**UNIVERSITÀ DEGLI STUDI DELLA TUSCIA DI VITERBO**

DIPARTIMENTO DI SCIENZE AMBIENTALI

CORSO DI DOTTORATO DI RICERCA

in GENETICA E BIOLOGIA CELLULARE

XX° CICLO.

# FUNCTIONAL GENOMICS ANALYSES OF A *DREB*-RELATED GENE IN DURUM WHEAT

(s.s.d. BIO/11)

Coordinatore: Prof. Giorgio Prantera

Firma ……..

Tutor: Dott. ssa Maria Sperandei

Firma……

Dottorando: Arianna Latini

Firma ……..

# DECLARATION

This PhD Thesis contains no material which as been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by any other person, except where due reference has been stated in the text.

I hereby give consent to this copy of my Ph.D. Thesis being made available in the University Library.

# CONTENTS

**INTRODUCTION** 1

**CHAPTER 1.** SUBJECT: WHEAT. PROBLEM: DROUGHT. SOLUTION: MOLECULAR ASSISTED BREEDING. 2
SUMMARY 2
COPING WITH WATER SCARCITY 3
THE GRASSES AND THE WHEAT 4
THE DROUGHT AND THE PLANTS 7
ASSISTED BREEDING TO DEVELOP STABLE DROUGHT TOLERANT LINES 9

**CHAPTER 2.** MOLECULAR RESPONSES TO DEHYDRATION AND THE INVOLVEMENT OF THE *DREB* GENES 10
SUMMARY 10
MOLECULAR BASIS OF DEHYDRATION TOLERANCE IN PLANTS 11
TRANSCRIPTIONAL REGULATION AND TRANSCRIPTION FACTORS 13
THE AP2 DOMAIN AND THE *DREB* GENE FAMILY 16
*DREB* HOMOLOGOUS GENES IN OTHER PLANT SPECIES 18

**CHAPTER 3.** TECHNICAL APPROACHES AND STRATEGIES USED IN THIS EXPERIMENTAL WORK 20
SUMMARY 20
QUANTITATIVE PCR (qPCR) 21
TRANSGENIC PLANTS 26

**INTRODUCTION TO THE EXPERIMENTAL SECTIONS** 30

**EXPERIMENTAL SECTION 1.** DEHYDRATION RESPONSIVE FACTOR 1 (*TdDRF1*) GENE: MOLECULAR ANALYSES UNDER GREENHOUSE CONDITIONS IN SOME DROUGHT TOLERANT DURUM WHEAT AND TRITICALE CULTIVARS 32
SUMMARY 32
MATERIALS AND METHODS 33
Plant materials and water stress treatment 33
Relative Water Content (RWC) 33
Qualitative RT-PCR 34
Time-course RT-PCR 34
Quantitative real-time RT-PCR 35
Cloning and sequence analysis of *TdDRF1* cDNAs 37
AP2 structural model building 37
RESULTS 39

Physiological analysis 39
Qualitative RT-PCR analysis 39
Time-course RT-PCR analysis 40
Real-time RT-PCR analysis 41
Cloning and sequence analyses 44
DISCUSSION 49

**EXPERIMENTAL SECTION 2.** EVALUATION OF THE EXPRESSION LEVELS OF *TdDRF1* TRANSCRIPTS IN DROUGHT TOLERANT AND SUSCEPTIBLE DURUM WHEAT CULTIVARS IN A FIELD TRIAL UNDER WELL AND REDUCED IRRIGATION CONDITIONS 54
SUMMARY 54
MATERIALS AND METHODS 55
Plant materials 55
Field experimental design 56
Crop management 57
Physiological analyses 57
Leaves harvesting for the RNA bank ensemble 58
Reverse transcription 59
Real-time PCR 59
Quantitative data analysis 60
RESULTS AND DISCUSSION 61
Physiological analyses 63
Molecular analyses 65

**EXPERIMENTAL SECTION 3.** CLONING OF THE THREE *TdDRF1* TRANSCRIPTS IN TRANSIENT EXPRESSION VECTORS 69
SUMMARY 69
ADDENDUM: Potato Virus X (PVX) and pP2C2S-mediated transient expression 70
MATERIALS AND METHODS 72
5' RACE 72
3' RACE 73
Cloning of the complete codifying sequences (CDSs) of *TdDRF1.1*, *TdDRF1.2* and *TdDRF1.3* 74
Engineering of viral constructs 75
Transcription of viral RNA and infection of *Nicotiana benthamiana* and *tabacum* plants 76
RNA extraction, reverse transcription, PCR and real-time PCR analyses 77
RESULTS AND DISCUSSION 79
5' and 3' UTR analysis 79
Full-length *TdDRF1* gene transcript cDNAs cloning 82
Transient expression of *TdDRF1* gene transcripts in *Nicotiana benthamiana* and *Nicotiana tabacum* plants infected with recombinant Potato Virus X (PVX) 85
CONCLUSIONS 90

**REFERENCES** 91

SUMMARY 106

RIASSUNTO 108

LITERATURE PRODUCTION 110

POSTER PRESENTATIONS AT CONGRESSES 111

ACKNOWLEDGEMENTS 112

FUNDINGS 113

# INTRODUCTION

# CHAPTER 1

# SUBJECT: WHEAT
# PROBLEM: DROUGHT
# SOLUTION: MOLECULAR ASSISTED BREEDING


## SUMMARY

Unmitigated climate changes due to increasing greenhouse gas emissions will have an adverse impact on crop yields and water resources. The cereals are our dominant source of food, providing more than 60 % of the calories and proteins for our daily life, with maize, rice and wheat vying in the number one position. The importance of crop resistance to water stress is further increasing, as the world population continues to expand and there is a resilient necessity to increase food production both through traditional breeding and through biotechnology-based germplasm.

## COPING WITH WATER SCARCITY

According to the Food and Agriculture Organization of the United Nations (FAO, http://www.fao.org/), global water use has been growing at more than twice the rate of population growth in the last century. Water scarcity already affects every continent and more than 40 % of the people on our planet.

The water scarcity situation is being exacerbated by climate changes, especially in the driest areas of the world. The human impact on the earth's environment and climate must be addressed in order to protect the world's water resources, but there are other factors involved, such as increases in the amount of water needed to grow the food for a growing population. Also, the trend towards urbanization and increases in domestic and industrial water use, by people who live in more developed areas, are factors that lead to growing water use (http://www.fao.org/newsroom/en/focus/2007/1000521/index.html).

More facts on water issues:

- The world population is currently increasing by around 100 million per year and it is expected to exceed 10 billion by 2050, with a concomitant requirement to double the food produced in a decreasing amount of arable land.
- Actually, around 1.2 billion people, or almost one-fifth of the world's population, live in areas of physical water scarcity, and 500 million people are approaching this situation.
- The daily drinking water requirement per person is 2-4 litres, but it takes 2000 to 5000 litres of water to produce one person's daily food.
- It takes 1000-3000 litres of water to produce just one kilo of rice and 13000 to 15000 litres to produce one kilo of grain-fed beef.
- Since 1950, the area of the earth under irrigation has doubled and water withdrawal for agricultural, domestic and industrial purposes has tripled.
- Agriculture is the number-one user of water worldwide, accounting for about 70 percent of all freshwater withdrawn from lakes, rivers and aquifers.

To tackle water scarcity and to ensure global, political and social stability, increasing food production and equitably, without compromising environmental integrity, is necessary to support initiatives that produce more food with proportionally less water.

## THE GRASSES AND THE WHEAT

The grasses plants belong to the Poaceae family (Gramineae family) and agricultural grasses grown for their edible seeds are called cereals. Cereals are the major source of carbohydrate for humans.
The three most important cereal crops - wheat, maize and rice - evolved independently from a common ancestor ~50-70 million years ago (Kellogg 2001), but despite this long period of independent evolution, cereal genes and genomes display high conservation. Comparison of physical and genetic maps of the grass genomes shows conservation of gene order and orientation, or synteny, and the concept of grasses as a single genetic system emerges (Keller & Feuillet 2000; Bhalla 2006) (Fig. 1).

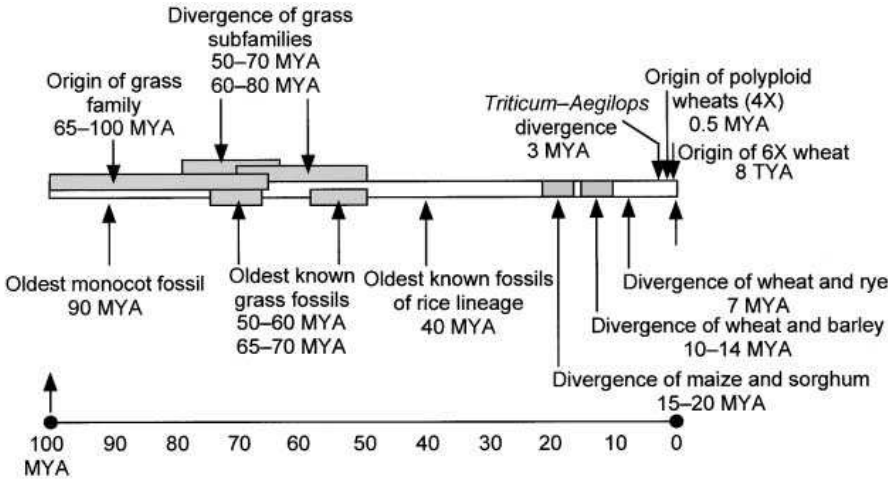



**Figure 1.** Grasses as a single genetic system.
Proceedings from Gill *et al.* 2004.

However these crops differ greatly in genome size, depending on the chromosome number and the ploidy level. The genome of bread wheat (2n=6x=42) is estimated at 16.000 megabase pairs (Mbp), the genome of durum wheat (2n=4x=28) at 12.000 Mbp, the genome of barley (2n=2x=14) at 5.000 Mbp, the genome of maize (2n=2x=20) at 3.000 Mbp, while rice (2n=2x=24) has a much smaller genome at 430 Mbp. The small genome and predicted high gene density of rice rendered it an ideal target for cereal gene discovery and genome sequence analysis (Meyers *et al*. 2004). A draft sequence of the rice genome (420 Mb) was published in Science in 2002 (Goff *et al.* 2002) and, together with the model plant Arabidopsis genome (Arabidopsis Genome Initiative 2000), comparative analysis for plants is now moving ahead rapidly and gene discovery in all cereals is greatly facilitated. Homologs of 98% of the known maize, wheat, and barley proteins are found in

rice; synteny and gene homology between rice and the other cereal genomes are extensive (Ware & Stein 2003), whereas synteny with *Arabidopsis* is limited.

Common or bread wheat, *Triticum aestivum*, is hexaploid (genome AABBDD) and durum wheat, *Triticum durum*, is tetraploid (AABB). The large genome of these two species is attributable to the polyploidy nature of wheat and the high content of repetitive elements within its genome. The ancestors of modern *Triticum turgidum* are the diploid species *Triticum urartu* (genome AA) and *Aegilops speltoides* (genome BB). *Triticum aestivum*, originated from the hybridization of *Triticum turgidum* and *Aegilops tauschii* (genome DD) which contributed the D genome (Fig. 2).

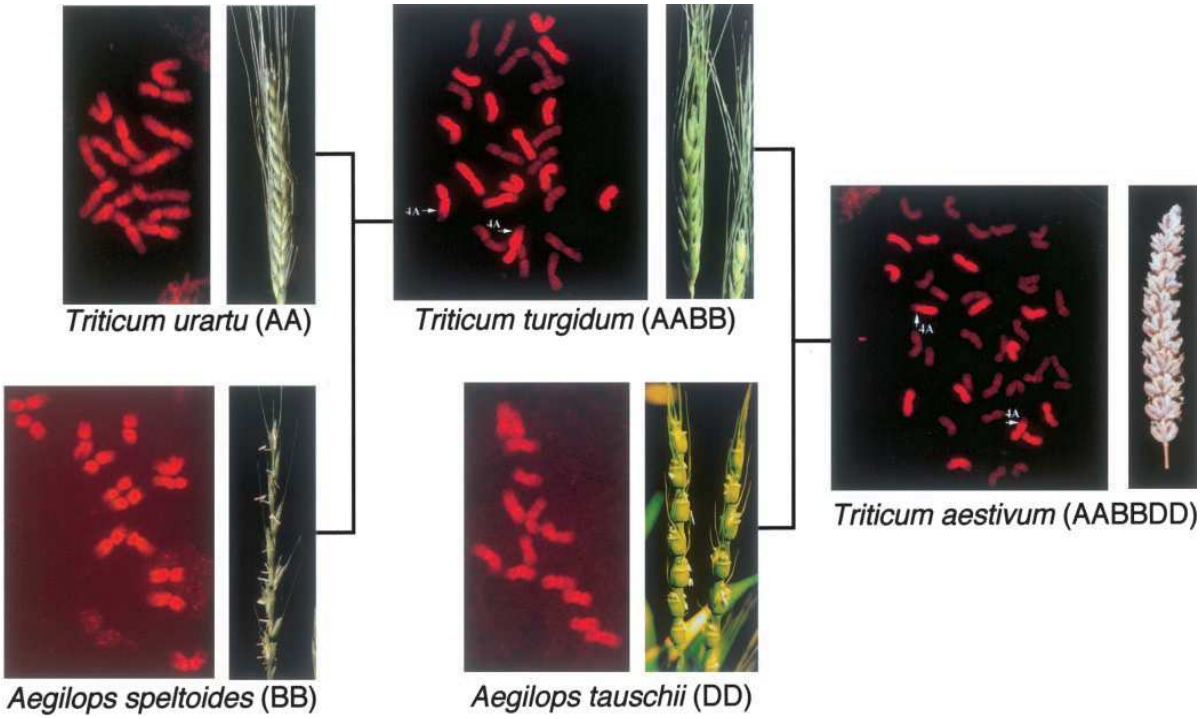



**Figure 2.** Wild relatives of modern cultivated wheats.
Proceedings from Gill *et al.* 2004.

Recently, a particular interest is addressed to evolutionary aspects of genomic redundancy in allopolyploids species, because of the significant content of repetitive DNA, representing more than 80% of the common wheat genome (Paux *et al.* 2006), and the simultaneous presence of homoeologous genomes. Actually polyploidy is recognized as a prominent speciation process in plants (Wendel 2000; Doležel *et al.* 2007) and allopolyploidy accelerates genome evolution in wheat (Feldman & Levy 2005).

Significant advances in the understanding of the wheat plant and grain biology must be achieved to increase absolute yields and protect the crop from an estimated average annual loss of 25 % caused by biotic (pests) and abiotic stresses (drought, heat, salinity, frost) (Gill *et al*. 2004).

In particular this PhD thesis focuses on durum wheat. *Triticum durum*, also called Macaroni wheat, is the only tetraploid species of wheat cultivated today. Durum wheat is cultivated on approximately 17 million hectares worldwide. About half of that area is in developing countries; its production is concentrated in the Middle East, Central India, and the Mediterranean region of West Asia and North Africa (WANA), where 80% of the durum area is found. Other production areas include Ethiopia, Argentina, Chile, Russia, and Kazakhstan, as well as Mexico, the United States, Italy, Spain, and Canada. Durum wheat production in developing countries is generally low because it is grown using low levels of inputs (e.g., fertilizer, water) in semi-arid regions and other marginal areas characterized by sharp annual fluctuations in cropping conditions. In addition, yields may be reduced by insects, poor crop management, and deficient weed control. In irrigated environments where moisture and other resources are not limiting, high yields are obtained, but these favourable environments represent just a small portion (28.2%) of the total durum wheat area (http://www.cimmyt.org/Research/wheat/map/developing_world/wheat_developing_world.htm).

Durum is the hardest of all wheats; its high protein content and gluten strength make durum good for pasta and bread.

## THE DROUGHT AND THE PLANTS

Drought is one of the most severe abiotic stresses limiting crop productivity and it plays a major role in determining the distribution of plant species across different types of environments. The effect on plant growth depends on both severity and duration of the stress (Araus *et al*. 2002; Bray 1997). Furthermore, a plant may experience several distinct abiotic stresses either concurrently or at different times through the growing season (Tester & Bacic 2005).

Drought can be defined as a period of below normal precipitation that limits plant productivity in both natural and agricultural systems (Boyer 1982). The response to this abiotic stress is a complex phenomenon, comprising a number of physio-biochemical processes at both cellular and whole organism levels activated at different stages of plant development, and several mechanisms have been adopted by plants to adapt to water stress.

The stress component that defines drought is a decrease in the availability of soil water and this reduction can be quantified as a decrease in water potential ($\psi_w$). To avoid a low tissue $\psi_w$, plants increase water uptake or limit water loss such that the rates of water loss and water uptake remain balanced. Such a balance is achieved in the short term mainly by stomatal closure. In the longer term, changes in root and shoot growth, tissue water storage capacity by accumulation of osmolytes, cuticle thickness and water permeability are some mechanisms of adaptation to water stress (Verslues *et al*. 2006).

In particular, for crop plants changes in root growth to maximize water uptake are of the greatest importance (Bohnert *et al*. 1995; Mahajan & Tuteja 2005). Such adjustments are termed "dehydration avoidance". In the case of mild water stress or water stress of limited duration, avoidance mechanisms by themselves can be sufficient to maintain plant performance. The trade-off in this case is the lost photosynthesis caused by reduced stomatal $CO_2$ uptake or by a shift of resources into root growth at the expense of photosynthetic and reproductive tissue.

On the other hand, if the stress becomes more severe and the plant is no longer able to maintain a good water potential, additional mechanisms become important for its survival. Dehydration tolerance is defined as the relative capacity to sustain or conserve plant function in a dehydrate state. This is sometimes seen as the second defence line after dehydration avoidance (Fig. 3).

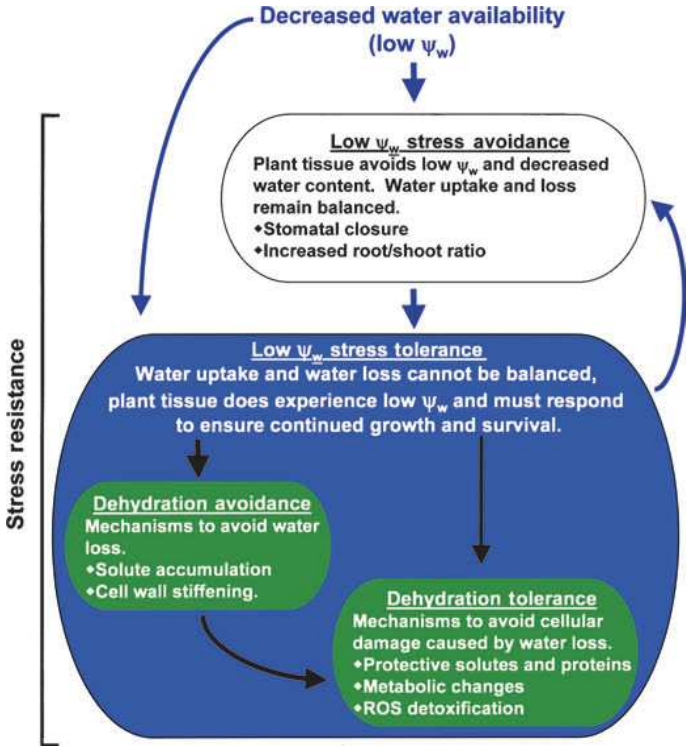



**Figure 3.** Conceptual diagram of the stress tolerance and stress avoidance model of low-$\psi_w$ responses. In low-$\psi_w$ stress avoidance the plant balances water uptake and water loss to avoid an effect of the stress on tissue $\psi_w$ or water content (essentially, the stress is kept outside the plant tissue). If this cannot be achieved and the plant tissue does experience low $\psi_w$ (the stress becomes internalized to the plant tissue), stress responses occur that maintain a high water content despite a decreased $\psi_w$ (dehydration avoidance) or tolerate reduced water content (dehydration tolerance). The term "stress resistance" is used when it is not possible or not desirable to refer to a more specific mechanism.
Proceedings from Verslues *et al*. 2006.

## ASSISTED BREEDING TO DEVELOP STABLE DROUGHT TOLERANT LINES

The conventional breeding programs are enough slow, it takes about 11 years to arrive to a pure new line coming from breeding, but in some cases they have been successful in developing abiotic stress-tolerant lines; however in most of these cases the tolerant plant showed an inverse relationship with yield. With the advent of recent genomic resources, the major approaches in exploiting the gene pool for imparting abiotic stress tolerance in crops of interest include, firstly, the identification of stress-tolerant genes and Quantitative Trait Loci (QTLs) conferring tolerance to stress in germplasm collections and, secondly, the development of respective molecular markers and use in marker-assisted programs. The increasing number of studies reporting QTLs for drought-related traits and yield in drought-stressed crops indicates a growing interest in this approach (Tuberosa & Salvi 2006).

Although transgenic plants overexpressing drought-responsive genes have shown sometimes to enhance stress tolerance in laboratory conditions by activating stress-responsive signal transduction and downstream genes (Cheng *et al*. 2002; Pellegrineschi *et al*. 2004; Fujita *et al*. 2005; Tang *et al*. 2006), their success in field conditions is rather poor (Langridge *et al*. 2006; Sreenivasulu *et al*. 2006). Hence, to develop stable population with improved stress tolerance and yield, it is equally important to integrate developed knowledge as an outcome of functional genomics into real knowledge-based breeding programs via genomics-assisted breeding.

Wild species and landraces have unique resistance alleles that are not found in the gene pool of modern cultivars (wild species possess considerably more variation than the cultivated species) and are known to be especially potent sources of abiotic stress tolerance traits for genomic improvement (Forster *et al*. 2000; Sreenivasulu *et al*. 2006). Therefore the genetic loci involved in the control of abiotic stress in cultivated species can now be targeted by investigating in wild gene pool. First, tolerant accessions/wild species can be backcrossed to an elite cultivar showing higher yield under abiotic stress conditions. With this approach the discovery and transfer of desiderable QTLs from unadapted to elite germplasm is simultaneous (Tuberosa & Salvi 2006).

In conclusion, close collaboration between molecular geneticists, plant physiologists and breeders is needed to critically assess the contribution of specific genes and application of molecular genetics to breeding for drought resistance in wheat and other crops.

# CHAPTER 2

# MOLECULAR RESPONSES TO DEHYDRATION AND THE INVOLVEMENT OF THE *DREB* GENES


## SUMMARY

Plant growth and productivity are strongly affected by various abiotic stresses, especially dehydration, and plant must respond and adapt to these stresses in order to survive. Exposure to unfavourable environmental conditions induces various biochemical and physiological changes in the process of acquiring stress tolerance. A number of genes have been described that respond to water stress at the transcriptional level and the *cis*- and *trans*-acting factors involved in the expression of dehydration-responsive genes have been extensively analyzed.

## MOLECULAR BASIS OF DEHYDRATION TOLERANCE IN PLANTS

Plants respond and adapt to abiotic stresses with an array of biochemical and physiological alterations (Ingram & Bartels 1996; Thomashow 1999). Deciphering the mechanisms by which plants perceive environmental signals and their transmission to cellular machinery to activate the adaptive responses is of critical importance for the development of rational breeding to ameliorate stress tolerance in crops.

The functions of the stress-inducible genes are classified into two major groups: proteins that probably function in stress tolerance (functional proteins) and protein factors that probably function in stress response (regulatory proteins) involved in further regulation of signal transduction. Figure 4 is a scheme of the functional classification of drought stress-inducible genes (Shinozaki & Yamaguchi-Shinozaki 1997; Bray 2004; Nakashima & Shinozaki 2006; Shinozaki & Yamaguchi-Shinozaki 2007).

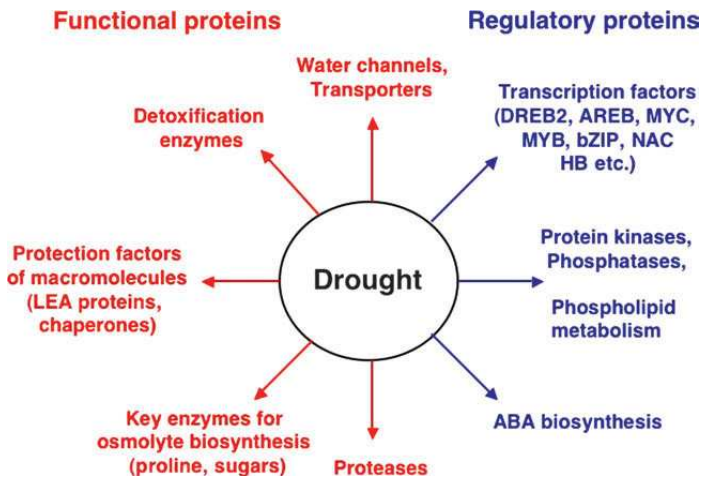



**Figure 4.** Functional and regulatory proteins responsive to dehydration.
Proceedings from Shinozaki & Yamaguchi-Shinozaki 2007.

Functional proteins are thought to function in protecting cells from stress by the production of important metabolic proteins. This group of protein includes water channel proteins and membrane transporters, chaperones, protease, late embryogenesis-abundant (LEA) proteins, enzymes that are involved in the synthesis of osmoprotectants (compatible solutes such as sugars, proline, betaine, etc.) and detoxification enzymes, such as glutathione-S-transferase, hydrolase, catalase, superoxide dismutase, ascorbate peroxidase, etc. (Valliyodan and Nguyen 2006).

In particular, the LEA class of proteins has been thought to function in the detoxification and alleviation of cell damage during dehydration (Close 1996). Overexpression of some LEA genes has been reported to result in enhanced tolerance to dehydration (Vinocur & Altman 2005; Bartels & Sunkar 2005; Umezawa *et al*. 2006) and recent biochemical analysis demonstrated that LEA proteins can prevent protein aggregation induced by desiccation as well as freezing (Goyal *et al*. 2005). Taken together with recent computational studies (Wise 2003; Wise & Tunnacliffe 2004), LEA proteins are proposed to function as chaperone-like protective molecules.

On the other side, regulatory proteins include transcription factors (Chen & Zhu 2004), kinases, enzymes for phosphoinositide turnover and enzymes for the synthesis of the plant hormone abscisic acid (ABA).

Recently, 299 drought-inducible genes, 213 high-salinity stress-inducible genes, and 54 cold-inducible genes were identified using a cDNA microarray containing approximately 7000 independent full-length *Arabidopsis* cDNA clones (Seki *et al*. 2002). Function of their gene products have been predicted from comparison of sequence homology with known proteins.

## TRANSCRIPTIONAL REGULATION AND TRANSCRIPTION FACTORS

A major level at which gene expression is regulated is the initiation of transcription and this is reflected in the percentage of genome dedicated to transcription factors in plants and other eukaryotes (Singh 1998; Riechmann & Ratcliffe 2000).

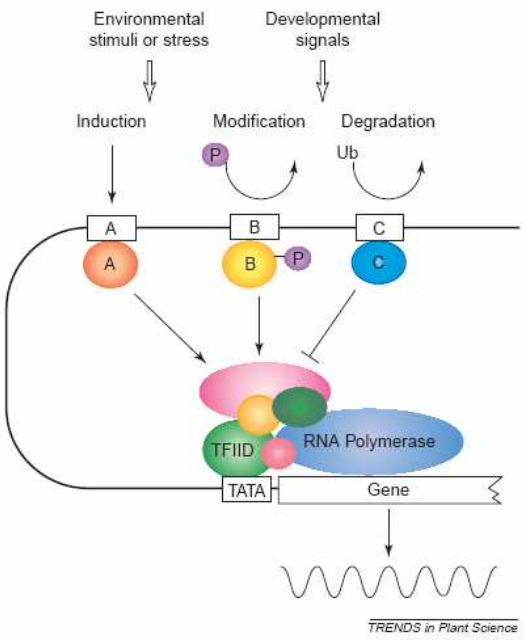



**Figure 5.** Scheme of transcriptional regulatory networks. The transcriptional initiation complex is regulated by transcription factors that are activated or repressed by environmental stimuli and/or developmental signals. The rectangular boxes labelled A, B and C represent the *cis*-acting factors and the ellipses labelled A, B and C represent the transcription factors. Ub, ubiquitin.
Proceedings from Yamaguchi-Shinozaki & Shinozaki 2005.

Typically activators have a modular structure consisting of discrete domains responsible for specific DNA binding, transcriptional activation, a nuclear localization sequence (NLS) and, in some cases, dimerization and/or other forms of protein-protein interaction (Liu *et al*. 1999). Plant transcription factors contain a wide variety of structural motifs that allow for binding to specific DNA sequences. Transcriptome analysis using microarray technology and conventional approaches have revealed that dozen of transcription factors (TFs) are involved in plant response to drought stress (Bartels & Sunkar 2005; Umezawa *et al*. 2006). Most of these TFs fall into several large families, such as AP2/ERF (Weigel 1995; Okamuro *et al.* 1997; Kizis *et al*. 2001; Magnani *et al.* 2004), bZIP, NAC, MYB, MYC, Cys2His2 zinc-finger and WRKY.

A single transcription factor (*trans*-acting element) is able to control the expression of many target genes through the specific binding to the *cis*-acting element within their promoter. This type of transcription unit is called "regulon". In Arabidopsis the analysis of the expression mechanisms of abiotic stress-responsive genes revealed apparent presence of multiple regulons. The regulons involved in the osmotic and in the cold stresses are shown in Figure 6.

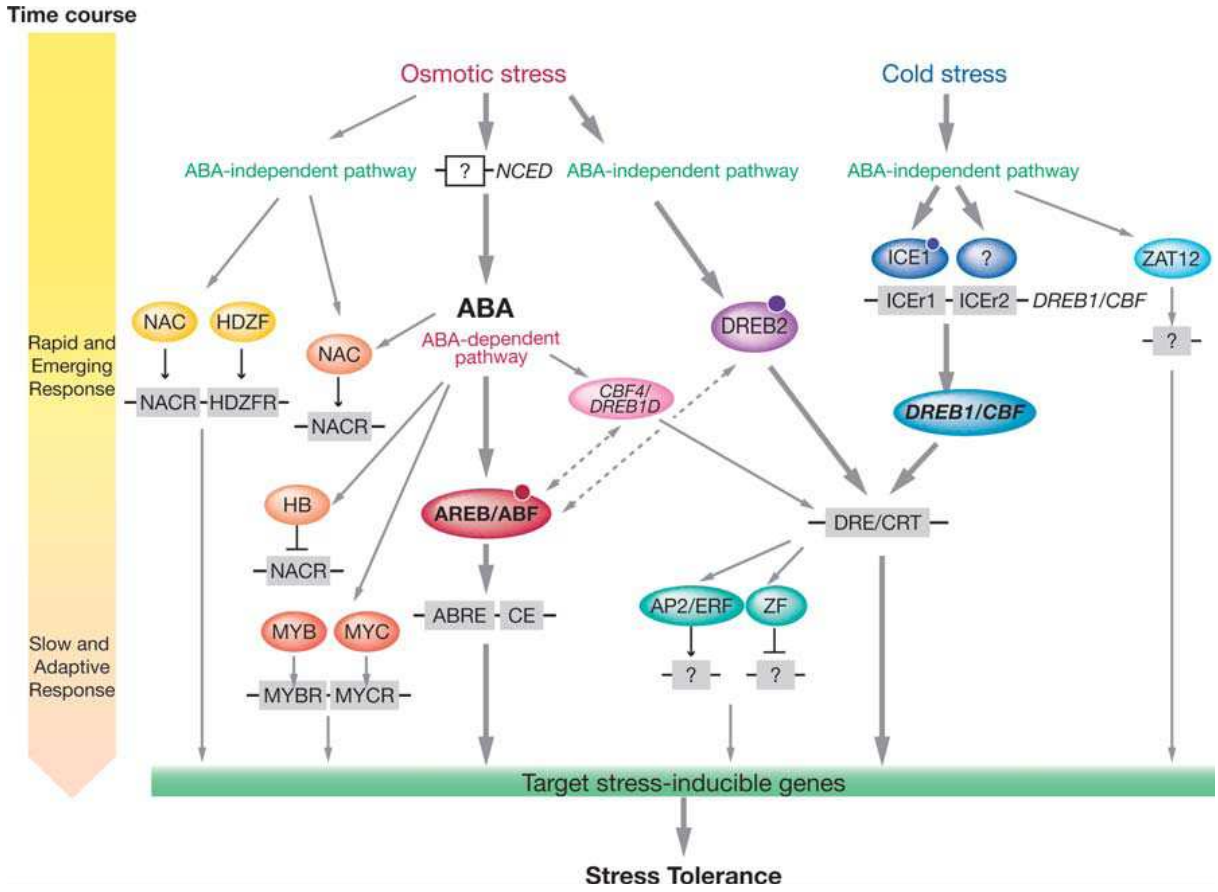



**Figure 6.** Transcriptional regulatory networks involved in osmotic and cold stress-responsive gene expression in Arabidopsis. Transcription factors controlling stress-inducible gene expression are shown in coloured ellipses and *cis*-acting elements involved in stress-responsive transcription are shown in boxes. Small filled circles reveal modification of transcription factors in response to stress signals for their activation, such as phosphorylation. Regulatory cascade of stress-responsive gene expression is shown from top to bottom. Early and emergency responses of gene expression are shown in the upper part, and late and adaptive responses in the bottom. Thick grey arrows indicate the major signalling pathways and these pathways regulate many downstream genes. Broken arrows indicate protein-protein interactions. Proceeding from Yamaguchi-Shinozaki & Shinozaki 2006.

One of the better known regulons involved in osmotic stress-responsive gene expression is dependent on abscisic acid (ABA). This molecule plays an important role in adapting vegetative tissues to abiotic stresses such as drought and high salinity, as well as in seed maturation and dormancy (Bray 1993; Shinozaki *et al.* 2003; Yamaguchi-Shinozaki & Shinozaki 2005). Promoter

regions of several ABA-inducible genes were compared with each other and the conserved sequence ACGTGG/TC was found and called ABA Responsive Element (ABRE). ABRE is the major *cis*-acting element that functions to regulate ABA-responsive gene expression (Fig. 6), but many other kinds of *cis*-acting elements also function in ABA-responsive gene expression. ABRE-binding proteins (AREB) or ABRE-binding factors (ABFs) are basic-domain leucine zipper (bZIP) transcription factors; binding to ABRE, they activate ABA-dependent gene expression (Fujita *et al*. 2005).

The other most known regulon related to the abiotic stresses, the Dehydration Responsive Element/C-repeat (DRE/CRT) regulon, is independent on ABA and will be addressed in the next chapter.

