UNIVERSITÀ DI PARMA

UNIVERSITA' DEGLI STUDI DI PARMA

Ph.D in Biotechnology and Life Sciences

XXXI cycle

Studying durum wheat genetic diversity:
A molecular approach to identify new alleles for drought resilience

Coordinator

Prof. Simone Ottonello

Supervisor:

Prof. Nelson Marmiroli

Prof. Mariolina Gullì

Dr. Michela Janni

Ph.D candidate: Valentina Buffagni

2015/2018

Index

List of abbreviation

ABA: Abscisic Acid

ABRE: ABA-Responsive Elements

AP2/ERF: APETALA2/Ethylene Response Element Binding Factors

AR5: fifth IPCC Assessment Report

AR6 sixth IPCC Assessment Report

AREB/ABF: ABA-Responsive Element-Binding Protein/ABA-Binding Factor

BAC: Bacterial Artificial Chromosome

BCAA: Branched-Chain Amino Acids

BCAT: Branched-Chain Amino Acid Aminotransferase

CDS: Coding Sequence

DAS Days After Sowing

das: Days After Stress

DBR: Digital Biovolume ratio

DI: Dry Index

DIN: Dry Index Non-stress plants

DIS: Dry Index of Stressed plants

DREB: Dehydration-Responsive Element Binding Protein

DS: Drought Stress

EMS: Ethyl Methane Sulphonate

FAO: Food and Agriculture Organization

FM: Functional Marker

FRET: Fluorescence Resonance Energy Transfer

GHG: Greenhouse Gas

GMO: Genetically Modified Organism

GOI: Gene of Interest

HRM: High-Resolution Melt

HTP: High-Throughput Phenotyping

HTPP: High-Throughput Phenotyping Platform

IPCC: Intergovernmental Panel on Climate Change

KASP: Kompetitive Allele Specific PCR

LEA: Late Embryogenesis Abundant

MAS: Marker-Assisted Selection

NGS: Next-Generation Sequencing

NIR: Near Infra-Red

RCP: Representative Concentration Pathways

RDI: Relative Dry Index

RWC: Relative Water Content

SNP: Single Nucleotide Polymorphism

TILLING: Targeting Induced Local Lesions IN Genomes

TSS: Transcription Starting Site

UCS: Union of Concerned Scientists

UN: United Nations

UNFCCC: United Nations Framework Convention on Climate Change

ABSTRACT

Mediterranean area is one of the main region affected by the increasing drought and heat-wave expected by the impact of global climate change (GCC), with strong drawbacks on agricultural productions and thus on food security.

Durum wheat (*Triticum turgidum* ssp. *durum*) is a rain-fed crop mainly cultivated in the Mediterranean basin, which despite being able to survive various environmental stresses, it is threatened by the impending GCC. In particular drought stress (DS) is one of the major constraints that can negatively affect crops growth and yield.

Identifying genotypes with higher adaptability to DS is thus a primary goal for researchers and breeders.

In this research, we studied genes known to be involved in DS response in other crop species, and characterised them in *T. durum* at genomic and expression level. We isolated the A and B homoeologs of two genes: i) *TdBCAT*, a putative chloroplast branched-chain aminotransferase (BCAT); ii) *TdABA7* coding for an 18 kDa protein which has still an unknown function. The expression analyses following the DS revealed a clear involvement of *TdBCAT* during the anthesis and post-anthesis events, while for *TdABA7* a major response was related to the germination stage. Moreover, both sequences had a tissue-specific and homoeologs-specific expression.

To elucidate the involvement of *TdBCAT* and *TdABA7* in DS response, a reverse genetic strategy was followed taking advantage of the wheat TILLING resource (www.wheat-tilling.com), thanks to the collaboration with the John Innes Centre (Norwich, England, UK) which lead to the generation of 11 and 15 double mutant lines for *TdBCAT* and *TdABA7* respectively. In parallel, according to the climate smart agriculture paradigm suggested by the Food and Agriculture Organization (FAO) to exploit the natural variation occurring in the durum wheat genetic resources, an allele mining approach has been performed to identify the presence of natural variation of the selected genes within a durum wheat collection of Single Seed Descent (SSD) lines representative of the Mediterranean cultivation area. The method was implemented in collaboration with the John Innes Centre, and it consisted of an innovative cost-effective targeted-resequencing approach by Next Generation Sequencing (NGS) coupled with the KASP assay to assess the presence of any identified SNP. By doing this, a total of 21 SNP in *TdBCAT* gene promoters were discovered in the SSDs under analysis, which lead to the discovery of 5 different haplotypes in both A and B genes, of which some uniquely represented. To link the genotype to the phenotype, the same subset of SSD genotypes were analysed in DS condition by high throughput phenotyping and the digital biovolume and dry index indices were analysed. This allows the identification of five SSD lines as highly resilient to drought and good candidates for wheat breeding.

RIASSUNTO

La regione mediterranea è una delle aree climatiche maggiormente colpite dalle crescenti ondate di caldo e siccità causate dal cambiamento climatico globale, con gravi ripercussioni sulla produzione agricola e quindi sulla sicurezza alimentare.

Il frumento duro (*Triticum turgidum* ssp. *durum*) è coltivato principalmente nel bacino Mediterraneo, tradizionalmente senza l'ausilio d'irrigazione artificiale. Nonostante la naturale capacità di questa coltura di resistere a diversi stress ambientali, essa è attualmente minacciata dall'aggravarsi dei cambiamenti climatici. In particolare lo stress idrico da siccità che le colture subiscono in mancanza di acqua, è uno dei fattori ambientali di maggior impatto sulla crescita e le rese produttive finali per il frumento duro.

L'identificazione di genotipi con una maggior adattabilità allo stress da siccità è quindi un obiettivo primario sia per i ricercatori sia per i *breeder* che lavorano nel campo del miglioramento genetico del frumento.

In questa ricerca, ci siamo focalizzati su geni noti per essere coinvolti nella risposta allo stress idrico in altre specie cerealicole, i quali sono stati quindi caratterizzati in fumento duro sia a livello genomico che trascrizionale. Sono stati isolati entrambe le copie omeologhe (A e B) di due geni: 1) *TdBCAT*, un'aminotransferasi degli amino acidi a catena ramificata con putativa funzione plastidica; 2) *TdABA7*, codificante per una proteina a basso peso molecolare (18kDa), la cui funzione molecolare e biologica è tuttora ignota. L'analisi di espressione in condizioni di stress idrico ha rivelato un chiaro coinvolgimento di *TdBCAT* durante l'antesi e i primi eventi di post-antesi, mentre per *TdABA7*, l'espressione in risposta allo stress sembra essere più relazionata alla fase di germinazione. Inoltre entrambi i geni presentano un'espressione tessuto e omeologo specifica.

Per capire meglio il ruolo di *TdBCAT* e *TdABA7* nella risposta allo stress idrico da siccità, abbiamo seguito un approccio di genetica inversa, sfruttanto uno strumento bioinformatico chiamato *in silico* wheat TILLING (www.wheat-tilling.com) grazie al quale è stato possibile generare 11 linee doppio mutanti per *TdBCAT* e 15 linee doppio mutanti per *TdABA7*. In parallelo, in linea con il paradigma di *"climate smart agriculture"* suggerito dalla FAO di sfruttare la variabilità naturale presente nelle risorse genetiche di frumento duro, un approccio di *allele mining* ha permesso di identificare la presenza di varianti alleliche nei geni selezionati analizzando una collezione di germoplasma di frumento rappresentativa delle colture di frumento del bacino mediterraneo. E' stato quindi definito un metodo innovativo ma soprattutto economico, che sfrutta l'unione di diverse tecniche (sequenziamento di zone target mediante NGS e tecnologia KASP) per scoprire varianti alleliche naturali nei geni in esame. In particolare, quest'approccio ha permesso di identificare 21 SNP nei promotori di *TdBCAT* nel set di genotipi di frumento duro analizzati. Quest'analisi ha quindi portato all'identificazione di 5 aplotipi per ognuno delle due copie omeoologhe del gene *TdBCAT*, alcuni dei quali rappresentati in un singolo

genotipo. Per correlare il genotipo al fenotipo, lo stesso set di genotipi è stato analizzato in condizioni di stress idrico attraverso una piattaforma di fenotipizzazione high-throughput. I due indici fenotipi considerati (digital biovolume e dry index), ha permesso l'identificazione di 5 genotipi come altamente resistenti allo stress da siccità e quindi a identificarli come ottimi candidati per futuri programmi di *breeding.*

Chapter 1 INTRODUCTION

1.1 Climate Change impact on Global Food Security

1.1.1 The climate system

It is a common mistake to confuse climate with weather. Climate refers to the conditions of the earth's lower surface atmosphere at a specific location, while weather refers to a day-to-day fluctuation of these conditions at the same location. Meteorologists use common variables to measure daily weather phenomena. The systematic measuring of the weather phenomena all around the world over several years, leads to a record of observations that can be computed and used to describe the global climate. According to the World Meteorological Organization, climate is a statistical description in terms of the mean and variability of relevant quantities over a period of 30 years. These quantities are most often surface variables such as temperature, precipitation, and wind. In a wider sense we can consider climate as the statistical description of the climate system state. The climate system (Figure 1.1) is a highly complex network of five major interacting components: the atmosphere, the hydrosphere, the cryosphere, the land surface and the biosphere. This system can evolve in time under the influence of its own internal dynamics and because of external natural forcing (e.g. volcanic eruptions, solar variations) and human-induced forcing such as the changing composition of the atmosphere or the land-usage (http://www.ipcc.ch/ipccreports/tar/wg1/040.htm).

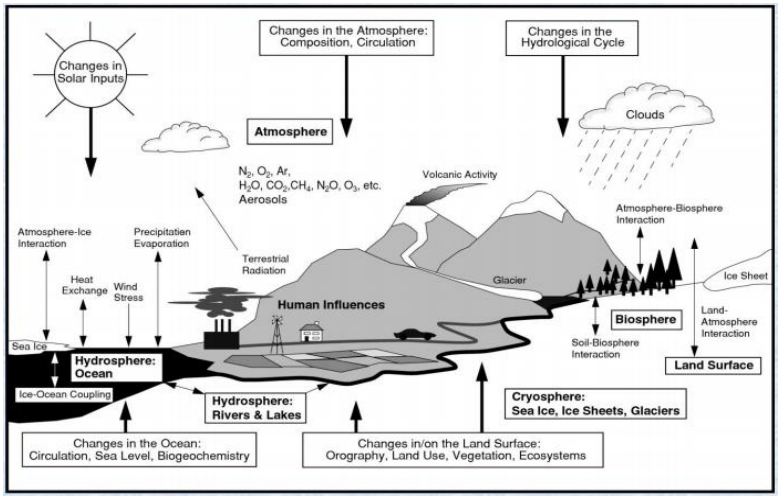


Figure 1.1 Representation of the components of the global climate system. Changes predicted in the early years coming are in bold while their processes and interactions are represented with thin arrows and some aspects that may change (reproduced from http://www.ipcc.ch/ipccreports/tar/wg1/040.htm**)**

1.1.2 Climate change: a public major issue

At the end of 2017 the Union of Concerned Scientists (UCS) signed a document in which they declare that "a great change in our stewardship of the Earth and the life on it, is required, if vast human misery is to be avoided". They expressed concern about current or potential damage on planet Earth, including freshwater availability, ozone depletion, forest loss, biodiversity destruction, the increasing human population growth and climate change (Ripple et al., 2017). A huge number of scientific organizations are actually interested on describing the climate dynamics and its consequences on life on this planet (http://www.opr.ca.gov/facts/list-of-scientific-organizations.html).

Today define what "c*limate change*" means still generate a lot of debates, especially in the public opinion. According to the United Nations Framework Convention on Climate Change (UNFCCC) we should refer to climate changes as only human-induced changes in the climate system. The UNFCCC is an international membership that opened for signature at the Earth Summit in Rio de Janeiro in June 1992 and entered into force on 21 March 1994 for preventing dangerous human interference with the climate system (web archieve: unfccc.int/essential_background/convention/items/6036.php). They recognize the effect of greenhouse gas (GHG) emission such as carbon dioxide, as a primary cause of the actual global warming. UNFCCC is responsible of the Kyoto Protocol adoption in 1997 and in 2015, as a proof that climate change should be considered as a major issue daring our existence on this planet, it drafted "The Paris Agreement", which deals GHG emissions, mitigations policy and finance (https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement). The Paris Climate Agreement is a pact between nearly 200 nations to voluntarily reduce their GHG emissions in an effort to fight climate change. The Paris Agreement's long-term goal is to keep the increase in global average temperature to well below 2 °C above pre-industrial levels; and to limit the increase to 1.5 °C, as this would substantially reduce the risks and effects of climate change.

UNFCCC makes a clear distinction between "climate change" attributable to human activities and "climate variability" attributable to natural causes. The public opinion frequently disagree about the attributions of human and natural causes to climate changes, as it is generally contested the hardness to make a clear distinguish between them. In 2013, a nationally representative survey of the Americans public opinion stated that 41% of Americans say a global warming is happening and it is human-caused and that one third (33%) believe that "there is a lot of disagreement among scientists about whether or not global warming is happening" (Leiserowitz et al., 2013). By contrast, in the same year, an extended analysis of 12,000 peer-reviewed papers from 1991–2011 matching the topics 'global climate change' or 'global warming', show that 97% of the papers that took a stand on the reality of human-caused global warming, said that global warming is happening and is human-caused (Cook et al., 2013).

