthesizing fructans and enhancing chilling tolerance, J. Exp. Bot. 59 (2008) 793–802.

Please cite this article in press as: S.R. Sandve, et al., Molecular mechanisms underlying frost tolerance in perennial grasses adapted to cold climates, Plant Sci. (2010), doi:10.1016/j.plantsci.2010.07.011