Different *cis*-acting elements in the promoters of stress-inducible genes are involved in the cross-talk of different stress signals. The major example is the promoter of an Arabidopsis drought-, high-salinity- and cold-inducible gene *Responsive to Dehydration 29A* (*RD29A*) (Yamaguchi-Shinozaki & Shinozaki 1993), encoding a LEA-like protein. This promoter contains two major *cis*-acting elements: ABRE and DRE/CRT (Yamaguchi-Shinozaki & Shinozaki 1994). It was originally thought that these *cis*-acting elements would function independently from one another, but then it was revealed that DRE/CRT functions cooperatively with ABRE as a coupling element in ABA-responsive gene expression in response to drought stress (Narusaka *et al*. 2003). This observation indicates that there are crosstalks between DREB and AREB/ABF regulons (Fig. 6): DRE/CRT functions in early stress signalling, whereas ABRE functions after the accumulation of ABA during drought and high salinity stress response. There are many ABA-inducible transcription factors that function downstream of ABA responses and stress responses. These transcription factors are involved mainly in late and adaptive processes during stress responses (Yamaguchi-Shinozaki & Shinozaki 2005).

## THE AP2 DOMAIN AND THE *DREB* GENE FAMILY

The highly conserved AP2/ERF (APETALA2/Ethylene Responsive Element Binding Factor) domain is a DNA-binding motif unique in plants and has been found in a large number of proteins in Arabidopsis and many other plants (Liu *et al*. 1998; Sakuma *et al*. 2002). The Arabidopsis genome, encode for 145 AP2/ERF related proteins, which are divided into three classes. The 1st class includes 14 proteins that have two AP2/ERF domains. The 2nd class comprises 6 proteins that have two different DNA-binding domains, AP2 and B3 (for example RAV1). The 3rd class comprises 125 proteins that have only one AP2/ERF domain. Of these, 121 proteins contain a conserved WLG motif in the middle of their AP2/ERF domains (Tab. 1). These 121 proteins are further classified into two groups on the basis of the similarity of the amino acid sequences of the DNA-binding domain: 56 proteins encode DREB/CBF-like proteins (group A) and 65 encode ERF (Ethylene Responsive Factor)-like proteins (group B). All proteins having a Val (V) in the 14$^{th}$ position and a Glu (E) in the 19$^{th}$ position of their AP2/ERF domains belong to group A, whereas those having a Ala (A) and a Asp (D) in the two corresponding positions belong to group B. Both group A and B proteins are further divided into six small subgroups, from A-1 to A-6 and B-1 to B-6, based on similarity in the AP2/ERF domains (Sakuma *et al*. 2002).

| DNA-binding domain type | | Subgroup | Total genes of *Arabidopsis* | Number of expressing genes[a] | Representative genes |
|---|---|---|---|---|---|
| Double ERF/AP2 | APETALA2 subfamily | | 14 | 7 | APETALA2, AINTEGUMENTA, RAP2.7 |
| ERF/AP2 + B3 | RAV subfamily | | 6 | 3 | RAV1, RAV2 |
| Single ERF/AP2 | DREB subfamily | A-1 | 6 | 4 | DREB1/CBF |
| | | A-2 | 8 | 2 | DREB2 |
| | | A-3 | 1 | 1 | ABI4 |
| | | A-4 | 16 | 3 | TINY |
| | | A-5 | 16 | 9 | RAP2.1, 9, 10 |
| | | A-6 | 9 | 5 | RAP2.4 |
| | | Subtotal | 56 | 24 | |
| | ERF subfamily | B-1 | 15 | 8 | AtERF3, 4, 7, 8, 9, 10 |
| | | B-2 | 5 | 3 | AtEBP, RAP2.2 |
| | | B-3 | 18 | 11 | AtERF1, 2, 5, 6, ERF1 |
| | | B-4 | 7 | 2 | RAP2.6 |
| | | B-5 | 8 | 2 | |
| | | B-6 | 12 | 3 | RAP2.11 |
| | | Subtotal | 65 | 29 | |
| | Others | | 4 | 2 | |
| | Single ERF/AP2 subtotal | | 125 | 55 | |
| Total | | | 145 | 65 | |

**Table 1.** Classification of Arabidopsis genes containing the AP2/ERF DNA-binding domain. Proceedings from Sakuma *et al*. 2002.

Transcription factors belonging to the AP2/ERF family that bind to DRE/CRT have been isolated and named Dehydration Responsive Element Binding factors (DREB). Two main classes of DREB factors were identified: DREB1/CBF and DREB2 (Liu *et al.*1998; Stockinger *et al.* 1997). The conserved DNA-binding motif of DREB1/CBF and DREB2 is A/GCCGAC (Kasuga *et al.* 1999; Sakuma *et al.* 2002). The DREB1A/CBF3, DREB1B/CBF1 and DREB1C/CBF2 are highly similar in amino acid sequences, and the genes occur in tandem within the Arabidopsis genome. These DREB1/CBF genes are quickly and transiently induced by cold stress (Shinwari *et al.* 1998), and their products activate the expression of >40 stress-inducible target genes (Fowler & Thomashow 2002; Maruyama *et al.* 2004, Gilmour *et al.* 2004). The *DREB2* genes are induced by dehydration, leading to the expression of various genes that are involved in drought-stress tolerance (Liu *et al.* 1998, Nakashima *et al.* 2000).

The DREB2A protein is expressed under normal growth conditions and is activated in the early stage of the osmotic and drought stress response through post-translational modification (Fig. 6). Very recently, a domain analysis of DREB2A revealed the existence of a negative regulatory domain in the central section of DREB2A, and deletion of this region renders DREB2A a constitutive active form (DREB2A-CA). Transgenic Arabidopsis plants overexpressing the DREB2A-CA showed growth retardation and improved tolerance to drought stress, like DREB1 overexpressing plants (Sakuma *et al.* 2006).

A comparative analysis between the transcriptome of Arabidopsis plants overexpressing DREB1A and DREB2A revealed that only 8 genes are expressed in common by the two types of transgenic plants. The target genes of DREB2A consist of at least 9 *LEA* members thought to confer dehydration tolerance. Moreover, by performing promoter analysis and gel mobility shift assay of the DREB1A and DREB2A up-regulated genes, it was underlined that DREB1A and DREB2A proteins have different binding specificities, therefore genes downstream DREB1A and DREB2A trigger different molecular effectors set conferring drought and freezing tolerance, respectively (Sakuma *et al.*, 2006).

## *DREB* HOMOLOGOUS GENES IN OTHER PLANT SPECIES

Multiple research efforts have been addressed to isolate drought- and osmotic- stress-inducible DREB-like transcription factors in several species of agronomic interest, due to the significance of gene regulation under dehydration and other abiotic stresses such as high salinity and cold. After that, through functional analyses, insights were gained into the *DREB* genes role and mechanism of action in model plants, such as Arabidopsis (Sakuma *et al*. 2006), and in species of agro-economic interest, such as rice, maize, bread wheat, barley (Skinner *et al*. 2005), soybean (Li *et al*. 2005), potato (Behnam *et al*. 2006), rape (Zhao *et al*. 2006) and spinach (Shen *et al*. 2003b).

In rice, five DREB homologs, *OsDREB1A-D* and *OsDREB2A* that show homology in their ERF/AP2 domains have been cloned. As Arabidopsis *DREB2A*, *OsDREB2A* was gradually induced by dehydration and high salinity stresses, but hardly increased under cold stress (Dubouzet *et al*. 2003).

In maize, a DREB1-type transcription factor, namely ZmDREB1A, was isolated by yeast one-hybrid screening (Qin *et al*. 2004; Jiang *et al*. 2004). Afterwards, the DREB2-type transcription factor ZmDREB2A from maize was isolated and two transcription forms, *ZmDREB2A-L* (long) and *ZmDREB2A-S* (short) were identified (Qin *et al*. 2007).

In barley, the gene for DREB2-type protein, *HvDRF1*, was reported to be accumulated under drought and salt stress and involved in ABA-mediated gene regulation (Xue & Loveridge 2004).

In bread wheat, the *TaDREB1* gene was found to be induced by drought, salinity and cold and was classified as a DREB2-type transcription factor by phylogenetic analysis (Shen *et al*. 2003a). The *wdreb2* gene was also identified (Egawa et al. 2006). Furthermore, in durum wheat, the *TdDRF1* gene was identified and preliminary characterized under my degree thesis (Latini *et al*. 2007).

A cDNA coding for a transcription activation factor DREB2A was also cloned in pearl millet (*Pennisetum glaucum*), namely *PgDREB2A*, and resulted to be more close to DREBs isolated from monocots. PgDREB2A is a phosphoprotein and the phosphorylated PgDREB2A does not bind to the DRE element (Agarwal *et al*. 2007).

Focusing on *DREB2*-type genes, homologous to Arabidopsis *DREB2A*, it has to be noted that these may differ for the gene structure: the Arabidopsis *AtDREB2A*, the rice *OsDREB2A* and the pearl millet *PgDREB2A* genes consist of two exons and one intron; the barley *HvDRF1*, the bread wheat *wdreb2* and the durum wheat *TdDRF1* genes consist of four exons and three introns, subjected to alternative splicing (AS).

Alternative splicing is the major contribute to the diversity of the transcriptome (Kriventseva *et al*. 2003). In human, genome-wide analyses indicate that most genes (up to 70%) generate different

transcripts with different exon contents by using alternative promoters, alternative polyadenylation sites and alternative splice sites (Brett *et al*. 2002; Stamm *et al*. 2005). About 40% of human genes produce at least five different splicing variants (SVs) and up to 10% of them produce more than 10 alternative transcripts each (de la Grange *et al*. 2007). The number of alternative spliced genes reported so far in plants is much smaller that in mammals, but it is increasing (Zhou *et al*. 2003). Large-scale AS analyses in plant were performed on the model *Arabidopsis,* thanks to the determination of its complete genomic sequence (Arabidopsis Genome Initiative 2000) and to the availability of extensive large-scale, full-length cDNA collections (Seki *et al*. 2002).

In Figure 7, I reported a phylogram, obtained using some isolated genes belonging to the AP2/ERF family from different plant species. I put in evidence that among the DREB2-like genes, as expected, the crop *DREB2* genes are distant from Arabidopsis and that, in particular, the genes of barley and wheat (*HvDRF1*, *TdDRF1* and *wdreb2*) are very closely related.

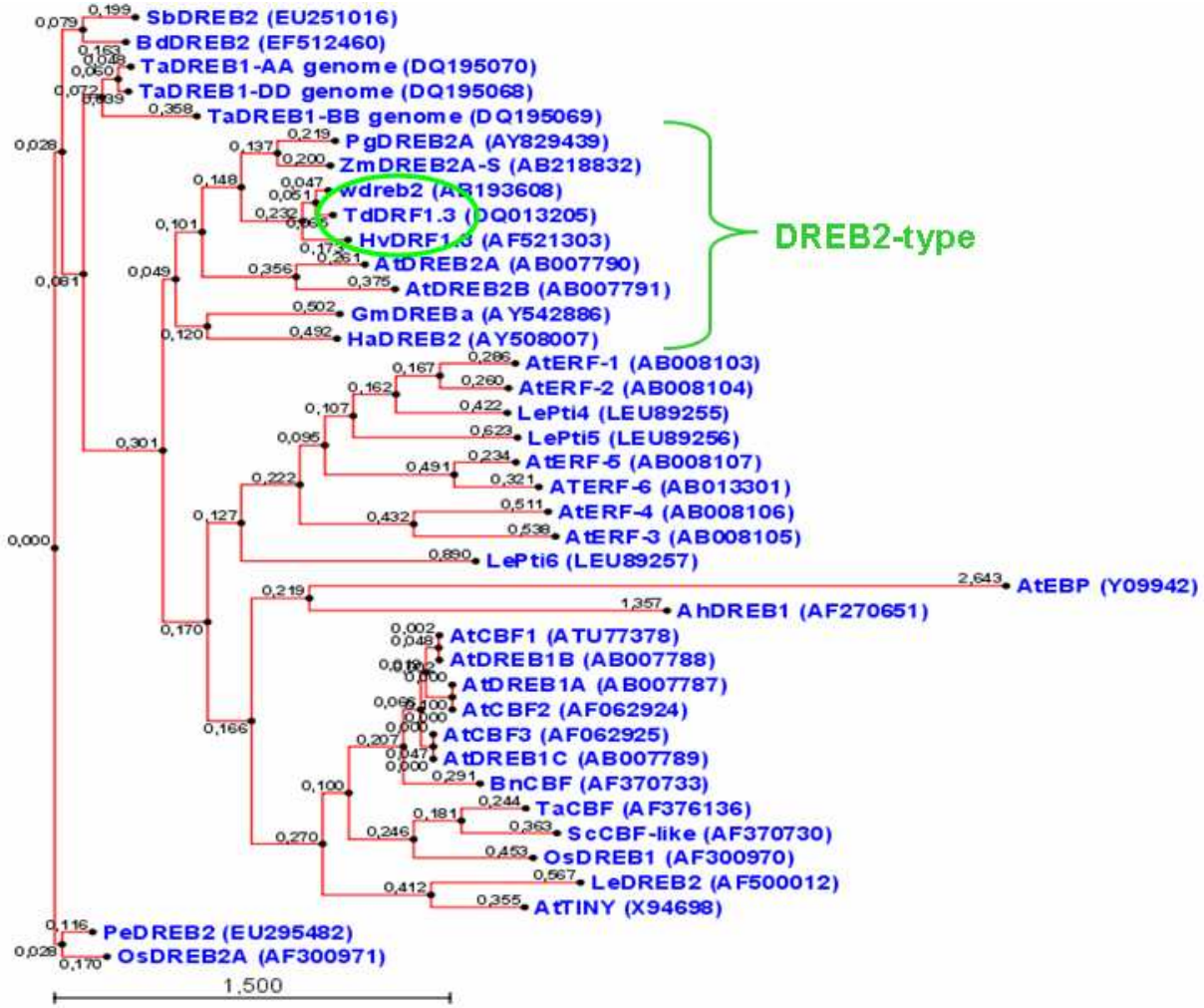



**Figure 7.** Phylogenetic tree of ERF/AP2 factors from different plant species. The tree was generated using the CDSs amino acid sequences with CLC Combined Workbench.

# CHAPTER 3

# TECHNICAL APPROACHES AND STRATEGIES USED IN THIS EXPERIMENTAL WORK


## SUMMARY

Functional genomics is the field of molecular biology that attempts to make use of the vast wealth of data produced by genomic projects (such as genome sequencing projects) to describe gene (and protein) functions and interactions. Unlike genomics and proteomics, functional genomics focuses on the dynamic aspects such as gene transcription, translation, and protein-protein interactions, as opposed to the static aspects of the genomic information such as DNA sequence or structures (http://en.wikipedia.org/wiki/Functional_genomics). Functional genomics technologies provide tools for detection and definition of cellular networks through which stress perception, signal transduction and defensive responses are mediated.

In this chapter there is a description of two of the major techniques used in functional genomics analyses and applied in the following experimental work: the real-time RT-PCR for quantifying gene expression and the genetically modified plants for the understanding of gene function and molecular networks.

## QUANTITATIVE PCR (qPCR)

The study of gene function requires the ability to accurately quantify temporal and spatial patterns of gene expression. The advent of real-time PCR, also named quantitative PCR (qPCR), has enabled rapid and reproducible high-throughput RT-PCR quantification, with an unparalleled dynamic range and extremely high sensitivity. For the above reason this technology has rapidly been adopted as a standard method for in-depth expression studies, including studies of alternative splicing, verification of microarray expression results, pathogen quantification, transgenic copy number determination and drug therapy studies (Bustin 2002; Bustin & Nolan 2004).

The kinetics of a PCR amplification reaction can be broken up into three phases:

1) *Exponential phase*: Exact doubling of product is accumulating at every cycle (assuming 100 % reaction efficiency), because the reagents are fresh and available.

2) *Linear phase*: The reaction components are being consumed, the reaction is slowing, and products are starting to degrade. In this phase there is a high variability.

3) *Plateau*: The reaction has stopped and no more products are being made. The plateau corresponds to the end-point of traditional PCR.

The amplification phases can be viewed in a logarithmic scale view (Fig. 8A) and in a linear scale view (Fig. 8B).

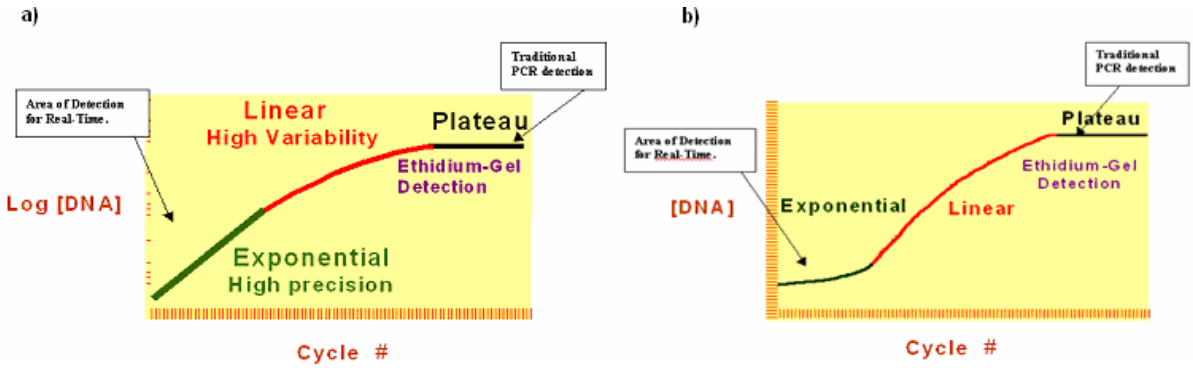



**Figure 8.** The kinetics of a PCR reaction. (a) Logarithmic scale view; (b) Linear scale view.

Traditional PCR analyses involve agarose gels for detection of the amplification products at the final phase (end-point) of the reaction. The result obtained by end-point PCR is independent on the plateau, corresponding to the saturation of the reaction and leading to inaccurate quantification.

Furthermore the agarose gel resolution is very poor, about 10 fold, and ethidium bromide for staining is not very quantitative.

On the other hand, the real-time PCR using a variety of different fluorescent dyes allows detecting the amplification during the exponential phase, while reaction is occurring. The principle standing at the base of the qPCR is that, during the exponential phase, PCR product will ideally double during each cycle if efficiency is perfect, i.e.100 %. It is possible to make the PCR amplification efficiency close to 100 % if the PCR conditions, primer characteristics, template purity and amplicon lengths are optimal. Together with the reaction efficiency, the other fundamental parameter for calculating qPCR results is the Threshold Cycle (Ct), defined as the fractional cycle number at which the fluorescence passes the fixed threshold. This value is inversely proportional to the starting copy number in the original template. An amplification plot is the plot of fluorescence signal versus cycle number.

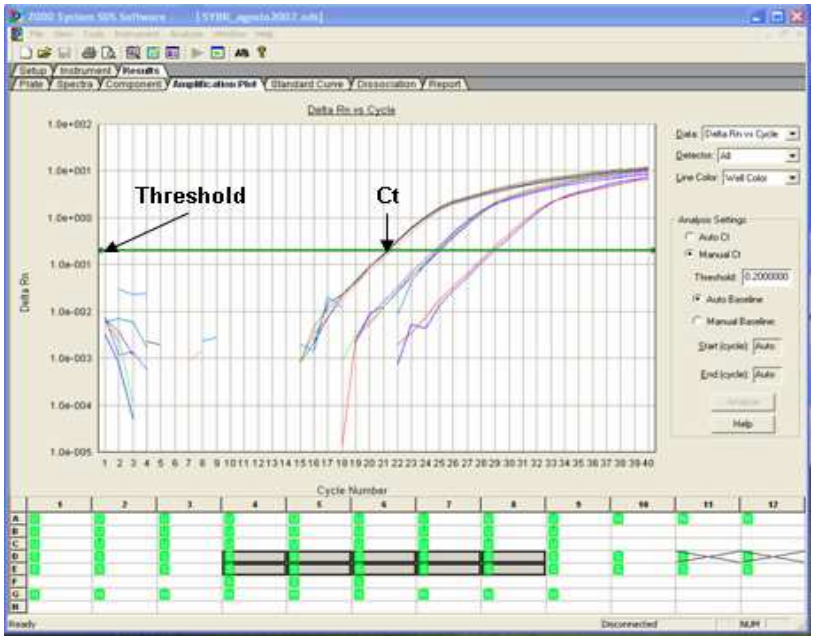



**Figure 9.** Output amplification plot of a serial dilution at three points from an ABI 7000 SDS (Applied Biosystems) real-time PCR instrument.

Different real-time PCR detection methods can be used for different applications. The SYBR Green I is an intercalating dye that binds the minor groove of double-stranded DNA and emits a fluorescence signal. As more double stranded amplicons are produced, SYBR Green signal will increase. Its major limitation is that it detects all dsDNA, including non-specific reaction products.

A dual-labeled fluorogenic probe (TaqMan probe) is a sequence-specific, single-stranded oligonucleotide labeled with two different dyes: a reporter dye is located on the 5'end and a quencher dye is located on the 3'end. The quencher dye inhibits the natural fluorescence emission of the reporter dye by Förster-type energy transfer.

The TaqMan chemistry is based on the 5' exonuclease activity of the Taq Polymerase. In the 5' nuclease assay, the probe is designed to anneal to a specific sequence of template between the forward and reverse primers. During the elongation step of the amplification, the probe is excited by light from the real-time qPCR instrument ($h\gamma_1$). The primer is elongated and the Taq 5' exonuclease activity hydrolyzes the probe bound to the specific DNA template, releasing the reporter dye from the target/oligonucleotide-quencher hybrid and causing an increase of fluorescence. The energy emitted by the reported dye ($h\gamma_2$) is detected by the real-time qPCR system. The significant increase of measured fluorescence signal is directly proportional to the amount of target DNA (Fig. 10). The greatest advantages of the TaqMan chemistry are that only the specific hybridization between probe and target allows the generation of the fluorescent signal and that the probes can be labeled with different, distinguishable reporter dyes, which allows amplification of one or more sequences in one reaction tube (multiplex PCR).

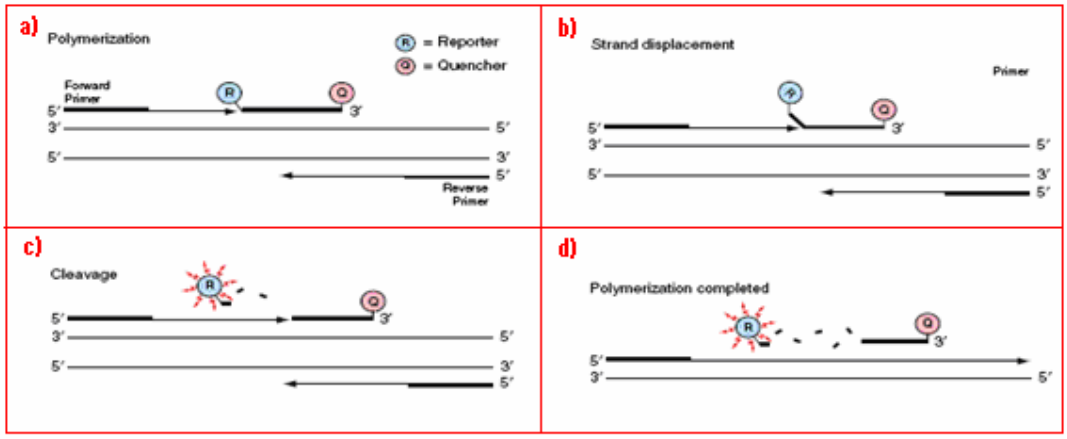



**Figure 10.** Schematic representation of how TaqMan detection chemistry works.

SYBR Green and TaqMan probes are the most commonly used detection techniques in real-time PCR, even though the desire to improve problematic quantitative assays is leading to continuous innovation and new technologies have appeared recently (Wong & Medrano 2005). Molecular Beacons (Tyagi & Kramer 1996; Tan *et al.* 2004), Scorpions (www.dxsgenotyping.com; Solinas *et al.* 2001) and Locked Nucleic Acids (LNA) probes (http://lna-tm.com/; Jepsen *et al.* 2004) are other

sequence-specific fluorescent probe systems engineered for real-time PCR (see Fluorescent Probes for real-time quantitative PCR of Proligo Primers and Probes at www.proligo.com). Some of their advantages rely in improved sensitivity and specificity, improved mismatch discrimination, when used for the detection of single nucleotide polymorphisms (SNPs), and shorter size that gives greater flexibility for design.

When quantitative RT-PCR (qRT-PCR) technique is used for the analysis of gene expression, there are several steps of the analytical process requiring a particular attention, such as assay design, template preparation, standardization of the reverse transcription and PCR protocols (Wolffs *et al*. 2004). Moreover the statistical analysis and interpretation of real-time data, is particularly important when the aim is the quantification of very low copy number targets. The biological sample acquisition and purification of its RNA mark the initial step of every qRT-PCR assay, and the quality of the template is arguably the most important determinant of the reproducibility and biological relevance of subsequent results. Regarding the reverse transcription, cDNA priming can be carried out using random primers, oligo-dT, or target-specific primers. Each of the three methods differs significantly with respect to cDNA yield (Zhang & Byrne 1999). To avoid bias, RT-PCR is typically referenced to an internal control gene, whose expression should not be influenced by the experimental conditions (Schmittigen & Zakrajsek 2000). However many studies showed that internal standards, mainly housekeeping genes used for the quantification of mRNA expression, could vary with the experimental conditions (Thelling *et al*. 1999; Stürzenbaum & Kille 2001; Radonic *et al*. 2004; Nicot *et al*. 2005).

Real-time PCR allows for two types of quantification: absolute quantification (AQ) and relative quantification (RQ). Absolute quantification uses serially diluted standards of known concentrations to generate a calibration curve. This curve produces a linear relationship between Ct and initial amounts of total RNA or cDNA, allowing the determination of the concentration of unknown samples based on their Ct values. This method assumes all samples have approximately equal amplification efficiencies.

Relative quantification measures the expression change of a gene target in comparison with either an external standard or a reference sample, also known as a calibrator (Livak & Schmittgen 2001). Examples of RQ applications are comparing expression levels of wild types vs. mutated alleles, a gene across different tissues or among different biological conditions, and are validating microarray results. There are two methods used to perform the RQ application: the $\Delta\Delta C_T$ method and the relative standard curve method.

The $\Delta\Delta C_T$ method requires an endogenous control, which is a gene target used to normalize quantitative PCR results. The endogenous control should consistently express in all samples and the

expression level should not be altered by any experimental condition. In order to use $\Delta\Delta C_T$ method, a validation experiment is necessary to ensure that the amplification efficiencies of endogenous control and target sequences are the same.

The relative standard curve method requires running a standard curve for every experiment. This method can be used when amplification efficiency of the target sequences and the endogenous control differs.

Since relative quantification is the goal of most for real-time PCR experiments, several data analysis procedures have been developed. There are numerous mathematical models for calculating the mean normalized gene expression from relative quantification assays (Pfaffl 2001; Muller *et al*. 2002; Simon 2003).

## TRANSGENIC PLANTS

In plant functional genomics studies, the gene overexpression or inactivation represent a very useful strategy to determine the function and molecular partners of a gene of interest (Bajaj *et al.* 1999; Bhalla 2006; Christensen & Feldmann 2007).

To reveal if a gene could play a role in drought tolerance, transgenic plants overexpressing the target gene have been developed and studied in greenhouse under different water lack settings and the advancement of DNA microarray technology allowed to identify several hundred stress-inducible genes as candidate genes for genetic engineering (Vinocur & Altman 2005; Umezawa *et al*. 2006). Multiple studies have attempted to augment plant stress tolerance by overexpressing various kinds of functional proteins such as enzymes for osmolyte biosynthesis and ion transporters (Zhang *et al*. 2004). However, the introduction of any single gene may not give sustained tolerance to abiotic stresses, because multiple stress responses are necessary for plant to endure severe stress conditions; moreover, the constitutive expression of these genes by a strong constitutive promoter may have serious implication with respect to energy loss and other deleterious effects (Shimamura *et al*. 2006).

On the other hand, the regulated expression of stressed-inducible transcription factors would regulate the expression of a large number of relevant downstream genes (Holmberg & Bulow 1998). For example, to a great extent, the actual knowledge regarding the *Arabidopsis* transcription factors DREB1A and DREB2A, involved in the response to drought, salt and cold stresses, was achieved by this way (Yamaguchi-Shinozaki & Shinozaki 1994; Stockinger *et al*. 1997; Shinozaki & Yamaguchi-Shinozaki 2000) and the manipulation of the expression of DREB TFs opens an excellent opportunity to develop stress tolerant crops in future (Agarwal *et al*. 2006). Transgenic *Arabidopsis* plants overexpressing DREB1A/CBF3 under the control of the cauliflower mosaic virus (CaMV) 35S promoter showed an increased tolerance to drought, high-salinity and freezing stress, besides growth retardation and alterations in basic metabolism under normal growth conditions (Liu *et al*. 1998; Kasuga *et al*. 1999; Zhang 2003). Use of stress-inducible *rd29A* promoter instead of the constitutive CaMV 35S one for the overexpression of DREB1A/CBF3 minimized the negative effects on plant growth. In a different way, overexpression of DREB2A neither caused growth retardation, nor improved stress tolerance, suggesting that DREB2A protein requires post-translational modifications such as phosphorylation for its activation (Yamaguchi-Shinozaki & Shinozaki 2006; Sakuma *et al*. 2006).

As regards the attenuation or annulment of the gene expression, the gene silencing is the approach to specifically and efficiently down-regulate a gene activity. This procedure makes use of constructs

expressing double-stranded RNA (dsRNA), usually in the form of self-complementary hairpin RNA (hpRNA) (Waterhouse & Heliwell 2003).

Plant genetic manipulation can be achieved by two possible ways: transient and stable transformation (Izquierdo Rojo 2001).

**Stable *vs* transient expression**

*Advantages*:

- the transgene is generally expressed in all the cells when it is under the control of a constitutive promoter;
- plant tissues are not stressed or damaged by the transformation procedure (no vacuum, syringes, bombardment, viruses, etc.).

*Disadvantages*:

- the generation of transgenic lines is time-consuming and, with some plant host, this can be a long process lasting some years;
- it requires a considerable investment because of the high number of plants that has to be handled and analyzed;
- it involves the chromosomal integration of the gene of interest and, consequently, the potential disruption of some plant function;
- the transgenic plant approach is complicated by ecological safety concerns, particularly if the transgenic plants are grown outdoors. However, when this technology is applied to study objectives, in a controlled and limited space, it does not represent a danger for human, animal and environmental health.

**Transient *vs* stable expression**

*Advantages*:

- rapid and convenient engineering;
- very high levels of gene expression can be achieve within a short-time (few days) and, however, the expression levels can be manipulated by increasing/decreasing the amount of the inoculated DNA;
- the gene of interest, after cloning in a suitable vector, is expressed without integration in the plant genome and therefore is not transferred to the progeny;
- the mRNA that has to be over-expressed can be easily introduced into the plant, where it is

able to spread autonomously across the plasmodesmata and the vascular system, miming a viral infection.

*Disadvantages*:

- the viral symptoms showed by the infected plant can interfere with the gene expression;
- once introduced into the plant, the viral nucleic acids present a high velocity of spontaneous mutations and the instability of the foreign gene in the viral genome can present some problems (Scholthof *et al*. 1996).

There are three major ways to introduce DNA for transient expression systems (Horser *et al*. 2002):

1. Delivery of "naked" DNA by particle bombardment.
2. Infection with modified viral vectors.
3. Infiltration of plant tissue with recombinant *Agrobacterium* (agro-infiltration).

Particle bombardment usually delivers the DNA only to few cells and the DNA has to be moved to the cell nucleus, whereas, when the DNA delivery is assigned to a modified viral vector, infection results in the systemic spread of the virus throughout most of the plant.

Autonomously replicating viruses offer numerous advantages for use as vehicles for transient expression of foreign genes, including their characteristic high levels of multiplication and concomitant levels of transient gene expression. Because of the nature of virus infections, the maximum level of foreign gene expression from viral genomes is predicted to occur within a short time-period, usually within one or two weeks after inoculation (Scholthof *et al*. 2002).

Agro-infiltration targets a significantly higher number of cells compared to particle bombardment and the T-DNA region harbouring the gene of interest is actively transported into the nucleus with the aid of several bacterial proteins. In the case of particle bombardment and agro-infiltration, the protein of interest is synthesized within 3–5 days following the treatment. Agro-infiltration also has the advantage that it yields good amounts of proteins and is not limited to small proteins, but can accommodate large complex proteins (Negrouk *et al*. 2005).

Most viruses of the vegetable kingdom present a RNA genome. Some plant viruses can be utilized as vectors for foreign DNA expression in a plant host, by employing a gene replacement strategy (Ahlquist & Pancha 1990; Porta & Lomonossoff 1996). Viruses offer several advantages in the transient expression of a gene of interest and, in particular, they are very efficient in the multiplication, generating a high copy number of the transgene and its encoded protein.

The major problem in the use of plant viruses as DNA vectors is the genetic instability and the insert removal, especially when the gene of interest is longer than 1 kb. Anyway this common attitude to remove the insert, manifested by plant viruses, reflects a general strategy, selected during the evolution, against the incorporation of DNA segments belonging to the plant, which does not give any advantage to the virus.

# INTRODUCTION TO THE EXPERIMENTAL SECTIONS

In response to water deficit, plants trigger qualitative and quantitative changes in their gene expression patterns and resulting proteins. Improved tolerance to water deficit and therefore adaptation to drought may depend on efficient signal perception and adequate responses at the transcriptional level, the extent of which can vary not only among species but even among genotypes of the same species (Sreenivasulu *et al.* 2007; Grzesiak *et al.* 2007).

The DREB proteins are important transcription factors that induce a set of abiotic stress-related genes and impart stress endurance to plants. By providing a better understanding of a crop's response to water deficit at the molecular level, this area of research can eventually lead to the development of effective tools and strategies for improving drought tolerance in commercially important crops (Agarwal *et al.* 2006), in particular for wheat which is grown to a great extent in drought-prone environments.

In this context, in a previous study, my collaborators at ENEA Research Center (Rome, Italy) and I have isolated a DREB2-related gene in durum wheat, designated as *TdDRF1* (*Triticum durum* Dehydration Responsive Factor 1), and established that it is expressed in response to dehydration (Latini *et al.* 2007). This gene produces three transcript variants, namely, *TdDRF1.1*, *TdDRF1.2* and *TdDRF1.3***,** through alternative splicing (Fig. 11).