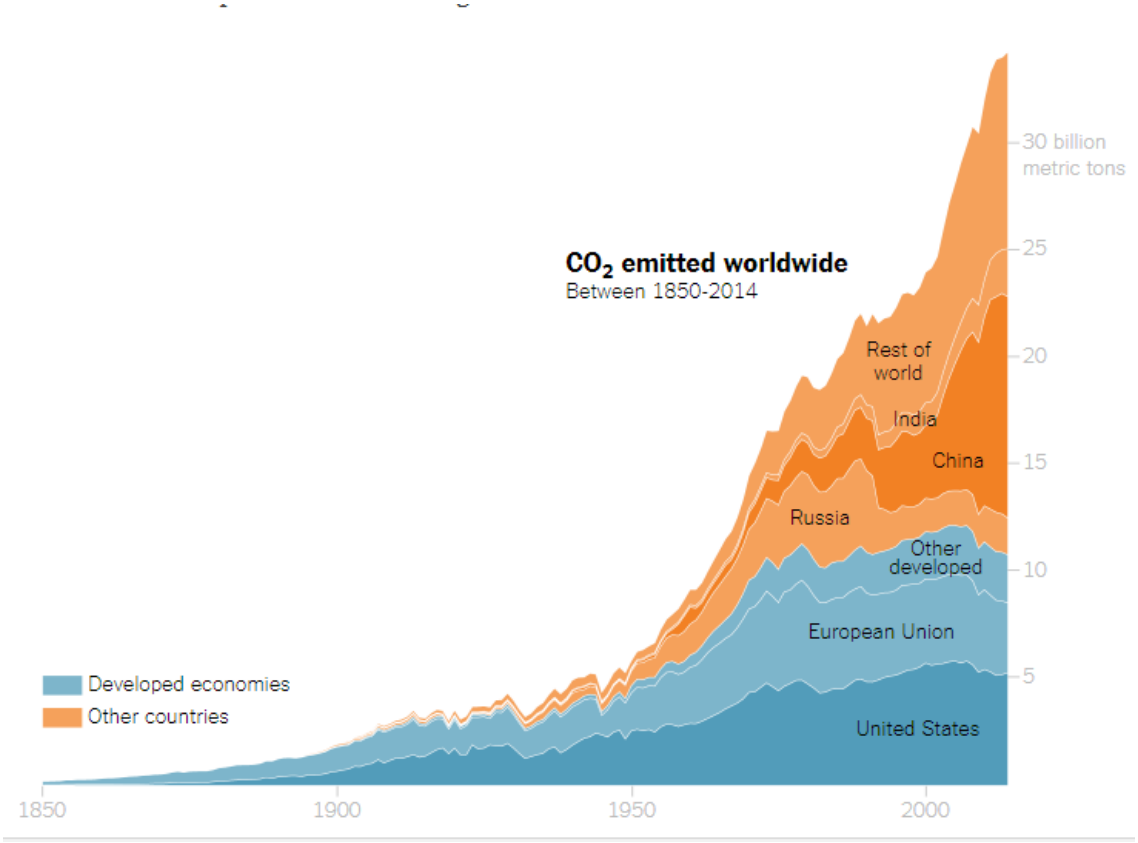


Figure 1.2 The global carbon dioxide emissions worldwide. With the advent of the industrial revolution, carbon dioxide emissions constantly increased in the western world (i.e. Europe and U.S.A) by the second half of 19th century. But after the Second World War, many other countries such as Russia rapidly followed the increasing trend of western countries. Nowadays developing countries like China or India registered a consistent boom of CO2 emissions, according to their rapid industrialization (reproduced from https://www.nytimes.com/interactive/2017/06/01/climate/us-biggest-carbon-polluter-in-history-will-it-walk-away-from-the-paris-climate-deal.html)

1.1.3 The Intergovernmental Panel on Climate Change

The Intergovernmental Panel on Climate Change (IPCC) is the leading international body for the assessment of climate change. Established in 1988 by the World Meteorological Organizaiton and the United Nations (UN) Environment Programme, the IPCC evaluates the state of climate science as a basis for informed policy actions and mitigation strategies, primarily on the basis of peer-reviewed and published scientific literature. The IPCC identifies where there is agreement or not in the scientific community and where further research is needed. It does not conduct its own research, but it assesses existing scientific knowledge on a politically sensitive and economically significant issue.

One of the main IPCC activities is the drafting of comprehensive Assessment Reports about the state of scientific, technical and socio-economic knowledge on climate change, its causes, potential impacts and response strategies. The IPCC drafted five Assessment Reports so far, by the cooperation of its three working groups constituting the scientific Commission: 1) one dealing with the physical science basis of climate change; 2) another dealing with impacts, adaptation and vulnerability; and 3) the last dealing with the mitigation of climate change. The next Assessment Report will be the sixth IPCC

Assessment (AR6) and it is expected to be published in 2021 (https://www.ipcc.ch/news_and_events/PR17-IPCC46_Press.shtml).

1.1.3.1 The fifth IPCC Assessment Report

In the fifth IPCC Assessment Report (AR5), human influences on climate system are clearly described. The presented scenario is far from being heartening: warmer atmosphere and ocean, diminished amounts of snow and ice, risen sea level, ocean acidifications, increasing extreme weather events. The main observed changes on climate are unprecedented and will have widespread impacts on human and natural system. The main driven force of climate changes in the 20th century has been the anthropogenic GHGs that have increased since the pre-industrial era and are now higher than ever (IPCC, 2014). Anthropogenic GHG emissions are driven by socioeconomic factors mainly related to population size, demand/offer supply chain, energy use and land use patterns, technology and climate policy.

By mid-21st century, the magnitude of the projected climate change is substantially affected by the choice of emissions scenario. The Representative Concentration Pathways (RCPs) are used to describe possible climate futures, by the use of factors like GHG emissions and atmospheric concentrations, air pollutant emissions and land use. Four different 21st century RCPs have been presented in this AR5: a stringent mitigation scenario (RCP2.6), two intermediate scenarios (RCP4.5 and RCP6.0) and one scenario with very high GHG emissions (RCP8.5). RCP2.6 represents the aim to keep global warming likely below 2°C above pre-industrial temperatures (IPCC, 2014). Nevertheless, surface temperature is projected to rise over the 21st century under all assessed emission scenarios (Figure 1.3a). It is very likely that heat waves will occur more often and last longer, and that extreme precipitation events will become more intense and frequent in many regions, while many others will experience a reduction of average precipitations. Indeed, changes in precipitation will not be uniform, as the high latitudes and the equatorial Pacific will experience an increase in annual mean precipitation, whereas in many mid-latitude and subtropical dry regions mean precipitation will likely decrease (Figure 1.3b).

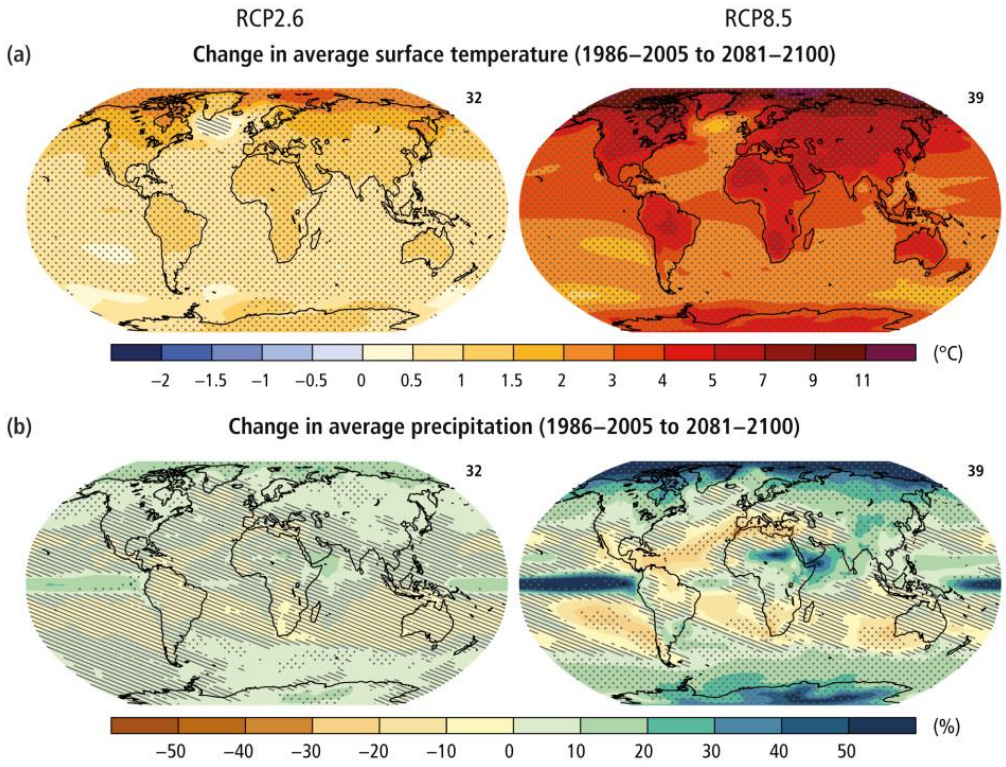


Figure 1.3 Change in (a) average surface temperature and (b) in average precipitation, based on multi-model mean projections for 2081–2100 relative to 1986-2005 under the RCP2.6 (left) and RCP8.5 (right) scenarios. The number of models used to calculate the multi-model mean is indicated in the upper right corner of each panel. Stippling (i.e., dots) shows regions where the projected change is large compared to natural internal variability, and where at least 90% of models agree on the sign of change. Hatching (i.e., diagonal lines) shows regions where the projected change is less than one standard deviation of the natural internal variability (reproduced from IPCC, 2014).

Impacts from recent climate-related extremes, such as heat waves, droughts, floods, cyclones and wildfires, reveal significant vulnerability and exposure of some ecosystems and many human systems to current climate variability (IPCC, 2014).

Many radical effects of climate change are already tangible and nevertheless the policy actions of the next years, it is a matter of fact that climate change will amplify existing risks and create new risks for natural and human systems during this century (IPCC, 2014).

1.1.4 The State of the World's Land and Water Resources for Food and Agriculture: Managing Systems at Risk

Water can be considered as a renewable resource because of the existence of the natural water cycle, which is strongly dependent on climate conditions. The water cycle is the process of continuous movement of water on, above and below the surface of the Earth and it is fundamental to determine the availability of water resources for human usage. Humans need a small amount of freshwater for basic needs (i.e. drinking and sanitation) and the fact that today these needs are not satisfied for many people is only a matter of access to, and quality of, the available water resources (Schewe et al., 2014). As Figure

1.4 shows, only 2.5% of the total is freshwater, which of about 30% is ground water, mostly used for agriculture (https://water.usgs.gov/edu/earthwherewater.html). Human activities affect the global water cycle by withdrawing water from different reservoirs for agricultural (70% of total water withdrawal), urban (11%), and industrial uses (19%) (FAO – AQUASTAT, 2014). Water withdrawals by different sectors can be extremely variable depending on the region of the world (Figure 1.4).

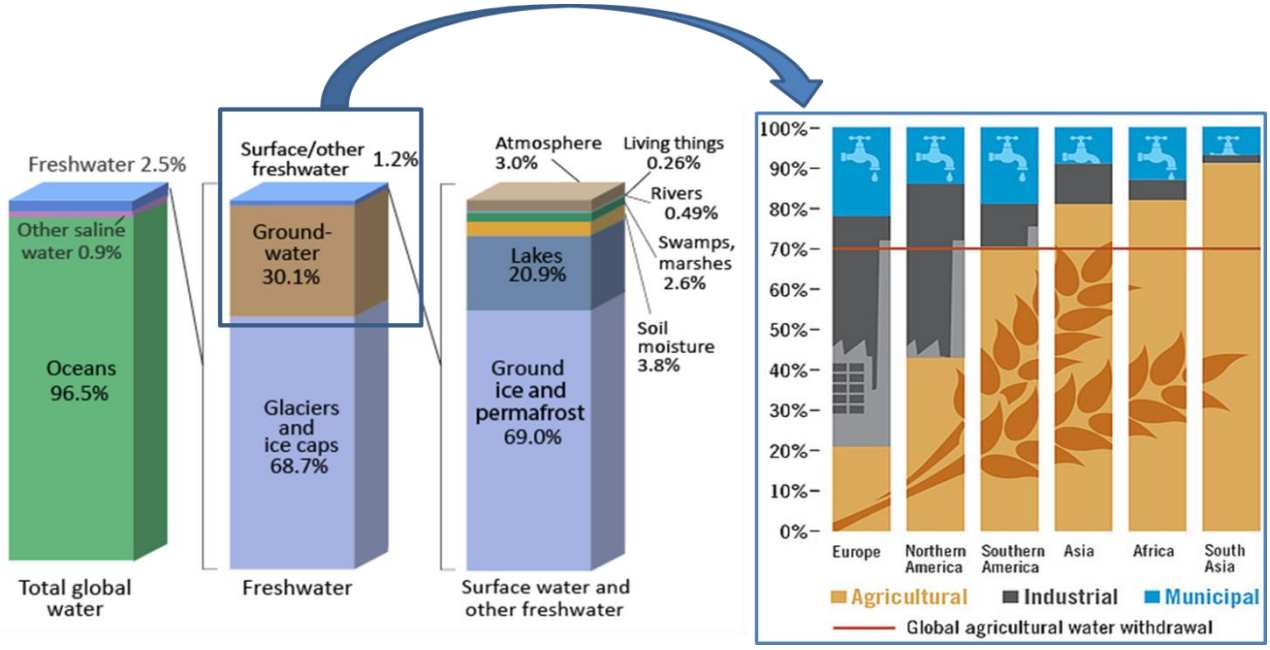


Figure 1.4 Distribution of water resources on the Earth (on the left) and freshwater withdrawals by sector (on the right). Data on water resources were reproduced and adapted from https://water.usgs.gov/edu/earthwherewater.html, while data on freshwater withdrawals are adapted from https://www.globalagriculture.org/report-topics/water.html, based on FAO-AQUASTAT 2014 data).