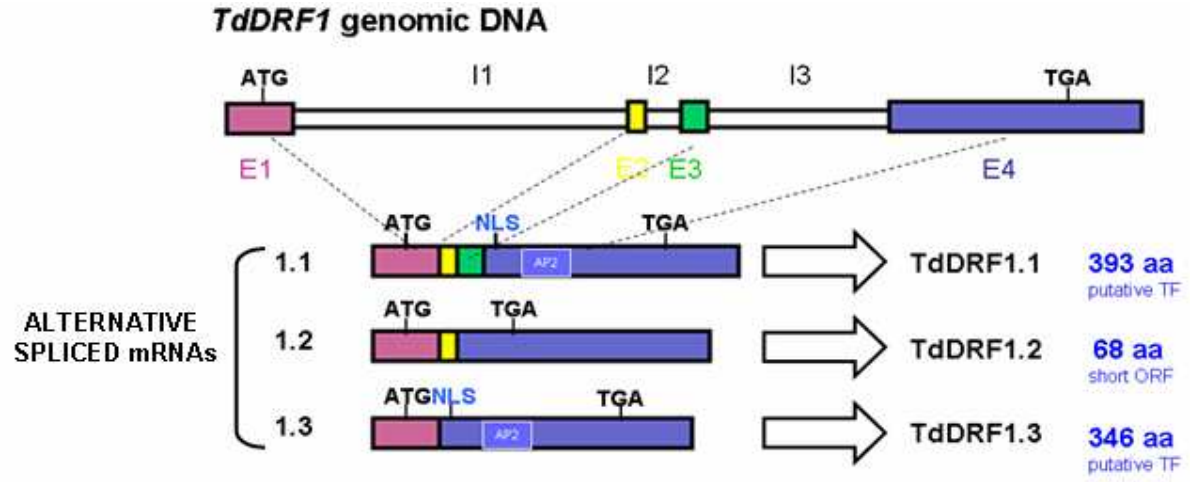



**Figure 11.** The genomic structure of *TdDRF1* gene and its alternatively spliced transcripts.

Homologous genes and transcripts were also described in *Hordeum vulgare, HvDRF1* (Xue & Loveridge 2004), and in *Triticum aestivum*, *TaDRF1/wdreb2* (Egawa *et al*. 2006). The transcript isoforms *TdDRF1.1* and *1.3* encode transcriptional activators and contain an AP2/ERF DNA-binding domain and a nuclear localization signal (NS) close to the N terminus. *TdDRF1.2* encodes a putatively truncated protein, lacking both the AP2 DNA-binding domain and the nuclear localization signal. The downstream target genes under the regulation of these transcription factors belong to the *Cor/Lea* (Cold Responsive/Late Embryogenesis Abundant) family genes (Curry *et al*. 1991; Rampino *et al*. 2006).

# EXPERIMENTAL SECTION 1

# DEHYDRATION RESPONSIVE FACTOR 1 (*TdDRF1*) GENE: MOLECULAR ANALYSES UNDER GREENHOUSE CONDITIONS IN SOME DROUGHT TOLERANT DURUM WHEAT AND TRITICALE CULTIVARS


## SUMMARY

The specific objectives of the present study were to explore the existence of genotype-dependent *TdDRF1* gene transcripts expression patterns upon dehydration in four durum wheat and one triticale cultivars, under controlled greenhouse conditions, and to observe genotypic differences in the sequence of one of the transcription factor proteins produced by this gene.

# MATERIALS AND METHODS

### Plant materials and water stress treatment

Four *Triticum durum* varieties, namely Ciccio, Atil C2000, Karalis and Vitromax and the Triticale cultivar Pollmer TCL 2003, were used in this study (Tab. 2). Ciccio (released in 1996) is an Italian variety adapted mostly to southern Italy. Atil C2000 (released in 2000) is a high yielding cultivar for the irrigated areas of northern Mexico and selected from the International Maize and Wheat Improvement Centre (CIMMYT) program. Karalis (released in 2002) is a promising new cultivar from the Centro Regionale Agrario Sperimentale (CRAS-Italy) adapted to the dry conditions of Sardinia. Vitromax (released in 2001) is a Spanish commercial durum wheat cultivar with good adaptation to rainfed environments. Pollmer TCL 2003 is a Triticale cultivar from Mexico developed by the CIMMYT program and selected for its adaptation to both irrigated and dry areas.
Seeds were put on moist filter paper in Petri dishes for 3-4 days at room temperature, in the dark, and transferred after germination into pots containing 1 l of soil (one seed per pot). Plants were grown in a greenhouse under a constant photoperiod of 16 hours at 20-22 °C during the day and at 18-20 °C during dark. Pots were watered daily with 50 ml of water. The time-course water stress experiments initiated on the last watering day after plants reached the heading stage (time zero, T0). For each genotype, the leaves were collected at T0, T0+2 days (T2), T0+4 days (T4) and T0+7 days (T7) from both control (always watered) and stressed (not watered after heading) plants.

| Genotype | Species | Genome Composition | Origin | Drought-Responsive Phenotype |
|---|---|---|---|---|
| **1. Ciccio** | *T. durum* | AA BB (Tetraploid) | Italy | ~ tolerant |
| **2. Atil C2000** | *T. durum* | AA BB (Tetraploid) | Mexico | ~ tolerant |
| **3. Karalis** | *T. durum* | AA BB (Tetraploid) | Italy | ~ tolerant |
| **4. Vitromax** | *T. durum* | AA BB (Tetraploid) | Spain | ~ tolerant |
| **5. Pollmer TCL 2003** | *Triticale* | AA BB RR (Hexaploid) | Mexico | ~ tolerant |

**Table 2.** List of the genotypes used in the following analyses.

### Relative Water Content (RWC)

In order to monitor the physiological dehydration status of the plants, leaf relative water content (RWC) was determined. Ten fully expanded leaves were collected at T0, T4 and T7 from each genotype, under both control and water stress conditions, and the fresh weight (FW) was immediately recorded. The turgid weight (TW) was recorded after re-hydration of the leaves by floating on distilled water in a close Petri dish for 24 h at 4 °C in the dark. The dry weight (DW)

was recorded after desiccation for 24 h at 80 °C. At the end, leaf RWC was measured for each sample according to the commonly used formula (Barrs & Weatherley 1968):

RWC (%) = [(FW-DW) / (TW-DW)] x 100.

Two independent measurements were performed for each sample.

**Qualitative RT-PCR**

To qualitatively assess the presence of the *TdDRF1* transcripts in all genotypes under uniform analyses conditions, total RNA was extracted from leaves collected at the 4$^{th}$ day of water stress (T4) with the RNAfast$^{TM}$-II Isolation System (Molecular Systems, USA). One step RT-PCRs were performed using Ready-to-Go RT-PCR Beads (GE Healthcare, UK) and a unique couple of primers, with the forward primer located in exon 1 (*E1for:* 5'-CATGACGGTAGATCGGAA-3') and the reverse primer in exon 4 (*E4rev*: 5'-GCCTGATCGTTTCAGAGA-3'). The relative location of these exons within the genomic sequence of *TdDRF1* and its mRNAs is shown in Figure 11. The reaction mix (50 µl), containing 500 ng of total RNA and 20 pmol of each forward and reverse primer, was run in the TaKaRa PCR Thermal Cycler Dice$^{TM}$ (Takara Bio Inc., Japan). Samples were held 30 min at 42 °C for the reverse transcription, 5 min at 95 °C for the RT inactivation and initial denaturation, and then amplified with 40 cycles, each of 40 s at 95 °C, 40 s at 50 °C and 70 s at 72 °C, followed by a final step of 5 min at 72 °C. Amplification products were run in 3% agarose and stained with ethidium bromide. Densitometric analysis of the bands was done using the UVIBandMap software (Uvitec, UK).

**Time-course RT-PCR**

Samples from the cultivar Atil C2000 collected at T0, T2, T4 and T7, for both control and water stressed conditions, were analysed by end-point RT-PCR using three different primer pairs designed to differentially amplify each of the three transcripts, *TdDRF1.1*, *TdDRF1.2* and *TdDRF1.3,* individually (Tab. 3).

| Target | Gene Bank Accession No. | Forward Primer (5' to 3') | Reverse Primer (5' to 3') | Amplicon Size (bp) | Rxn Ta (°C) |
|---|---|---|---|---|---|
| ***TdDRF1.1*** | DQ013204 | AATGAATATGCACTCTTGGCGC | CGTTGCCTTACACCACGGAAT | 278 | 61 |
| ***TdDRF1.2*** | DQ011890 | TCAAACCAATAGGAAAAAGCGACC | CGTTGCCTTACACCACGGAAT | 222 | 62 |
| ***TdDRF1.3*** | DQ013205 | CGCTCCAGCCTGGAAGGAAAAA | CGTTGCCTTACACCACGGAAT | 226 | 63 |
| ***actin*** | CJ967245 | GACCCAGACAACTCGCAACT | CTCGCATATGTGGCTCTTGA | 502 | 58 |

**Table 3.** Primers used in the time-course and expected length of the amplicons.

Reactions used Ready-to-Go RT-PCR Beads (GE Healthcare) in a TaKaRa PCR Thermal Cycler Dice™ (Takara Bio Inc.). In Table 3, the reaction annealing temperatures ($T_a$) used in the PCR as well as the expected lengths of the amplicons are reported. Reaction samples (50 µl) containing 1 µg of total RNA and 30 pmol of each primer were run using the same thermal cycling conditions as above. To account for differences in the relative amount of RNA used, wheat *actin* fragment was used as internal control and was amplified in separate tubes using a primer pair described by Pellegrineschi (2004). PCR products were run in a 3% AGAROSE High Resolution (Sigma-Aldrich, USA) gel and stained with ethidium bromide. Direct calculation of the band intensity (expressed as peak volume) was done using the UVIBandMap software (Uvitec).

**Quantitative real-time RT-PCR**

*RNA preparation and reverse-transcription:* For each sampling time (T0, T4 and T7), flag leaves from three representative individual plants were pooled together, immediately frozen in liquid nitrogen and stored at -80 °C until extraction. Total RNA was extracted using the TRI® Reagent (Sigma-Aldrich) according to the manufacturer's instructions. RNA concentrations were measured by NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, USA). The quality and purity of the preparations were evaluated using the $OD_{260}/OD_{280}$ and $OD_{260}/OD_{230}$ nm absorption ratios and the integrity of the preparations was ascertained by agarose gel electrophoresis (2%).

Total RNA was reverse transcribed with oligo $d(T)_{16}$ using TaqMan® Reverse Transcription reagents (Applied Biosystems, USA). Multiple 10 µl reactions were performed using 400 ng of total RNA for each sample.

*Experimental design:* Real-time PCR reactions were carried out in the ABI PRISM® 7000 Sequence Detection System (Applied Biosystems) using the TaqMan® MGB (Minor Groove Binder) probe-based technology. Primers and probes were designed using the Assay-By-Design software (Applied Biosystems) according to the manufacturer's instructions. Figure 12 schematically shows the localization of primers and probes for the differentiation of the three transcripts and Table 4 includes the sequence details of the oligos used. Initially, four amounts of reverse transcribed total RNA (25, 50, 150 and 250 ng) were tested for each transcript and each sampling time, to establish the efficiencies of the amplification reactions. Finally, the quantitative PCR analyses were carried out using 50 ng total RNA/sample.

For each PCR reaction a final volume of 20 µl was used, containing cDNA, 2X TaqMan® PCR MasterMix (Applied Biosystems) and 20X TaqMan® Gene Expression Assay (Applied Biosystems). Standard thermal cycling conditions included initial 2 min at 50 °C and 10 min at 95 °C followed by 40 two-step cycles including denaturation of 15 s at 95 °C and

annealing/denaturation of 1min at 60 °C. The amplified bands corresponding to each of the three transcripts were cloned in pCR®II-TOPO® (Invitrogen, USA) and then sequenced.

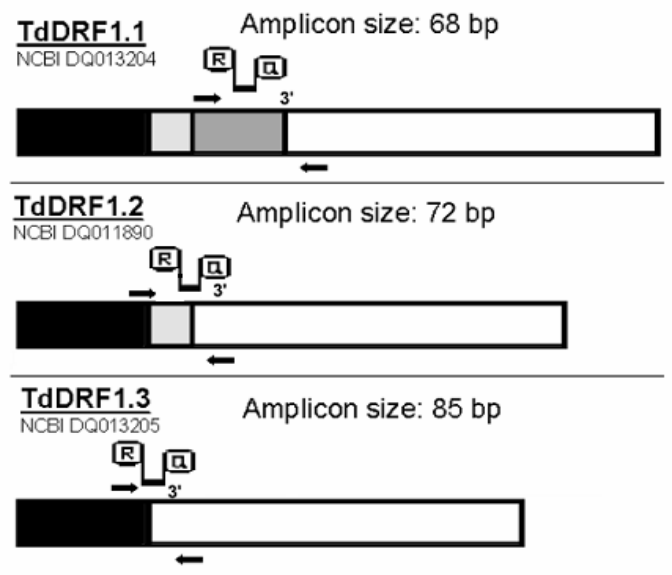



**Figure 12.** Experimental design of TaqMan primers and probes for the selective amplification and detection of the three *TdDRF1* gene transcripts. Accession numbers refers to Ciccio cultivar, whose cDNAs proceeded from leaves at the 4th day without water.

| Target | Forward Primer (5' to 3') | Amplicon Size (bp) |
|---|---|---|
| ***TdDRF1.1*** | Forward Primer GCAAAACCCCAAGGGAGATG<br>Reverse Primer GTGATCTCCGAGGTCGCTTTT<br>TaqMan® Probe CCTGTGCCATCTATTCTG | 68 |
| ***TdDRF1.2*** | Forward Primer AGAACTTGTGGAGCAGAGGAAAG<br>Reverse Primer TCTCCGAGGTCGCTTTTCCTA<br>TaqMan® Probe TCATGTTCTCGTCAAACCAA | 73 |
| ***TdDRF1.3*** | Forward Primer GCCGCACCCTTCGAGAT<br>Reverse Primer CAGAGACTGAATTAGGCCCATCAC<br>TaqMan® Probe CAGCCTGGAAGGAAAA | 85 |

**Table 4.** Primers and probes used in real-time. Each probe includes the FAM (490 nm) reporter dye at the 5' terminus ad a no fluorescent quencher (NFQ) at the 3' terminus.

Wheat *actin* amplicon (see Table 3) using the SYBR® Green reagent chemistry was chosen as endogenous control in quantitative PCR analyses. Each 20 μl reaction for the endogenous control

contained 10 ng of reverse transcribed cDNA, 10 μl of SYBR® Green PCR Master Mix (Applied Biosystems), 20 pmol of each *actin-for* and *actin-rev* primers. Initially, four serial dilutions of cDNA template (100, 10, 1 and 0.1 ng) were utilized to determine the PCR efficiency for the *actin* fragment.

All targets, endogenous controls and no template controls (NTC) reactions were run in triplicate.

*Post-run data analysis:* The values of Threshold Cycle (Ct) were exported from the ABI PRISM 7000 SDS (Applied Biosystems) and processed in Microsoft Excel. Linear regression analyses using the least squares method were used in plotting average Ct values versus log of template concentration. The amplification efficiencies were calculated from the slopes of these straight lines, according to the equation $E_x=10^{(-1/slope)}$ (Ramakers *et al.* 2003; Rutledge & Cote 2003). Given that the efficiencies of amplification for targets and reference transcripts are very similar, the comparative Ct method was used for the relative quantification. The expression level of the three target sequences, normalized to that of the endogenous control (*actin*) and relative to that for sampling time T0, was estimated using the formula $2^{-\Delta\Delta Ct}$ (Applied Biosystems User Bulletin #2; Livak & Schmittigen 2001).

**Cloning and sequence analysis of *TdDRF1* cDNAs**

For each of the five genotypes, the complete codifying sequences (CDSs) of *TdDRF1* gene were amplified by RT-PCR, using the following primers: *E1for-upstream* (5'-CACGACTCTCCCAACCTCTC-3') located in the 5' untranslated region (UTR) and *E4rev-downstream* (5'-GGTCCACCATTTGATCTTCATT-3') located in the 3' UTR.

The bands produced were extracted from gel, purified with MinElute Gel Extraction Kit (Qiagen, USA) and cloned in pCR®II-TOPO® (Invitrogen). For each genotype, 6 positive clones were sequenced and sequences were analysed by EMBL-EBI (www.ebi.ac.uk), Expasy (www.expasy.org), NCBI (www.ncbi.nlm.nih.gov) and CLC Combined Workbench 3.0.3 (CLC bio, USA) tools.

**AP2 structural model building**

The three-dimensional (3D) structures of the AP2 DNA-binding domain of Ciccio (ap2_c), Atil C2000 (ap2_a), Pollmer TCL 2003/Karalis (ap2_pk), Vitromax type *a* (ap2_va) and Vitromax type *b* (ap2_vb) were derived by homology modelling using as template the GCC-box binding domain of Arabidopsis AtERF1 structure (PDB id: 1GCC(A)), the only available AP2 structure determined by heteronuclear multidimensional NMR (Allen *et al.* 1998). The sequence homology between the 1GCC template with the ap2 targets was about 80% with sequence identity of about 65%. Sequence

alignments were derived using ClustalW (Thompson *et al.* 1994). Structure alignments were obtained and analysed by Protein Structure Comparison Service (http://www.ebi.ac.uk/msd-srv/ssm) (Krissinel & Henrick 2004). Both aligned results were inspected and adjusted to minimize the number of gaps and insertions.

Homology models of the ap2 structures were built using Modeller v9 (Fiser & Sali 2003), a comparative protein modelling program that generates three-dimensional coordinates through satisfaction of spatial restraints. The program generates a set of distances and dihedral angles restraints from the structure used as template and produces a model for the target sequence that minimizes the violations of the set of restraints. A set of ten model structures were generated for each ap2 target and ranked by their molecular probability density function (pdf) values obtained after highest optimization level. From these sets, a single structure for each AP2 domain was selected on the basis of the lowest pdf value, the stereochemistry and energy quality as assessed by Procheck (Laskowski *et al.* 1993) and Prosall (Sippl 1993).

# RESULTS

## Physiological analysis

In order to ensure that the various leaf tissue samples used in the analyses were under the expected physiological water stress level, the relative water content (RWC) was determined for all genotypes and monitored as stress progressed. In the present study, all cultivars, in the absence of water-stress, exhibited a RWC characteristic for non-stressed tissues (90-100%). RWC values decreased significantly at the 4$^{th}$ day of withholding water application and even further at the 7$^{th}$ day. This clear trend was observed for all genotypes with no significant differences among them. The evolution over time of the plants' dehydration status as measured by their RWC is shown and discussed together with the real-time RT-PCR analyses.

## Qualitative RT-PCR analysis

Qualitative RT-PCR using RNA from plants at T4 and primers *E1for* and *E4rev* was performed to assess the presence of *TdDRF1* transcripts in the selected genotypes and to reveal any genotype-dependent qualitative or semi-quantitative differences in the expression patterns of the three transcripts. The signal for *TdDRF1.1* was the faintest among those for the three splicing variants, in all genotypes. The highest peak volume was that of splicing variant *1.2* for all cultivars, with *1.3* showing intermediate values. For a given splicing variant, peak volumes were similar among durum genotypes. The peak volumes observed for the triticale cultivar were substantially lower than those from durum wheat, for all splicing variants (Fig. 13).

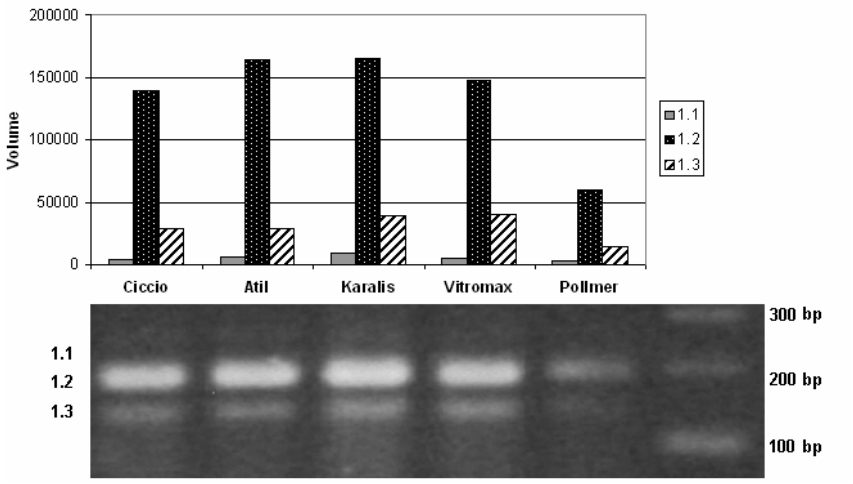



**Figure 13.** Qualitative RT-PCR with *E1for* and *E4rev* in the different genotypes at the 4$^{th}$ day without water. Densitometric analysis of the gel bands is shown in the upper part of the figure.

**Time-course RT-PCR analysis**

In order to investigate the relative abundances of three *TdDRF1* transcripts over time as dehydration occurred, a preliminary end-point RT-PCR analysis was carried out in one variety, Atil C2000, following the T0, T2, T4 and T7 time-course in both control and stressed conditions. In this case, each splicing product was amplified separately.

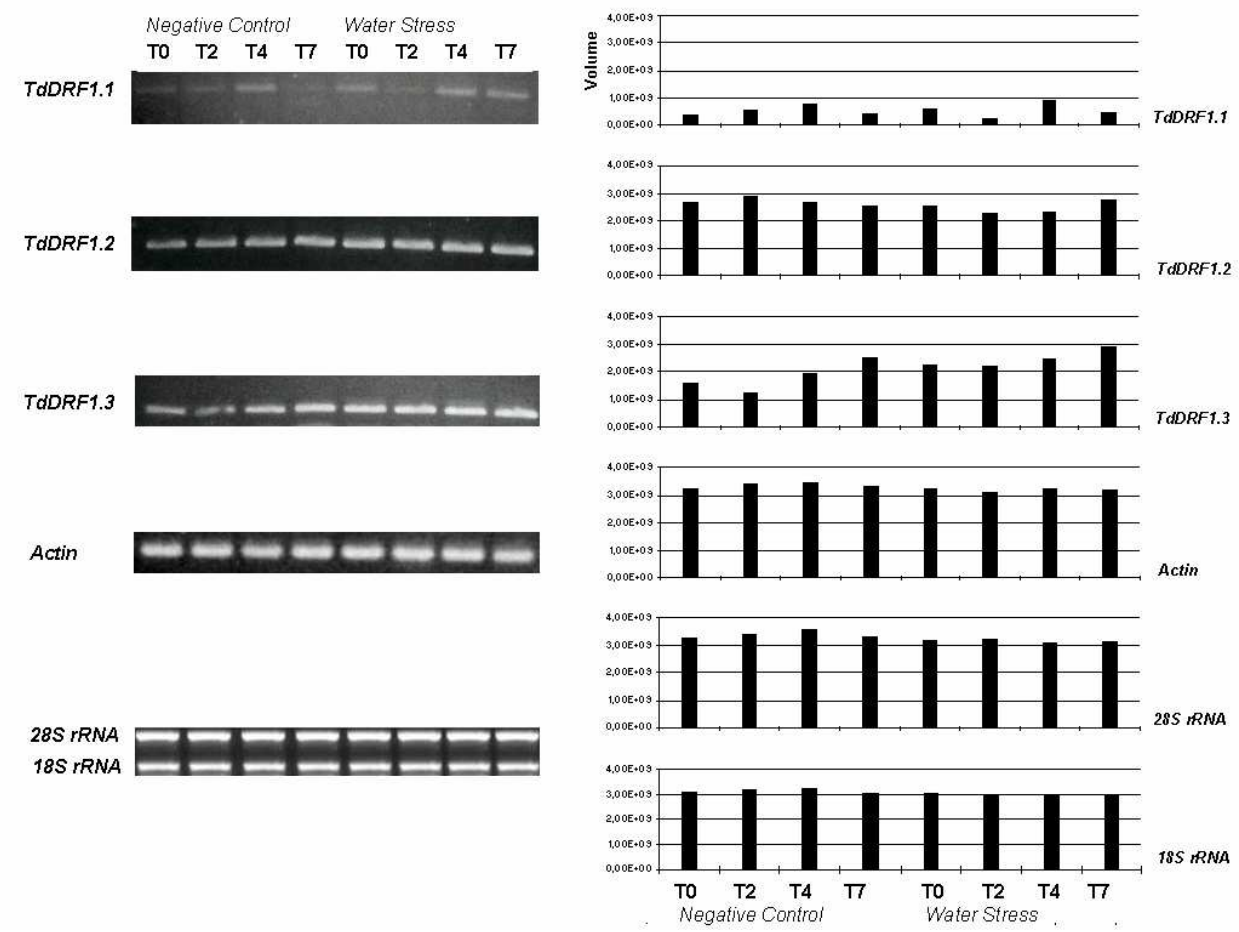



**Figure 14.** Time-course RT-PCR in Atil C2000 cultivar. Gel bands on the left side and their densitometric analysis on the right side. Lane 1-4) well-watered conditions: T1, T2, T3 and T4, respectively. Lane 5-8) water stress conditions: T1, T2, T3 and T4, respectively. From the top to the bottom: *TdDRF1.1*, *TdDRF1.2*, *TdDRF1.3*, *actin* as internal control and RNA (500 ng/lane).

The results shown in Figure 14 confirm that the *TdDRF1.1* transcript is the least abundant one over all the stress time-course. A slight increase in its intensity is observed at T4, especially under water stress. Differences between *TdDRF1.2* and *TdDRF1.3* are not as large as shown by the qualitative RT-PCR analysis, especially under water-stress conditions. Considered individually, it is apparent that *TdDRF1.2* is present in all samples (control and stressed) as a relatively constant and very

intense band, whereas *TdDRF1.3* is visible more intensely in the stressed samples, in particular at T4 and T7.

The *actin* fragment amplification and the extracted RNA (500 ng) results are also shown in Figure 14. They demonstrate that equal quantities of RNA template were used in the RT-PCR reactions.

**Real-time RT-PCR analysis**

To obtain a better quantification of the relative abundances of the three transcripts, real-time RT-PCR experiments were performed (Baek & Skinner 2004). Under ideal conditions, real-time PCR allows following the amplification throughout its exponential phase, phase during which the quantity of amplified PCR product is proportional to the quantity of initial template present in the reaction (Freeman *at al.* 1999). A simplified time-course at three points was selected: T0 (last watering and calibrator for this relative quantification), T4 and T7. We used the TaqMan® MGB probe-based technology to obtain a selective amplification and detection of each transcript with a good reliability. The design of real-time PCR primers and probes was based on the sequence of the three transcripts from the cultivar Ciccio, previously submitted to NCBI (accession numbers DQ013204, DQ011890, DQ013205 for *TdDRF1.1*, *1.2* and *1.3*, respectively), as shown in Figure 12. The amplification/detection specificity for each target was confirmed by electrophoresis of the samples at the final PCR stage (Fig. 15). Furthermore, the band corresponding to each of the three transcripts was cloned and sequenced.

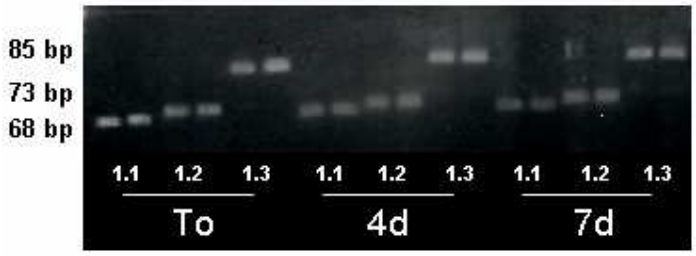



**Figure 15.** Confirmation of primers and probes specificity. Results refer to Ciccio cultivar. Electrophoresis was carried out on 3% agarose gel stained with ethidium bromide.

The amplification efficiencies (E) were high for all amplification products ($1.80<E<2.05$; $0.98<R^2<0.99$). The *actin* gene, used herein as endogenous control, had an amplification efficiency ($E=1.85$ $R^2=0.998$) similar to those for the other target sequences and showed no genotype- and/or time-dependent variation. This result implies the $\Delta\Delta Ct$ calculation is valid. The Ct values of *actin*

and *TdDRF1.2* were characterized by the best reproducibility, with standard deviations of the average Ct values ≤0.1.

Results, shown in Figure 16 (average Ct values exported from the ABI Prism 7000 SDS software), indicate that *TdDRF1.2* exhibited the highest expression levels (lowest Ct values) and *TdDRF1.1* the lowest ones at all sampling times/RWC and for all genotypes. *TdDRF1.1* could not be detected, under our experimental conditions, after T4 for the cultivar Vitromax and was absent in the triticale cultivar Pollmer TCL 2003 during the whole time course, including at T0. These quantitative results are in good agreement with those obtained with the previously described qualitative RT-PCR analyses (Fig. 14).

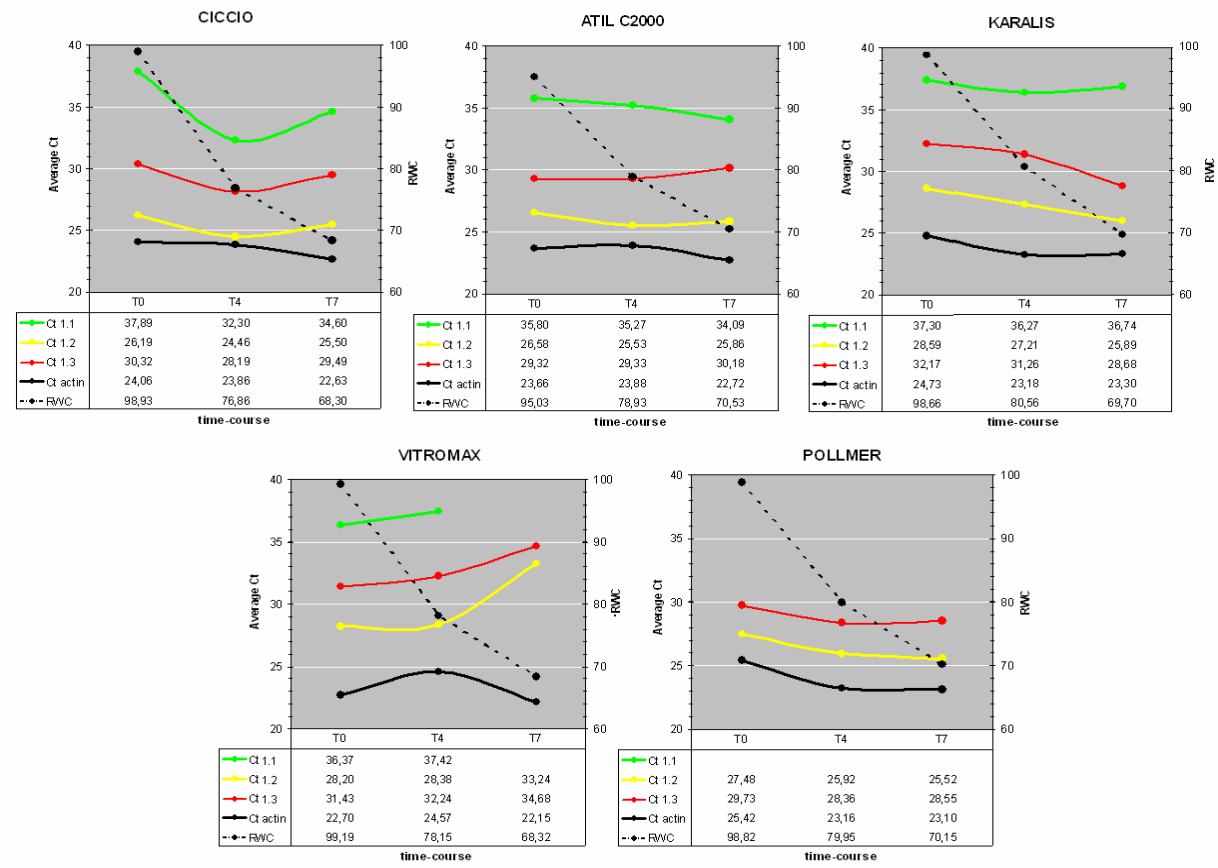



**Figure 16.** Expression profiles of the *TdDRF1* gene transcripts by real-time RT-PCR analysis in the genotypes under analysis in comparison to RWCs. Each double Y-axis graph displays the $C_t$ trends for targets and *actin* and RWCs vs. water stress time-course in one genotype. $SD_{Ct} \leq 0.250$ and $SD_{RWC} \leq 0.22$.

Standardized changes in expression levels of the three transcripts can be best represented as Mean Fold Changes (MFCs) as shown in Figure 17. All three transcripts exhibit changes in expression levels that differ from one genotype to the other. General expression patterns, considering the three

transcripts together, are also genotype-dependent. Cultivars Ciccio, Atil C2000 and Vitromax can be grouped together based on their overall expression pattern similarity. These cultivars were characterized by a surge in overall expression at T4 followed by a decrease at T7, generally below the T0 level. This decrease in gene expression at T7 was most pronounced for Vitromax, reaching undetectable levels. Karalis, on the other hand, exhibited a drastically different expression pattern and was characterized by a very slight decrease in overall gene expression at T4, followed by a very sharp surge at T7, not involving *TdDRF1.1*. Finally, the triticale Pollmer TCL 2003, which demonstrated a third type of pattern characterized by the absence of detectable expression levels of *TdDRF1.1* as mentioned previously, exhibited a slight decrease in gene expression at T4, which was not followed by any recovery at T7.