Agriculture is not only the largest consumer of the Earth's available freshwater, but even the largest user of habitable land. Indeed the agriculture accounts for 37.5% of the total land area which cannot be easily expanded. In Europe agricultural land (including natural grassland) accounts for almost half of the territory (FAOSTAT, 2016).

Agriculture actually provides more than enough calories for all people on the planet, but yet 800 million people remain undernourished, and approximately 2 billion suffer from micronutrient deficiencies (Ramankutty et al., 2018). Agriculture is also a major livelihood for 40% of the world's population, providing not only food but even fibre, biofuels, and many other products for the current human population. The actual 7.6 billion population is growing fast, and it's predicted to reach 9.8 billion by 2050 (UN World Population Prospect, 2017). As a direct consequence of this and the changing of habits of many developing countries, the food demand is increasing. To meet the growing demand, agriculture in 2050 will need to produce almost 50% more food, feed and biofuel than it did in 2012 (FAO, 2017).

It is in this context that the terms of water security and food security are closely intertwined.

Actually about 40% of total agricultural production relies on irrigation, while rain-fed cropland produce 60 to 70% of the global food (Lal, 2015; UNESCO, 2016). Without improved efficiency measures by 2050, the global agricultural water consumption (including both rain-fed and irrigated agriculture) is estimated to increase by a further 20% (UNESCO, 2016), thus an increase in irrigation water scarcity is expected (Haddeland et al., 2014). An increase in irrigation water scarcity means a likely decrease in agricultural food productivity; hence adaptation measures need to be addressed to increase food production and better water management.

Crop yield is affected by various agronomic and environmental factors, but water availability and temperature are the most critical. As we have already described, extreme weather conditions will be more frequent in the next decades. A recent analysis of the effects of droughts and extreme heat during 1964–2007, shows a significantly reduced national cereal production by 9–10% (Lesk et al., 2016). In this paper authors explain how production losses due to drought events were associated with a reduction in both harvested area and yields, whereas extreme heat mainly decreased cereal yields.

A "climate-smart agriculture" could help to sustainably increase food security and incomes, and to adapt and build resilience to climate change, while capturing potential mitigation co-benefits. It should connect several innovations, such as conservation agriculture, agroecology, agroforestry and the development of crop varieties that are more tolerant to pests, diseases, drought, waterlogging and salinity (FA0, 2017). Sustainability has become a key aspect of this century. On September 25th 2015, the UN nations published the 2030 Agenda as "(…) a plan of action for people, planet and prosperity to encourage policy for a sustainable development" (www.un.org/sustainabledevelopment). In Figure 1.5 the 17 goals of the Agenda 2030 are presented, a guide for the mitigation policy to undermine the current global-scale problems affecting humanity and eradicate poverty in all its forms and dimensions.

Figure 1.5 The 17 goals of the 2030 Agenda for Sustainable Development (reproduced from www.un.org/sustainabledevelopment

1.1.5 Europe under climate change: perspective on water availability and cereal production in the drying Mediterranean region

Our current understanding of European climate leads to projected temperature increase from 2 to 4°C and precipitation changes of 10 to −50% by the 2080s (Iglesias and Garrote, 2015). As IPCC already stated in the AR4 (IPCC, 2007), changes in both precipitation and temperature will not be equally distributed across different regions or seasons of Europe. Changes are likely to be more pronounced in Southern Europe than Northern Europe, with average temperature increases reaching +5 °C by 2080 and an alarming increase of extreme temperature (i.e. hot-wave). Northern Europe will experience an increase in annual mean precipitation while in the Mediterranean countries precipitations will decrease. In Figure 1.6 a comparison between the end of the last century and the projections for the end of the current one is presented.

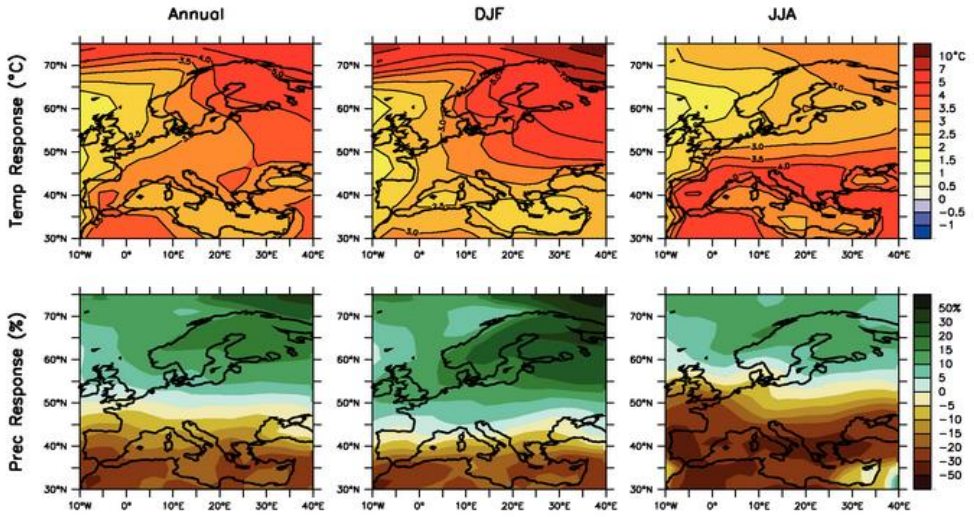


Figure 1.6 Temperature and precipitation changes over Europe. Top row: Annual mean, DJF (December, January, February) and JJA (June, July, August) temperature change between 1980 to 1999 and 2080 to 2099, averaged over 21 models. Bottom row: same as top, but for fractional change in precipitation (Reproduced from IPCC 2007).

Climate change will significantly reduce renewable surface water and groundwater resources and this is expected to exacerbate competition among the water users and sectors, i.e. agriculture, ecosystems, settlements, industry and energy producers (Döll et al., 2015). Meteorological (rainfall) and agricultural (soil moisture) droughts have already become more frequent since 1950 in Southern Europe (IPCC, 2014). Climate change will affect the near-surface soil moisture content as well as the evapotranspiration rates, increasing the occurrence of drought episodes with severe consequences for farming, natural ecosystems and forestry especially in Southern Europe (Iglesias and Garrote, 2015; Ruosteenoja et al., 2018).

The climate changes will bring inevitably to an increasing demand of water for agriculture in this area. Agriculture is nowadays a significant water user in Europe, accounting for around 21% of total water use (FAO–AQUASTAT 2017). There are considerable differences in the amounts of freshwater abstracted within each of the European Union (EU) Member States (Figure 1.7), partially reflecting the size of each country and the resources available, but also withdrawal practices, climate and the industrial and agricultural structure. Changes on freshwater availability can be natural or human related, but in the Mediterranean area where irrigation of crops accounts for virtually all agricultural water use, the climate impact dominates (Eulalia and Dessislava, 2015; Haddeland et al., 2014).

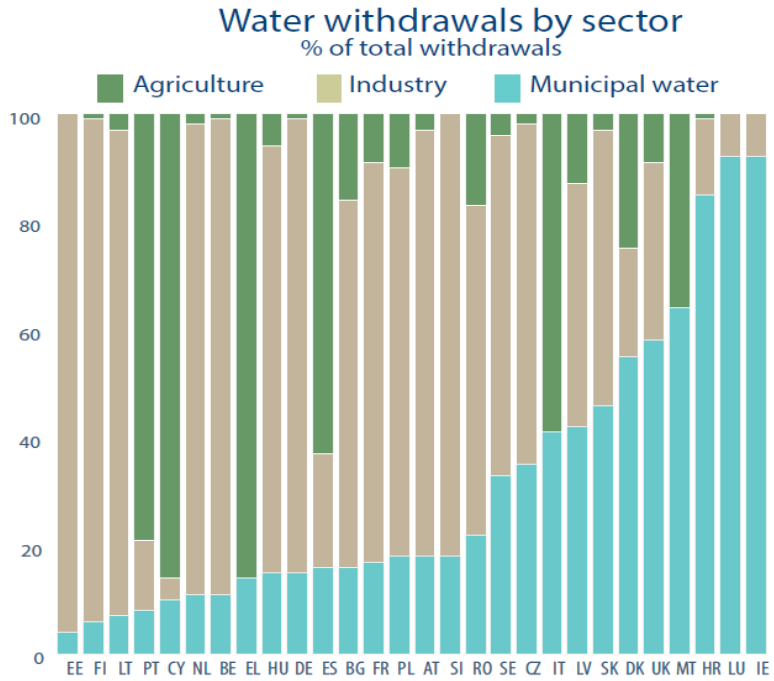


Figure 1.7 Water withdrawals by sector in EU-28 countries. Data are presented as a percentage of total withdrawals for the latest available FAO – AQUASTAT data resource (reproduced from www.europarl.europa.eu). Country codes: Austria (AT), Belgium (BE), Bulgaria (BG), Croatia (HR), Cyprus (CY), Czech Republic (CZ), Denmark (DK), Estonia (EE), Finland (FI), France (FR), Germany (DE), Greece (EL), Hungary (HU), Ireland (IE), Italy (IT), Latvia (LV), Lithuania (LT), Luxembourg (LU), Malta (MT), Netherlands (NL), Poland (PL), Portugal (PT), Romania (RO), Slovakia (SK), Slovenia (SI), Spain (ES), Sweden (SE), United Kingdom (UK). (Reproduced from Eulalia and Dessislava, 2015).

Actually cereals satisfy a large proportion of the global food demand, but to meet the needs of the growing population, cereal production needs to increase by at least 1.1% /year by 2050 (Peña-Bautista et al., 2017). In Europe, cereals account for 21% of total crop production (EUROSTAT 2016). For the major cereal crops such as wheat, rice and maize in temperate regions like the Mediterranean, climate change without adaptation strategies is projected to negatively impact both production and yield (Daryanto et al., 2017).

The projected climate trend could hence affect food security even in Europe by the end of this century. Many adaptation strategies have already been proposed for a better agricultural water management in this continent (Iglesias and Garrote, 2015) and among all, biotechnology has been proposed as an important player to support food security goals. Biotechnology can operate in the molecular genetics field, introducing climate change resilient crops to face the impending droughts events and the increasing irrigation requirements. Moreover Iglesias and Garrote (2015) suggested improving crop diversification to respond to the expected loss of biodiversity due to climate change.

Traditionally rainfed crops are the first target of the predicted water scarcity, which an example in Southern Europe is represented by durum wheat. Durum wheat is a crop cultivated only in arid and semiarid region where drought and high temperature often occur together at the end of the growing season, reducing crop yield potential of about 50% (Altenbach, 2012).

1.2 Wheat

Among the major cereal crops (i.e. maize, wheat and rice), the importance of wheat is undeniable. Indeed wheat can be considered the largest primary commodity worldwide, as it is grown on more land area than any other commercial crop (FAOSTAT, 2016). Wheat is our most important staple food grain, providing humans with 20% of the total calories, many essential nutrients and more protein than all types of meat combined (Uauy, 2017a).

The fame and fortune of wheat are not only related to its ability to grow almost everywhere on the globe, but also to the unique properties of the gluten protein fraction, which makes wheat the most versatile cereal for the preparation of diverse foods. Today there are two major cultivated wheat species: common wheat (*Triticum aestivum* L.) and durum wheat (*Triticum turgidum* L. subsp. *durum* Desf.), which differ from each other in genomic makeup, grain composition, processing quality, and main food uses (Peña-Bautista et al., 2017).

Common (or bread) wheat can be used for baked goods and noodles, while durum (or pasta) wheat is used to make semolina-based products (i.e. pasta, bulgur and couscous) accounting for 5% of the total wheat production (Shewry, 2009).

Whereas the uniqueness of this crop, it is not surprising that the global wheat demand is increasing worldwide (Shewry and Hey, 2015) By contrast, wheat production is expected to be strongly affected by climate changes such as the increasing frequency of days with extreme temperature together with the occurrence of drought, late spring frosts and severe winter frosts associated with inadequate snow cover. In addition, overly wet and cool weather enhances disease occurrence, contribute to lodging and complicate crop management increasing yield variability in places like European countries (Trnka et al., 2014).

1.2.1 Wheat cultivation in EU: an insight into the Italian durum wheat production

In 2016, the harvest of cereals in the EU decreased by about 4.4 % compared to the previous year (EUROSTAT 2017), which was largely explained by unfavourable climatic conditions (ec.europa.eu/eurostat/statistics-explained/index.php/Agricultural_production_-_crops#Cereals).

EU-28 is the first global producer for both common and durum wheat, accounting for about 20% and 25% respectively of global wheat production (International Grain Council 2015/2016 data, https://www.igc.int/en/markets/marketinfo-sd.aspx). While common wheat is cultivated in almost all European countries, durum wheat is mainly harvested in the Southern countries (Figure 1.8), and the top 4

producers are: Italy (56.8% of the total), France (18.2%), Spain (11.4%) and Greece (9.1%) (EUROSTAT 2016).

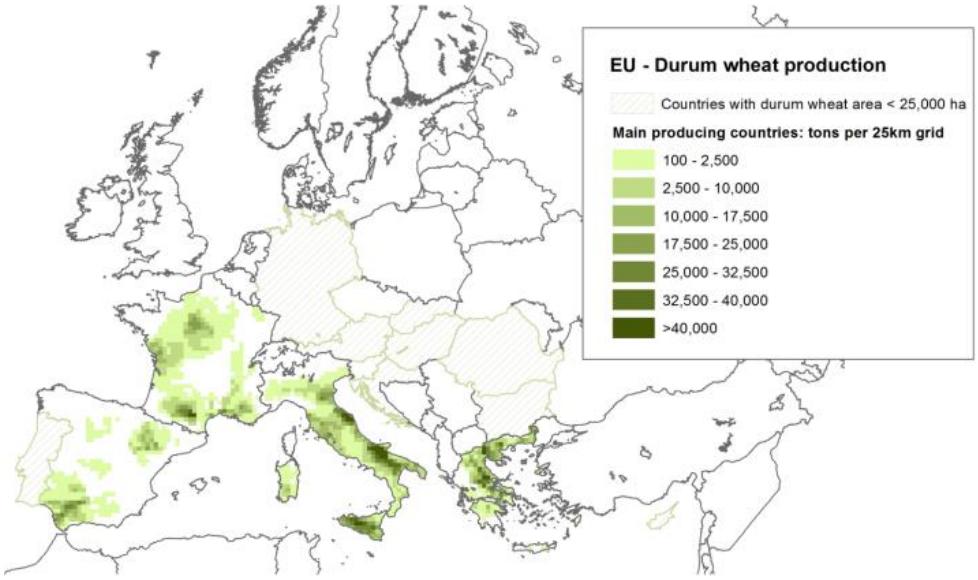


Figure 1.8 EU durum wheat production (reproduced from http://www.italmopa.com/wp-content/uploads/2017/05/144_all_2.pdf)

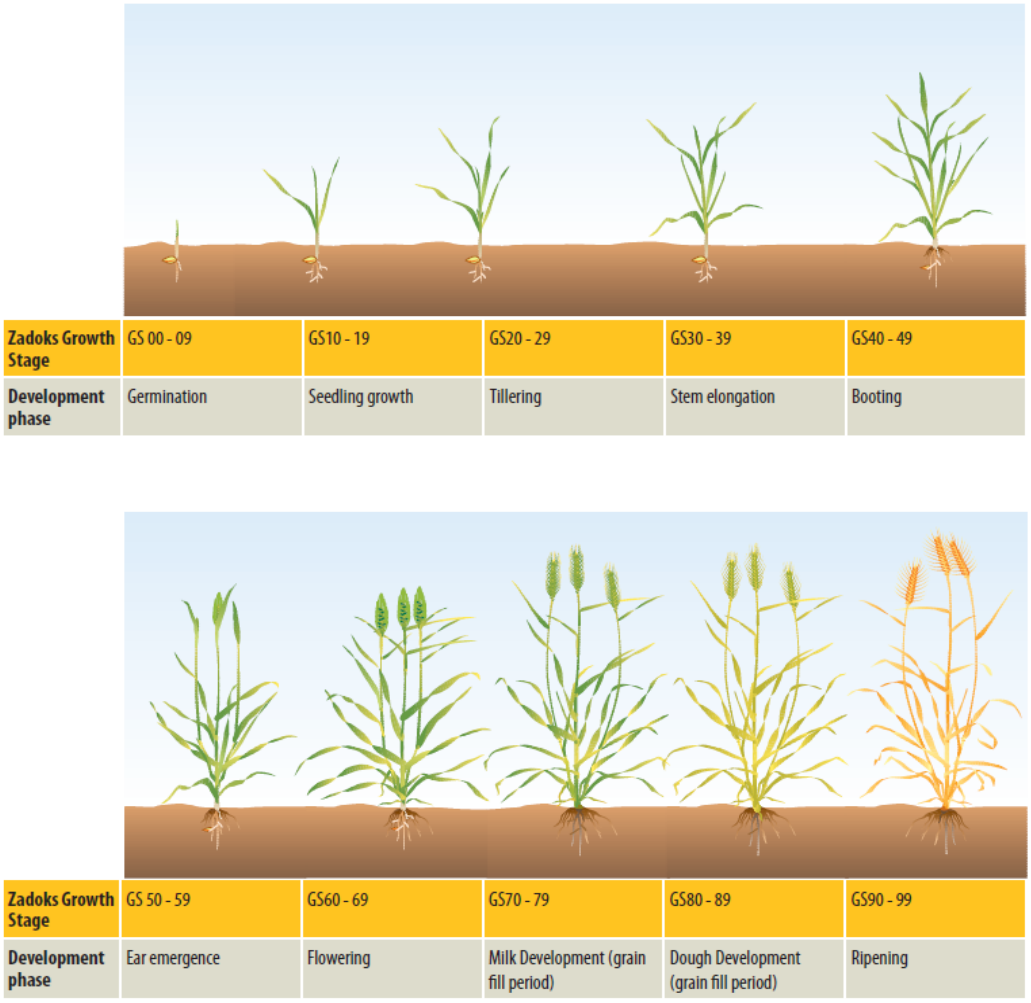
Italy is not only the first EU producer of durum wheat but also the top seller, accounting for 39% of the total exports of Europe (EUROSTAT, 2016). However, EU actually imports half of the total incomes from the first producer worldwide, Canada. Italy holds on strongly this market but nevertheless its strong production, it is responsible for the 80% of the total imports of Europe (EUROSTAT, 2016), which again mostly come from Canada.

1.2.2 Generalities about wheat: growth cycle and reproduction

Bread wheat and durum wheat are monocotyledonous plants of the *Poaceae* (or *Gramineae*) family, of the *Triticeae* tribe and belong to the genus *Triticum*. Wheat belongs to the same family that includes the major crop plants such as barley (*Hordeum* vulgare L.), oat (*Avena sativa* L.), rye (*Secale cereale* L.), maize (*Zea mays* L.) and rice (*Oryza sativa* L.). The *Triticeae* tribe, within the *Pooideae* subfamily, contains more than 15 genera and 300 species (Soreng et al., 2015).

The growth cycle of wheat is divided in different phases: germination, seedling establishment and leaf production, tillering and head differentiation, stem and head growth, head emergence and flowering, and grain filling and maturity (Figure **1.9**). Several systems have been developed to provide numerical designations for growth and developmental stages. The Feekes, Zadoks, and Haun scales are used the most frequently. The Zadoks Decimal Code (Zadoks et al., 1974) is internationally used to describe growth stages of cereals and it is the most widely accepted system for describing wheat development.

This system is based on two-digit code: the first digit refers to the principal stage of development beginning with germination (stage 0) and ending with kernel ripening (stage 9); the second digit, between 0 and 9 subdivides each principal growth stage. An adapted Zadoks Decimal Code (Stapper, 2007) is provided in the Annex (table 6.1).



Zadoks Growth Stage	GS 00 - 09	GS10 - 19	GS20 - 29	GS30 - 39	GS40 - 49
Development phase	Germination	Seedling growth	Tillering	Stem elongation	Booting

Zadoks Growth Stage	GS 50 - 59	GS60 - 69	GS70 - 79	GS80 - 89	GS90 - 99
Development phase	Ear emergence	Flowering	Milk Development (grain fill period)	Dough Development (grain fill period)	Ripening

Figure 1.9 Growth stages of wheat and Zadoks cereal growth stages (reproduced from GRDC Cereal Growth Stages Guide) (Crofts, 1989).

Generally referring to the sowing season, wheat can be classified into spring or winter wheat (Crofts, 1989): spring wheat is sown in the spring and matures in late summer, while winter wheat is sown before winter and thus seedlings (Zadoks code 10-19) are exposed to a prolonged cold period (vernalization) which is necessary to trigger reproductive stage in the early spring and to produce good yields. A definition by Crofts (1989) suggests that the distinction between winter and spring wheat is based on the presence or absence respectively, of genes controlling the vernalization requirement (*Vrn*).

In Italy, wheat is mainly sown in late-autumn and harvested in the early summer. The spring sowing is unusual for the Italian wheat crops, but it can be used as an emergency plan if the autumn sowing failed for any reason, like insufficient soil moisture. Flowering time (Zadoks code 60-69) is a spring event into

the Italian wheat fields and it is followed by the grain filling stage. Grain filling along with flowering stage, are most likely exposed to the possible drought and heat-wave event that we have registered so far.

Wheat is predominantly a self-pollinating species. Wheat breeders have a perfect knowledge of the biology of wheat reproduction, since they have to work in a narrow period of time which corresponds to few days between the emergence of spike and the anthesis event. During the anthesis, flowers generally remain closed (cleistogamous flowers), and the three anthers burst and release pollen directly on the stigma during the anthesis (Zadoks code 60-69). Once the anthers dehisce, wheat pollen generally remains viable for 15-30 minutes. Flowering generally begins in the basal florets just above the centre of the spike and proceeds towards the apex and base of the spike. Unfertilized florets usually open, exposing the receptive stigma to foreign pollen. The duration of wheat stigma receptivity depends on variety and environmental conditions, but ranges from 6 to 13 days (De Vries, 1971). Flowering of the entire spike can last for three to six days, depending on weather conditions. Once, fertilized, the ovary rapidly increases in size. Two to three weeks after fertilization, the embryo is physiologically functional, and able to generate a new seedling.

1.2.3 The origins and evolution of durum wheat

Wheat originated in the Fertile Crescent thousands of years ago alongside einkorn and barley, and nowadays it is grown in nearly all regions of the world on an area equivalent to that of Greenland (220 million hectares; FAOSTAT 2016) (Figure 1.10).

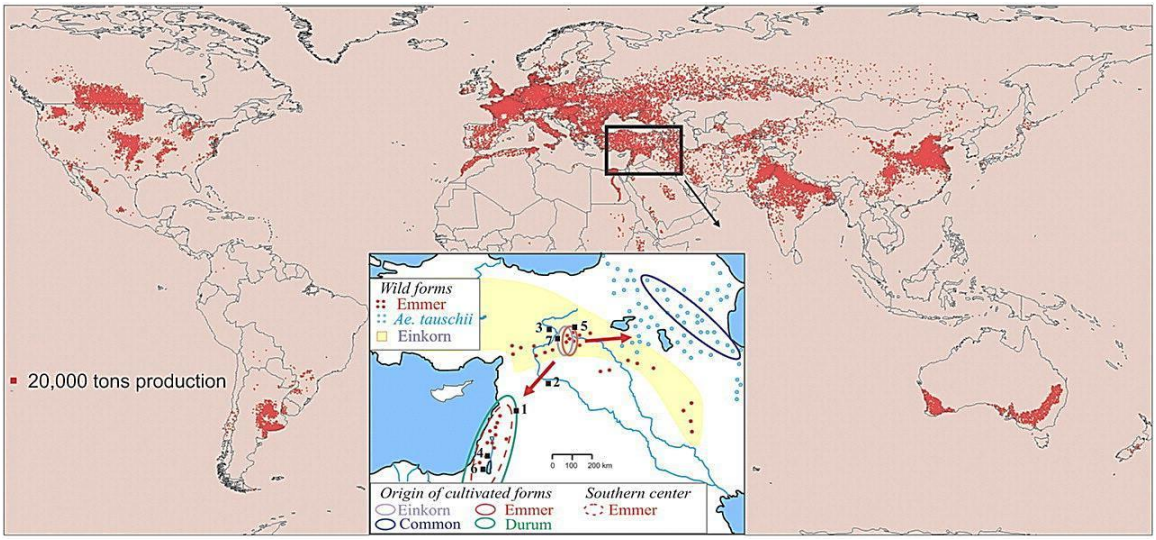


Figure 1.10 The origin and current distribution of wheat. In the insert the origin area is represented, in which the solid line ovals indicate the putative geographic regions of origin of the cultivated forms (einkorn, emmer, common and durum species), whereas the dotted red line indicates a southern center of domesticated emmer diversity. The approximate distributions of wild emmer and *Ae. tauschii* (a progenitor of wheat) are indicated by dots, and that of wild einkorn (the wild ancestor of durum wheat) by yellow shading. Numbers indicate archaeological sites where remains of domesticated cereals dating back more than 9000 years BP were found (reproduced from Dubcovsky and Dvorak, 2007)

Modern durum wheat (*Triticum turgidum* ssp. *durum*) evolved from its primitive ancestors both through natural events and through traditional and modern breeding technologies, to become the modern varieties we all eat today (Peña-Bautista et al., 2017). The history of its origin deals with the ability of this crop to grow in regions with temperate and dry climates, with hot days and cool nights during the growing season, typical of Mediterranean region.