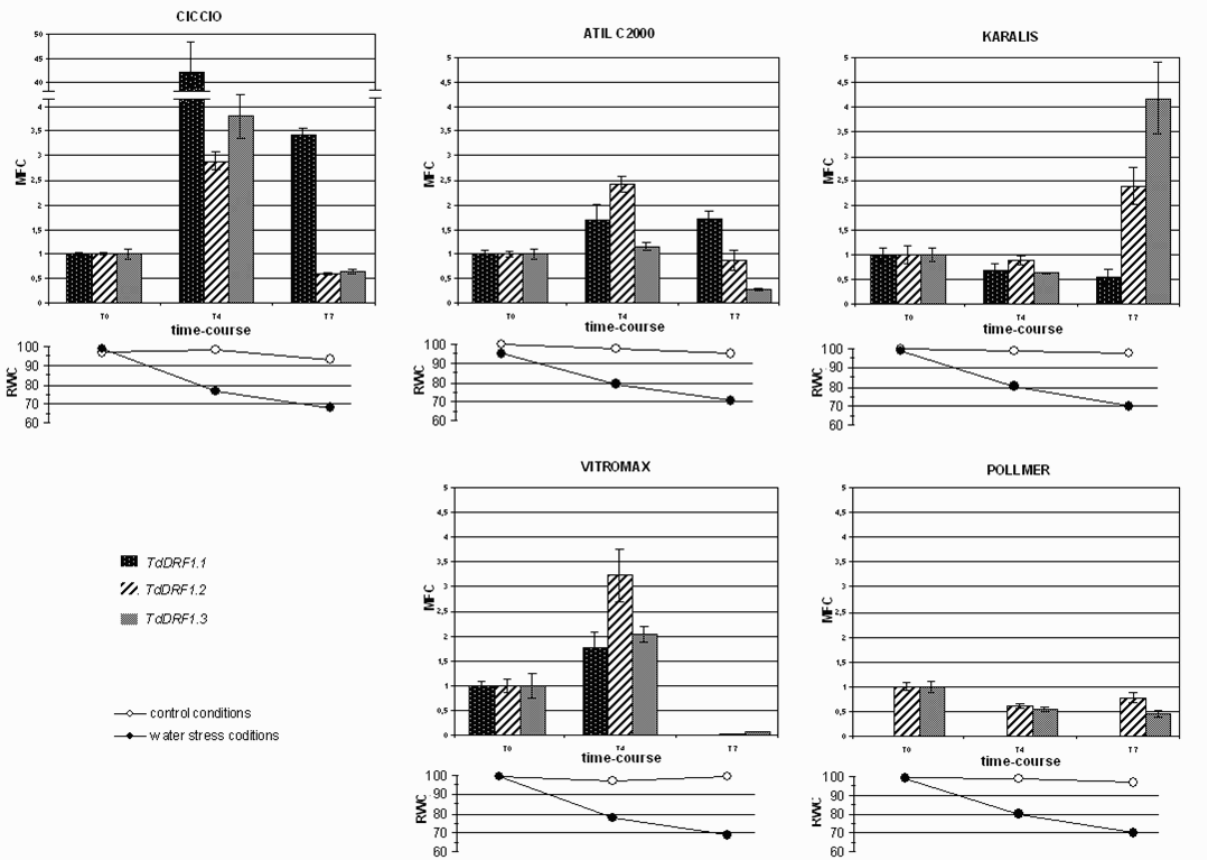



**Figure 17.** Expression profiles of the *TdDRF1* gene transcripts by real-time RT-PCR analysis in the genotypes under analysis. Each graph displays the Mean Fold Change (MFC) in expression of the three transcript variants, normalized to *actin* and relative to the time zero (T0 represents the 1x expression for the three targets), vs. water stress time-course in one genotype. Starting RNA template was 50 ng/sample. Each sample was treated separately: results were averaged after $2^{-\Delta\Delta Ct}$ calculation and the variance (SD) are shown as error bars in the graphs. In the lower part the trends of the RWCs are shown in both control and water stress conditions.

In addition to these general trends in patterns of overall expression, some differences with regards to individual transcripts are noteworthy and make the picture more complex. *TdDRF1.1* is the transcript that deviates the most from the general trends just mentioned for overall expression. In the cultivar Ciccio, levels of *TdDRF1.1* at T7 decrease but remain greatly above those observed for T0. In Atil C2000 its levels remain constant from T4 to T7. In Karalis, *TdDRF1.1* does not follow the other transcripts in their late surge at T7. *TdDRF1.2*, on the other hand, shows an individual expression very much in accordance with the overall expression, that is, a sharp peak at T4 followed by a sharp decrease at T7 (below T0 levels) for the group of Ciccio, Atil C2000 and Vitromax, hardly any change between T0 and T4 followed by a sharp surge for Karalis, a slight decrease between T0 and T4 and low the level maintained from T4 to T7 for Pollmer TCL 2003. *TdDRF1.3* also follows the general expression pattern trend, except that its increase between T0 and T4 in Atil C2000 is not significant.

**Cloning and sequence analyses**

Real-time RT-PCR results have put in evidence that *TdDRF1* gene expression profile is genotype-dependent. For this reason we analysed the genetic variability of *TdDRF1* sequences among the five varieties, to investigate the presence of genotype-specific polymorphisms.

The complete CDSs of *TdDRF1* gene were amplified and six clones for each genotype were selected, sequenced and deposited to NCBI (accession numbers are reported in the legend of Figure 18). Among the clones, few of them were deriving from *TdDRF1.1* isoform (only some genotypes represented), some from *TdDRF1.3* (all genotypes represented) and many from *TdDRF1.2* (all genotypes represented). Because the *1.1* clones were not available from all genotypes and, as mentioned previously, the *TdDRF1.2* transcript does not encode any AP2 transcription factor, we have focused our sequence analyses on the variations in the *TdDRF1.3* transcript.

All nucleotide sequences related to each genotype were translated into their corresponding amino acid primary sequences and aligned (Fig. 18). The AP2 DNA-binding domain is located between Gly90 and Ala153. Among the deduced amino acid (aa) sequences corresponding to TdDRF1.3, we found two polymorphic sequences for Karalis (type *a* and *b*) and Vitromax (type *a* and *b*) and a unique sequence for the remaining cultivars. The putative polypeptide chain is 335 aa long in most cases, with the exception of Atil C2000 (341 aa) and Ciccio (346 aa). These differences are due to two insertions common to Atil C2000 and Ciccio, the first involving a single amino acid at position 16 and the second involving five amino acids at position 210-216, and to a third insertion unique to Ciccio, involving five amino acids at position 286-290. None of these are within the AP2 domain. The first two insertions indicate a common origin between these two cultivars, but their amino acid

sequences differ overall by more than 20 aa, mostly concentrated in the range 291-346, suggesting a rapid divergence in the sequence of their *TdDRF1* gene. This is also revealed graphically by the phylogenetic tree in Figure 19.

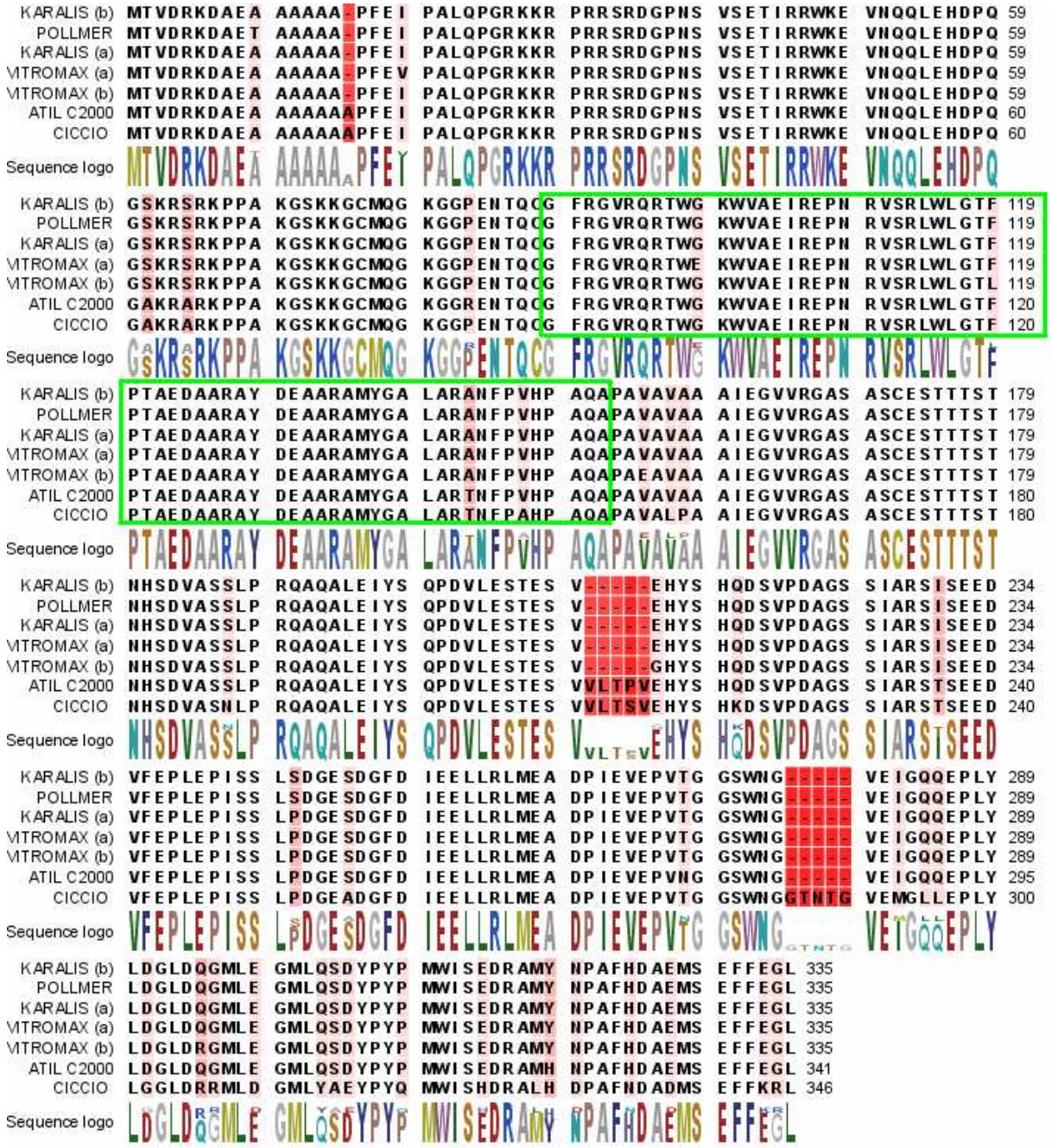



**Figure 18.** Alignment of the codifying amino acid sequences of TdDRF1.3. Red gradient background displays the level of conservation at each position of the alignment and completely conserved residues are in white background. The AP2 DNA-binding domain region is put in evidence in the green frame. The NCBI accession numbers of the sequences used here are: Ciccio (DQ013205), Atil C2000 (…), Karalis type *a* (…), Karalis type *b* (…), Vitromax type *a* (…), Vitromax type *b* (…) and Pollmer TCL 2003 (…). This graphic was exported from CLC Combined Workbench 3.0.3.

In general, among the other five aligned sequences, only 8 aa substitutions are present and spread over the whole primary sequence, two of them in the region codifying the AP2 domain. The sequences of Karalis (type *a* and *b*) and that of Pollmer TCL 2003 are very similar, with a single substitution at position 10 differentiating the triticale from the two Karalis isoforms and another single substitution at position 246 differentiating the two Karalis isoforms. This similarity is also graphically demonstrated in Figure 19. In the case of Vitromax, the two isoforms (type *a* and *b*) differ from each other as a result of 6 aa substitutions, including two in the AP2 domain.

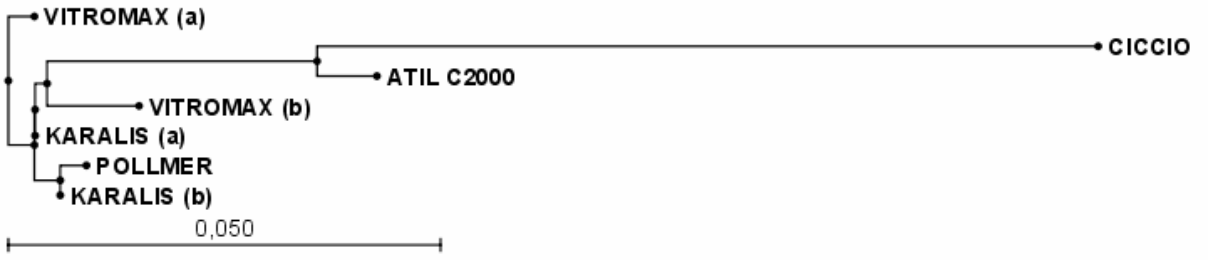



**Figure 19.** Phylogenetic tree of the codifying amino acid sequences of TdDRF1.3. The sequences used for the tree construction are those of the alignment in Figure 18. This graphic was exported from CLC Combined Workbench 3.0.3.

Being the domain responsible for the DNA recognition and binding, the AP2 DNA-binding domain is thought to be highly conserved amongst various plant-specific transcription factors. However, Figure 18 reveals the presence of different residues in some critical positions between the genotypes. These differences are better highlighted in the following structural alignment (Fig. 20) and scheme (Fig. 21).

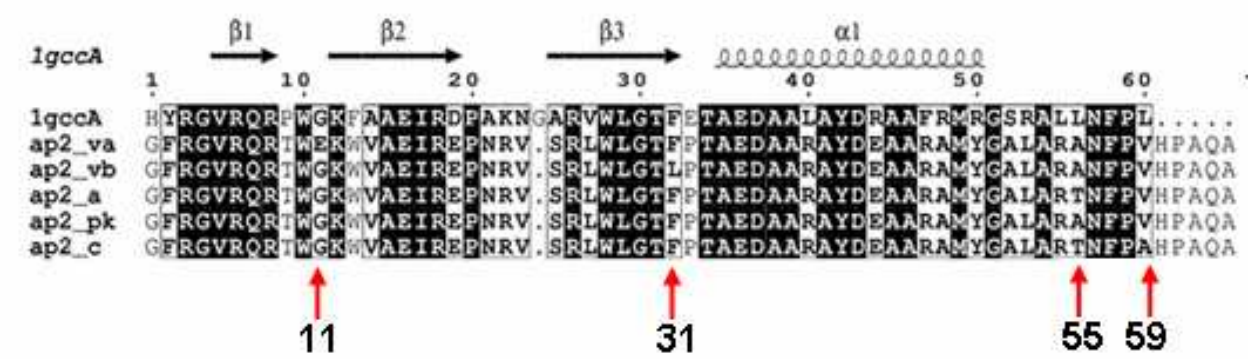



**Figure 20.** Structural sequence alignment of the AP2 domains of Ciccio (ap2_c), Pollmer TCL 2003 and Karalis (ap2_pk), Atil C2000 (ap2_a), Vitromax type *a* (ap2_va) and Vitromax type *b* (ap2_vb) with the structural template 1GCC(A). Conserved residues are boxed. Black background: strictly conserved residues; black boldface: well conserved residues. Dots represent deletions. The secondary structure and sequence number are assigned according to the template. The figure was generated by using ESPript2.2 (Gouet *et al*. 1999).

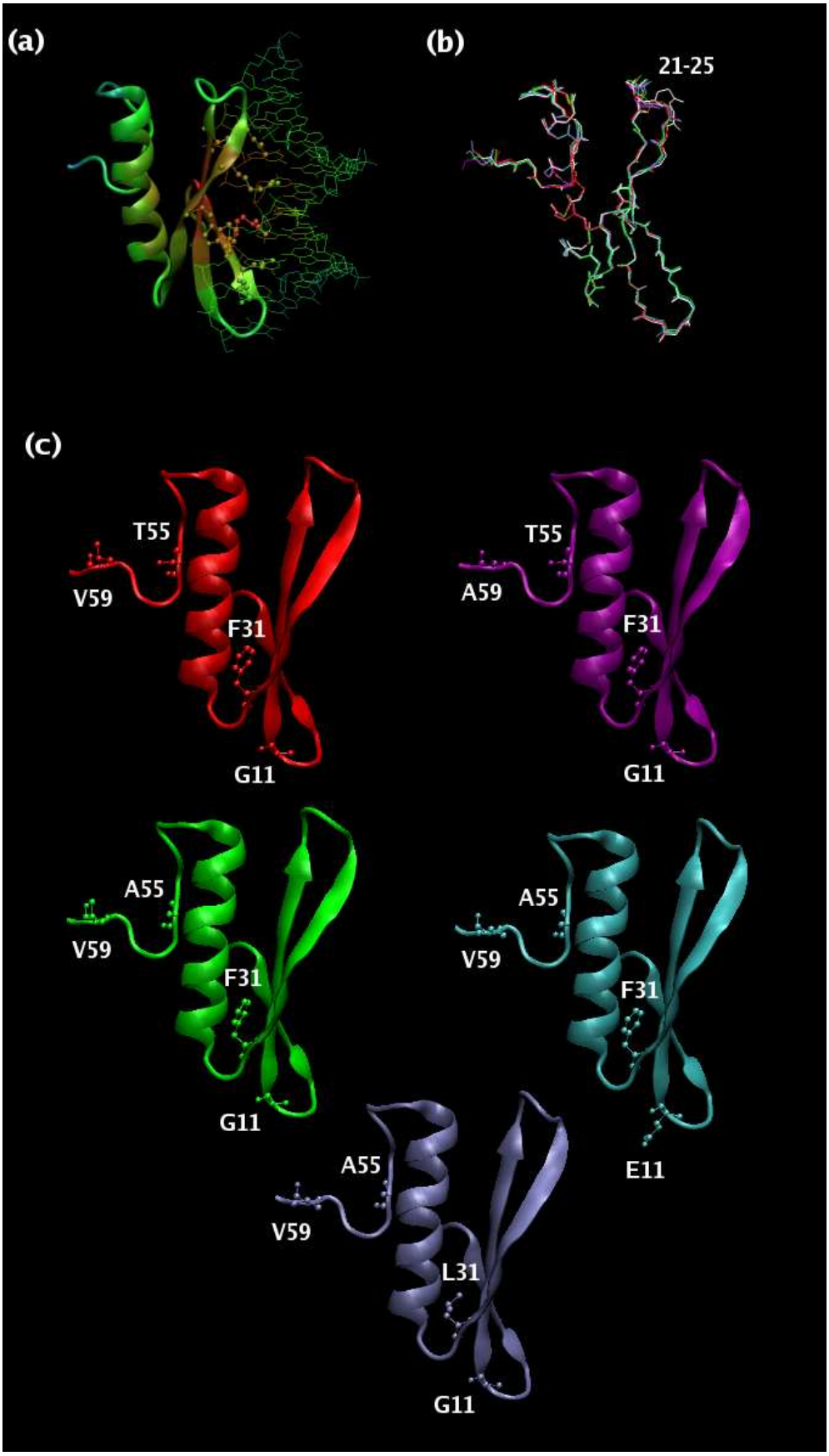
| | | | | |
|---|---|---|---|---|
| **CICCIO** | 11(G) | 31(F) | 55(T) | 59(V) |
| **POLLMER/KARALIS** | 11(G) | 31(F) | 55(A) | 59(V) |
| **ATIL C2000** | 11(G) | 31(F) | 55(T) | 59(V) |
| **VITROMAX type *a*** | 11(E) | 31(F) | 55(A) | 59(V) |
| **VITROMAX type *b*** | 11(G) | 31(L) | 55(A) | 59(A) |

**Figure 21.** Scheme reporting the genotype-specific amino acids in the variable positions (11, 31, 55 and 59)

The highly conserved Gly at position 11 of the AP2 domain (residue 100 of Figure 18) is replaced by Glu in Vitromax type *a*. On the other hand, Vitromax type *b* shows a Leu at position 31 of the same domain (residue 119) instead of the conserved Phe/Tyr as reported in literature (Liu *et al.* 1998; Magnani *et al.* 2004; Sakuma *et al.* 2006). Additional differences are observed at position 55 (residue 144) where Atil C2000 and Ciccio show a Thr, while the other sequences have an Ala, and at position 59 (residue 148) where the conserved Val is substituted by an Ala in Ciccio. These differences are visualized on the homology modelled structures of AP2 domains (Fig. 22) obtained according to the structural alignment reported in Figure 20.

With regards to the three-dimensional structure inferred for the AP2 domain, it consists of a three-stranded β-sheet (β1, β2, β3) and one α-helix (α1), which contacts the DNA through Arg and Trp residues located in the β-sheet (Fig. 22a). All the modelled structures show strictly similarity among themselves (root mean square deviation (rmsd) values < 0.3 nm) and with the template (rmsd value < 0.4 nm), with the exception of the loop connecting the β2 and the β3 (Fig. 22b). The residues at position 11 and 31 are structurally located close to the β2 and β3, respectively, whereas both the residues 55 and 59 are located far from the β-sheet, the crucial region for the binding to the target DNA sequences (Fig. 22c).

**Figure 22.** Homology modelling of the AP2 domain in the studied genotypes. (a) Sequence alignment of the AP2 domains of Ciccio (ap2_c), Pollmer TCL 2003 and Karalis (ap2_pk), Atil C2000 (ap2_a), Vitromax type *a* (ap2_va) and Vitromax type *b* (ap2_vb) with the structural template 1GCC(A). Conserved residues are boxed. Black background: strictly conserved residues; black boldface: well conserved residues. Dots represent deletions. The secondary structure and sequence number are assigned according to the template. The figure was generated by using ESPript2.2 (Gouet *et al.* 1999). (b) Ribbon representation of the 1GCC template. The Arg and Trp residues which contact the DNA (lines) are depicted in ball and sticks. (c) Superposition of the main-chain atoms of Atil C2000 (red), Ciccio (purple), Pollmer/Karalis (green), Vitromax type *a* (cyan) and Vitromax type *b* (iceblue) onto those of template (white). The loop connecting the β2 and the β3 (21-25 residues) is labelled. (d) Ribbon representation of the AP2 domain modelled structures. The polymorphic residues are indicated and labelled.


(a)
(b)
21-25
(c)
T55
V59
F31
G11
T55
A59
F31
G11
A55
V59
F31
G11
A55
V59
F31
E11
A55
V59
L31
G11

## DISCUSSION

The identification and characterization of genes coding for transcription factors involved in response to dehydration in agricultural crops, the study of genotypic differences in their expression as water stress occurs and the comparative analysis of their protein sequences can result in a better understanding of the molecular mechanisms underlying a crop's response to water stress.

In a manuscript published in the Annals of Applied Biology journal, the presence of a Dehydration Responsive Factor 1 gene in *Triticum durum* (*TdDRF1*), which belongs to the *DREB* gene family, was reported (Latini *et al*. 2007). This gene produces, through alternative splicing, three transcripts, namely, *TdDRF1.1* which consists of four exons (E1-E4), *TdDRF1.2* which consists of E1, E2 and E4 and *TdDRF1.3* which consists of E1 and E4 (Fig. 11). The *TdDRF1.2* transcript, albeit the most abundant of the three regardless of the water stress conditions, does not include a sequence for an AP2 DNA-binding domain and therefore may not play a role as transcriptional activator. Its function, if any, has yet to be elucidated. On the other hand, *TdDRF1.1* and *TdDRF1.3* transcripts include a sequence for an AP2 domain and so their encoded proteins can act as transcriptional activators regulating the expression of downstream genes directly protecting plant cells against dehydration, as reported for the highly similar *HvDRF1.1* and *HvDRF1.3* transcripts in barley (Xue & Loveridge 2004) and for the *wdreb2α* and *γ* transcripts in bread wheat (Egawa *et al*. 2006).

In this study, genotypes were selected based on their adaptation to environments or conditions where water deficit is common. The Italian cultivars Ciccio and Karalis and the Spanish variety Vitromax are grown in rainfed areas where rainfall is highly variable, often insufficient for optimal development, and irregularly distributed during the growing season. The Mexican cultivar Atil C2000, although released primarily for its excellent yield potential under irrigation, has shown outstanding resilience when water availability is severely limited in the favourable soils and weather conditions of northern Mexico. This is also the case for the triticale cultivar Pollmer TCL 2003, which was also included to explore how different this synthetic hybrid species may be from its durum ancestor with whom it shares two genomes (AA and BB). Whereas these genotypes' performances will certainly vary with the nature of water stress (severity, timing, other interacting soil or weather circumstances) under field conditions in different environments, all them were considered likely to exhibit some type of active defence response as water becomes limiting, especially under controlled conditions such as those of the present study. These controlled conditions were designed to provide uniform water stress levels for all genotypes in order to ensure that any genotypic differences observed in the *TdDRF1* gene expression were not confounded by erroneous variation in stress intensity. This was conclusively verified through the monitoring of the

leaf Relative Water Content (RWC), a physiological parameter considered to be a good indicator of plant's water status (Diaz-Pérez *et al.* 1995; González & González-Vilar 2001). It provides an integrated estimation of the variations in water potential (WP), turgor potential (TP) and osmotic adjustment (OA). Moreover, genetic association found between RWC and plant performance under drought stress (Tahara *et al.* 1990; Rodriguezmaribona *et al.* 1992) suggests that this parameter may also provide an estimation of genetic differences between genotypes with regards to their response to water stress. We did not find any statistically significant difference in the RWC of the five genotypes used in this study as water stress progressed which could indicate that they have similar cellular water retention capacity upon dehydration, at least under the controlled conditions of the present experiment. All genotypes show the same initial RWC level, a constant decrease as stressed occurs with similar values during the entire time-course experiment.

Prior to undertaking the truly quantitative real-time RT-PCR experiments, which uses transcript specific primers and probes, we have conducted a semi-quantitative RT-PCR analysis, with a single primer pair amplifying all 3 transcripts under the same amplification conditions, to ascertain that all genotypes included in the study qualitatively exhibited the expected *TdDRF1* transcripts. This was done at the T4 sampling time only as preliminary analyses indicated that the amplification signals were most intense at that stage. Under these experimental conditions, it was previously established (Latini *et al.* 2007) that when just one band appears, it corresponds to *TdDRF1.2* cDNA (193 bp); when two bands appear, they correspond to *TdDRF1.2* and *TdDRF1.3* cDNA (193 bp and 140 bp, respectively) and when three bands appear, very rarely, they correspond to the three *TdDRF1.2*, *TdDRF1.3* and *TdDRF1.1* cDNA (281 bp for *TdDRF1.1*). This analysis confirmed the presence of all three transcripts in all genotypes and, based on the semi-quantitative densitometric quantification, indicated that the most abundant transcript was *TdDRF1.2* followed by *TdDRF1.3*, with the rarest being *TdDRF1.1*, especially in the triticale Pollmer TCL 2003 which revealed barely detectable levels of this transcript. In addition, we have verified, using a semi-quantitative end-point RT-PCR analysis with transcript-specific primer pairs and conditions, that variation could be detected over the time course of the experiment for all transcripts. While describing the relative abundances of each transcript, it is recognized that these results can only be considered as semi-quantitative, being limited by the plateau phase of the amplification's dynamic, which corresponds to the saturation of the reaction (Breljak *et al.* 2005). However, it is also worth stating that these semi-quantitative analyses provided results that are generally consistent with those obtained by real-time RT-PCR analyses with regards to the relative abundances of the three transcripts and to the existence of significant variation during the dehydration time-course.

Real-time RT-PCR results suggest that the production of the three transcripts is related to the dehydration status of the plant and that the expression patterns are rather different among the genotypes analysed, even if these exhibited very similar rates of dehydration as indicated by the non significant differences in their RWC at all stages of the time-course. Considering overall gene expression, the durum genotypes could be classified into two distinct groups according to the timing at which a surge in *TdDRF1* expression occurred after the onset of dehydration. Ciccio, Atil C2000 and Vitromax exhibited a rapid activation with a substantial increase in the production of all three *TdDRF1* transcripts prior to reach the 80% RWC level. Karalis, on the other hand, was characterized by a significant slower activation, showing increase in *TdDRF1* transcripts production only after passing the 80% RWC level. Given the fact that all these genotypes were dehydrating at the same rate (similar RWC over the time-course), these results reveal a genotype-dependent dehydration threshold for *TdDRF1* activation. If one considers *TdDRF1.1* and *TdDRF1.3* individually, these being the only two transcripts coding for transcriptional activators, a somewhat different grouping could be inferred among the four durum wheat genotypes. Whereas Ciccio and Vitromax seem to respond to dehydration by activating both *TdDRF1.1* and *TdDRF1.3*, Atil C2000 does so primarily through production of *TdDRF1.1* and Karalis activates exclusively, and in a late fashion, *TdDRF1.3*. The preferential reliance on one or the other transcript, or lack thereof, indicate an additional level of genotype-related differences in the *TdDRF1* gene expression in response to dehydration. A third level of genotype-related expression differences could be observed in how sustained the expression was over time, that is, in the duration of transcript production maintenance as dehydration progressed. While the quantity of each transcript and its resulting transcription activator, beyond a certain level, may not be of consequence for the activation of genes involved in protection against dehydration, the extended duration of transcript production may well be.

A final level of genotypic differences was observed in the amino acid sequence of TdDRF1.3. There were notable differences among the cultivars, in the form of three insertions and more than 20 substitutions spread over the entire primary sequence, including the region corresponding to the AP2 DNA- binding domain. In particular, the AP2 sequence was essentially conserved among all genotypes, including the triticale cultivar, and in relation to published sequences from other species (Liu *et al*. 1998; Magnani et al. 2004; Sakuma *et al*. 2006), with the exception of four positions. Among these, the highly conserved Gly at position 11 is replaced by Glu in one isoform of Vitromax and a Leu at position 31 replaces the conserved Phe/Tyr in the second isoform of the same cultivar.

The identification of isoforms for the sequence of the *TdDRF1.3* gene product in two of the five genotypes, namely Vitromax and Karalis, adds another level of the genotypic differences related to

*TdDRF1*. Since the tissue used to clone the *TdDRF1.3* transcripts and deduce the corresponding protein sequence came from a bulk of leaves from several plants, we cannot conclusively eliminate the possibility that these within-genotype polymorphisms are due to the lack of genetic uniformity of Vitromax and Karalis at the *TdDRF1* locus. Genetic purity at any given locus is not necessarily excepted to occur if there was no selection pressure applied to do so, especially for loci controlling traits that are not readily measurable, or have not purposely been monitored. Within cultivar variation is not uncommon in durum wheat and has been reported for some glutenin subunits of the grain storage proteins for example (Upelniek *et al*. 1994). The other possible explanation for the presence of different isoforms of the *TdDRF1.3* gene product within a cultivar may be the existence of several copies of the *TdDRF1* gene, some of which have mutated in Vitromax and Karalis, thereby allowing their detection. The final, but less likely, possibility is that the isoforms are produced from homeologuous loci, one from the A and the other from the B genomes. This seems unlikely because of the great similarity in amino acid sequences between the two isoforms of Vitromax (six substitutions, with two in the AP2 DNA-binding domain) and Karalis (only one substitution). One would except more polymorphisms from homeologous sequences. To investigate whether the amino acid substitutions observed in the AP2 DNA-binding domains of the five genotypes used in this study could affect its structural architecture and therefore alter its DNA-binding capacity, we have used a preliminary *in silico* approach, based on homology modelling, to generate likely 3D structural models of all the AP2 DNA-binding domain variants. Results from this model structural analysis did not suggest that any of the observed polymorphisms would affect the structural arrangement of the DNA-binding domains and therefore should not result in a major functional alteration of the *TdDRF1.3* transcription factor. However, functional experiments and finer structural analyses extended even to the entire transcription factor could assess the actual impact, if any, of the observed polymorphisms on the protein functionality.

The high yielding, relatively drought tolerant triticale Pollmer TCL 2003 was included in this study to explore differences in *TdDRF1* expression upon dehydration between this species, in which the A and B genomes it has in common with durum wheat co-exist with the rye genome R, and its durum wheat ancestor. In terms of the sequence of its *TdDRF1.3* gene product, the triticale was identical to one isoform of the same gene product from Karalis and differed with the other isoform of the same durum cultivar by only one substitution, outside of the AP2 DNA-binding domain. It did not at all cluster in a separate group based on the *TdDRF1.3* gene product sequence. Nevertheless, it distinguished itself drastically from the durum wheat cultivars in the expression pattern of all the transcripts produced by its *TdDRF1* gene. In addition to the undetectable levels of *TdDRF1.1*, there was hardly any response in the form of change in expression of the other two transcripts at any stage

of dehydration during the time course of the present experiment. This may suggest that plants from this cultivar do not rely in a significant manner on the transcription factor produced by *TdDRF1* to activate downstream genes responsible for protecting them from dehydration. In that case, it is possible that the role of the homeologous DRF1-type system from the rye genome R, not detected with the molecular assay used in this study, has taken a more dominant and determinant role in the plant's response to dehydration. To my knowledge, the DRF1 ortholog system has not been yet characterized in rye, but its putative expression could have resulted in the more or less complete silencing of its homeologous system of durum wheat. Homeologous gene silencing in polyploids has been reported in several crops such as cotton (Adams *et al.* 2003) and even in triticale with regards to disease resistance genes (Ren *et al*. 1997).

Field trials described in the second section are still in progress using several drought tolerant and susceptible genotypes to evaluate and compare the *TdDRF1* gene expression profile under open field conditions. Furthermore, functional analyses by means of gene silencing or overexpressing systems followed by transcriptomics analyses, will be set up to better investigate the effect of this gene on the other inducible-genes related to drought response. These experiments are thought to give more information on the drought induced molecular mechanisms and to be useful in a future molecular assisted breeding program.

# EXPERIMENTAL SECTION 2

# EVALUATION OF THE EXPRESSION LEVELS OF *TdDRF1* TRANSCRIPTS IN DROUGHT TOLERANT AND SUSCEPTIBLE DURUM WHEAT CULTIVARS IN A FIELD TRIAL UNDER FULL AND REDUCED IRRIGATION CONDITIONS


## SUMMARY

In the second section, I report preliminary results of an experimental work for assessing the expression profile of the *TdDRF1* gene and its relation to the tolerance of durum wheat cultivars in open field conditions. Two groups of durum cultivars were chosen for this aim: one group exhibiting good drought tolerance and the other one exhibiting strong drought susceptibility. Plants were grown at CIMMYT's experimental fields (Obregón, Mexico) under water stressed (reduced irrigation) and non-stressed (full irrigation) conditions and the *TdDRF1* transcripts expression levels were monitored by real-time RT-PCR.

## MATERIALS AND METHODS

### Plant materials

The present study included 6 genotypes of durum wheat, three tolerant and three susceptible ones. The genotypes are listed in Table 5.

| | GENOTYPE | PEDIGREE | DATE OF RELEASE | DEVELOPMENT COUNTRY |
|---|---|---|---|---|
| TOLERANT | **Duilio** | Cappelli//Anhinga/Flamingo | 1984 | Italy |
| | **Altar 84** | Ruff/Free Gallipoli/2/ Mexicali75/3/Shwa | 1985 | Mexico |
| | **Gediz//Fgo//Gta/3/Srn_1/4/ Totus/5/Ente/Mexi_2//Hui/3/ Yav_1/Gediz/6/Sombra_20/ 7/Stot//Altar 84/Ald** | Gediz//Fgo//Gta/3/Srn_1/4/ Totus/5/Ente/Mexi_2//Hui/3/ Yav_1/Gediz/6/Sombra_20/ 7/Stot//Altar 84/Ald | not yet released | Mexico |
| SUSCEPTIBLE | **Creso** | CpB144× [(Yt54N10B)$Cp^{2}$63$Tc^{3}$] | 1974 | Italy |
| | **Colosseo** | Mexa/Creso Mutant | 1994 | Italy |
| | **Arment//Srn_3/Nigris_4/3/ Canelo_9.1** | Arment//Srn_3/Nigris_4/3/ Canelo_9.1 | not yet released | Mexico |

**Table 5.** List of the three analysed tolerant and of the three susceptible *Triticum durum* genotypes, their origin, year of release and development site.

Three genotypes for each group (tolerant and susceptible) were included in the current study to provide adequate diversity in plant lines and to ensure that more reliable information could be generated. The three tolerant cultivars show a very similar early development and the three susceptible cultivars show a very similar late development, on the basis of CIMMYT breeder's work and experience in Mexico. Indeed, the genotypes belonging to the same group do not manifest variability in germination and heading times when grown under the same environmental conditions. Duilio, Creso and Colosseo are Italian commercial varieties; Altar 84 is a high yielding variety; Gediz/(…) is an advanced improved line and Arment/(…) is an experimental line, all the last ones are from CIMMYT.