Everything started in the "core-area" of the Fertile Crescent (southern-eastern Turkey to northern Syria), where the tetraploid durum wheat (2n = 4x = 28, AABB genome) evolved through a first polyploidization event between the donor of the A genome (*Triticum urartu*; 2n = 2x =14, AA genome) and the donor of the B genome (*Aegilops speltoides*-related species; 2n = 2x =14, BB genome) (El Baidouri et al., 2017). This event (0.5 million years ago) generated the tetraploid wild emmer ancestor (*Triticum turgidum* ssp. *diccocoides*; 2n = 4x = 28, AABB genome). The wild emmer afterwards became the modern durum wheat through a series of domestication events led by the very first farmers, which gave birth to the western agriculture about 12,000 years ago (Kilian et al., 2010). The earliest cultivated forms were essentially landraces, selected by ancient farmers from wild populations likely because of their superior yield (e.g. seed size) and other architectural and physiological traits (e.g. free-threshing state, plaint height, seed dormancy, heading date) (Brown, 2010; Kilian et al., 2010).

From the Fertile Crescent, farming spread throughout Europe, Asia and Africa supporting the cultivation of domesticated emmer wheat (*Triticum turgidum* spp. *dicoccum*) that followed human migration flows, and became the most important crop in the Fertile Crescent until the early Bronze Age, 10,000 years BC (El Baidouri et al., 2017). Three centre of diversification and domestication were proposed as the Fertile crescent, N. Africa and the highlands of Ethiopia (Kabbaj et al., 2017, Janni et al 2018 submitted) and support the hypothesis that wheat was likely brought to southern Italy from N. Africa (Oliveira et al., 2012).

Free-threshing tetraploid wheats (*Triticum turgidum* spp. *turgidum*) originated from the cultivated emmer (*Triticum turgidum* spp. *dicoccum*), furtherly selected to supply the born of durum wheat and its spread in the Mediterranean region as a crop specialized for the production of semolina-products. A further polyploidization event between the ancient emmer wheat (AABB genome) and *Aegilops tauschii* (DD genome) generated the hexaploid ancestor (AABBDD) of modern bread wheat about 8,000 years ago (Figure 1.11).

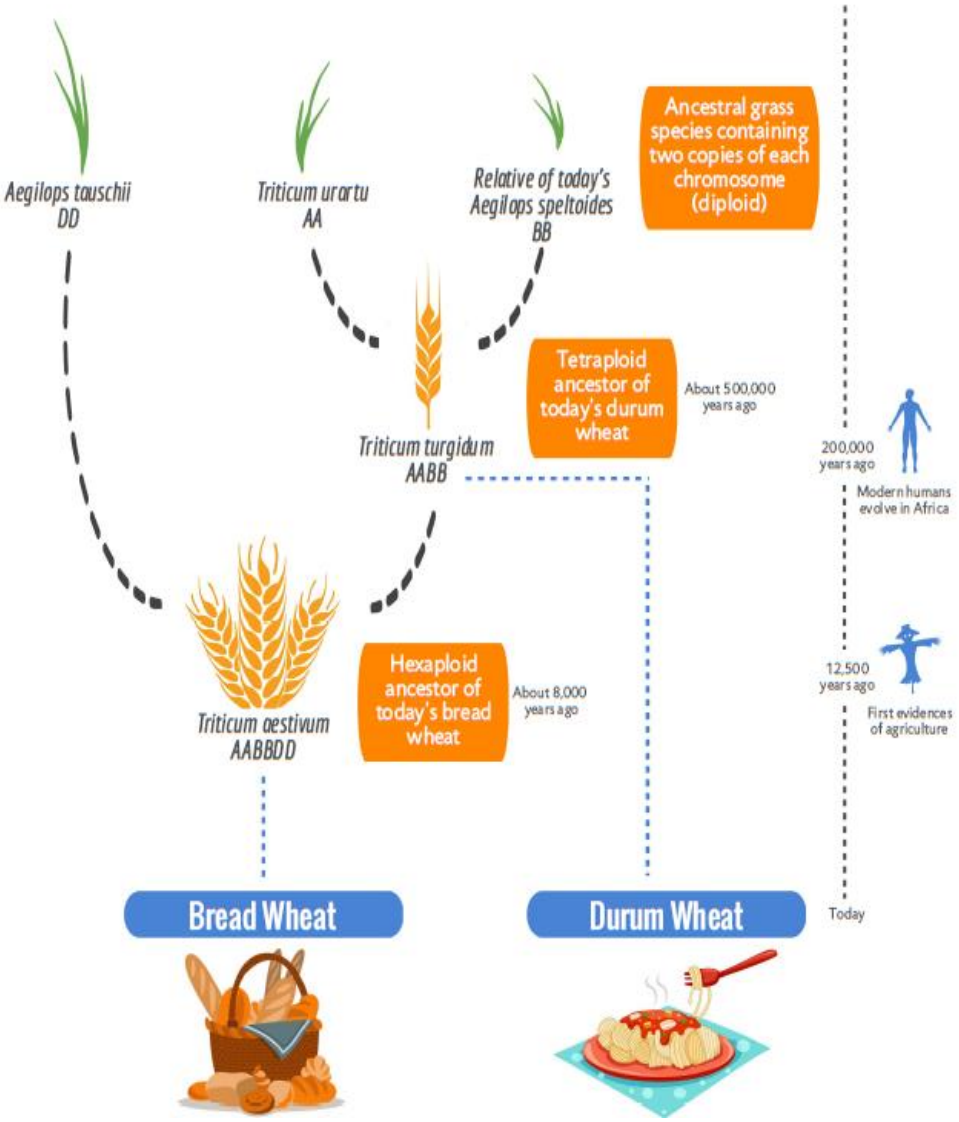


Figure 1.11 A schematic reproduction of the evolutionary origin of the two most cultivated modern wheat species: bread and common wheat (reproduced from https://www.wheatgenome.org/News/Latest-news/Origin-of-Wheat)

1.2.4 Wheat genomic

Wheat species can be divided into three groups based upon their ploidy level: (1) diploid 2n = 14 = einkorn wheat; (2) tetraploid 4n = 28 = emmer wheats, the class durum wheat belongs to; (3) hexaploid = 6n = 42 = bread wheats (Kilian et al., 2010). The allopolyploid structure of the wheat genome is considered one of the key factors in the success of wheat as a global staple crop. The presence of two or three copies of a gene may confer more plasticity and thus allow adaptation to changing environmental conditions (IWGSC, 2014).

A deep knowledge of wheat genome is important to undermine the key traits for wheat adaptation to environmental stresses, but its complexity has always been the main obstacle. Wheat genome has a very complex structure, a huge dimension — more than five times the size of a single copy of the human

genome — , and a high percentage of repetitive sequences (Callaway, 2017). Common wheat is a hexaploid species, which genome is composed of 21 pairs of chromosome of 16 billion bases and about 85% of repetitive sequences. Wheat genome can be further divided into A, B and D subgenomes, composed of seven chromosomes each. The evolution of wheat leads to the sharing of A and B genomes in the two species of bread and durum wheat.

The tetraploid durum wheat genome is about 12 Gb. Considering the relatively recent origin of polyploid wheat species, most genes in tetraploid and hexaploid wheat are present in two to three functional copies, called homoeologs, with similarities of over 97% across coding sequences (Uauy, 2017b).

The lack of a fully sequenced reference genome have slowed down wheat geneticists for many years, but this year (2018) bread wheat genome has been finally considered as completed (IWGSC, 2018). Initial efforts focused on individual chromosome-based approaches using long-insert Bacterial Artificial Chromosome (BAC) libraries, but thanks to the advent NGS technologies everything changed rapidly (Uauy, 2017b).

In the over past six years, many genome assemblies of wheat were released, differing for annotation and completeness: the CSS (IWGSC, 2014), the TGAC v.1 (Clavijo et al., 2017) and the very recent RefSeq. v.1.0 (IWGSC, 2018) (Table 1.1).

	CSS	**TGAC v.1**	**RefSeq v1.0**
Publication/Release date	IWGSC, 2014	Clavijo et al, 2017	IWGSC, 2018
Contigs/Chromosomes	> 1 million contigs	753,943 contigs	21 chromosomes + ChrU
Mean scaffold size	7.7 kb	88.7 kb	Full Chromosomes
Assembly Size	10.2 Gb	13.4 Gb	14.6 Gb
Resources	Archive EnsemblPlants TILLING mutants expVIP, wheatEXP	Archive EnsemblPlants expVIP	EnsemblPlants expVIP

Table 1.1 Comparison of the currently available annotated genome assemblies in bread wheat. The "Resources" row indicates wheat database interlinked with the genome assemblies: all the three assemblies are available on EnsemblPlants (plant.sensembls.org/Triticum_aestivum/Info/index) or on the archive of the same database; TILLING mutant database is available at www.wheat-tilling.com; expression data are available at www.wheat-expression.com (expVIP) and at https://wheat.pw.usda.gov/WheatExp/ (wheatExp). (Reproduced and adapted from Adamski et al.2018).

Since June 2016, the IWGSC has integrated all chromosome-based resources (physical maps and genetic maps, whole-genome-profiling-WGPTM sequence tags, optical maps, and markers) and released in January 2017 the "IWGSC RefSeq v1.0", the first version of the high-quality reference sequence of bread wheat (IWGSC, 2018). All the data are now public available (https://wheat-urgi.versailles.inra.fr/Seq-Repository/Annotations) and integrated into the EnsemblPlants platform (https://plants.ensembl.org/Triticum_aestivum/Info/Index).

The last reference genome of bread wheat gives an easy access to sequence-level information, which along with the emerging transcriptomic information (Ramírez-González et al., 2018) could help researchers to speed-up the functional annotation of both durum and bread wheat genes.

For durum wheat, two main reference assemblies are actually available. One in the cultivar Kronos (https://opendata.earlham.ac.uk/opendata/data/Triticum_turgidum/EI/v1.1/) and the other is in the cultivar Svevo (https://www.interomics.eu/durum-wheat-genome). Nevertheless, the deeper annotation level in the genomic and transcriptomic of bread wheat makes these data the most precious resource for durum wheat researchers, especially for functional genomic studies.

1.2.5 Durum wheat genetic resources

From the very beginning of agriculture, when the domestication of wild emmer started, humans exploited natural genetic variability to meet their own needs. The subsequent breeding of domesticated durum and bread wheat, have led to narrowed scenery of genetic diversity. The loss of genetic diversity is a phenomenon known as "genetic erosion" that lead, upon domestication, to a reduction of the initial diversity by 84% in durum wheat and by 69% in bread wheat (Jaradat, 2013). Reduced genetic variability is a key step in the extinction vortex, therefore, the current genetic composition of a crop species influences how well its member will adapt to the future environmental changes (Govindaraj et al., 2015). Nowadays a limited number of elite varieties of durum and common wheat are cultivated, while landraces are pushed into the background because of their lower commercial value and mainly conserved in germplasm collection (Giraldo et al., 2016). The term "landrace" has generally been defined as a cultivated, genetically heterogeneous variety that has evolved in a certain eco-geographical area and is therefore adapted to local agro-environmental conditions (Casañas et al., 2017). Landraces are: i) adapted to local environmental conditions and local management practices, hence associated with small-scale farming (Jaradat, 2013); ii) highly adaptable to both biotic and abiotic stresses (Zeven, 1998); iii) an intermediate stage in domestication between wild wheat and elite cultivars. Indeed, allelic variation gradually lost through domestication and breeding, has been recovered only by going back to landraces (Giraldo et al., 2016).

Most of the exotic material, including landraces, are preserved in gene banks as germplasm and constitute an easily transferable and valuable source of genetic variation for agronomical, morphological, adaptive and quality traits (Ceoloni et al., 2014; Dwivedi et al., 2016). Gene banks are, thus, responsible for developing and maintaining large numbers of collections to ensure the availability of genetic diversity of different species as a public good (Wang et al., 2017). Regarding durum wheat genetic resources, the EURISCO database (http://eurisco.ecpgr.org) reports about 7,000 accessions of traditional varieties and/or landraces on a total of 17,257 accessions. Such a high number

of genotypes for breeding studies are a limiting factor: consequently, new approaches are needed to overcome this constraint.

Lopes and colleagues (2015) in their review described how the genetic diversity of landraces can be exploited for wheat breeding, especially with respect to the field of adaptation to climate change. Notwithstanding the main advantages of modern cultivars, such as a greater productivity and higher nutritive values (Migliorini et al., 2016), the current yield grain trend is insufficient to meet the rising food demand and climate changes are expected to strongly affect this trend. Moreover, the modern varieties are also characterized by narrow genetic basis which, in times, results in a greater vulnerability to new biotic and abiotic stresses, as well as to unpredictable environmental conditions (Wang et al., 2017).

Considering the oncoming climate changes, it becomes of high priority to boost the rate of genetic improvement, especially as regard traits like abiotic stress tolerance (Jaradat, 2013; Pignone and Hammer, 2013), using germplasm landraces collection and, more in general, any wheat genetic resource.

1.3 Plant response to drought stress

1.3.1 Plant strategy against drought stress

The water scarcity expected in the ongoing century due to climate changes, will limit the water availability for agriculture and will expose crops to more frequent and severe drought events.

Drought can be defined as a protracted period of deficient precipitation resulting in extensive damage to crops, and a consequential loss of yield. DS alone is considered the single most devastating environmental stress, able to affect crop productivity more than any other (Basu et al., 2016).