## Field experimental design

The field trials were set up at CIMMYT's experimental station, located near Ciudad de Obregón (27.33' latitude, 109.09'W longitude and 38 m above sea level) in northwestern Mexico. Two distant areas were used for the full-irrigation (FI) and the reduced-irrigation (RI) trials.

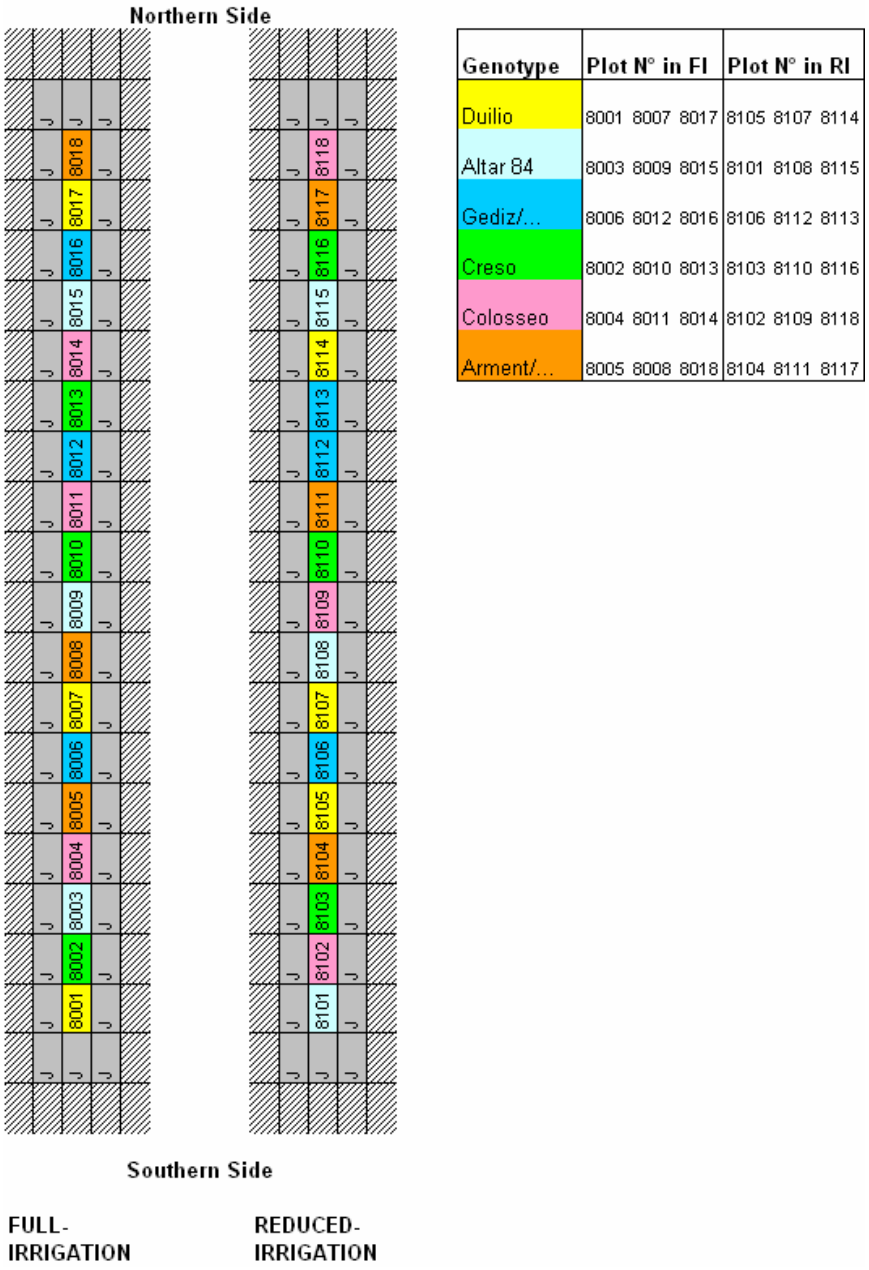



**Figure 23.** Plot map of the full- (FI) and the reduced- (RI) irrigation trials and numbers of the blocks corresponding to each genotype. "J" stands for Jupare durum wheat cultivar.

The experimental unit in field experiments is the "field plot" or "block". In this experiment, the 6 genotypes were arranged in three randomized complete blocks (RCB) within the two irrigation treatments (Fig. 23). Each experimental block consisted of two beds 80 cm wide and 3 m long. Plot sizes were uniform and equal numbers of plants were grown in each row. The FI and the RI trials were both surrounded by a border of Jupare durum wheat variety to reduce edge effects.

**Crop management**

Field experiment was conducted during the 2006-2007 winter-spring growing season, being the seeding date at the end of November 2006.

During the planting time, irrigation was applied to maintain an above-ground wet level of 3-5 cm in all treatments. The control treatment received in average 8 hours irrigations at 2-weeks intervals until harvest, whilst the stressed treatment received irrigations at 1-month intervals (in addition to that for germination).

Grain was harvested after maturity on June 2007.

**Physiological analyses**

The following physiological and optical properties were assessed on the plants.

*Relative water content (RWC):* Leaf RWC was measured in both full- and reduced-irrigated plants, at T5, T6, T7 and T8 (see Table 6). It was calculated as described in the first section, according to the equation from Barrs & Weatherley (1968):

$$RWC\ (\%) = [(FW-DW)/(TW-DW)] \times 100$$

Where FW is the sample fresh weight, TW is the sample turgid weight and DW is the sample dry weight. Each sample was made up of 10 uniform flag leaves.

*Canopy temperature:* The measurement of canopy temperature is one of the several criteria used to assist the selection by breeders in the field. Canopies emit a long-wave infrared radiation dependent on their temperature, which is sensed by an infrared thermometer (IRT). The hand-held IRT (Crop Trak Mini IR Thermomenter by Spectrum Technologies, Inc., USA) was used.

As the correlation between canopy temperature and plant water status becomes stronger as plant water status is reduced, the collection of canopy temperature data began under well-developed drought stress. Measurements were done just after solar noon when plant water deficit is maximized, during relatively stable atmospheric conditions. Measurements were performed several times and, on every occasion, 6 temperature readings were recorded (3 values from each of the two beads constituting one block) and then averaged, resulting in a single value to represent every single block.

*Measurement of chlorophyll content:* The Minolta SPAD-502 chlorophyll hand-held meter (Spectrum Technologies Inc., USA) was used to acquire a rapid estimation of flag leaf chlorophyll content. The measurements were taken on the uppermost collared leaf halfway from the leaf base to the tip and halfway from the midrib to the leaf margin. The chlorophyll measurements were performed several times at T6 and T7 (Tab. 6) and, on every occasion, 6 temperature readings were recorded (3 values from each of the two beads constituting one block) and then averaged, resulting in a single value to represent every single block.

**Leaves harvesting for the RNA bank ensemble**

Ten representative flag leaves for each block were harvested, pooled together, immediately frozen in liquid nitrogen and subsequently stored at -80 °C prior to RNA extraction.

The sampling schedule is reported in Table 6.

| TIME-COURSE | COLLECTION DATE | DAYS AFTER T1 |
|---|---|---|
| T1 (1st harvesting date) | 01-29-2007 | - |
| T2 (2nd harvesting date) | 02-05-2007 | 7 |
| T3 (3rd harvesting date) | 02-12-2007 | 14 |
| T4 (4th harvesting date) | 02-19-2007 | 21 |
| T5 (5th harvesting date) | 02-26-2007 | 28 |
| T6 (6th harvesting date) | 03-05-2007 | 35 |
| T7 (7th harvesting date) | 03-12-2007 | 42 |

**Table 6.** Time-course leaves collection schedule.

Total RNA was extracted from leaves using the TRIzol® Reagent (Invitrogen, USA), according to manufacturer's instructions, and lyophilized to allow transportation from Mexico to Italy. Three independent extractions, each yielding 25-35 µg of total RNA, were prepared for each sample: two of them were resuspended in nuclease-free water and stored at -80 °C and one of them was stored at room temperature in a dry area.

**Reverse transcription**

RNA extractions were qualitatively assessed by agarose gel electrophoresis and quantified by NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, USA).

The High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, USA) was used for the reverse transcription (RT) reactions. The final RT reactions (20 µl) contained 2 µl of 10X RT Buffer, 0.8 µl of 25X dNTP Mix (100mM), 2 µl of 10X RT Random Primers, 1 µl of RNase Inhibitor (20 U/µl) and 2 µg of total RNA in nuclease-free water. The thermal cycling conditions were as suggested by the kit protocol: 10 min at 25 °C, 2 h at 37 °C, 5 min at 85 °C and at 4 °C.

**Real-time PCR**

To increase the quantity of the three gene specific targets, cDNAs were firstly preamplified for 14 cycles using TaqMan® PreAmp Master Mix (Applied Biosystems). The resulting preamplified reactions were then diluted (1:20) in 1X TE buffer and used as the starting material for the subsequent singleplex real-time PCRs. Below a schematization of the procedure is presented.

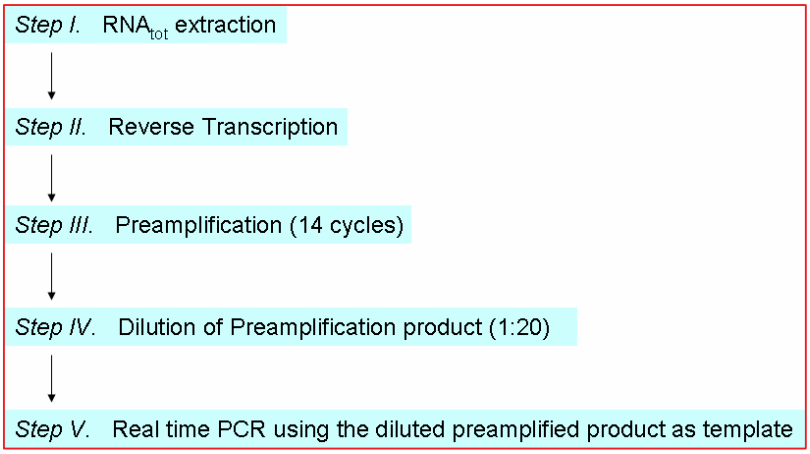



**Figure 24.** Experimental real-time PCR workflow.

Equal volumes of the three 20X Custom TaqMan Gene Expression Assays (Applied Biosystems) for the three *TdDRF1* gene transcripts were combined and diluted to a final concentration 0.2X. Being highly expressed, the assay for the 18S was omitted in the assay pool.

Preamplification reactions (50 µl) contained 2 µl of TaqMan® PreAmp Master Mix 2X (Applied Biosystems), 12.5 µl of pooled assay mix (0.2X, each assay) and 250 ng of cDNA sample in nuclease-free water. Reactions were hold at 95 °C for 10 min and then at 95 °C for 15 sec and 60 °C for 4 min for 14 times.

Three separate preamplification reactions for each single transcript were performed to check the preamplification uniformity of the three targets (Applied Biosystems TaqMan® PreAmp Master Mix Kit Protocol).

Real time PCR reactions were carried out in the Applied Biosystems 7300 Real Time PCR System. The primers and the probes used in the amplification reactions have been already described in the Materials and Methods of the Experimental Section I (see Figure 12 and Table 4).

The final volume (20 µl) of a single PCR reaction contained 10 µl of 2X TaqMan® Universal PCR Master Mix with AmpErase® UNG (Applied Biosystems), 1 µl of 20X Custom TaqMan® Gene Expression Assay (Applied Biosystems) and 2 µl of pre-amplified product as template, for the three transcripts target and for the endogenous control.

The Eukaryotic 18S rRNA Endogenous Control (FAM MGB Probe, Non-Primer Limited) (Applied Biosystems) was used as endogenous control for the relative quantification. The 18S assay, designed on the X03205.1 GenBank accession, generates an amplicon of 187 bps.

To determine the PCR efficiencies, different template amounts corresponding to 500, 250, 100, 50 and 10 ng of total initial RNA were tested. Concerning the 18 S rDNA, 5 ng were examined, too.

All targets (unknowns and standards), endogenous controls, and no template controls (NTC) reactions were performed in triplicate.

The standard thermal cycling conditions were set and runs were incubated 2 min at 50 °C for the AmpErase UNG activation, 10 min at 95 °C for the Taq activation and, then, 15 sec at 95 °C and 1 min at 60 °C, for 40 cycles.

**Quantitative data analysis**

Once the amplification efficiencies for the three *TdDRF1* targets and the 18S rRNA endogenous control were calculated (see the post-run data analysis described in the Experimental Section I) and assessed to be very similar, the relative quantification of expression levels of the three targets, normalized to the endogenous control (18S) and relative to sampling T1 (here chosen as RQ calibrator), was lead with the comparative Ct method (Livak & Schmittigen 2001), applying the formula $2^{-\Delta\Delta Ct}$.

A different threshold setting was established for each amplicon, this was feasible because this was a RQ assay and each target was treated separately. The optimum thresholds were 0.03 for *TdDRF1.1*, 0.3 for *TdDRF1.2*, 0.2 for *TdDRF1.3* and 0.4 for 18S rRNA. The Ct values at the target-specific threshold were exported from the Applied Biosystems 7300 instrument and processed in Microsoft Excel. MFC results, proceeding from the three randomized replicate blocks (biological experimental replicas) for each durum wheat cultivar, were lastly averaged. The final standard deviations (SDs), according to the theory of error propagation, were calculated as the sum of the three single SDs.

.

## RESULTS AND DISCUSSION

The key drought evaluation site of the International Maize and Wheat Improvement Center (CIMMYT) is located at the Centro de Investigación Agricolas de Noroeste (CIANO) in northwestern Mexico, near Ciudad de Obregón (Sonora, Mexico). This site is arid, and wheat is growing using irrigation. Drought is generated using a combination of gravity and drip irrigation methods to generate controlled moisture-stress scenarios. The "genotype x year" interaction under moisture stress is low and the relevance of germplasm selected at this site, under limited and optimal irrigation, to global wheat-growing environments has been demonstrated (Trethowan *et al*. 2001, 2003). When germplasm selected at this site is tested globally, significant rates of improvement in productivity have been observed (Trethowan & Crossa 2006). The soils at this location have been carefully characterized for biotic stresses (nematodes, root rots) and other abiotic stresses (micronutrient imbalances) thereby ensuring that the observed differentiation of genotypes is due to water and not to other confounding factors.

Preliminary field experiments were performed, during the growing seasons 2004-05 and 2005-06 not only at CIMMYT but also at ENEA-Casaccia RC (Posters presented at the international conferences "*Plant GEMs 2006*" and "*The Future of Drylands* 2006"; Manuscript in press in *Biosaline Agricolture and High Sainity Tolerance in Plants*) and for set up the better experimental conditions of the current research. In those experiments were included about twenty genotypes characterized by a wide spectrum of phenotypic response to drought and different yields. The relations among potential and actual grain yields, drought tolerance, stress severity and plant phenology is illustrated in Figure 25. The choice of several genotypes had the notable disadvantage that plant growth was not synchronized among the different cultivars (different flowering dates), leading to tricky situations at the moment of the interpretation of molecular results.

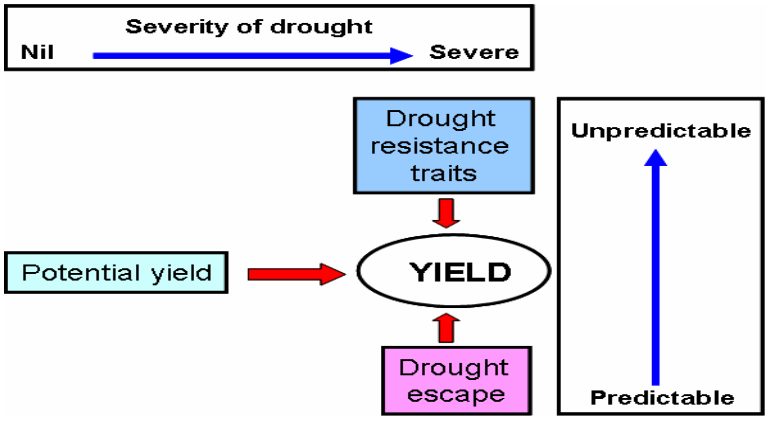



**Figure 25.** Importance of genetic yield potential and flowering time in determining grain yield under various drought conditions.
Proceedings from Ouk *et al*. 2006.

Furthermore, the extension of real-time analysis to several genotypes did not allow performing the molecular analysis of the biological replicates, for obvious economical issues, and it was necessary to pool the leaves material proceeding from the three randomized blocks, thus affecting the reliability of the results because of the many sources of error, at any step of the procedure, which must be taken into account in such kind of experimentation, and eliminating most of the variability that is intrinsic to field conditions.

For all the above reasons, six durum wheat cultivars were chosen for this current experiment: three tolerant and three susceptible genotypes, the former displaying a very similar early development and the latter a very similar late development. These cultivars were grown under two water availability settings, full- and reduced- irrigation (Fig. 26), in field trials with the least restrictive experimental design, which is the completely randomized design.


**Figure 26.** Experimental field at CIMMYT (Obregón, Mexico). (a) Full-irrigation trial at young developmental stage. (b) Full-irrigation trial at maturation stage. Creso plants under full- (c) and reduced-irrigation (d) conditions.

**Physiological analyses**

Obregón is a region with lower than average rainfall, every year; thus, drought is the major factor affecting yields. Physiological measures were performed for evaluating the drought tolerance of the cultivars.

*Relative water content (RWC)*: RWC measurements are reported in Table 7.

| | | | T5 | | | T6 | | | T7 | | | T8 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FI | Genotype | Plot Number | RWC | Mean RWC | SD | RWC | Mean RWC | SD | RWC | Mean RWC | SD | RWC | Mean RWC | SD |
| | Duilio | 1 | 89.7 | 92.4 | 2.43 | 88.7 | 91.3 | 2.25 | 86.0 | 89.2 | 2.82 | 86.2 | 88.7 | 2.41 |
| | | 7 | 94.4 | | | 92.2 | | | 90.1 | | | 91.0 | | |
| | | 17 | 93.1 | | | 92.9 | | | 91.4 | | | 88.9 | | |
| | Altar 84 | 3 | 96.2 | 97.8 | 6.69 | 94.4 | 93.9 | 1.19 | 88.9 | 91.0 | 3.21 | 87.0 | 84.0 | 4.04 |
| | | 9 | 92.0 | | | 92.5 | | | 94.7 | | | 79.4 | | |
| | | 15 | 105.1 | | | 94.7 | | | 89.4 | | | 85.6 | | |
| | Gediz/... | 6 | 98.4 | 95.0 | 5.66 | 86.5 | 88.3 | 1.70 | 88.5 | 86.0 | 1.56 | 83.4 | 82.8 | 0.57 |
| | | 12 | 90.4 | | | 88.9 | | | 86.3 | | | 82.6 | | |
| | | 16 | 96.2 | | | 89.4 | | | 83.1 | | | 82.3 | | |
| | Creso | 2 | 93.0 | 94.0 | 2.84 | 94.0 | 92.2 | 3.59 | 93.3 | 85.2 | 7.70 | 90.0 | 85.1 | 5.54 |
| | | 10 | 97.2 | | | 94.6 | | | 78.0 | | | 79.1 | | |
| | | 13 | 91.8 | | | 88.1 | | | 84.2 | | | 86.3 | | |
| | Colosseo | 4 | 91.6 | 97.7 | 5.83 | 87.0 | 93.0 | 6.16 | 82.6 | 79.6 | 2.76 | 84.5 | 84.1 | 2.08 |
| | | 11 | 103.2 | | | 92.6 | | | 78.9 | | | 81.8 | | |
| | | 14 | 98.4 | | | 99.3 | | | 77.2 | | | 85.9 | | |
| | Arment/... | 5 | 93.5 | 91.4 | 2.65 | 90.6 | 87.5 | 2.67 | 78.1 | 81.1 | 4.28 | 92.0 | 88.2 | 3.55 |
| | | 8 | 92.2 | | | 85.7 | | | 79.2 | | | 87.5 | | |
| | | 18 | 88.4 | | | 86.3 | | | 86.0 | | | 85.0 | | |

| | | | T5 | | | T6 | | | T7 | | | T8 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RI | Genotype | Plot Number | RWC | Mean RWC | SD | RWC | Mean RWC | SD | RWC | Mean RWC | SD | RWC | Mean RWC | SD |
| | Duilio | 5 | 71.4 | 71.3 | 1.90 | 65.3 | 67.3 | 2.05 | 64.2 | 63.9 | 1.03 | 63.4 | 61.7 | 1.88 |
| | | 7 | 69.3 | | | 67.3 | | | 64.8 | | | 62.1 | | |
| | | 14 | 73.1 | | | 69.4 | | | 62.8 | | | 59.7 | | |
| | Altar 84 | 1 | 67.3 | 68.2 | 1.23 | 71.0 | 68.2 | 2.43 | 60.5 | 62.0 | 1.62 | 59.3 | 59.8 | 0.55 |
| | | 8 | 67.7 | | | 66.7 | | | 61.7 | | | 59.8 | | |
| | | 15 | 69.6 | | | 66.9 | | | 63.7 | | | 60.4 | | |
| | Gediz/... | 6 | 73.5 | 72.8 | 2.26 | 68.3 | 69.7 | 0.99 | 65.3 | 66.5 | 0.85 | 62.7 | 63.3 | 0.64 |
| | | 12 | 70.3 | | | 69.7 | | | 66.5 | | | 61.8 | | |
| | | 13 | 74.6 | | | 71.2 | | | 67.7 | | | 65.4 | | |
| | Creso | 3 | 67.5 | 66.3 | 1.72 | 65.4 | 64.7 | 1.51 | 60.2 | 61.3 | 1.15 | 58.4 | 55.8 | 2.60 |
| | | 10 | 67.0 | | | 65.8 | | | 62.5 | | | 53.2 | | |
| | | 16 | 64.3 | | | 63.0 | | | 61.2 | | | 55.8 | | |
| | Colosseo | 2 | 69.7 | 68.1 | 3.27 | 69.0 | 66.0 | 2.91 | 63.8 | 61.7 | 2.10 | 57.2 | 56.4 | 1.24 |
| | | 9 | 70.2 | | | 65.7 | | | 61.8 | | | 57.1 | | |
| | | 18 | 64.3 | | | 63.2 | | | 59.6 | | | 55.0 | | |
| | Arment/... | 4 | 68.2 | 68.4 | 0.72 | 64.4 | 65.1 | 1.54 | 61.4 | 60.2 | 1.53 | 56.3 | 55.3 | 1.05 |
| | | 11 | 67.8 | | | 64.1 | | | 60.8 | | | 55.4 | | |
| | | 17 | 69.2 | | | 66.9 | | | 58.5 | | | 54.2 | | |

**Table 7.** RWC (%) measurements of leaves in full- (FI) and reduced-irrigation (RI) field trials. Plot numbers refer to Figure 23. RWC was recorded from T5 (5$^{th}$ collection date) to T8 (8$^{th}$ collection date). Leaves at T8 were used for the RWC estimation but not for the RNA bank ensemble and real-time RT-PCR analysis

As expected, under full-irrigation, all cultivars exhibit high RWC values along the four sampling times (from T5 to T8) and, under reduced-irrigation, leaves appeared to undergo a significant decrease in RWC. Furthermore, there is a significant difference between the two groups of cultivars, with the tolerant ones retaining more water in the leaf tissue with respect to the susceptible ones.

*Canopy temperature*: Canopy temperature may give an indication of the interaction between canopy and environment, even though the relationships between canopy temperature, air temperature and transpiration are not simple, involving atmospheric conditions (vapor pressure deficit, air temperature and wind velocity), soil (mainly available soil moisture) and plant (canopy size, canopy architecture and leaf adjustments to water deficit). In particular, as canopy architecture differs, varieties may differ in their canopy temperature (Blum *et al.* 1989). Relatively lower canopy temperature in drought stressed crop plants indicates a relatively better capacity for taking up soil moisture and for maintaining a relatively better plant water status.

Canopy temperature was recorded several times during this work: among the analyzed cultivars, it was very similar in the reduced-irrigation and slightly different in the full-irrigation. As reported in Figure 27, relatively to T5, every time Duilio and Altar 84 had lower temperatures than most of the three susceptible cultivars.

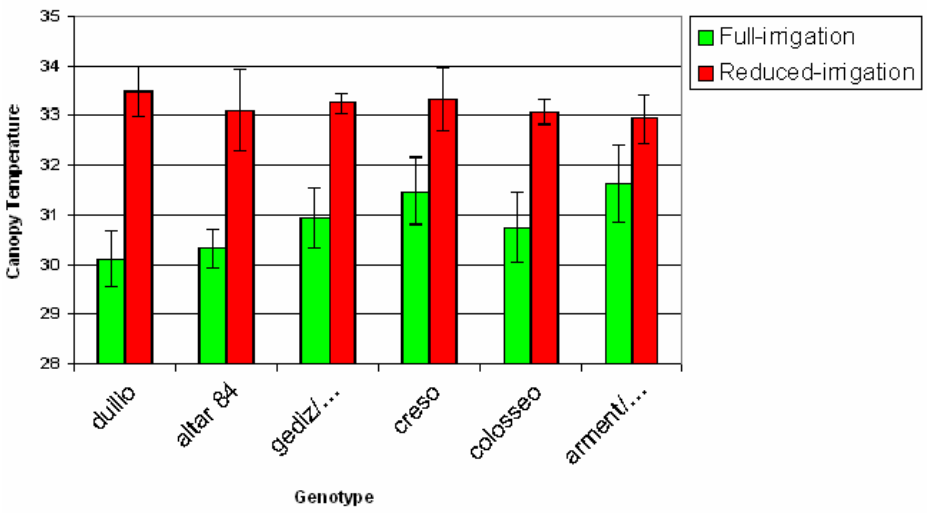



**Figure 27.** Canopy temperature at T5 in full- and reduced-irrigated plants.

*Measurement of chlorophyll content:* The Minolta SPAD measures light conductance of the leaf, thereby indicating chlorophyll concentration per unit area of the leaf. A decrease in chlorophyll content means reduction in photosynthesis. SPAD measurements were carried out at T6 and T7. In

the susceptible cultivars the chlorophyll content was lower under stressed conditions than under well watered ones (see Figure 28, relatively to T7).

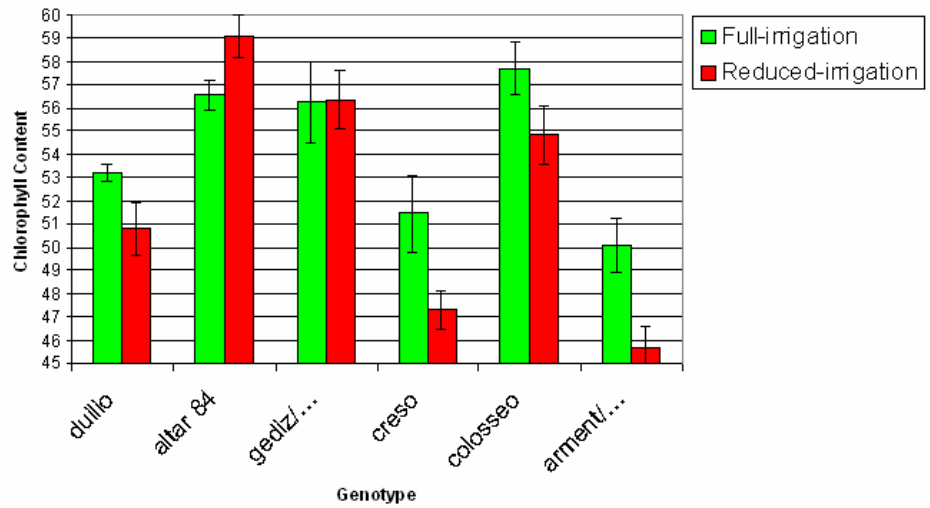



**Figure 28.** Chlorophyll content at T7 in full- and reduced-irrigated plants.

In conclusion, physiological measurements prove that Duilio, Altar 84 and Gediz/(…) are wheat lines more drought resistant than Creso, Gianni and Arment/(…); moreover, the tolerant lines gave higher yield than the susceptible ones (data not shown). All these data remark previous information collected by the researchers at CIMMYT.

**Molecular analyses**

The work reported here may be considered as a pioneer experimental work, being the first time that the molecular expression profile of a gene that codes for transcription factors related to drought stress was analysed using plant material coming from experimental field, under control and water stress conditions.

An RNA bank was made up of 252 samples, deriving from 6 genotypes (3 tolerant and 3 susceptible), two watering regimes (FI and RI), 7 time points (from T1 to T7) and 3 plots (biological replicates). The RNA bank was used for the current experiment and will be used also in the future to study the expression profiles of other drought-related genes, so that it will be possible to compare the behaviours of several genes.

In the last few years, real-time RT-PCR has become one of the most frequently used methods for quantification of specific mRNAs in a vast variety of samples. Notwithstanding the present extensive use, it represents a relatively new method for plant studies and few examples are reported

in literature. Plant material is a very complex matrix and both the experimental design and the normalization techniques are more critical than in other organisms in view of biological accurate conclusions.

As already mentioned, the *TdDRF1* transcript targets of quantification, with the exception of *TdDRF1.2*, are transcripts codifying for transcription factors known to be very low abundant in the total RNA. In this experiment, the cDNA preamplification was performed as intermediated multiplex step between reverse transcription and real time PCR, allowing the enrichment for the three cDNA targets using the pool of the three TaqMan Gene Expression Assays. In theory the pre-amplification for 14 cycles generates approximately 16.000-fold amplification of each gene-specific target. In the amplification plots showed in Figure 29, the result improvements are well noticeable: by using the preamplification (on the right), Ct values stay in the optimal detection range of the instrument; smoothing and amplitude normalization of the curves are not necessary to reduce the random cycle-to-cycle noise; matrix effects are eliminated and curves are more highly reproducible.

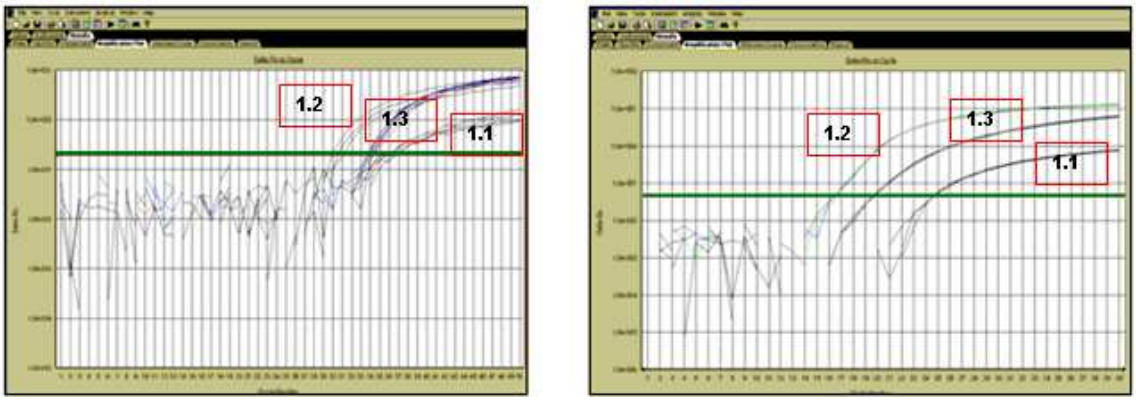



**Figure 29.** Amplification plots obtained by direct PCR (on the left) and by means of an intermediate preamplification step (on the left).

Here I report the molecular results obtained for the two widely spread Italian varieties: the tolerant Duilio and the susceptible Creso. The RQ data proceeding from the other four durum wheat cultivars is still under analysis. In Figure 30 the relative quantification results are displayed as MFC in expression of each transcript relative to the T1. Each bar of a histogram derives from 9 measures (real-time PCR reactions); in fact, each sample was performed in triplicate (technical replicates) and for each sample there were three plots (biological replicates). The obtained results were very consistent and reproducible, as can be evinced by the error bars representing the SDs.

Concerning the relative amounts of the three transcripts, the *TdDRF1.1* resulted the less expressed, the *TdDRF1.2* was the most abundant at each time and genotype and the *TdDRF1.3* was

intermediate (see amplification plot on the right in Figure 29). This behaviour is in agreement with the previous results obtained from greenhouse experiments.

For these two varieties it is quite evident that all the three transcripts present a similar trend, but, unexpected are the expression profiles in relation to the water stress. The tolerant Duilio variety shows increasing expression levels of all the transcripts under full-irrigation conditions, particularly at T5, and decreasing expression levels under reduced-irrigation, particularly at T4 for *TdDRF1.1* and at T2-5 and 7 for both *TdDRF1.2* and *TdDRF1.3*. On the other side, the susceptible Creso variety does not seem to present any significant change in expression during the time-course and comparing the full- and the reduced-irrigation treatments. Hence, contrary to the general view, drought tolerance could be related to an increasing of the gene expression when water is available and to a decreasing when water is lacking.

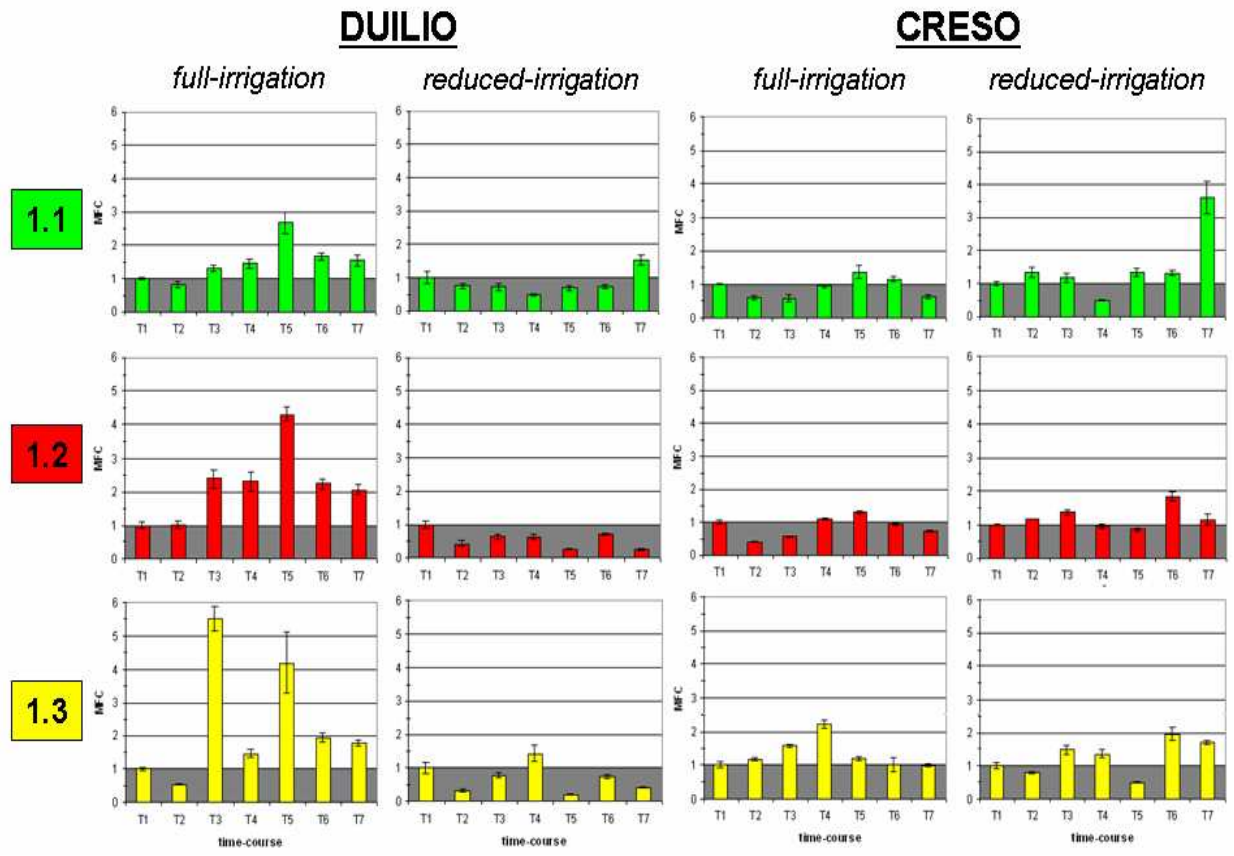



**Figure 30.** Expression profiles of the *TdDRF1* gene transcripts by real-time RT-PCR analysis in the tolerant Duilio and in the susceptible Creso durum wheat varieties, under full- and reduced-irrigation field conditions. Each graph displays the Mean Fold Change (MFC) in expression of one transcript variant, normalized to 18S rRNA and relative to the 1st collection date (T1 represents the 1x expression for the three targets), during the time-course (Tab. 6). The bars over the grey area represent an increased expression level while those under the grey area represent a decreased expression level.

When the data related to the other four cultivars will be analysed, it will be possible, hopefully, to put in correlation the *TdDRF1* gene expression profile with the tolerance/susceptibility to drought. However, as it often happen in the research work, these first results open other questions: the mechanism of action of TdDRF1.1 and TdDRF1.3 transcription factors in durum wheat has not yet been established and it is required to understand if they play as activators or as repressors.

# EXPERIMENTAL SECTION 3

# CLONING OF THE THREE *TdDRF1* TRANSCRIPTS IN TRANSIENT EXPRESSION VECTORS


## SUMMARY

Information on DREB2 is mainly limited to *AtDREB2A*, whose function has been dissected by the group of Shinozaki and Yamaguchi-Shinozaki, and other few species. *TdDRF1* is a *DREB2A*-homologous gene and it was isolated by using a primer couple derived from the Arabidopsis sequence. At the moment, for the durum wheat gene, no direct information regarding the *TdDRF1* encoded proteins and their actual function is available. Hence, while studying the *TdDRF1* transcripts expression profile, dependent on the particular plant water status and on the genotype, many efforts have been concentrated to attempt clarify the function of the three gene products. TdDRF1.1 and TdDRF1.3 should be putative transcriptional activators and, as it is well known, regulatory proteins are low abundant into the cell. For this reason, to carry out a proteomic analysis and expect to visualize these proteins in a 2DE gel, can be useful to make use of an overexpressing system. On the other hand, *TdDRF1.2* mRNA should encode just a short ORF of 68 amino acids or its translation could be inhibited.

Transgenic plants, since they allow the expression of drought-related genes *in vivo*, are an excellent system to better understand the mechanisms governing water stress tolerance and to examine the function of a gene/protein(s). Being the stable transformation of wheat very complex and time-consuming, a transient expression system was considered more suitable for getting preliminary indications regarding the *TdDRF1* gene function and, at this regard, a Potato Virus X (PVX)-derived expression vector was chosen. In this section, the isolation of the full-length cDNAs and the codifying sequences for the three *TdDRF1* gene transcripts, the engineering of the plasmid constructs for the transient overexpression in *Nicotiana benthamiana* and *Nicotiana tabacum* are described. An analysis of the mRNA levels expressed by the infected plants is also reported.

**ADDENDUM: Potato Virus X (PVX) and pP2C2S-mediated transient expression**

Potato virus X (PVX), belonging to the potexvirus group (Koening & Laseman 1989), is a single-stranded ssRNA virus, which is capped at the 5' end (Sonenberg *et al*. 1978) and polyadenylated at the 3' end (Morozov *et al*. 1981). Each viral particle is composed of a single RNA molecule and about 1300 subunits of coat protein (CP).

The positive sense genome of the PVX contains five large open reading frames (ORFs) under the control of the T7 RNA polymerase promoter (Fig. 31). The most 5' of these ORFs encodes an RNA replicase (RdRp, RNA-dependent RNA polymerase) of 166 kDa, involved in the viral replication. It is followed by three partially overlapping ORFs, known as triple block. They encode proteins of 25, 12 and 8 kDa, respectively, and are also called movement proteins because they are involved in the viral infection and in its systemic propagation. Downstream these ORFs, there is the coat protein, necessary for efficient RNA accumulation and viral systemic movement (Chapman *et al*. 1992).

A foreign protein can be expressed using the vector pP2C2S, derived from PVX (Baulcombe *et al*. 1995), through the insertion of the gene of interest between the duplicated regions of the CP promoter (Fig. 31 and 32). The green fluorescent protein (Baulcombe *et al*. 1995), a flower-specific Myb factor (Sablowski *et al*. 1995), a fungal elicitor (Hammond-Kosack *et al*. 1995), a tomato disease resistance gene (Rommens *et al*. 1995), an anti-microbial defensin (Saitoh *et al*. 2001), the *Arabidopsis thaliana* ascorbate peroxidase 3, APX3 (Escobar *et al*. 2003) and functional single chain antibodies (scFvs) in plants (Ziegler *et al*. 2000; Galeffi *et al*. 2005) are some of the proteins expressed by pP2C2S.

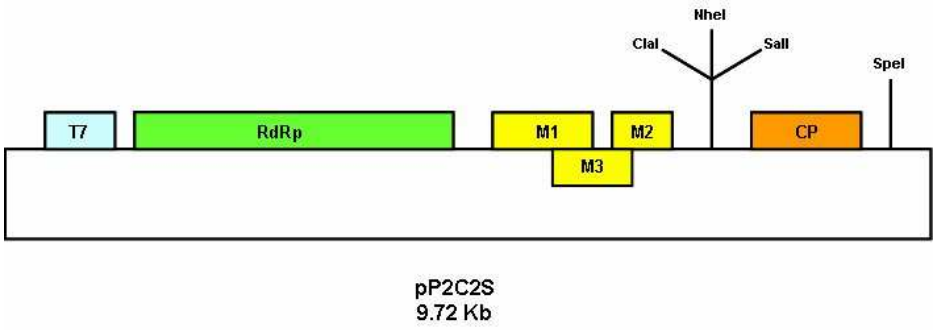



**Figure 31.** pP2C2S, the PVX-derived vector for transient expression. RdRp is the RNA-dependent RNA polymerase; M1 (25 kDa), M2 (8 kDa) and M3 (12 kDa) are the movement proteins of the triple block; ClaI and SalI are the more useful restriction enzymes to clone the target codifying sequence (CDS); CP is the coat protein. In the pP2C2S, the strong promoter of the CP was duplicated and inserted upstream the cloning site.

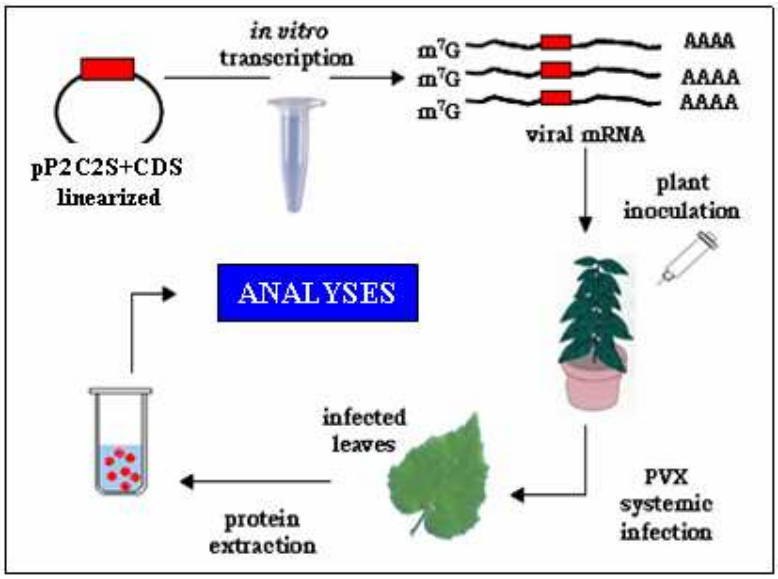



**Figure 32.** Methodological workflow of the transient expression system by means of the pP2C2S vector. First step is the engineering of the pP2C2S construct containing the CDS that has to be expressed. The recombinant pP2C2S has to be linearized (for example by using SpeI restriction enzyme which presents a single recognition site in the recombinant plasmid) and then *in vitro* transcribed. The resulting mRNAs are directly inoculated into young leaves of *Nicotiana benthamiana* plants. Plant begins to express mRNAs and to produce the codifying proteins 4-5 days post-inoculation (dpi). When viral symptoms (chlorotic spots and tissue curling) are clearly visible on leaves (4-7 dpi), the proteins and/or the complete set of mRNAs are extracted from the leaves and further analyzed.

### 5' RACE

Ciccio was the durum wheat variety used for the amplification of both the 5' and 3' mRNA ends. A schematic representation for the 5' RACE is shown in Figure 33. Flag leaves proceeding from a greenhouse experiment at the 4th day of water stress were the source of total RNA, which was extracted by the TRI® Reagent (Sigma-Aldrich, USA) according to the manufacturer's instructions. The 5'/3' RACE Kit, 2nd Generation (Roche Applied Science, Germany) was used for the next applications. Total RNA (5 µg) was the template of the reverse transcription with the *drebrev2* primer (5'-GCTCCCGAATTTCAGCAA-3') and the reaction was performed in a final volume of 20 µl, incubating at 55 °C for 1 h and then at 85 °C for 5 min. The new synthesized cDNA was then purified with MinElute PCR Purification Kit (Qiagen, USA) and the poly(A)-tailing of first-strand cDNA was done using dATP and terminale transferase in a 25 µl reaction volume. Afterwards, 5 µl of the dA-tailed cDNA were amplified and the first round of PCR was led with oligo(dT)-anchor primer (5'-GACCACGCGTATCGATGTCGACTTTTTTTTTTTT TTTTV-3') as forward one and *rev1.2* primer (5'-CGTTGCCTTACACCACGGAAT.3') as reverse one. The thermal cycling was as following: 2 min at 94 °C; (30 s at 94 °C, 40 s at 62 °C, 1 min at 72 °C) x 10 times; (30 s at 94 °C, 40 s at 62 °C, 1 min + 20 s EXT at 72 °C) x 25 times; 7 min at 72 °C. EXT corresponds to an elongation of each successive cycle by additional 20 s. The PCR was run in a 1.2% (w/v) agarose gel and the resulted bands were purified and directly sequenced.

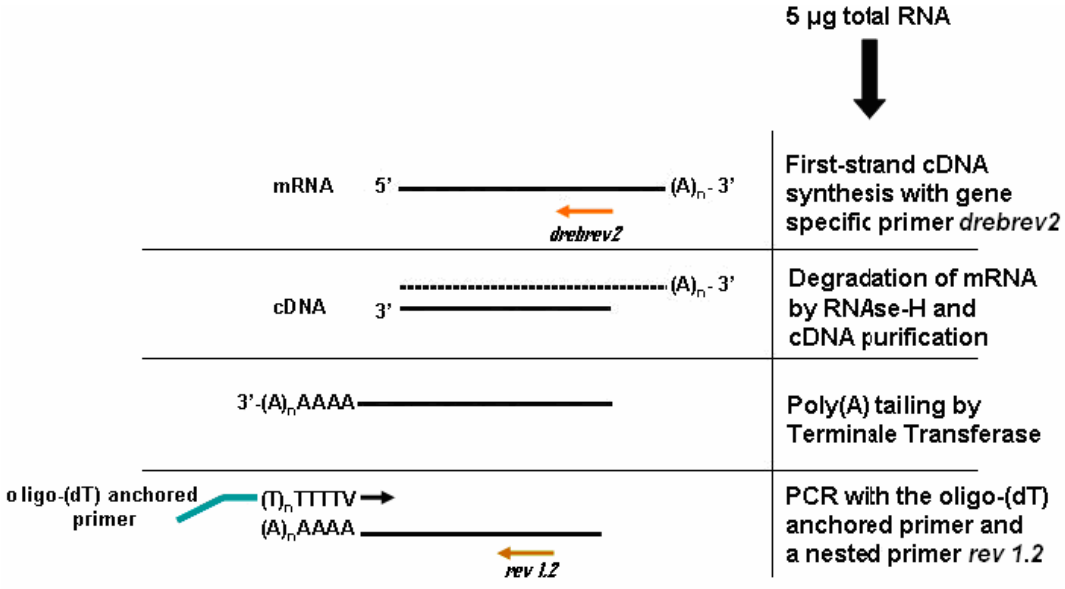



**Figure 33.** Experimental workflow of the 5' RACE.

### 3' RACE

A schematic representation for the 3' RACE is shown in Figure 34. First-strand cDNA synthesis was initiated at the poly(A)$^+$ tail of 5 µg total RNA using oligo(dT)-anchor primer (5'-GACCACGCGTATCGATGTCGACTTTTTTTTTTTTTTTV-3') in a 20 µl reaction volume. The incubation for the reverse transcription was 1 hour at 55 °C and 5 min at 85 °C. Then 1.5 µl of the first strand cDNA (correspondent to 375 ng of reverse transcribed RNA) were directly amplified by PCR using several forward primers coupled with the PCR anchor primer (5'-GACCACGCGTATCGATGTCGAC-3') as reverse one. In order to amplify segments deriving from each transcript variant, the following forward primers were used: *for1.1new* (5'-AATGAATATGCACTCTTGGCGC-3') for specific selection of the *TdDRF1.1* cDNA; *for 1.2* (5'-TCAAACCAATAGGAAAAAGCGACC-3') for specific selection of the *TdDRF1.2* cDNA and *for1.3* (5'-CGCTCCAGCCTGGAAGGAAAAA-3') for specific selection of the *TdDRF1.3* cDNA. Moreover, a primer in the middle of exon 4, *E4for-monte* (5'- TGCCCGTGCTTATGACGA-3'), was used because, annealing with all the three isoforms and giving a minor size PCR product, it was expected to give a more visible band in the gel. The primers used, their respective localization in the mRNA, their specificity, the annealing temperature used in the PCR reaction and length of the expected bands, calculated on the base of *HvDRF1* codifying sequences (NCBI accession: AY223807), are shown in Table 8.

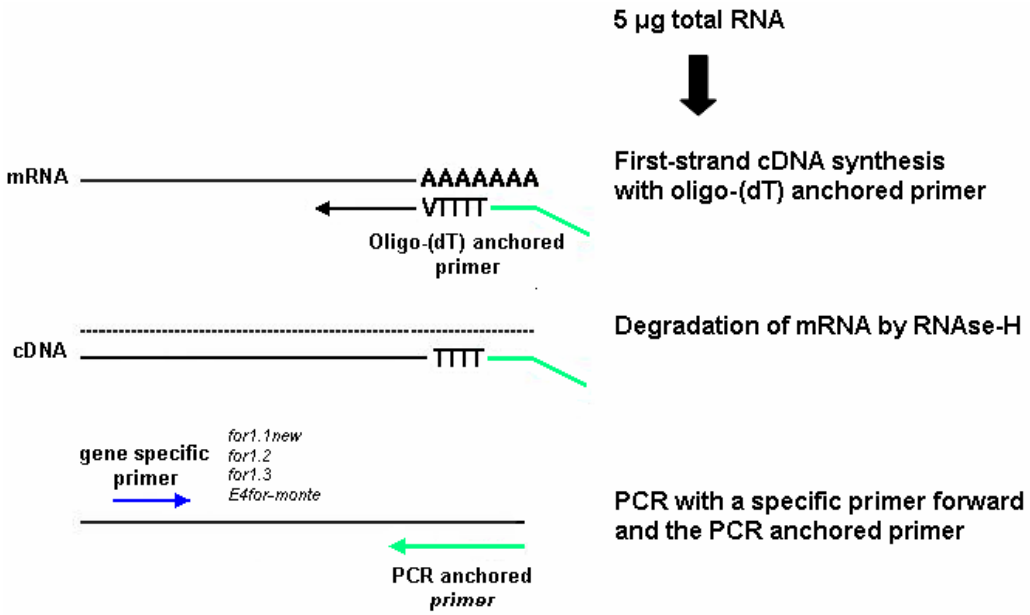



**Figure 34.** Experimental workflow of the 3' RACE.

| Forward Primer | Primer Localization | Specificity | $T_{annealing}$ | Expected Length |
|---|---|---|---|---|
| *E4for-monte* | E4 (~300 bp after the beginning of E4) | for the three transcript variants | 60.0 °C | 1002 bp |
| *for1.1new* | exon 3 | *TdDRF1.1* | 61.4 °C | 1367 bp |
| *for1.2* | E2-E4 junction | *TdDRF1.2* | 62.2 °C | 1311 bp |
| *for1.3* | E1-E4 junction | *TdDRF1.3* | 64.7 °C | 1315 bp |

**Table 8.** Features of the forward primers used in the 3'-RACE PCR, coupled with the PCR anchor primer as reverse primer.

Being the expected lengths of the amplification products >1 Kb, the High Fidelity PCR Enzyme Mix (Fermentas Inc., USA) was used because offering high fidelity for amplification up to 3-6 kb. The thermal cycling was as reported for the 5'-RACE. The PCR reactions were visualized in a 1.2% (w/v) agarose gel and the bands of interest were purified and directly sequenced.

Both 5' and 3' UTR sequences were mainly handled and managed with ClustalW (http://www.ebi.ac.uk/Tools/clustalw/index.html; Thompson *et al*. 1994) and Geneious (http://www.geneious.com). The major data resources for the functional analysis of the 5' and 3' untranslated regions are UTRdb and UTRsite (Mignone *et al*. 2005), that were exploited together with UTRblast and UTRscan utilities (http://www.ba.itb.cnr.it/UTR/).

**Cloning of the complete codifying sequences (CDS) of *TdDRF1.1, TdDRF1.2* and *TdDRF1.3***

The full-length sequences of the three transcripts of the Ciccio durum wheat cultivar were translated into the correspondent amino acid sequences by the Expasy tool at http://expasy.org/tools/dna.html web site. Based on the inferred protein sequences, the primers pair, made up of the *E1F-up* (5'-CACGACTCTCCCAACCTCTC-3') in the 5'UTR and *E4R-down* (5'-GGTCCACCATTTGATCTTCATT-3') in the 3'UTR, was designed in order to amplify the complete CDSs of the three transcripts for the further cloning.

To selective amplify *TdDRF1.1* two segments were separately amplified and then jointed. The 5' terminal segment was amplified by using *E1F-up* and *E3R* (5'-GCATCTCCCTTGGGGTTT-3') primers and the 3' terminal one by using *E3F* (5'-GCTGCCTGT GCCATCTATTCT-3') and *E4R-down* primers. Both *E3R* and *E3F* primers anneal in exon 3, with *E3F* beginning exactly one nucleotide after the end of the *E3R*, as schematized in Figure 35.

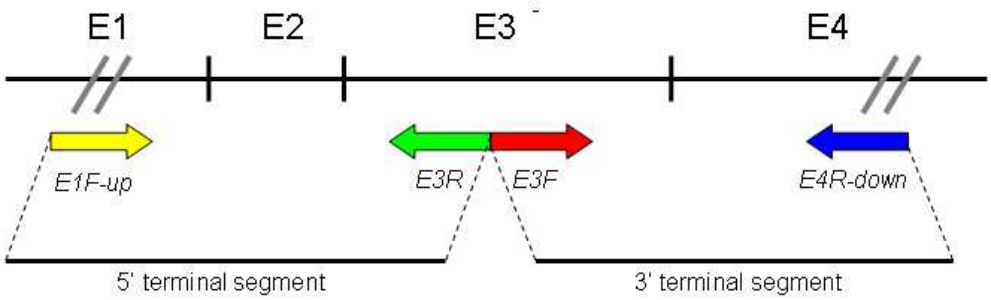



**Figure 35.** *TdDRF1.1* complete CDS was cloned in several steps as described in the text.

Taking advantage of the TA/UA cloning technology for the *Taq* polymerase-amplified PCR products, each of the two segments of *TdDRF1.1* transcript was cloned into pCR®II-TOPO® (Invitrogen, USA), according to the manufacturer's instructions, and then re-amplified with Phusion™ DNA Polymerase (FINNZYMES, Finland), which does not add 3'-A overhangs to the amplified fragment. The blunt end PCR products were checked in 1% agarose gel and equimolar amounts of the two cDNA segments were jointed together using Rapid DNA Ligation Kit (Fermentas Inc.). A PCR reaction with a common Taq polymerase was then performed using 2 µl of ligation reaction as DNA template and *E1F-up* and *E4R-down* primers. The band of the expected size (sum of the lengths of the 5' terminal segment and the 3' terminal one, equivalent to 299 + 1081 = 1380 bp), containing four possible alternatives of joining of the 5' and 3' terminal segments, was cloned into pCR®II-TOPO® (Invitrogen). After the screening, by digestion with EcoRI restriction enzyme, six recombinant clones were sequenced.

**Engineering of the viral constructs**

The CDSs coding TdDRF1.1, TdDRF1.2 and TdDRF1.3 were amplified from the pCR®II-TOPO® clones of the complete CDSs by PCR, using *For-ClaI* (5'-ATCGATATGACGGTAGATCGGAAGGA-3') as forward primer and *Rev-SalI* (5'-GTCGACGCCGACCAAACACCAT-3') as reverse one. In 5', *For-ClaI* primer contains the ATCGAT ClaI recognition sequence, immediately followed by the ATG starting codon of transcription, and *Rev-SalI* primer contains the GTCGAC SalI recognition sequence. Since the forward primer was engineered in exon 1 (in +1 position) and the reverse primer in exon 4 (after the TGA stop codon), these two primers were used to clone the three wheat CDSs into the viral vector, PVX-derived, pP2C2S (Baulcombe *et al.* 1995) and generate the *PVX:TdDRF1.1*, *PVX:TdDRF1.2* and *PVX:TdDRF1.3* constructs (Fig. 36).

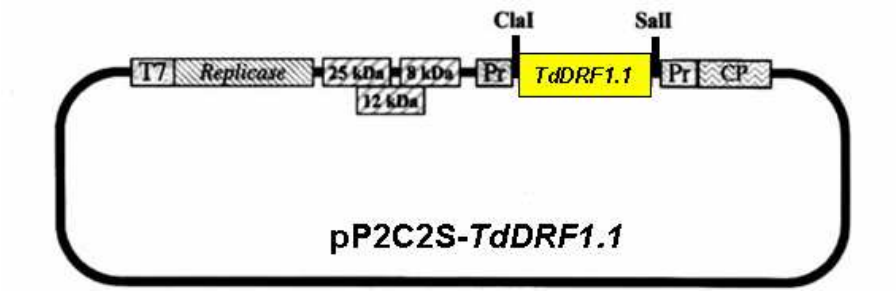



**Figure 36.** A schematic representation of the hybrid plasmid *pP2C2S:TdDRF1.1* for the infection of *N. benthamiana* and *N. tabacum* plants to overexpress the recombinant TdDRF1.1. During the vector cassette *in vitro* expression, the RNAs carrying *TdDRF1*open reading frames are synthesized. Pr is Coat protein duplicated promoter regions. Homolog hybrid plasmids were constructed also for the overexpression of TdDRF1.2 and TdDRF1.3.

The three *For-ClaI*/*Rev-SalI* PCR fragments were again cloned into pCR®II-TOPO®. Since ClaI is a methylation-sensitive restriction enzyme, both TOPO clones and pP2C2S vector were grown in SCS110 competent cells (Stratagene, USA), that are deficient for the Dam and Dcm methylases, and then cut with Fast-Digest ClaI e SalI enzymes (Fermentas Inc.). Ligation reactions were performed with Rapid DNA Ligation Kit (Fermentas Inc.), using a 1:3 vector:insert molar ratio, and were finally transformed in DH5α strain of *E. coli* under ampicillin selection. The screening of the clones was done by ClaI and SalI restriction analysis.

The positive recombinant clones were sequenced with the primers *pP2C2S-for* (5'-CGATCTCAAGCCACTCTCC-3') and *pP2C2S-rev* (5'-ACCCTATGGGCTGTGTTGT-3'), designed respectively upstream and downstream the insertion site. Moreover, to get the complete sequence of these constructs, other primers annealing in the insert were used.

**Transcription of viral RNA and infection of *Nicotiana benthamiana* and *tabacum* plants**

Before the *in vitro* transcription, the recombinant plasmids were linearized with SpeI (NEB, UK). Overnight digestions at 37 °C were performed using 25 μg DNA template and 20 U of enzyme in a 50 μl volume. SpeI enzyme was inactivated by incubation at 65 °C for 20 min. DNA was finally purified with Wizard® SV Gel and PCR Clean-Up System (Promega, USA).

Synthesis of the transcripts was performed as described by Chapman and co-workers (1992) and Baulcombe and co-workers (1995). Transcription reactions were set up in a final volume of 50 μl/each. They contained 5 μl of 10X NEbuffer for RNA Polymerase (NEB), 2.5 μl of 100 mM DTT (Promega), 5 μl of 10X A/C/G/U mix (composed of 20 mM of each ATP, CTP, UTP ad 2 mM of GTP), 40 U of RNase Inhibitor (Applied Biosystems, USA), 5 μl of 5 mM 7mG(5')ppp(5')G RNA Cap Structure Analog (NEB) and 5 μg of linear DNA. The cap structure protects the 5'end of the transcripts from degradation and enables their *in vivo* translation. After 5 min of incubation at 37

°C, 250 U of T7 RNA polymerase (NEB) were added to each reaction and, after 25 min of incubation at 37 °C, GTP was added till reaching the concentration of the other nucleotides. Reactions were again incubated at 37 °C for 35. Finally, the mRNA produced in a transcription reaction was purified by phenol:chloroform:isoamilic alcohol (25:24:1; Sigma-Aldrich) extraction and dissolved in 20 µl of DEPC water.

The transcribed mRNA was used to infect two opposite young leaves of a plant with six true leaves (i.e. 20 cm height). Plants were manually inoculated with infectious PVX-RNAs (1 µg mRNA per leaf). For each transcript, 3 individual plants of *N. benthamiana* and 3 of *N. tabacum* were infected. Both directly infected leaves and systematically ones were harvested 7 days post-infection (dpi) for downstream analyses.

**RNA extraction, reverse transcription, PCR and real-time PCR analyses**

Total RNA was extracted using TRI® Reagent (Sigma-Aldrich) from systemically infected leaves at 7 dpi, from leaves of mock inoculated plants at 7 dpi and from leaves of wild-type plants at the same developmental stage. RNA was quantified with BioPhotometer (Eppendorf, Germany) and its quality and integrity were electrophoretically verified by ethidium bromide staining and by $OD_{260}/OD_{280}$ and $OD_{260}/OD_{230}$ nm absorption ratios.

The reverse transcription with random hexamers of 2 µg total RNA was performed using the RevertAid H Minus M-MuLV Reverse Transcriptase (Fermentas Inc.), according to the manufacturer's guidelines.

PCR analyses were performed by using 100 ng of cDNA as template. To amplify *TdDRF1.1* the primer pair made of *for1.1new* (5'-AATGAATATGCACTCTTGGCGC-3') and *rev1.1* (5'-TCCGCAGAATAGATGCCACAG-3') was used; to amplify *TdDRF1.2* the primer pair made of *for1.2* (5'-TCAAACCAATAGGAAAAAGCGACC-3') and *rev1.2* (5'-CGTTGCCTTACACCACGGAAT-3') was used and, at last, to amplify *TdDRF1.3* the primer pair made of *for1.3* (5'-CGCTCCAGCCTGGAAGGAAAAA-3') and *rev1.2* (5'-CGTTGCCTTACACCACGGAAT-3') was used. The ReadyMix™TaqPCR Reaction Mix with $MgCl_2$ (Sigma-Aldrich) was used for 25 µl reactions. The thermal cycling conditions were as follows: denaturation for 2 m at 94 °C; amplification for 38 cycles, with denaturation for 1m at 94 °C, annealing for 2 m at 58 °C, extension for 2 m and 30 s a 72 °C, and final extension for 5 m at 72 °C. PCR products were run in 2% (w/v) agarose gel and visualized on an UV transilluminator after ethidium bromide staining.

Relative quantification by real-time PCR was carried out on *PVX:TdDRF1.2* overexpressing plants (one individual of *N. benthamiana* and one of *N. tabacum*) and on *PVX:TdDRF1.3* overexpressing

plants (one individual of *N. benthamiana* and one of *N. tabacum*), in comparison with wild type plants (again, one individual of *N. benthamiana* and one of *N. tabacum*). Real time PCR reactions were performed in the Applied Biosystems 7300 Real Time PCR System, by means of the TaqMan assays for *TdDRF1.2* and *TdDRF1.3* described in the previous experimental sections (for details see Figure 12 and Table 4, Section 1). Each 20 µl reaction contained 100 ng of cDNA, 2X ¡Taq™ SuperMix with ROX sample (Bio-Rad, USA) and 20X TaqMan® Gene Expression Assay (Applied Biosystems). Thermal cycling was made up of 10 m at 95 °C and then 15 s at 95 °C and 1 m at 60 °C, for 40 times. Data analysis was done applying the comparative Ct ($2^{-\Delta\Delta Ct}$) method (Livak & Schmittgen 2001). Unknown and no template control (NTC) reactions were run in duplicate.

# RESULTS AND DISCUSSION

## 5’ and 3’ UTR analysis

In the field of the analysis of mRNA structure and expression, with the aim of understanding the function of a gene, the RACE (Rapid Amplification of cDNA Ends) is generally a fundamental step. RACE is the PCR-based technique that allows obtaining the complete cDNA of a gene (full-length cDNA), including the transcriptional start site and the 5’ and 3’ untranslated regions (UTRs), for following cloning and expression.

The five prime untranslated region (5' UTR), also known as the leader sequence, is a particular section of mRNA and the DNA that codes for it. It starts at the +1 position (where transcription begins) and ends just before the start codon (AUG) of the coding region. It usually contains a ribosome binding site (RBS) and information for regulation of translation and mRNA stability. The 5' UTR may be a hundred or more nucleotides long (http://en.wikipedia.org/wiki/5’_UTR; Sonenberg 1994).

For the Ciccio cultivar through the 5’ RACE, without the necessity of a second round of amplification (nested PCR), a ~500 bp band was amplified and then sequenced (Fig. 37a). It corresponded to a 5’ terminal fragment of *TdDRF1.2* transcript and, actually, its total length of 513 bp was found to be the sum of the following segments: 39 bp of the oligo dT-anchor primer, 132 bp of the 5’UTR in exon 1, 80 bp of the codifying sequence (CDS) in exon 1, 53 bp of exon 2 and 210 bp of exon 4, from the beginning to the reverse primer (*rev1.2*).

The 5’ UTR of *TdDRF1.2* mRNA was compared with some *DREB2A*-related genes. A significant similarity was found with the reported 5’ UTR mRNA sequences of *wdreb2* (NCBI AB193608) and *HvDRF1* (NCBI AY22380) (Fig. 37b), while no significant alignments were obtained using 5’UTR of *ZmDREB2A* (NCBI AB218832), *OsDREB2A* (NCBI AF300971) and *AtDREB2A* (NCBI AB007790).

The UTRdb was used to approach a study of the structural and functional features of the isolated 5’ UTR sequence (Pesole *et al*. 1999; Pesole *et al*. 2002). UTRdb is a specialized non-redundant database of 5’ and 3’ UTR sequences from eukaryotic mRNAs in which the entries are annotated for the presence of repetitive elements and functional sequence patterns. According to the results of the UTRscan of the 5’UTR, no functional patterns among those included in UTRsite was found.


**Figure 37.** Results of the RACE 5'.
(a) PCR with oligo(dT)-anchor primer and *rev1.2* primer gives a band of ~500 bps (lane 1); Negative Control (NTC, no template control) (lane 2). GeneRuler 1kb DNA Ladder (Fermentas Inc.) is on the left. (b) Alignment of the 5' UTR regions of Ciccio *TdDRF1.2*, bread wheat *wdreb2* (NCBI AB190608; Egawa *et al*. 2006) and barley *HvDRF1* (NCBI AY223807, Xue & Loveridge 2004), showing nucleotide conservation and gap fraction. Red colour indicates maximum conservation.

The three prime untranslated region (3' UTR) is the mRNA section that follows the coding region. It usually contains several regulatory sequences, such as the polyadenylation signal, AAUAAA or a slight variant, which marks the site of cleavage of the transcript approximately 30 bps past the signal, followed by the addition of several hundred adenine residues (poly(A)$^+$ tail). Other *cis*-regulatory elements often found in the 3'UTR are binding sites for proteins, which may affect mRNAs stability or localization in the cells, and binding sites for miRNAs. 3' UTRs result to be longer than 5'UTR and sometimes can reach up to several kilobases (http://en.wikipedia.org/wiki/3'_UTR; Sonenberg 1994).

The gel in Figure 38 shows the results of the 3'RACE PCR reactions. There was no needing for further nested PCR. From the A band in lane 1 of Fig. 38 the 3' terminal sequence of *TdDRF1.2* cDNA was obtained: it corresponded to the exon 4, from the *E4for-monte* primer till the poly(A)-tract. From the B band in lane 2 the 3' terminal sequence of *TdDRF1.1* was obtained: it was a segment from the *for1.1new* primer till the poly(A) tail. From lane 3 two bands were apparently interesting: the upper C band corresponding to the 3' terminal segment of *TdDRF1.1* from the junction between exon 2 and exon 3, that is where the *for1.2* primer anneals when reaction conditions are not very stringent; lower band D corresponding to the 3' terminal segment of *TdDRF1.2* from the junction between exon 2 and exon 4, that is where the primer specifically anneals. Actually this is possible because the *for1.2* primer was designed at level of the junction between exon 2 and exon 4 and the nucleotides at the beginning of exon 3 and those at the

beginning of exon 4 show a certain level of identity. In lane 4 three bands appeared: E band corresponding to the *1.1* splice variant, F band corresponding to the 1.2 splice variant and G band corresponding to the *1.3* splice variant. As explained for the *for1.2* primer, also the primer *for1.3* was designed at level of an exon-exon junction and, usually, the junctions show a sequence identity among the exons in the regions flanking the splice sites. The gel intensities of each band are also in line with the above explanation: *TdDRF1.1* and *1.2* bands are similarly less intense than *TdDRF1.3* band, being the last the specific product (Fig. 38).