When plants experience DS, an early response is the rapid closure of stomata to reduce transpiration rate, triggered by abscisic acid (ABA) phytohormone. The increase in ABA level under water deficit conditions activates many downstream stress responses, especially at cellular and molecular level, that can induce drought tolerance (Cheng et al., 2016). As soon as plant perceives soil drought, ABA is synthetized in the roots and channelled through the xylem to the shoots where it suppresses leaf expansion and causes stomatal closure to lower the transpiration rate and thus prevent water loss (Murtaza et al., 2016). As a consequence, leaves are deprived of CO_2 and thus photosynthetic carbon assimilation is decreased in favour of photorespiration.

Plants have three main adaptation strategies to drought resistance (Figure 1.12):

1) Drought escape, related to the ability of completing the life cycle and reproduce before the onset of drought, thus the stress is not experienced by the plant. (Basu et al., 2016)
2) Drought avoidance, which is related to the ability of sustaining the water cellular status under drought despite the reduced water content in the soil (Basu et al., 2016). This is achieved through a variety of adaptive traits to increase the water-use efficiency (WUE) involving the minimization of water loss (water savers) and optimization of water uptake (water spenders), either by minimizing water loss through transpiration or by withdrawing more water from the soil (Kooyers, 2015).
3) Drought tolerance is the ability of plants to endure low tissue water content through adaptive traits as soon as the DS event is perceived by the plant to withstand dehydration through osmotic adjustment and production of molecules that stabilize proteins (Kooyers, 2015). These adaptive traits involve the activation of a large number of drought-responsive genes and many regulators of the signalling pathway and transcription (Cheng et al., 2016, Basu et al., 2016).

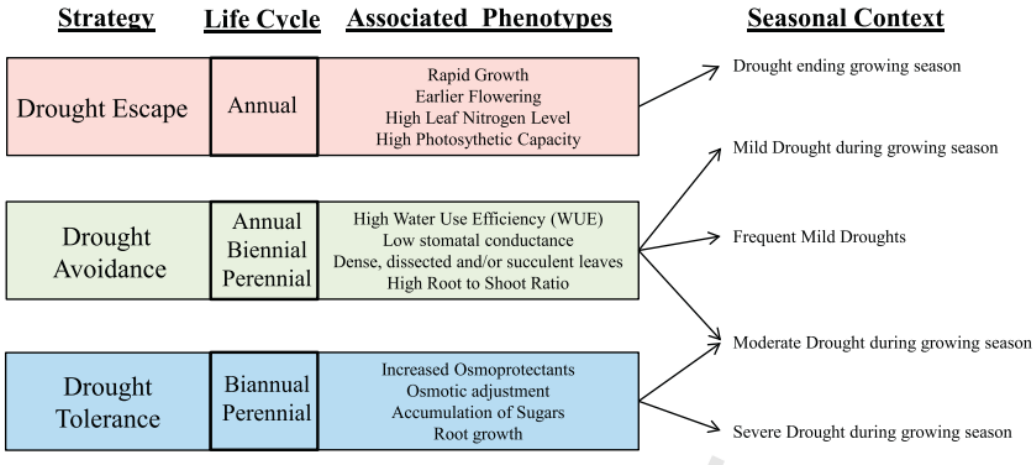


Figure 1.12 Schematic diagram describing the three ways plants evolve drought resistance. Each strategy has specific traits and phenotypes that have historically been associated, which of some examples are listed. Arrows designate the environmental context that would most favor each strategy. Reproduced from Kooyers (2015)

DS can affect plants development at any growth stage, but the effect varies depending on the severity and duration of the stress itself, as well as on the plant growth stage at the time of the stress occurrence. In particular, DS can have a detrimental effect on: (1) plant growth and productivity, since both cell division and cell enlargement are affected; (2) plant water relations (i.e. Relative Water Contents –RWC- , leaf water potential, osmotic potential, pressure potential and transpiration rate); (3) mineral uptake and assimilation; (4) light harvesting and carbon fixation (Aroca, 2012).

The stress response is a very complex phenomenon and plants display a wide range of mechanisms to overcome DS, at all level of complexity: morphological/anatomical, physiological, biochemical and molecular. A list of the major examples of mechanism of response and defence is presented below.

1) Morphological adaptations. Examples are the plasticity of root architecture to improve the water and nutrient uptake, the leaf area and its correlation with the transpiration rate and water loss from the leaves (Aroca, 2012).
2) Physiological adaptations that help plants to maintain metabolic activities at low tissue water potential (Bowne et al., 2012). Examples of this response are osmotic adjustment and the dynamics of phytohormones. Many hormones are involved in the physiological adaptations, but ABA plays a key role in the root-shoot signalling. Another example is the osmotic adjustment, by which organic and inorganic solutes are accumulated in the cytosol to lower the osmotic potential thereby asserting cell turgor. These compatible solutes (e.g. sugars and free amino acids) do not damage the cellular structures even at higher concentration.
3) Biochemical adaptation includes the generations of reactive oxygen species (ROS) as one of the earliest biochemical response to biotic and abiotic stresses. The production of ROS in plants, known as the oxidative collapse, is an early event, and acts as a secondary messenger to trigger subsequent defence reactions. The increased ROS level results in oxidative damage to macromolecules (proteins, DNA and lipids). To ensure normal cellular function under these conditions, plants developed an internal protective enzyme-catalyzed clean up system (superoxide dismutase, peroxidase, catalase, ascorbate peroxidase) which along with the non-enzymatic antioxidant mechanism keeps the integrity of the photosynthetic apparatus during the oxidative stress (Murtaza et al., 2016).
4) Molecular adaptations. The external drought *stimuli* are perceived by molecular sensors on the cell membrane which activate the signal transduction pathways, resulting in the expression of drought-responsive genes. The products of these genes can be metabolic proteins, transcriptional regulators of the signalling cascade, functional proteins for the protections of the membranes and transporters (aquaporins and sugar transporters) (Kaur and Asthir, 2017; Lata and Prasad, 2011). Further details about the gene regulatory networks involved in DS response are presented in paragraph 1.3.3.

1.3.2 The effect of drought on cereal crop

The interest on studying the molecular and genetic bases of drought response in cereal crops is mainly related to how severely drought can affect the final yield. The increasing food demand does not tolerate crop loss, thus traits like yield-stability has to be preserved even under the adverse climate changes.

The response to DS depends mainly on the growth stage and WUE which can varies among the same crop species (Daryanto et al., 2017). Species maintaining higher RWC under stress conditions are reported to be less susceptible to low water potential, and thus retain their growth and productivity under stress conditions (Joshi et al., 2016). Visible symptoms describing drought resistance in cereals include leaf rolling, stay green ability, epicuticular wax deposition, closing of stomata, enhanced root length leading

to higher WUE, photochemical quenching, photoinhibition resistance, osmotic adjustments, and membrane stabilization at the cellular level (Joshi et al., 2016).

Drought can affect final cereal yield depending on the period of occurrence. During vegetative phase, it is responsible of smaller yield loss compared to that during the reproductive phase, with a sensitivity that is crop-dependant (Daryanto et al., 2017).

Reproductive stage is considered critical in many crops, wheat included (Berger et al., 2010). Lack of water during the flowering can cause infertility of pollen, thus reduce the final yield by reducing the grain number (Farooq et al., 2014). A drought event during a pre-anthesis phase can shorten the anthesis itself, reducing the final number of grain (Fahad et al., 2017). But among all the growth stages, the grain-filling is the most sensitive to water deprivation. If severe drought occurs during the grain-filling, water scarcity curtails potential grain size by reducing the rate and duration of grain filling (Fahad et al., 2017; Farooq et al., 2014). The severity of water deficit during the reproductive stage, variably affects yield-related traits in durum wheat. While moderate drought (40% water reduction) only affects grain weight, severe water deficits (> 50% water reduction) can reduce both the number of fertile ears per unit area and the number of grains per ear (Daryanto et al., 2017).

1.3.3 Gene Regulatory networks in drought stress response

Most of the knowledge of the molecular mechanisms underneath the drought response comes from studies in the model plant *Arabidopsis thaliana*. Notable progress has been made so far, especially by the use of modern genetic and functional genomic approaches, and many drought related genes have been identified and characterized in plants. These genes include the primarily regulatory factors such as Transcription Factors (TFs), signalling protein kinases and protein phosphatases which play a major role in signal transduction and expression of downstream stress-responsive genes. In plants, up to 10% of genes in the genome potentially encode TFs, which are divided into different gene families based on the distinct structure of their DNA-binding domain (Joshi et al., 2016). These include AREB, DREB, MYB, WRKY, NAC, and bZIP.

The accumulation of ABA under water limiting conditions (i.e. drought, osmotic stress, freezing, high temperatures) plays a central role in stress responses and tolerance, as well as into the crosstalk in abiotic stress response (Nakashima et al., 2014).

Drought-responsive genes are generally divided into two groups, according to their dependency on ABA for their induction (Kaur and Asthir, 2017):

1) ABA-DEPENDENT GENE PATHWAY. The promoter regions of ABA-responsive genes are characterized by the presence of at least one copy of a conserved *cis*-element, named ABA responsive elements (ABRE, PyACGTGG/TC) in combination with coupling elements (CE)

(Joshi et al., 2016; Nakashima et al., 2014). ABA-dependent signalling consists of two main gene clusters regulated by ABA-responsive element-binding protein/ABA-binding factor (the AREB/ABF regulon) and the MYC/MYB regulon (Budak et al., 2015). The family of bZIP TFs AREB/ABFs, acts as the major TFs under ABA-dependent response to abiotic stress conditions in *A.thaliana*. The AREB/ABFs TFs are induced by DS and their transcriptional activities are controlled by ABA-dependent phosphorylation of multiple sites within conserved domains. Protein kinases SnRK2 (group A PP2Cs) and RCAR/PYR/PYL ABA receptors control the ABA signalling pathway including AREB/ABFs (Nakashima et al., 2014).

The second category of TFs regulated by ABA are MYBs, which have been characterized in recent years in *Arabidopsis* and in various crop species, testifying their involvement in drought response (Roy, 2015).

2) ABA-INDEPENDENT GENE PATHWAY. It involves the APETALA2/Ethylene Responsive Factor (AP2/ERF) superfamily which can be divided in two main groups, the Dehydration-Responsive Element Binding- proteins (DREBs) and the ERFs, induced respectively by dehydration and cold (Joshi et al., 2016; Licausi et al., 2013). DREB TFs bind the promoter of stress responsive genes that show the presence of a *cis*-element named DRE/CRT (A/GCCGAC) (Nakashima et al., 2014). There are two groups: DREB1/CBF TFs which specifically interact with the DRE/CRT and control the expression of a large number of stress-responsive genes in *Arabidopsis* involved in tolerance to DS, but also freezing and salinity; DREB2 TFs induced by osmotic stress and heat stress in *Arabidopsis* (Nakashima et al., 2014).

Several interactions between the AREB/ABFs and DREB/CBFs have been reported: the physical interaction between DREB1A/CBF3, DREB2A and DREB2C with AREB/ABF proteins. Moreover there is an interaction at the signalling pathway level, as ABRE/ABF TFs and SnRK2 are involved in the expression of DREB2A gene under osmotic conditions (Nakashima et al., 2014) (Figure 1.13).