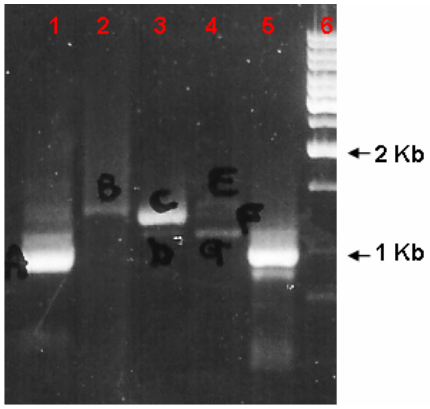



**Figure 38.** Results of the RACE 3'.
Lane 1) PCR with *Efor-monte* and PCR anchored primer. Band A corresponds to 940 bp of the 3' terminal exon 4.
Lane 2) PCR con *for1.1new* and PCR anchored primer. Band B corresponds to 1.305 bp of *TdDRF1.1* (68 bp in exon 3 + 1.237 bp of exon 4).
Lane 3) PCR con *for1.2* and PCR anchored primer. Band C corresponds to 1.390 bp of *TdDRF1.1* (12 bp in exon 2 + 88 bp in exon 3 + 1.237 in exon 4); band D corresponds to 1.249 bp of *TdDRF1.2* (12 bp in exon 2 + 1.237 bp in exon 4).
Lane 4) PCR con *for1.3* and PCR anchored primer. Band E corresponds to 1.394 bp of *TdDRF1.1* (16 bp in exon 1 + 53 bp in exon 2 + 88 bp in exon 3 + 1.237 bp in exon 4); band F corresponds to 1.306 bp of *TdDRF1.2* (16 bp in exon 1 + 53 bp in exon 2 + 1.237 bp in exon 4); band G corresponds to 1.253 bp of *TdDRF1.3* (16 bp in exon 1 + 1.237 bp in exon 4).
Lane 5) Kit Positive Control (C+).
Lane 6) GeneRuler 1kb DNA Ladder (Fermentas Inc.).

The 3'UTR sequence was 279 bp long. Interestingly an Internal Ribosome Entry Site (IRES) was found (nucleotides 185-279). Internal mRNA ribosome binding is a mechanism of translation initiation alternative to the conventional 5'-cap dependent ribosome scanning mechanism. The internal initiation mechanism, firstly observed in picornavirus, is also found in some cellular mRNAs and it is dependent on the IRES *cis*-acting element, located in the 5'UTR region of the mRNA. Comparative analysis of known IRES in some cellular mRNAs identified a common structural motif forming a Y-type strem-loop structure followed by the AUG triplet or followed by additional stem-loop structures and the AUG triplet (Le & Maizel 1997; López De Quinto *et al*.

2001; Koh *et al*. 2003). Here, an IRES element was found in the 3' UTR of all three *TdDRF1* gene products and a similar issue at the moment is not reported in the literature.

**Full-length *TdDRF1* gene transcript cDNAs cloning**

By means of the RACE experiments, for the durum wheat Ciccio genotype, the full-length sequences of the three transcripts were reconstructed and it was possible to deduce the complete codifying sequences (CDS) and the translations of each mRNA.

In this specific case, the three complete CDSs had to be individually cloned and this task was quite difficult and time-consuming because of the gene and transcripts structures (Fig. 11). It is worth to recall here that exon 1 and exon 4, which contain the 5' and the 3' mRNA extremities, respectively, are present in all the three transcript variants.

For the three *TdDRF1* gene transcripts a multiplex reaction with primers *E1F-up* and *E4R-down* was drafted (Fig. 39). The RT-PCR was repeated several times with different conditions and most of the factors that could affect the amplification were taken under consideration. Being desirable to start from total RNA enriched in the target mRNAs, the plant material water status was taken into account and, for this reason, Ciccio leaves at both 2$^{nd}$ and 4$^{th}$ day of water stress were used. Three bands – 1.380 bp for *TdDRF1.1*, 1.292 bp for *TdDRF1.2* and 1.239 bp for *TdDRF1.3* - were expected but, in any case, with this procedure, it has been possible to obtain only the complete CDSs of *TdDRF1.2* and *TdDRF1.3* (Fig. 39 and 41). Probably the efficiency of amplification and the product yield were strictly related to the relative abundances of the target mRNAs in the total RNA. In fact, as highlighted by the real-time experiments, the *TdDRF1.1* transcript is always the less abundant.

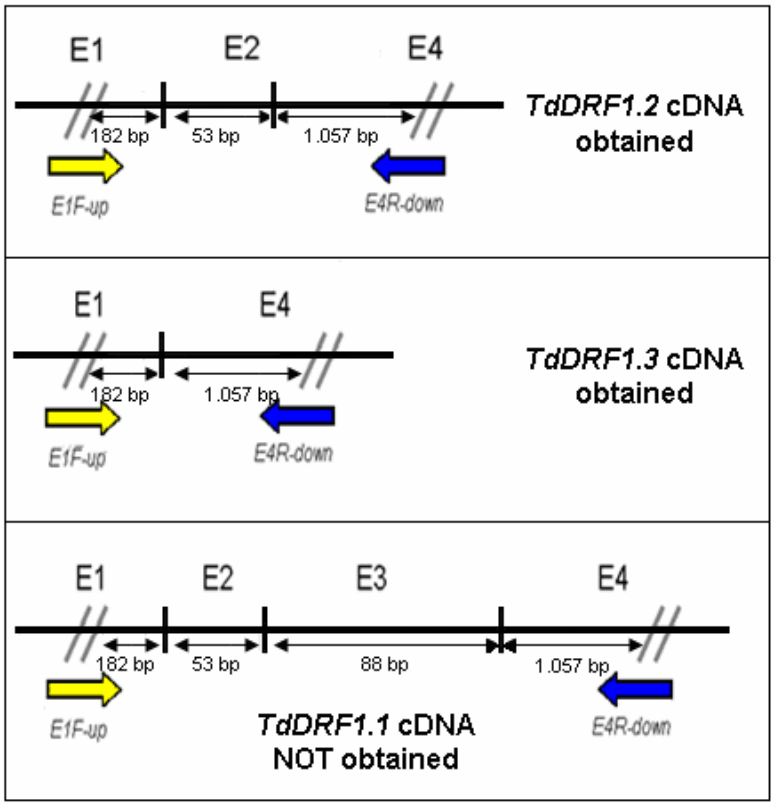



**Figure 39.** Scheme of the amplification of the *TdDRF1* CDSs.

The complete CDS of *TdDRF1.1* was finally amplified by the strategy described in Materials and Methods. In Figure 40 are shown the PCR reactions to amplify the 5' terminal and the 3' terminal segment. Only the bands having the expected molecular weight were cloned and the fragment identities were confirmed by sequencing.

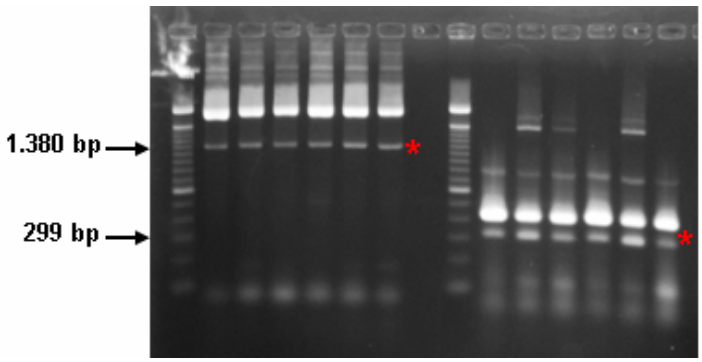



**Figure 40.** PCR reactions with primers pair *E1F-up* and *E3R* (lanes on the right) and with primers pair *E3F* and E4R-*down* (lanes on the left). Asterisks point at the bands of the expected size. Molecular weights marker is 100 bp DNA ladder (Invitrogen).

Afterwards, the two segments were joined together and cloned into the TOPO vector. The digestion with EcoRI, restriction enzyme, which releases the insert from the vector, revealed that many positive recombinant clones were obtained from this transformation. Six recombinant clones have been sequenced to discriminate between the four possibilities of joining two double stranded (ds) DNA fragments and two of them resulted to contain the complete CDS of *TdDRF1.1*, maintaining the right open reading frame (ORF).

Final results of the transcripts CDSs cloning are in Figure 41.

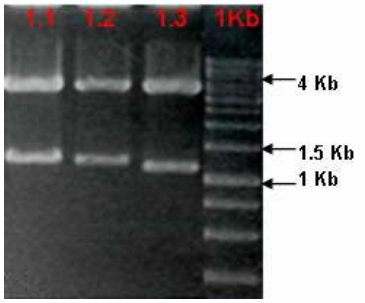



**Figure 41.** EcoRI restriction analysis of the TOPO clones containing the three transcript cDNAs amplified by the primers pair *E1F-up* and *E4R-down*. The linearized pCR®II-TOPO® vector is 3.973 bp; the *TdDRF1.1* insert is 1.380 bp, the *TdDRF1.2* is 1.292 bp; the *TdDRF1.3* is 1.239 bp.

In conclusion, for Ciccio cultivar the three full-length cDNAs having the characteristics reported in the Table 9 were obtained.

| Transcript | Full-length cDNA length (bp) | Full-length cDNA composition and UTRs/exons lenghts | CDS length (aa) | Predicted MW of the protein |
|---|---|---|---|---|
| *TdDRF1.1* | 1.590 bp | 5' UTR: 132 bp<br>E1: 80 bp<br>E2: 53 bp<br>E3: 88 bp<br>E4: 958 bp<br>3' UTR: 279 | 393 aa | 42831.9 |
| *TdDRF1.2* | 1.502 bp | 5' UTR: 132 bp<br>E1: 80 bp<br>E2: 53 bp<br>E4: 958 bp<br>3' UTR: 279 bp | 68 aa | 7086.0 |
| *TdDRF1.3* | 1.449 bp | 5' UTR: 132 bp<br>E1: 80 bp<br>E4: 958 bp<br>3' UTR: 279 bp | 369 aa | 40105.9 |

**Table 9.** Characteristics of the full-length cDNA and the CDS of the tree TdDRF1 gene products.

**Transient expression of *TdDRF*1 gene transcripts in *Nicotiana benthamiana* and *Nicotiana tabacum* plants infected with recombinant Potato Virus X (PVX)**

The heterologous overexpression of the durum wheat *TdDRF1.1*, *TdDRF1.2* and *TdDRF1.3* mRNAs in *N. benthamiana* and *tabacum* plants was achieved by using the PVX-derived viral vector pP2C2S. In the transient expression system based on PVX, the foreign sequences are cloned behind the duplicated coat protein subgenomic promoter which allows efficient transcription of the inserted gene, and the foreign proteins are produced as free proteins.

The three viral constructs (*PVX:TdDRF1.1*, *PVX:TdDRF1.2* and *PVX:TdDRF1*.3; Fig. 42) were engineered using the procedure described in Materials and Methods and sequenced before their use in the subsequent plants infection.

It was known that the *PVX:GFP* successfully replicates in the *N. benthamiana* (Baulcombe *et al*. 1995; Ruiz *et al*, 1998). In this experiment both *N. benthamiana* and *N. tabacum* were used. Plants were manually infected with transcripts generated *in vitro* from the recombinant virus vectors, according to Chapman *et al*. (1992).

Inoculated plants were monitored for their phenotype and analysed for the virus replication.

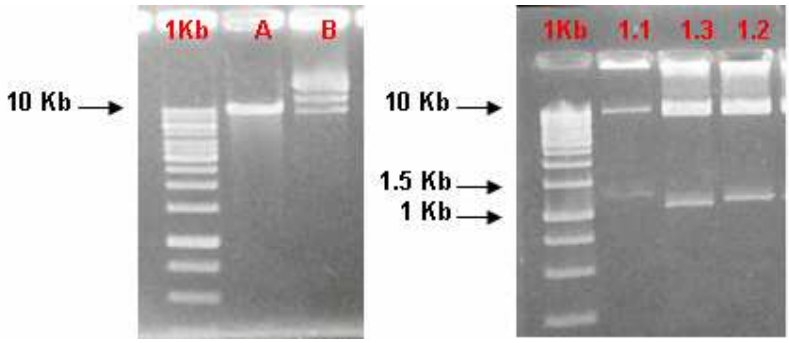



**Figure 42.** Cloning of the three *TdDRF1* CDSs in pP2C2S. Gel on the left: lane A) linearized pP2C2S plasmid by digestion with ClaI and SalI (9.72 Kb); lane B) pP2C2S miniprep. Gel on the right: ClaI and SalI restriction analysis of the *PVX:TdDRF1.1*, *PVX:TdDRF1.2* and *PVX:TdDRF1*.3.

Directly infected leaves and systematically ones were harvested 7 days post inoculation (dpi) for the molecular analysis. Both *Nicotiana* species, inoculated with *PVX:TdDRF1.1-1.2-1.3* viral RNA, developed typical viral symptoms on leaves after 7-10 dpi in contrast to the mock inoculated plants, indicating a systemic infection of plants (Fig. 43c).

To prove for virus replication and *TdDRF1.1-1.2-1.3* transcripts accumulation, samples were taken 7 dpi and total RNA for RT-PCR and real-time RT-PCR analyses was isolated from the samples. In

In Figure 44 some results of the RT-PCR analysis are shown, led with isoform specific primer pairs. In all inoculated plants was observed the specific amplification of the target mRNAs.

To discriminate the exogenous from the endogenous expression, the relative abundances of the transcript in wild type and in infected plants were compared by means of real-time RT-PCR. Quantitative PCR was carried out on wild type plants, plants infected by *PVX:TdDRF1.2* and plants infected by *PVX:TdDRF1.3*. In Figure 45 are shown the results of this analysis, which confirm the overexpression of *1.2* and *1.3* mRNAs with respect to endogenous homologues in the two species of *Nicotiana*. Both *N. benthamiana* and *tabacum* inoculated plants showed an enhancement of the target expression. Results are shown as mean fold change in expression with respect to wild-type plants, here chosen as reference of the experiment, and the maximum overexpression level obtained with this procedure was near a 4-fold increase.

The expression level achieved in a transient transformation should be directly related to the concentration of the inoculated viral mRNA. Actually, even if it is difficult to foresee when a limited increase in the expression of the target ($1.5 \leq MFC \leq 4$) produces the desired effects of an overexpression, with the aim of further analyses, the detected positive-fold changes in the mRNA targets are supposed to be meaningful because related to a master gene, that exerts its function upstream in the complex networks controlling the water stress response.


**Figure 43.** *Nicotiana benthamiana* (a) and *Nicotiana tabacum* (b) wild type plants before inoculation. (c) Symptoms of the viral disease in *N. bentamiana*. Systemically infected leaves appeared with chlorotic spots and leaf curling. The photograph was taken 10 dpi with *in vitro* transcripts of the pP2C2S.


**Figure 44.** RT-PCR results. Two individual infected plants were analyzed for each viral constructs. In the upper part of the gel the PCR reactions relative to six *N. bentamiana* (NB) plants (*a* and *d* infected with *PVX:TdDRF1.1*; *b* and *e* with *PVX:TdDRF1.2*; *c* and *f* with *PVX:TdDRF1.3*) were loaded and the last three lanes are no template controls (NTCs). In the downer part of the gel the PCR reactions relative to six *N. tabacum* (NT) plants (*a* and *d* infected with *PVX:TdDRF1.1*; *b* and *e* with *PVX:TdDRF1.2*; *c* and *f* with *PVX:TdDRF1.3*) were loaded and the last three lanes were dedicated to the positive controls (C+) using Ciccio cDNA at the 4$^{th}$ day of water stress as template. Molecular weights marker is 100 bp DNA Ladder (Fermentas Inc.). The obtained bands were 69 bp long for *1.1* mRNA, 220 bp for *1.2* and 225 bp for *1.3*, as expected.

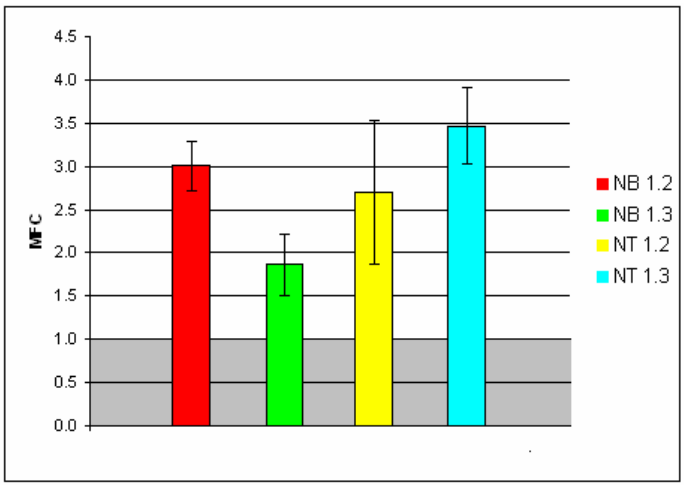



**Figure 45.** Real-time RT-PCR results. Relative quantification of *TdDRF1.2* and *1.3*. Wild type plants are intended as reference samples and the target expression level in the infected plants is reported as a target/reference ratio. MFC is equal to 1 in wild type plants. From the left to the right: *N. benthamiana* infected with *TdDRF1.2* (red) and with *TdDRF1.3* (green), *N. tabacum* infected with *TdDRF1.2* (yellow) and with *TdDRF1.3* (blue).

Hopefully, the overexpression of each *TdDRF1* isoform will allow determining the changes that each transcript produces in the total plant expression profile, by a comparative analysis of both control and overexpressing plants, at mRNA and/or protein level, by microarray (transcriptome analysis) and/or by 2 dimensional gel electrophoresis of proteins (proteome analysis). At the

moment, different protein extraction protocols are under examination with the aim to increase the resolution of SDS-PAGE and 2-D electrophoresis.

In the future, the parallel production of three types of stable transgenic wheat, each type overexpressing one transcript variant of the *TdDRF1* gene, would represent the best way to better determine and evaluate the trait/traits related to the water stress tolerance, conferred by each *TdDRF1* gene product. Wheat transformation with vectors that include the ectopic sequence of a gene for the resistance to water stress (such as *DREB1A* of *Arabidopsis thaliana*) has been already reported (Pellegrineschi *et al*. 2004). In the case under study, it will be possible to generate a durum wheat overexpressing a gene of its own and not derived from another species. The comparison of the phenotypes and the molecular behaviours of wild type and the transgenic plants, could give many insights into the key role of these factors in regulating the expression of other genes involved in the response to dehydration in wheat.

The cloning of the three CDSs in a vector for the heterologous expression in bacteria is another work in progress. For this purpose the pQE-30 Xa vector (Qiagen) was chosen, because it is provided of a tag (6xHis) for the purification of the exogenous expressed protein and, moreover, it presents a recognition site for a protease, which allows the elimination of the tag after the cleavage by the Xa Protease. As it is known, the most important advantage of the expression of a protein in bacteria is the possibility to produce and recover a great amount of protein in a short time, that is the basic requirement to set up the protein crystallization, in the view of a structural analysis based on the X-ray diffraction. This would be highly desirable in the near future, given the scarcity of deposited 3D plant protein structures. The disadvantage of the prokaryotic expression system lies on the lack of post-translational modifications in the resulting proteins.

As already said, the role of the *TdDRF1.2* mRNA has still to be elucidated, but some considerations can be done. Importantly, although the *TdDRF1.2* transcript can be predicted to produce a truncated protein of 68 aa, its translation could be inhibited by several mechanisms or it could be degraded. These mechanisms include nonsense-mediated mRNA decay (NMD) (Baker & Parker 2004) and targeting of microRNAs (miRNA) (Bentwich *et al*. 2005; Pillai 2005). The NMD is a surveillance pathway that leads to the degradation of mRNAs that contain a premature stop codon (PTC) (Chang *et al*. 2007). In addition to alternative splicing, short variations in nucleotide sequence, due to polymorphisms or mutations, single amino acid deletions, insertions or substitutions can lead to a frameshift and the consequent formation of a PTC. The NMD was also described in plants (Schwartz *et al*. 2006; Wu *et al*. 2007). Another pathway, evolved as post-transcriptional regulation system, involves microRNAs (miRNAs), which are short non-coding RNAs (21-25 bps) that can

either induce the degradation or suppress the translation of mRNAs (Carrington & Ambros 2003; Brennecke *et al*. 2005).

## CONCLUSIONS

In conclusion, the complete codifying sequences of the three *TdDRF1* transcripts were isolated and cloned. They could be transferred in any kind of vector to allow possible functional and/or structural analyses, both *in vitro* and *in vivo*.

Besides, in this third section of my PhD thesis, I describe an approach towards the investigation of the molecular function of all the three gene products, using a transient overexpression system. The PVX-derived constructs for the three *TdDRF1* CDSs were engineered and used for infecting *N. benthamiana* and *tabacum* plants. An increased mRNA expression level, in comparison to wild type plants, was detected in all inoculated plants. The immediate next step will be the protein analysis with the double aim of isolating the target proteins produced in plant and of investigating their molecular partners in the networks involved in the dehydration stress signalling and response.

All these information plus the experimental data reported in the previous sections of this PhD work, as basic research, will be useful to enrich the actual knowledge about the molecular basis of drought stress and, possibly, they will draw together to a future application in molecular assisted breeding.

# REFERENCES


Adams KL, Cronn R, Percifield R & Wendel JF (2003). Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proceedings of the National Academy of Science of the USA* **100**, 4649-4654.

Agawal P, Agarwal PK, Nair S, Sopory SK & Reddy MK (2007). Stress-inducible DREB2A transcription factor from *Pennisetum glaucum* is a phosphoprotein and its phosphorylation negatively regulates its DNA-binding activity. *Molecular Genetics and Genomics* **277**, 189-198.

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

Ahlquist P & Pancha RF (1990). Gene amplification and expression by RNA viruses and potential for further application to plant gene transfer. *Physiologia Plantarum* **79**, 163-167.

Allen MD, Yamasaki K, Ohme-Takagi M, Tateno M & Suzuki M (1998). A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *EMBO Journal* **17**, 5484-5496.

Applied Biosystems (2001). User Bulletin #2: ABI PRISM 7700 Sequence Detection System, P/N 4303859.

Applied Biosystems (2007). TaqMan® PreAmp Master Mix Kit Protocol, P/N 4366127.

Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815.

Araus JL, Slafer GA, Reynolds MP & Royo C (2002). Plant breeding and drought in C3 cereals: what should we breed for? *Annals of Botany* **89**, 925-940.

Baek KH & Skinner DZ (2004). Quantitative real-time PCR method to detect changes in specific transcript and total RNA amounts. *Electronic Journal of Biotechnology* **7**, 55-60.

Baker KE & Parker R (2004). Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Current Opinion in Cell Biology* **16**, 293-299.

Barrs C & Wheatherley PE (1968). A re-examination of the relative turgidity technique for estimating water deficit in leaves. *Australian Journal of Biological Science* **15**, 413-428.

Bartels D & Sunkar R (2005). Drought and salt tolerance in plants. *Critical Reviews in Plant Sciences* **24**, 23-58.

Bajaj S, Targolli J, Liu LF, Ho THD & Wu R (1999). Transgenic approaches to increase dehydration-stress tolerance in plants. *Molecular Breeding* **5**, 493-503.

Baulcombe DC, Chapman S & Santa-Cruz S (1995). Jellyfish green fluorescent protein as a reporter for virus infections. *Plant Journal* **7**, 1045–53.

Behnam B, Kikuchi A, Celebi-Toprak F, Yamanaka S, Kasuga M, Yamaguchi-Shinozaki K & Watanabe KN (2006). The *Arabidopsis DREB1A* gene driven by the stress-inducible *rd29A* promoter increases salt-stress tolerance in proportion to its copy number in tetrasomic tetraploid potato (*Solanum tuberosum*). *Plant Biotechnology* **23**, 169-177.

Bentwich I (2005). Prediction and validation of microRNAs and their targets. *FEBS Letters* **579**, 5904-5910.

Bhalla PL (2006). Genetic engineering of wheat – current challenges and opportunities. *Trends in Plant Biotechnology* **24**, 305-311.

Blum A, Shilper L, Golan G & Mayer J (1998). Yield stability and canopy temperature of wheat genotypes under drought stress. *Field Crop Research* **22**, 289-296.

Bohnert HJ, Nelson DE & Jensen RG (1995). Adaptation to environmental stresses. *Plant Cell* **7**, 1099-1111.

Boyer JS (1982). Plant productivity and environment. *Science* **218**, 443-448.

Bray E (1993) Molecular responses to water deficit. *Plant Physiology* **103**, 1035-1040.

Bray E (1997). Plant responses to water deficit. *Trends in Plant Science* **2**, 48-54.

Bray E (2004). Genes commonly regulated by water-deficit stress in *Arabidopsis thaliana*. *Journal of Experimental Botany* **55**, 2331-2341.

Breljak D, Ambriović-Ristov A, Kapitanović S, Čačev T & Gabrilovac J (2005). Comparison of three RT-PCR based methods for relative quantification of mRNA. *Food Technology and Biotechnology* **43**, 379-388.

Brennecke J, Stark A, Russell RB & Cohen SM (2005). Principles of microRNA-target recognition. *PLoS Biology* **3**, 404-418.

Brett D, Pospisil H, Valcarcel J, Reich J & Bork P (2002). Alternative splicing and genome complexity. *Nature Genetics* **30**, 29-30.

Bustin SA (2002). Quantification of mRNA using real-time reverse transcription (RT-PCR): trend and problems. *Journal of Molecular Endocrinology* **29**, 23-29.

Bustin SA & Nolan T (2004). Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *Journal of Biomolecular Techniques* **15**, 155-166.

Carrington JC & Ambros V (2003). Role of microRNAs in plant and animal development. *Science* **301**, 336-338.

Chang YF, Imam JS & Wilkinson MF (2007). The nonsense-mediated decay RNA surveillance pathway. *Annual Review of Biochem*istry **76**, 51-74.

Chapman S, Kavanagh T & Baulcombe D (1992). Potato virus X as a vector for gene expression in plants. *Plant Journal* **2**, 549–57.

Chen J & Zhu T (2004). Networks of transcription factors with roles in environmental stress response. *Trends in Plant Science* **9**, 591-596.

Cheng Z, Targolli J, Huang X & Wu R (2002). Wheat LEA genes, PMA80 and PMA1959, enhance dehydration tolerance of transgenic rice (Oryza sativa L.). *Molecular Breeding* **10**, 71-82.

Christensen CA & Feldmann KA (2007). Biotecnology approaches to engineering drought tolerant crop. In *Advances in molecular breeding toward drought and salt tolerant crops*. *Springer Netherlands* Publisher 333-357.

Close TJ (1996). Dehydrins: emergence of a biochemical role of a family plant of dehydration proteins. *Physiologia Plantarum 97*, 795-803.

Curry J, Morris CF, & Walker-Simmons MK (1991). Sequence analysis of a cDNA encoding a group 3 LEA mRNA inducible by ABA or dehydration stress in wheat. *Plant Molecular Biology* **16**, 1073-1076.

Diaz-Peréz JC, Shackel KA & Sutter EG (1995). Relative water content and water potential of tissue. *Journal of Experimental Botany* **46**, 111-118.

Doležel J, Kubalálová M, Paux E, Bartoš J & Feuillet C (2007). Chromosome-based genomic in the cereals. *Chromosome Research* **15**, 51-66.

Dubouzet JG, Sakuma Y, Ito Y, Kasuga M, Dubouzet EG, Miura S, Seki M, Shinozaki K & Yamaguchi-Shinozaki K (2003). *OsDREB* genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *Plant Journal* **33**, 751-763.

Egawa C, Kobayashi F, Ishibashi M, Nakamura T, Nakamura C & Takumi S (2006). Differential regulation of transcript accumulation and alternative splicing of a *DREB2* homolog under abiotic stress conditions in common wheat. *Genes and Genetic Systems* **81**, 77-91.

Escobar C, Hernández LE, Jiménez A, Creissen G, Ruiz MT & Mullineaux PM (2003). Transient expression of *Arabidopsis thaliana* ascorbate peroxidase 3 in *Nicotiana benthamiana* plants infected with recombinant potato virus X. *Plant Cell Reports* **21**, 699-704.

Feldman M & Levy AA (2005). Allopolyploidy – a shaping force I the evolution of wheat genomes. *Cytogenetic and Genome Research* **109**, 250-258.

Fiser A & Sali A (2003) Modeller: generation and refinement of homology-based protein structure models. *Methods in Enzymology* **374**, 461-491.

Forster BP, Ellis RP, Thomas WTB, Newton AC, Tuberosa R, This D, El-Enein RA, Bahri MH & Ben Salem M (2000). The development and application of molecular markers for abiotic stress tolerance in barley. *Journal of Experimental Botany* **51**, 19-27.

Fowler S & Thomashow MF (2002). *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to CBF cold response pathway. *Plant Cell* **14**, 1675-1690.

Freeman WM, Walker SJ & Vrana KE (1999). Quantitative RT-PCR: pitfalls and potential. *Biotechniques* **26**, 112-122.

Galeffi P, Lombardi A, Di Donato M, Latini A, Sperandei M, Cantale C & Giacomini P (2005). Expression of single-chain antibodies in transgenic plants. *Vaccine* **23**, 1823-1837.

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

Gilmour SJ, Fowler SG & Thomashow MF (2004). Arabidopsis transcriptional activators CBF1, CBF2 and CBF3 have matching functional activities. *Plant Molecular Biology* **54**, 767-781.

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhause J, Macalma T, Oliphant A & Briggs S (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp *japonica*). *Science* **296**, 92-100.

González L & González-Vilar M (2001). Determination of relative water content. In *Handbook of Plant Ecophysiology Techniques*, chapter **14**, 207-212.

Gouet P, Courcelle E, Stuart DI & Mtoz F (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305-308.

Goyal K, Walton LJ & Tunnacliffe A (2005). LEA proteins prevent protein aggregation due to water stress. *Biochemical Journal* **388**, 151-157.

de la Grange P, Dutertre M, Correa M & Auboeuf D (2007). A new advance in alternative splicing databases: from catalogue to detailed analysis of regulation of expression and function of human alternative splicing variants. *BMC Bioinformatics* **8**, 180.

Grzesiak M, Rzepka A, Hura T, Hura K & Skoczowski A (2007). Changes in response to drought stress of triticale and maize genotypes differing in drought tolerance. *Photosynthetica* **45**, 280-287.

Hammond-Kosack KE, Staskawicz BJ, Jones JDG & Baulcombe DC (1995). Functional expression of a fungal avirulence gene from a modified potato virus X genome. *Molecular Plant-Microbe Interactions* **8**, 181–185.

Holmberg N & Bulow L (1998). Improving stress tolerance in plants by gene transfer. *Trends in Plant Science* **3**, 61-66.

Horser C, Abbot D, Wesley V, Smith N and Waterhouse P (2002). Gene silencing – principles and applications. In *Genetic engineering: principles and methods*. Edited by Setlow JK. *Kluwer Academic/Plenum Publishers*, Volume **24**, 239-357.

Ingram J & Bartels D (1996). The molecular bases of dehydration tolerance in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**, 377-403.

Izquierdo Rojo M (2001). Ingeniería genética y transferencia génica. Ediciones Pirámide, S.A.

Jepsen JS, Sorensen MD & Wengel J (2004). Locked nucleic acid: a potent nucleic acid analog in therapeutics and biotechnology. *Oligonucleotides* **14**, 130-146.

Jiang Y, Qin F, Li Y, Fang X & Bai C (2004). Measuring specific interaction of transcription factor ZmDREB1A with its DNA responsive element at the molecular level. *Nucleic Acids Research* 32, e101.

Kasuga M, Liu Q, Miura S, Yamaguchi-Shinozaki K & Shinozaki K (1999). Improving plant drought, salt and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nature Biotechnology* **17**, 287-291.

Keller B & Feuillet C (2000). Colinearity and gene density in grass genomes. *Trends in Plant Science* **5**, 246-251.

Kellogg EA (2001). Evolutionary history of the grasses. *Plant Physiology* **125**, 1198-1205.

Kizis D, Lumbreras V & Pagès M (2001). Role of AP2/EREBP transcription factors in gene regulation during abiotic stress. *FEBS Letters* **498**, 187-189.

Koenig R & Laseman DE (1989). Potato virus X, potexvirus group. In *AAB Description of plant viruses*, No. 354 (Murant AF and Harrison BD, eds). Warwick: Association of Applied Biologists, 5 pp.

Koh DCY, Wong SM & Liu DX (2003). Synergism of the 3'-Untranslated Region and an Internal Ribosome Entry Sire differentially enhances the translation of a plant virus coat protein. *Journal of Biological Chemistry* **278**, 20565-20573.

Krissinel E & Henrick K (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallographica Section D Biological Crystallography* **60**, 2256-2268.

Kriventseva EV, Koch I, Apweiler R, Vingron M, Bork P, Gelfand MS & Sunyaev S (2003). Increase of functional diversity by alternative splicing. *Trends in Genetics* **19**,124-128.

Langridge P, Paltridge N & Fincher G (2006). Functional genomics of abiotic stress tolerance in cereals. *Briefings in Functional Genomics and Proteomics* **4**, 343-354.

Latini A, Sperandei M, Sharama S, Cantale C, Iannetta M, Dettori M, Ammar K, Galeffi P (in press). Molecular analyses of a dehydration related gene from DREB family in durum wheat and Triticale. In *Biosaline Agriculture and High Salinity Tolerance in Plants.* Published by Birkhäuser and Verlag.

Latini A, Rasi C, Sperandei M, Cantale C, Iannetta M, Dettori M, Ammar K & Galeffi P (2007). Identification of a DREB-related gene in *Triticum durum* and its expression under water stress conditions. *Annals of Applied Biology* **150**, 187-195.

Laskowski R, MacArthur M, Moss D & Thornton J (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **26**, 283-291.

Le SY & Maizel Jr JV (1997). A common RNA structural motif involved in the internal initiation of translation of cellulat mRNAs. *Nucleic Acids Research* **25**, 362-369.

Li XP, Tian AG, Luo GZ, Gong ZZ, Zhang JS & Chen SY (2005). Soybean DRE-binding transcription factors that are responsive to abiotic stresses. *Theoretical and Applied Genetics* **110**, 1355-1362.

Liu L, White MJ & MacRae TH (1999). Transcription factors and their genes in higher plants. *FEBS journal* **262**, 247-257.

Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K & Shinozaki K (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.

Livak KJ & Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* **25**, 402-408.

López De Quinto S, Lafuente E & Martínez-Salas E (2001). IRES interaction with translation initiation fators: functional characterization of novel RNA contacts with eIF3, eIF4B, and eIF4GII. *RNA* **7**, 1213-1226.