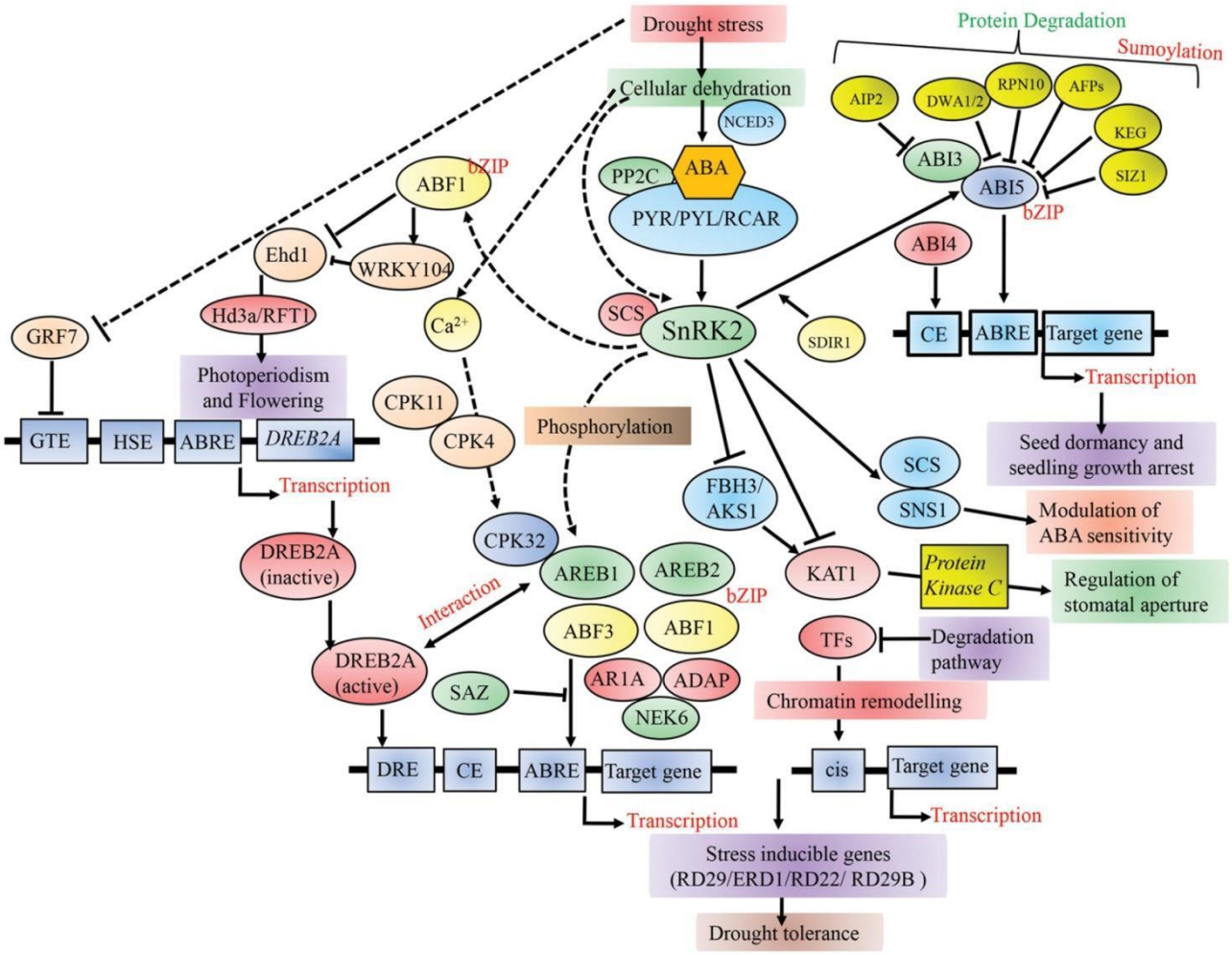


Figure 1.13 A schematic model of the transcriptional regulation of different TFs playing key roles in cellular dehydration in plants. PP2C-PYR/PYL/RCAR complex positively regulates AREB/ABF-SnRK2 pathway. SnRK2s further modulate other TFs downstream including AREB/ABFs, FBH3/AKS1, and SNS1 during stress as well as seed maturation. Phosphorylated AREB/ABF TFs, AREB1, AREB2, ABF3, and ABF1 bind to the promoter region of target genes and activate their expression in response to dehydration stress. GRF7 suppresses the expression of DREB2A, which is a key TF in ABA-independent gene expression. Broken lines indicate possible roles. PP2C-PYR/PYL/RCAR, pyrabactin resistance1/PYR1-like/regulatory components of ABA receptors; PP2C, protein phosphatase 2C; CE, coupling element; GTE, GRF7-targeting cis element; HSE, heat shock element; TFs, transcription factors (reproduced from Joshi et al., 2016).

1.4 Comparative genomics in durum wheat

Undermining the involvement of new target genes in drought tolerance is an important feature to understand the genetic and molecular basis of this complex phenomenon in wheat species.

Comparative plant genomics examines the similarities of, and differences in, genomes between plant species. By comparing genomes of evolutionarily divergent species, we can better understand the patterns and processes that underlie plant genome evolution as well as uncover functional regions of genomes (Caicedo and Purugganan, 2005). Comparative approaches are of particular interest in polyploid species such as wheat, because of their features that obstruct gene isolation (such as the complexity of the genome). For polyploid wheat species, closely related diploid species like barley (*Hordeum vulgare*), can really come to a hand. Indeed barley genome displays a mosaic of structural similarity to hexaploid bread wheat A, B, and D subgenomes, and most of the chromosome arms exhibit a well-conserved synteny between the two species (Mayer et al., 2011).

Plants react to DS in a complex manner. The "adaptation syndrome" begins with stress perception, which initiates signal transduction pathways that cause changes at the cellular, physiological and developmental levels, which involve hormone balance, modification of metabolites and repression or induction of many regulated genes (Blum, 1996).

Amino acid pools and concentration are influenced by abiotic stress: free amino acid levels are affected by increased protein degradation and some amino acids act as compatible solutes (i.e. proline) whose synthesis is strongly induced by drought. Amino acids can act in defence mechanisms, for example the BCAAs leucine, isoleucine, and valine serve as precursors of secondary metabolites involved in pathogen response, and their level is increased in response to drought and heat stress (Rizhsky et al., 2004).

We focused on two genes previously isolated by our co-workers in barley, which are involved in DS response: *HvBCAT-1* and *HvABA7*. The promoter regions of both *HvBCAT-1* and *HvABA7* contain many typical stress-related *cis*-acting elements we have already spoken about in the previous paragraph, including ABRE, DRE, MYC and MYB, which could explain the induced expression of these genes under water-limiting conditions, and thus classify them as drought-related genes.

HvBCAT-1 was first identified and isolated under drought condition in barley seedlings (Malatrasi et al., 2006). The gene was mapped on chromosome 4H of barley, and it encodes for an enzyme, a branched-chain-amino-acid aminotransferase (BCAT), and orthologues of AtBCAT-3.

HvABA7 encodes for a small protein (18 kDa), with no evidence of orthologues out of *Triticeae*. As *HvBCAT-1*, *HvABA7* was identified and isolated in barley seedlings under drought conditions (Gulli et al., 1995). In literature no other evidence of the involvement of this gene in DS has been presented in other *Triticeae* species so far, as well as functional annotation of its gene product.

BCATs are PLP-dependent enzymes which catalyse the last step of the synthesis and/or the initial step of the degradation of the three BCAAs (Isoleucine, Valine and Leucine) (Figure **1.14**). A unique feature of the BCAAs metabolism is that the four biosynthetic enzymes are common to all three amino acids, even though the substrates are different. The BCAA metabolism is closely related to the primary metabolism, as isoleucine belongs to aspartate-derived pathway, whereas valine and leucine are derived from pyruvate (Joshi et al., 2010).

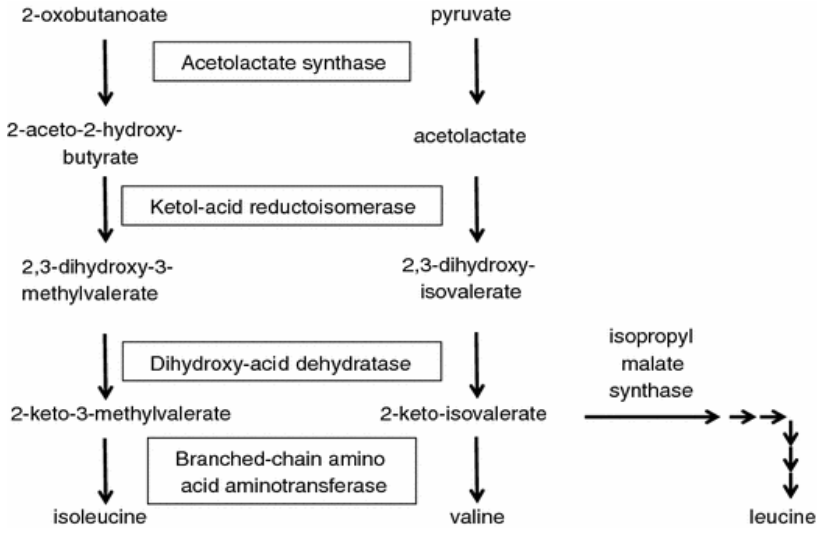


Figure 1.14 Branched-chain amino acid synthesis in plants. Enzymes in boxes are common to parallel pathways of isoleucine, leucine, and valine synthesis (reproduced from Joshi et al, 2010)

In *A.thaliana*, 6 members of the BCAT gene family have been isolated and characterized: *AtBCAT1* (At1g10060), *AtBCAT2* (At1g10070), *AtBCAT3* (At3g49680), *AtBCAT4* (At3g19710), *AtBCAT5* (At5g65780) and *AtBCAT6* (At1g50110). A pseudogene *AtBCAT7* has also been identified (At1g50090), with no evidence of transcription (Binder, 2010). Complementation studies of a yeast Δbat1/Δbat2 double knockout strain revealed that all AtBCATs but AtBCAT-4 can function as BCATs *in vivo* (Diebold et al., 2002).

Experiments in which N-terminal of the various BCATs were tagged with GFP, revealed three distinct subcellular localizations (Binder, 2010): BCAT1 is imported into mitochondria and is most likely involved in the degradation of Leucine. BCAT2, BCAT3, and BCAT5 are targeted to plastids and are assigned to the biosynthesis of BCAAs. While the cytosolic BCAT4 has a completely different function acting at the interface between primary and secondary metabolism (Knill et al., 2008). Biosynthesis of BCAAs occurs exclusively in plastids and many of the genes involved in their metabolism are found to carry plastid-targeting signal peptides (Joshi et al., 2010). However, the catabolism of BCAAs, which requires β-oxidation, might occur in mitochondria, peroxisomes, or both (Joshi et al., 2010).

Metabolite profiling studies revealed that the ABA accumulated during dehydration regulates the accumulation of various amino acids, including BCAAs (Obata and Fernie, 2012). The increased BCAAs

levels are correlated with the dehydration-inducible expression of their key biosynthetic genes, which are regulated by endogenous ABA (Urano et al., 2009). There are many evidences in literature which testify an increase of BCAAs levels under water-deficit conditions along with other compatible solute (Obata and Fernie, 2012), even in leaf tissues of different wheat cultivars (Bowne et al., 2012).Nevertheless, the role of BCAAs in water-deficit stress tolerance is still unknown, as well as the molecular involvement of BCATs enzyme.

The catabolic BCATs include the plastid BCAT3, and BCAT2, both interesting for their putative involvement DS. BCAT3 is not only involved in the BCAAs biosynthesis, but rather it has a dual function in both primary and in secondary metabolism. Indeed BCAT3 is clearly active in both the BCAAs and the Met-derived glucosinolate biosynthesis (Knill et al., 2008). In literature there are evidences of an alteration of glucosinolate composition of plants occurring under water-deficit stress and that the addition of exogenous glucosinolate hydrolysis products alleviates the adverse effects of this stress (Scarpeci et al., 2017). In *Arabidopsis*, the BCAT2 gene was induced in response to dehydration stress and this fact positively correlated with the accumulation of BCAAs (Urano et al., 2009).

The involvement of BCAT-3 in DS has been recently enforced in a study made in Arabidopsis (Scarpeci et al., 2017). The study regarded AtERF019, a TF member of the AP2/ERF super-family that is involved in growth, development and hormone responses, as well as in the regulation of responses to biotic and abiotic stress. Scarpeci and colleagues (2017) showed how plants which over-expressed AtERF019 have an increased tolerance to water deficiency, a delay in flowering time and a delay in senescence when compared with wild type plants Col-0. Authors underlined several differences in protein expression in those plants over-expressing *AtERF019*, and among the stress-related proteins they remarked the presence of BCAT3, which has been predicted in this study as a target of AtERF019 for the very first time. Moreover, the T-DNA insertional mutants BCAT3 plants were more sensitive to water-deficit than Col-0 plant. Since BCAT3 accumulated at higher levels under control conditions in plants over-expressing *AtERF019* compared to Col-0 plants, they suggested that glucosinolates and BCAAs contributed to the higher tolerance to water-deficit shown by transformed plants.

1.5 Functional genomics and molecular breeding in wheat

Functional genomics aims to acquire knowledge on gene function by linking the effect of different alleles to the change of the phenotype in the organism. Functional genomics relies on –omics data to reach this goal, thus the recent advances in wheat genomics and NGS techniques opened a wide range of possibilities in the functional genomics in wheat (Adamski et al., 2018). As soon as the gene of interest (GOI) has been selected, several resources can be used for its functional characterization. These includes resources based both on natural and induced variation and can involve both transgenic and non-transgenic approaches.

A valuable tool to create new variability in wheat is actually the Targeting Induced Local Lesions IN Genomes (TILLING). More recently, genome-editing technologies as the CRISPR/Cas9 techniques, provided additional tools to create DNA variation in polyploid wheat, but a further optimization is still needed to establish an efficient genome editing system in this species (Jia et al., 2018).

The identification of new sources of genetic variability are of great importance in wheat breeding programs since the actual priorities are mainly focused onto the final yield and yield stability, as well as the acquisition of resistance to disease, pests and of course, abiotic stress to face the impending climate changes.The so call prebreeding step in which researchers find new candidates for the introgression of genetic traits of interest in commercial cultivars is becoming mandatory due to the long process needed to obtain new wheat varieties (Shimelis and Laing, 2012).

Biotechnologies in this context could be used to manipulate plants genomes through insertion of gene(s) and generate the so called Genetically Modified Organisms (GMOs). Since a strict EU legislation prevent the use of GMOs as food resource, the genetic improvement of crops cannot rely on any techniques which lead to GMOs plants. Even if the development of GMOs is much faster compared to the non-transgenic breeding techniques, the Europe's highest court ruled on 25 July 2018 that even gene-edited crops by the mean of CRISPR/Cas9 editing techniques should be subject to the same stringent regulations as conventional GMOs (Callaway, 2018).

Thus, the non-transgenic breeding method actually represents the only valuable resource to improve the genetic of crops like wheat in Europe.

1.5.1 TILLING, a reverse genetic strategy for functional genomics in wheat

TILLING is a reverse genetics technique that combines random mutagenesis with high-throughput screening for single nucleotide mutations. It allows the isolation of mutations in GOIs and can be used to investigate gene function (Henikoff et al., 2004; Wang et al., 2010).

The advantage of TILLING in polyploids organisms is related to the redundancy effect of homoeologous copies on a single-gene knockout (Jia et al., 2018). The functional redundancy allows polyploids organisms to tolerate a high mutation rate, resulting in an higher probability to identify mutations in a GOI even in small population (Uauy et al., 2009).