Magnani E, Sjölander K & Hake S (2004). From endonucleases to transcription factors: evolution of the AP2 binding domain in plants. *Plant Cell* **16**, 2265-2277.

Mahajan S & Tuteja N (2005). Cold, salinity and drought stresses: an overview. *Archives of Biochemistry and Biophysics* **444**, 139-158.

Maruyama K, Sakuma Y, Kasuga M, Ito Y, Seki M, Goda H, Shimada Y, Yoshida S, Shinozaki K & Yamaguchi-Shinozaki K (2004). Identification of cold-inducible downstream genes of the *Arabidopsis* DREB1A/CBF3 transcriptional factor using two microarray systems. *Plant Journal* **38**, 982-993.

Meyers BC, Scalabrin S & Morgante M (2004). Mapping and sequencing complex genomes: let's get physical! *Nature Reviews Genetics* **5**, 578-581.

Mignone F, Grillo G, Licciulli F, Iacono M, Liuni S, Kersey PJ, Duarte J, Saccone C & Pesole G. (2005). UTRdb and UTRsite: a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs. *Nucleic Acids Research* **33**, D141-D146.

Morozov SY, Gorbulev VG, Novikov VK, Agranovskii AA, Kozlov YV, Atabekov JG & Baev AA (1981). Primary structure of 5'- and 3'-terminal regions of the potato virus X genomic RNA. *Doklady Akademii Nauk SSSR* **259**, 723-725.

Muller PY, Janovjak H, Miserez AR & Dobbie Z (2002). Processing of gene expression data generated by quantitative real-time RT-PCR. *BioTechniques* **32**, 1372-1379.

Nakashima K, Shinwari ZK, Sakuma Y, Seki M, Miura S, Shinozaki K & Yamaguchi-Shinozaki K (2000). Organization and expression of two *Arabidopsis DREB2* genes encoding DRE-binding proteins involved in dehydration- and high-salinity-responsive gene expression. *Plant Molecular Biology* **42**, 657-665.

Nakashima K & Yamaguchi-Shinozaki K (2006). Regulons involved in osmotic stress-responsive and cold stress-responsive gene expression in plants. *Physiologia Plantarum* **126**, 62-71.

Narusaka Y, Nakashima K, Shinwari ZK, Sakuma Y, Furihata T, Abe H, Narusaka M, Shinozaki K & Yamaguchi-Shinozaki K (2003). Interaction between two *cis*-acting elements, ABRE and DRE, in ABA-dependent expression of Arabidopsis *rd29A* gene in response to dehydration and high-salinity stresses. *Plant Journal* **34**, 137-148.

Negrouk V, Eisner G, Lee H, Han K, Taylor D & Wong HC (2005). Highly efficient transient expression of functional recombinant antibodies in lettuce. *Plant Science* **169**, 433-438.

Nicot N, Hausman JF, Hoffmann L & Evers D (2005). Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *Journal of Experimental Botany* **56**, 2907-2914.

Okamuro JK, Caster B, Villarroel R, Vam Montagu M & Jofuku KD (1997). The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in Arabidopsis. *Proceedings of the National Academy of Science of the USA* **94**, 7076-7081.

Ouk M, Basnayake J, Tsubo M, Fukai S, Fischer KS, Cooper M & Nesbitt H (2006). Use of drought response index for identification o drought tolerant genotypes in rainfed lowland rice. *Field Crop Research* **99**, 48-58.

Paux E, Roger D, Badaeva E, Gay G, Bernard M, Sourdille P & and Feuillet C (2006). Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. *Plant Journal* **48**, 463-474.

Pellegrineschi A, Reynolds M, Pacheco M, Brito RM, Almeraya R, Yamaguchi-Shinozaki K & Hoisington D (2004). Stress-induced expression in wheat of the *Arabidopsis thaliana* DREB1A gene delays water stress symptoms under greenhouse conditions. *Genome* **47**, 493-500.

Pesole G, Liuni S, Grillo G, Ippedico M, Larizza A, Makalowski W & Saccone C. (1999). UTRdb: a specialized database of 5' and 3' untranslated regions of eukaryotic mRNAs. *Nucleic Acids Research* **27**, 188-191.

Pesole G, Liuni S, Grillo G, Lacciulli F, Mignone F, Gissi C & Saccone C (2002). UTRdb and UTRsite: specialized databases of sequences and functional elements of 5' and 3' untranslated regions of eukaryotic mRNAs. Upadate 2002. *Nucleic Acids Research* **30**, 335-340.

Pfaffl MW (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research* **29**, e45.

Pillai RS (2005). MicroRNA function: multiple mechanisms for a tiny RNA? *Rna* **11**, 1753-1761.

Porta C & Lomonossoff GP (1996). Use of viral replicons for the expression of genes in plants. *Molecular Biotechnoogy.* **5**, 209–221.

Qin F, Sakuma Y, Li J, Liu Q, Li YQ, Shinozaki K & Yamaguchi-Shinozaki K (2004). Cloning and functional analysis of a novel DREB1/CBF transcription factor involved in cold-responsive gene expression in *Zea mays* L. *Plant Cell Physiology* **45**, 1042-1052.

Qin F, Kakimoto M, Sakuma Y, Maruyama K, Osakabe Y, Tran LSP, Shinozaki K & Yamaguchi-Shinozaki K (2007). Regulation and functional analysis of *ZmDREB2A* in response to drought and heat stress in *Zea mays* L. *Plant Journal* **50**, 54-69.

Radonic A, Thulke S, Mackay IM, Landt O, Siegert W & Nitsche A (2004). Guideline to reference gene selection for quantitative real-time PCR. *Biochemical and Biophysical Research Communications* **313**, 856-862.

Ramakers C, Ruijter JM, Lekanne Deprez RH & Moorman AFM (2003). Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience Letters* **339**, 62-66.

Rampino P, Pataleo S, Gerardi C, Mita C & Perrotta C (2006) Drought stress response in wheat: physiological and molecular analysis of resistant and sensitive genotypes. *Plant, Cell and Environment* **29**, 2143-2152.

Ren SX, McIntosh RA & Lu ZJ (1997). Genetic suppression of the cereal rye-derived gene *Pm8* in wheat. *Euphyca* **93**, 353-360.

Riechmann JL & Ratcliffe OJ (2000). A genomic perspective on plant transcription factors. *Current Opinion in Plant Biology* **3**, 423-434.

Rodriguezmaribona B, Tenorio JR, Conde JR & Ayerbe L (1992). Correlation between yield and osmotic adjustment of peas (*Pisum sativum* L) under drought stress. *Field Crop Research* **29**, 15-22.

Rommens CMT, Salmeron JM, Baulcombe DC & Staskawicz BJ (1995). Use of a gene expression system based on potato virus X to rapidly identify and characterise a tomato *Pto* homolog that controls fenthion sensitivity. *Plant Cell* **7**, 249–257.

Ruiz MT, Voinnet O & Baulcombe DC (1998). Initiation and maintenance of virus-induced gene silencing. *Plant Cell* **10**, 937-946.

Rutledge RG & Côte C (2003). Mathematics of quantitative kinetic PCR and the application of standard curves. *Nucleic Acids Research* **31**, e93.

Sablowski RWM, Baulcombe DC & Bevan M (1995). Expression of a flower-specific Myb protein in leaf cells using a viral vector causes ectopic activation of a target promoter. *Proceedings of the National Academy of Science of the USA* **92**, 6901–6905.

Saitoh H, Kiba A, Nishihara M, Yamamura S, Suzuki K & Terauchi R (2001). Production of antimicrobial defensin in *Nicotiana benthamiana* with a potato virus X vector. *Molecular Plant-Microbe Interaction* **14**, 111–115.

Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K & Yamaguchi-Shinozaki K (2002). DNA-binding specificity of the ERF/AP2 domain in Arabidopsis DREBs transcription factors involved in dehydration- and cold-inducible gene expression. *Biochemical and Biophysical Research Communications* **290**, 998-1009.

Sakuma Y, Maruyama K, Osakabe Y, Qin F, Seki M, Shinozaki K & Yamaguchi-Shinozaki (2006). Functional analysis of an *Arabidopsis* transcription factor, DREB2A, involved in drought-responsive gene expression. *Plant Cell* **18**, 1292-1309.

Schmittgen T & Zakrajsek BA (2000). Effect of experimental treatment on housekeeping gene expression: validation by reel-time quantitative RT-PCR. *Journal of Biochemical and Biophysical Methods* **46**, 69-81.

Scholthof HB, Scholthof KBG & Jackson AO (1996). Plant virus gene vectors for transient expression of foreign proteins in plants. *Annual Review Phytopatology* **34**, 299-323.

Scholthof KBG, Mirkow E & Scholthof HB (2002). Plant viruses gene vectors: biotechnology and applications in agriculture an medicine. In *Genetic engineering: principles and methods*. Edited by Setlow JK. *Kluwer Academic/Plenum Publishers*, Volume **24**, 239-357.

Schwartz AM, Komarova TV, Skulachev MV, Zvereva AS, Dorokhov YL & Atabekov JG (2006). Stability of plant mRNAs depends on the length of the 3'-untranslated region. *Biochemistry (Moscow)* **12**, 1377-1384.

Seki M, Narusaka M, Kamiya A , Ishida J, Satou M , Sakurai T, Nakajima M, Enju A, Akiyama K, Oono Y, Muramatsu M, Hayashikazi Y, Kawai J, Carnici P, Itoh M, Ishii Y, Arakawa T, Shibata K, Shinagawa A, Shinozaki K (2002). Functional annotation of a full-length *Arabidopsis* cDNA collection. *Science* **296**, 141-145.

Shen YG, Zhang WK. He SJ, Zhang JS, Liu Q & Chen SY (2003a). An EREBP/AP2-type protein in *Triticum aestivum* was a DRE-binding transcription factor induced by cold, dehydration and ABA stress. *Theoretical and Applied Genetics* **106**, 923-930.

Shen YG, Zhang WK, Yan DQ, Du BX, Zhang JS, Liu Q & Chen SY (2003b). Characterization of a DRE-binding transcription factor from a halophyte Atriplex hortensis. *Theoretical and Applied Genetics* **107**, 155-161. Shinagawa A & Shinozaki K (2002). Functional annotation of a full-length *Arabidopsis* cDNA collection. *Science* **296**, 141–145.

Shimamura C, Ohno R, Nakamura C & Takumi S (2006). Improvement of freezing tolerance in tobacco plants expressing a cold-responsive and chloroplast-targeting protein WCOR15 of wheat. *Journal of Plant Physiology* **163**, 213–219.

Shinozaki K & Yamaguchi-Shinozaki K (1997). Gene expression and signal transduction in water-stress response. *Plant Physiology* **115**, 327-334.

Shinozaki K & Yamaguchi-Shinozaki K (2000). Molecular responses to dehydration and low temperatures: differences and cross-talk between two stress signalling pathways. *Current Opinion in Plant Biology* **3**, 217-223.

Shinozaki K, Yamaguchi-Shinozaki K & Seki M (2003). Regulatory network of gene expression in the drought and cold stress response. *Current Opinion in Plant Biology* **6**, 410-417.

Shinozaki K & Yamaguchi-Shinozaki K (2007). Gene networks involved in drought stress response and tolerance. *Journal of Experimental Botany* **58**, 221-227.

Shinwari ZK, Nakashima K, Miura S, Kasuga M, Seki M, Yamaguchi-Shinozaki K & Shinozaki K (1998). An Arabidopsis gene family encoding DRE/CRT binding proteins involved in low-temperature-responsive gene expression. *Biochemical and Biophysical Research Communications* **22**, 161–170.

Simon P (2003). Q-Gene: processing quantitative real-time RT-PCR data. *Bioinformatics* **19**, 1439-1410. Singh KB (1998). Transcriptional regulation in plants: the importance of combinatorial control. *Plant Physiology* 118, 1111-1120.

Sippl M.J (1993). Recognition of errors in three-dimensional structures of proteins. *Proteins* **17**, 355-362.

Skinner JS, von Zitzewitz J, Szücs P, Marquez-Cedillo L, Filichkin T, Amundsen K, Stockinger EJ, Thomashow MF, Chen THH &.Hayes PM (2005). Structural, functional, and phylogenetic characterization of a large *CBF* gene family in barley. *Plant Molecular Biology* **59**, 533-551.

Solinas A, Brown LJ, McKeen C, Mellor JM, Nicol J, Thelwell N & Brown T (2001). Duplex Scorpion primers in SNP analysis and FRET applications. *Nucleic Acids Research* **29**, e96.

Sonenberg N, Shatkin AJ, Ricciardi RP, Rubin M & Goodman RM (1978). Analysis of terminal structures of RNA potato virus X. *Nucleic Acids Research* **5**, 2501-2512.

Sonenberg N. (1994) mRNA translation: influence of the 5' and 3' untranslated regions. *Current Opinion in Genetics and Development* **4**, 310-315.

Sreenivasulu N, Sopory SK & Kavi Kishor PB (2006). Deciphering the regulatory mechanisms of abiotic stress tolerance in plants by genomic approach. *Gene* **388**, 1-13.

Stamm S, Ben-Ari S, Rafalska I, Tang Y, Zhang Z, Toiber D, Thanaraj TA & Soreq H (2005). Function of alternative splicing. *Gene* **344**, 1-20.

Stockinger EJ, Gilmour SJ & Thomashow M (1997). *Arabidopsis thaliana CBF1* encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a *cis*-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proceedings of the National Academy of Science of the USA* **94**, 1035-1040.

Stürzenbaum SR & Kille P (2001). Control genes in quantitative molecular biological techniques: the variability of invariance. *Comparative Biochemistry and Physiology Part B* **130**, 281-289.

Tahara M, Carver BF, Johnson RC & Smith EL (1990). Relationship between relative water content during reproductive development and winter wheat grain yield. *Euphytica* **49**, 255-262.

Tan W, Wang K & Drake TJ (2004). Molecular beacons. *Current Opinion in Chemical Biology* **8**, 547-553.

Tang W, Newton RJ, Li C & Charles TM (2006). Enhanced stress tolerance in transgenic pine expressing the pepper *CaPF1* gene is associated with the polyamine biosynthesis. *Plant Cell Reports* **26**, 115-124.

Tester M & Bacic A (2005). Abiotic stress tolerance in grasses. From model plants to crop plants. *Plant Physiology* **137**, 791-793.

Thellin O, Zorzi W, Lakaye B, De Borman B, Coumans B, Henne G, Grisar T, Igout A & Heinen E (1999). Housekeeping genes as internal standards: use and limits. *Journal of Biotechnology* **75**, 291-295.

Thomashow MF (1999). Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. *Annual Review of Plant Physiology and Plant Molecular Biology* **50**, 571-599.

Thompson JD, Higgins DG & Gibson TJ (1994). CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680.

Trethowan RM, Crossa J, van Ginkel M & Rajaram S (2001). Relationships among bread wheat international yield testing locations in dry areas. *Crop Science* **41**, 1461-1469.

Trethowan RM, van Ginkel M, Ammar K, Crossa J, Payne TS, Cukadar B, Rajaram S & Hernandez E (2003). Associations among twenty years of bread wheat yield evaluation environments. *Crop Science* **43**, 1698–1711.

Trethowan RM & Crossa J (2006). Lessons learnt from forty years of international spring bread wheat trials. *Euphytica* **157**, 385-390.

Tuberosa R & Salvi S (2006). Genomics-based approaches to improve drought tolerance of crops. *Trends in Plant Science* **11**, 405-412.

Tyagi S & Kramer FR (1996). Molecular beacons: probes that fluoresce upon hybridization. *Nature Biotechnology* **14**, 303-308.

Umezawa T, Fujita M, Fujita Y, Yamaguchi-Shinozaki K & Shinozaki K (2006). Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future. *Current Opinion in Biotechnology* **17**, 113-122.

Upelniek VP, Novoselskaya AY, Sutka J, Galiba G & Metakovsky V (1994). Genetic variation at storage protein-coding loci of common wheat (cv 'Chinese Spring') induced by nitrosoethylurea and by the cultivation of immature embryos in vitro. *Theoretical and Applied Genetics* **90**, 372-379.

Valliyodan B & Nguyen HT (2006). Understanding regulatory networks and engineering for enhanced drought tolerance in plants. *Current Opinion in Plant Biotechnology* **9**, 189-195.

Verslues PE, Agarwal M, Katiyar-Agarwal S, Zhu J & Zhu JK (2006). Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. *Plant Journal* **45**, 523-539.

Vinocur B and Altman A (2005). Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations. *Current Opinion in Biotechnology* **16**, 123-132.

Ware D & Stein L (2003). Comparison of genes among cereals. *Current Opinion in Plant Biology* **6**, 121-127.

Waterhouse PM & Helliwell CA (2003). Exploring plant genomes by RNA-induced gene silencing. *Nature Reviews Genetics* **4**, 29-38.

Weigel D (1995). The APETALA2 domain is related to a novel type of DNA binding domain. *Plant Cell* **7**, 388-389.

Wendel JF (2000). Genome evolution in polyploids. *Plant Molecular Biology* **42**, 225-249.

Wise MJ (2003). LEAping to conclusions: a computational reanalysis of late embryogenesis abundant proteins and their possible roles. *BMC Bioinformatics* **4**, 52.

Wise MJ & Tunnacliffe A (2004). POPP the questions: what do LEA proteins do? *Trends in Plant Science* **9**, 13-17.

Wolffs P, Grage H, Hagberg O & Radstrom P (2004). Impact of DNA polymerases and their buffer systems on quantitative real-time PCR. *Journal of Clinical Microbiology* **42**, 408–411.

Wong ML & Medrano JF (2005). Real-time PCR for mRNA quantitation. *Biotechniques* **39**, 75-85.

Wu J, Kang JH, Hettenhausen C & Baldwin IT (2007). Nonsense-mediated mRNA decay (NMD) silences the accumulation of aberrant trypsin protease inhibitor mRNA in *Nicotiana attenuata*. *Plant Journal* **51**, 693-706.

Xue GP & Loveridge CW (2004). *HvDRF1* is involved in abscisic acid-mediated gene regulation in barley and produces two forms of AP2 transcriptional activators, interacting preferably with a CT-rich element. *Plant Journal* **37**, 326-339.

Yamaguchi-Shinozaki K & Shinozaki K (1993). Characterization of the expression of a desiccation-responsive *rd29A* gene of *Arabidopsis thaliana* and analysis of its promoter in transgenic plants. *Molecular Genes and Genetics* **236**, 331-340.

Yamaguchi-Shinozaki K & Shinozaki K (1994). A novel *cis*-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *Plant Cell* **6**, 251-264.

Yamaguchi-Shinozaki K & Shinozaki K (2005). Organization of *cis*-acting regulatory elements in osmotic- and cold-stress-responsive promoters. *Trends in Plant Science* **10**, 88-94.

Yamaguchi-Shinozaki K & Shinozaki K (2006). Transcriptional regulatory networks in cellular responses to dehydration and cold stresses. *Annual Review of Plant Biology* **57**, 781-803.

Zhang J & Byrne CD (1999). Differential priming of RNA templates during cDNA synthesis markedly affects both accuracy and reproducibility of quantitative competitive reverse-transcriptase PCR. *Biochemical Journal* **337**, 231-241.

Zhang JZ (2003). Overexpression analysis of plant transcription factors. *Current Opinion in Plant Biology* **6,** 430-440.

Zhang JZ, Creelman RA & Zhu JK (2004). From laboratory to field. Using information from Arabidopsis to engineer salt, cold, and drought tolerance in crops. *Plant Physiology* 135, 615-621.

Zhao TJ, Sun S, Liu Y, Liu JM, Liu Q, Yan YB & Zhou H. (2006). Regulating the drought-responsive element (DRE)-mediated signalling pathway by synergic functions of trans-active and trans-inactive DRE binding factors in *Brassica napus*. *Journal of Biological Chemistry* **281**, 10752-10759.

Zhou Y, Zhou C, Ye L, Dong J, Xu H, Cai L, Zhang L & Wei L (2003). Database and analyses of known alternatively spliced genes in plants. *Genomics* **82**, 584–595.

Ziegler A, Cowan GH, Torrance L, Ross HA & Davies HV (2000). Facile assessment of cDNA constructs for expression of functional antibodies in plants using the potato virus X vector. *Molecular Breeding* **6**, 327-335.

# SUMMARY

Adverse environmental stresses, such as drought, cold and high salinity, have a strong influence on agricultural production. The study of the mechanisms involved in plant response and tolerance to these stresses represents a major challenge for plant scientists especially in light of foreseen global climate changes. Plant genomics research is beginning to provide detailed information related to the molecular mechanisms and crosstalks of the signalling pathways involved in the different kinds of abiotic stress response.

Wheat, together with maize and rice, is the most important crop grown worldwide. Durum wheat is a minor crop, grown on only 8 to 10 % of all the wheat cultivated area, but it is better adapted to semiarid climates than bread wheat. Significant advances in understanding the molecular biology of this tetraploid species must be achieved to improve its tolerance to drought and other environmental insults and, actually, this PhD research work focuses on dehydration tolerance in durum wheat by means of functional genomics techniques.

A number of genes have been described that respond to water stress at the transcriptional level and several *cis*- and *trans*-acting factors involved in the expression of dehydration-responsive genes, such as the DRE element (Dehydration Responsive Element) and the DREB factors (DRE-Binding factors), have been extensively dissected. In a previous study, my collaborators at ENEA Research Center (Rome, Italy) and I have isolated a DREB2-related gene in durum wheat, designated as *TdDRF1* (*Triticum durum* Dehydration Responsive Factor 1), and established that it is expressed in response to dehydration. This gene produces three transcript variants, namely, *TdDRF1.1*, *TdDRF1.2* and *TdDRF1.3*, through alternative splicing. The transcript isoforms *TdDRF1.1* and *TdDRF1.3* encode AP2 transcriptional activators and *TdDRF1.2* encodes a putatively truncated protein, lacking the AP2 DNA-binding domain.

The current work is composed of three experimental sections. In the fist section, I report the analysis of the expression profiles of the three *TdDRF1* alternatively spliced transcripts upon dehydration, in four durum wheat and one triticale cultivars grown in greenhouse, and reveal that, even though these cultivars exhibit a very similar water retention in their leaves, their tolerance to dehydration may depend on a genotype-specific *TdDRF1* expression pattern together with many other additional genotype-specific traits. Furthermore, the genetic variability of *TdDRF1* sequences was explored among the five cultivars, in search of genotype-specific polymorphisms.

In the second section, I describe a notable experimental work, that is still in progress, for assessing the expression profile of the *TdDRF1* gene and its relation to the tolerance of durum wheat cultivars, grown under water stressed (reduced irrigation) and non-stressed (full irrigation) conditions at CIMMYT's experimental fields (Obregón, Mexico). Two groups of durum cultivars were chosen for this scope: one group exhibiting good drought tolerance and the other one exhibiting strong drought susceptibility. The *TdDRF1* transcripts expression levels were monitored by real-time RT-PCR and results of one tolerant (Duilio) and one susceptible (Creso) cultivars are reported.

While studying the *TdDRF1* transcripts expression profile, dependent on the particular plant water status and on the genotype, many efforts have been concentrated to attempt clarify the function of the three gene products, for which no direct information is available in literature. In the third section, I describe an approach towards the investigation of the function of all the three gene products, using a heterologous overexpression system PVX (Potato Virus X)-mediated. Initially, the complete codifying sequences (CDSs) of the three *TdDRF1* transcripts were isolated and cloned and, after that, the PVX-derived constructs for the transient overexpression in *Nicotiana benthamiana* and *Nicotiana tabacum* were engineered. An increased mRNA expression level, in comparison to wild type plants, was detected in all inoculated plants. The immediate next step will be the protein analysis with the double aim of isolating the target proteins produced in plant and investigating their molecular partners in the networks involved in the dehydration stress signalling and response.

All the information and experimental data reported in this PhD work, as basic research, will be useful to enrich the actual knowledge about the molecular basis of drought stress and, possibly, they will draw together to a future application in molecular assisted breeding.

# RIASSUNTO

Condizioni ambientali avverse, come la siccità, il freddo e l'elevata salinità, esercitano un'influenza negativa sulla produttività agricola. Lo studio dei meccanismi coinvolti nella risposta e nella tolleranza della pianta a questi stress rappresenta una grande sfida per i ricercatori vegetali, soprattutto in visione dei cambiamenti climatici globali attesi. La ricerca genomica delle piante sta iniziando a fornire informazioni dettagliate sui meccanismi molecolari e le interazioni tra le vie segnaletiche coinvolte nei differenti tipi di risposta agli stress abiotici.

Il grano, insieme al mais ed al riso, è il cereale più coltivato al mondo. Il grano duro è una coltivazione minore che rappresenta solo l'8-10 % di tutta l'area coltivato a grano, ma si adatta meglio ai climi semiaridi rispetto al grano tenero. Al fine di migliorare la sua tolleranza all'aridità e agli altri insulti ambientali, è necessario che vi siano avanzamenti significativi nella comprensione della biologia molecolare di questa specie tetraploide e, a questo proposito, il mio dottorato di ricerca si focalizza sullo studio della tolleranza alla disidratazione in grano duro per mezzo di tecniche di genomica funzionale.

Sono stati descritti numerosi geni che rispondono allo stress idrico a livello trascrizionale e sono stati analizzati estensivamente dei fattori *cis*- e *trans*-agenti coinvolti nell'espressione dipendente da disidratazione, come l'elemento DRE (Elemento che Risponde alla Disidratazione) ed i fattori DREB. Durante uno studio precedente, i miei collaboratori presso il Centro Ricerche ENEA-Casaccia (Roma) ed io abbiamo isolato in grano duro un gene del tipo DREB2, chiamato *TdDRF1* (Fattore 1 che Risponde alla Disidratazione in *Triticum durum*), e stabilito che la sua espressione viene modulata in condizioni di stress idrico. Questo gene produce tre tipi di trascritti per splicing alternativo: *TdDRF1.1*, *TdDRF1.2* e *TdDRF1.3*. I trascritti *TdDRF1.1* e *TdDRF1.3* codificano per attivatori trascrizionali di tipo AP2, mentre il trascritto *TdDRF1.2* codifica per una putativa proteina tronca priva del dominio di legame al DNA di tipo AP2.

Questa Tesi di Dottorato è composta di tre sezioni sperimentali. Nella prima sezione riporto l'analisi dei profili di espressione dei tre trascritti alternativi del gene *TdDRF1* durante disidratazione, in condizioni controllate di serra, di quattro coltivazioni di grano duro ed una di triticale. I risultati rivelano che, nonostante questi genotipi trattengano la stessa quantità d'acqua nei tessuti e pertanto siano similmente tolleranti, la loro tolleranza alla disidratazione può dipendere da un pattern di espressione genotipo-specifico del gene *TdDRF1* oltre che da ulteriori caratteristiche specifiche sempre relative al genotipo. Inoltre è stata esaminata la variabilità genetica presente nelle sequenze

di *TdDRF1* delle cinque varietà di frumento col fine di evidenziare dei polimorfismi genici genotipo-specifici.

Nella seconda sezione descrivo un lavoro sperimentale di rilievo, ancora in corso, concepito per valutare il profilo di espressione del gene *TdDRF1* e la sua relazione con la tolleranza alla siccità nelle coltivazioni di grano duro, cresciute sotto condizioni ottimali (irrigazione completa) e condizioni di stress idrico (irrigazione ridotta), presso i campi sperimentali del CIMMYT (Obregón, Messico). Due gruppi di genotipi di grano duro sono stati selezionati: un gruppo di arido-tolleranti ed un gruppo di arido-suscettibili. I livelli di espressione dei trascritti di *TdDRF1* sono stati monitorati via *real-time* RT-PCR e vengono presentati i risultati conseguiti per una varietà tollerante (Duilio) ed una suscettibile (Creso).

Parallelamente allo studio del profilo di espressione dei trascritti del gene *TdDRF1*, dipendente dal particolare stato idrico della pianta e dal genotipo, sono stati eseguiti degli esperimenti indirizzati al chiarimento della funzione di questi tre trascritti, dei quali in letteratura non è disponibile alcuna informazione diretta. In particolare, nella terza sezione, descrivo un approccio rilevare la funzione molecolare dei tre trascritti genici, avvalendosi di un sistema eterologo di over-espressione mediato dal virus X della patata (PVX). Il primo passo è stato l'isolamento ed il clonaggio delle sequenze codificanti complete (CDSs) per i tre trascritti *TdDRF1*; successivamente, sono stati ingegnerizzati i costrutti derivati dal PVX per l'over-espressione transiente in piante di *Nicotiana benthamiana* e *Nicotiana tabacum*. In tutte le piante infettate è stato rilevato un aumento degli mRNA dei trascritti, rispetto alle piante *wild-type*. Il passo immediatamente successivo sarà l'analisi delle proteine con il doppio scopo di isolare le proteine bersaglio prodotte in pianta e di determinare le molecole con cui queste interagiscono nelle reti coinvolte nella segnalazione dello stress idrico e nella conseguente riposta della pianta.

Tutte queste informazioni ed i dati sperimentali riportati in questa Tesi di Dottorato sono utili come ricerca di base, arricchendo le attuali conoscenze sui meccanismi molecolari dello stress idrico e, possibilmente, gettano le basi per una futura applicazione nel miglioramento genetico assistito da marcatori molecolari.

# LITERATURE PRODUCTION

**Latini A**, Sperandei M, Sharama S, Cantale C, Iannetta M, Dettori M, Ammar K, Galeffi P (in press). Molecular analyses of a dehydration related gene from DREB family in durum wheat and Triticale. In *Biosaline Agriculture and High Salinity Tolerance in Plants.* Published by Birkhäuser and Verlag.

**Latini A**, Rasi C, Sperandei M, Cantale C, Iannetta M, Dettori M, Ammar K & Galeffi P (2007). Identification of a DREB-related gene in *Triticum durum* and its expression under water stress conditions. *Annals of Applied Biology* **150**, 187-195.

# POSTER PRESENTATIONS AT CONGRESSES

Thiyagarajan K, Tesfay E, **Latini A**: Identification of the homeologous, drought-related *TdDRF1* gene in wild ancestor wheats, *Aegilops speltoides* and *Triticum urartu*. National Congress "*Società Italina di Genetica Agraria (SIGA) LI Annual Congress*", September 23-26, 2007. Riva del Garda (Italy).

Lucretti S, Cantale C, De Vita P, Fiocchetti F, Galeffi P, Giuliano G, **Latini A**, Nardi L, Roselli M, Sperandei M, Turchi L: Nuove risorse genetiche per decifrare la struttura del genoma del grano duro – Genetica e genomica della risposta di Triticum durum allo stress idrico e salino. National Congress "*Frumisis – Nuove tecnologie genetiche per compensare gli effetti dei cambiamenti climatici sulla coltura grano duro nei paesi mediterranei*", June 12, 2007. Rome (Italy).

**Latini A**, Sperandei M, Rasi C, Cavicchioni G, Cantale C, Iannetta M, Dettori M, Ammar K, Galeffi P: Molecular analyses of a DREB-related gene in durum wheat. International Congress "*ICBA&HST*", November 3, 2006. Tunis (Tunisia).

**Latini A**, Rasi C, Sperandei M, Cavicchioni G, Sharama S, Cantale C, Iannetta M, Dettori M, Ammar K, Galeffi P: Expression profile analyses of a dehydration-responsive transcription factor gene in durum wheat. International Congress "*International Scientific Conference: The Future of Drylands*", June 19-21, 2006. Tunis (Tunisia).

**Latini A**, Sharma S, Sperandei M, Cantale C, Dettori M, Ammar K, Galeffi P: Molecular responses in controlled field conditions: expression of a DREB-related gene in durum wheat and triticale. International Congress "*5rd Plant GEMs (Genomics European Meetings)*", October 11-14, 2006. Venis (Italy).

Galeffi P, **Latini A**, Rasi C, Sperandei M, Cavicchioni G, Palmieri E, Cantale C, Iannetta M, Dettori M, Pfeiffer W, Pellegrineschi A, Ammar K: Molecular responses to drought stress: expression of TdDRF1 gene in several durum wheat varieties in controlled greenhouse and field conditions. International Congress "*INTERDROUGHT-II. 2nd International Conference on Integrated Approaches to Sustain and Improve Plant Production Under Drought Stress*", September 24-28, 2005. Rome (Italy).

**Latini A**, Rasi C, Cavicchioni G, Chiaretti D, Sperandei M, Cantale C, Iannetta M, Dettori M, Ammar K., Galeffi P: Expression study of a gene for a dehydration-responsive transcription factor in durum wheat. National Congress "*Società Italiana di Genetica Agraria (SIGA) XLIX Annual Congress*", September 12-15, 2005. Potenza (Italy).

# ACKNOWLEDGEMENTS

This PhD thesis is dedicated to my family, to honour their love, patience and never-ending support.

First of all I would like to acknowledge my research group at ENEA-Casaccia RC (Rome). A few lines are too short to make a complete account of my deep appreciation for my PhD advisor, Dr. Maria Sperandei, and the leadership of my laboratory team, Dr. Patrizia Galeffi. I am also very grateful to Dr. Cristina Cantale. I wish to express my gratitude to them for their continued encouragement and invaluable scientific suggestions during this work. Many thanks for the large amount of time they spent teaching me how to develop my experiments in an efficient way. Moreover, I wish to thank Dr. Caterina Arcangeli for her precious collaboration on the homology modelling work as well as for her friendship.

I want to thank Dr. Karim Ammar, for his relevant support during this work and for having accorded to me the opportunity to work at CIMMYT laboratories (Mexico). I will always treasure his helpful suggestions and precious scientific advice. Special thanks to all the CIMMYT researchers and technical staff of El Batán (Mexico D.F) and Obregón (Sonora).

Thanks to Mrs. Giovanna Cavicchioni and Mrs. Eleonora Palmieri for the greenhouse work and to Mr. Domenico Chiaretti for the field trial at ENEA-Casaccia.

Thanks to the PhD coordinator, Prof. Giorgio Prantera, for his interest in my research.

Special thanks to Dr. Emanuele Tosi for his precious help in real-time PCR experiments.

Thanks to Dr. Kazuo Shinozaki for inspiring this work.

Thanks to Dr. Massimo Iannetta, Dr. Vittorio Rosato and Dr. Massimo Cristofaro for their economical support.

Thanks to Dr. Marco Dettori for his useful agronomic suggestions about the field experiments.

Many thanks to the “transient” people at my laboratory.

Last but not least, lots of thanks to ALL my friends!

# FUNDINGS

This work was partially supported by the Italian RIADE (Integrated Research for Applying new technologies and processes for combating DEsertification) project; FRUMISIS (MIPAF); fellowships of the "Secretaría de Relaciones Exteriores" offered by the Mexican Government for my training stages at CIMMYT (Mexico); IMAGE Project (CNR) and FIRB-adart (MIUR).