In TILLING populations the average mutation density in diploid species is 1/380 kb, whereas it increase to 1/49 kb in tetraploids (as durum wheat and tobacco) and to 1/32 kb in hexaploids (as bread wheat and oat) (Wang et al., 2012). This high mutation frequency observed in polyploidy TILLING populations, facilitates the identification of large allelic series in target genes by screening a relatively small number of individuals and increases the probability to identify at least one nonsense mutation per target gene. According to the data of published TILLING population, the probability of identifying at least one truncation mutation by screening 2000 individuals have been estimated as 98% and 91% in hexaploid and tetraploid population respectively, whereas only 27% in diploids (Wang et al., 2012).

The most used chemical mutagen agent is ethyl methane sulphonate (EMS) which alkylates guanine bases creating mainly G–A and C-T transitions on the DNA and can often induce novel stop codons (McCallum et al., 2000). EMS-mutagenized populations are generated by exposing seeds (M0) to the mutagen and then allowing the resultant M1 plants to self-fertilize to originate M2 seeds where mutations will be stably inherited (Figure 1.15). The seeds must be exposed to a proper dose of EMS to ensure a high level of mutations without affecting viability and fertility. To ensure the greatest number of unique novel mutations one seed should be taken from each M1 plant. DNA is extracted from M2 plants leaf material and used for mutation detection. M3 seeds are archived and stored so that gene function can be studied in any plants in which mutations are identified. The M3 population is still segregating therefor the SSD method should be used to identify the desired mutation and to generate the near-homozygous material (with estimate 97% of the mutations being homozygous at M6), although up to half of all mutations present in the M1 are lost in the process (Figure 1.15) (Parry et al., 2009).

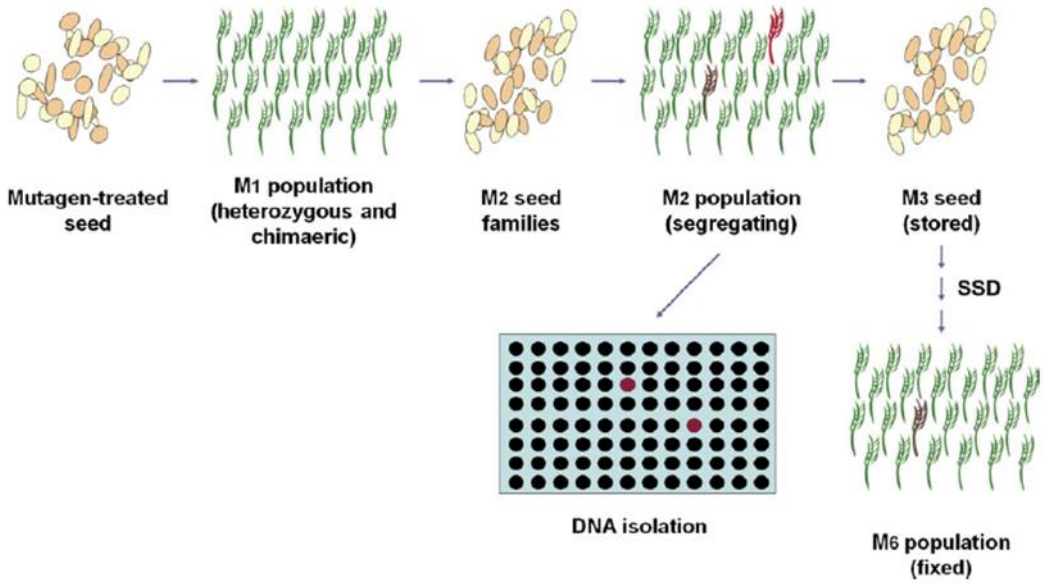


Figure 1.15 The general strategy used to create a mutant population in self-fertilizing crops (reproduced from Parry et al. 2009).

A number of TILLING resources have been developed in wheat, including tetraploid and hexaploid libraries in several varieties (Krasileva et al., 2017; Parry et al., 2009; Uauy et al., 2009). A TILLING library have been developed in the cultivar Svevo, providing variation in a high-quality genetic background, particularly suitable for breeders and researchers (Bovina et al., 2014). Uauy and colleagues (2009) produced a TILLING population for both bread and durum wheat, with mutation densities of 1/38kb for hexaploid and 1/51 kb for tetraploid. The tetraploid TILLING population can be used to generate mutants for basic research project due to the relatively easy possibility to generate null mutants through a single generation of crosses between A and B genome mutations, followed by selection of homozygous double mutants in the F2 populations. Hexaploid population, instead, is the most relevant for breeding application in bread wheat but requires an additional cross to combine the three mutant alleles within a single plant (Borrill et al., 2015).

The TILLING approach is effective for structure-function studies since it generates allelic series including null alleles at low frequency. In the recent years TILLING have been widely used to discover gene function and to mine new alleles in complex genomes like wheat (Botticella et al., 2011; Dong et al., 2009; Hazard et al., 2012) and it is emerging as important tool suitable for both research and breeding in polyploid species like wheat (Krasileva et al., 2017; Sestili et al., 2015).

TILLING have been successfully applied to develop a high-amylose durum and bread wheat varieties through combining null mutation in each of the three *SbeIIa* homoeologous (Slade et al., 2012).

The high mutation frequency is an important drawback for the mutation-phenotype correlation, as it is required to clean the background mutations through backcross prior to gene characterization. The improvements of techniques to reveal multiple deleterious mutations helped to partially overcome this issue leading to the possibility to score most phenotype after only one or two generation (Krasileva et al., 2017). This can be done by looking for the segregation after the self-pollination of heterozygous plant or

by single cross between homozygous alleles, where the maintenance of the phenotype in the F1 could guide to the individuation of the correct phenotype (Henikoff et al., 2004). Moreover, the use of homozygous sibling lines segregating for wild type and mutant alleles provides a powerful tool to reduce interference from background mutations and correlate targeted mutation with function (Parry et al., 2009).

1.5.2 TILLING by sequencing and in silico TILLING

During the years, the conventional TILLING platforms became tedious and time consuming, as mutation screening was based on the PCR amplification of a target region from pooled DNA of mutagenized plants followed by eteroduplex detection. Techniques of detection by the use of endonuclease CEL I, and the high-resolution melt analysis (HRM) have been so far used for the mutation detection, however the recent advances in sequencing technology allowed the use of NGS also for mutation detection (Jia et al., 2018), increasing throughput and accurate recovery of induced mutations and extending the analysis to multiple amplicons within one screen (Borrill et al., 2015; Tsai et al., 2011). The application of NGS, opened the possibility to a fast genome-wide discovery of mutations, however the whole genome sequencing costs are still limiting the possibility for whole-genome re-sequencing of mutant populations. A reduced representation methodology such as exome-capture, which allow to restrict sequencing to the coding part of the genome, has been demonstrated to be effective in animal and plant genomes and could constitute a powerful tool for mutation discovery when applied to mutagenized populations of wheat (Krasileva et al., 2017).

Recently, exome capture has been demonstrated to be a suitable method to describe the types and frequency of mutations present in EMS-mutagenized rice and wheat plants (Henry et al., 2014).

The wheat exome is estimated to account for less than 2% of the complete genome and exome capture in wheat allows to reduce the genome size from 17 Gbp to <150 Mbp suitable for NGS (Figure 1.16a). Application of NGS technology allowed sequencing of exome-captured DNA from 2735 tetraploid Kronos and hexaploid Cadenza EMS-induced mutants (Krasileva et al., 2017). Exome capture and re-sequencing of TILLING populations is a resource to identify mutations in the target gene with a predicted effect on the protein function leading TILLING in wheat toward an *in silico* screening activity (Figure 1.16b) (Borrill et al., 2015; Wang et al., 2012). The feasibility of using exome capture for genome re-sequencing as a method of mutation detection in polyploid wheat have been recently confirmed by other authors (King et al., 2015).

An exome capture resource also provides users with online access (http://www.wheat-tilling.com/) to conveniently identify knockout alleles for over 90% of wheat genes. Moreover, predesigned SNP-based primers are available to validate the mutations (Ramirez-Gonzalez et al., 2015a), which can be combined to develop double or triple null mutants in order to overcome the problem of redundant homoeologs. Such a resource makes stocks of mutant lines permanent and easy to access and should facilitate characterization of wheat gene function (Adamski et al., 2018).

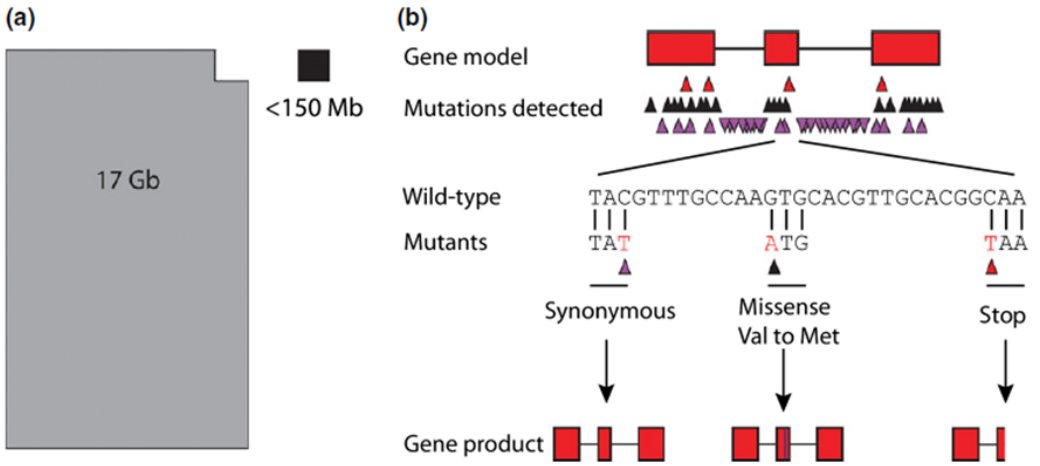


Figure 1.16 *In silico* TILLING. (a) Exome capture reduces the wheat genome from 17 Gb to < 150 Mb for next-generation sequencing. (b) Variant calling identifies mutations within gene models. Amongst the mutations detected, the first row of triangles depicts truncation mutations (red), the second depicts missense mutations (black), and the third depicts silent mutations (purple). Upward-pointing triangles indicate an exonic change and downward-pointing triangles indicate an intronic change. The sequences of wild-type and mutant individuals within the second exon are shown, along with the effects on amino acid usage. The resulting gene products are indicated at the bottom. (Reproduced from Borrill et al. 2015).

1.5.3 Conventional and modern breeding in wheat

Plant breeding is an important component of agriculture and can be considered the basis of human civilization. Breeding methods have evolved throughout time along with the development of civilizations, essentially to meet the human need of food resources.

We can distinguish plant breeding essentially in two main categories: the conventional and the innovative methods.

Conventional breeding relies on sexual crossing of one parental line with another parental line in order to achieve a desirable trait (e.g. resistance to environmental stresses). To reach this goal, breeders choose the best progeny and back-cross it to a recurrent parent to dilute the irrelevant or undesired trait (Hou et al., 2014). Obviously this process takes time, several years for many crop species, to finally get a commercial use. Prerequisite to traditional breeding strategies is the existence of genetic variation for the desired traits, and sexual compatibility of the organisms carrying those traits.

Modern breeding includes the marker-assisted breeding techniques, the transgenic technologies and the precision-engineering editing techniques such as CRISPR/Cas9 (Savadi et al., 2018). Modern techniques basically can assist breeders in improving the selection methods. These include development of molecular marker technologies and identification of genetic markers linked to traits of interest (Shimelis and Laing, 2012). Genetic markers are closely related with the target gene and they act as flags.

Wheat breeding programmes across the world are dependent on a limited crop germplasm and currently, worldwide grown wheat cultivars capture only 10% of the existing wheat diversity (Savadi et al., 2018).

Landraces can come to an hand to overcome this phenomenon of genetic erosion and genetic bottlenecks imposed on crop plants during domestication (Lopes et al., 2015). Indeed, exotic germplasm such as landraces and wild relatives possess high levels of genetic diversity for valuable traits, including adaptation to stressful environments and more efficient nutrient utilization (Wang et al., 2017).

The advent of affordable high-throughput genotyping and phenotyping technologies, together with omics-based systematic genetic technologies and emerging statistical genomic methods, provide new avenues for efficient management, characterization, and utilization of exotic germplasm (Wang et al., 2017). Thus, landraces can provide alleles, which has to be introgressed into modern varieties by hybridization, once identified.

1.5.4 Mining wheat genetic resources by SNP-based molecular markers

Molecular markers, and especially DNA-based markers, are widely used for different purposes: i) characterize genetic diversity in germplasm collections; ii) develop high-density genetic maps to identify genomic regions associated with important traits; iii) facilitate the identification of superior alleles and selection of desirable genes or traits in breeding programmes (Kadirvel et al., 2015). DNA-based markers are the variations observed in a particular DNA portion among different individuals of a species; these variations may be due to different mutations (insertion, deletion, substitution or replication errors of tandem repeated DNA). The DNA markers can be considered an ideal genetic marker system as they meet the characteristics to be unlimited in number, insensitive to environmental conditions, highly reliable and easily automatable. Several DNA marker technologies have been developed and improved over the years from restriction based to PCR and to sequence based technologies; they differ on the basis of the polymorphism and detection methodology. Markers can also be classified on the basis of their ability to distinguish between heterozygotes and homozygotes genotypes. A codominant marker is able to distinguish the different alleles present in heterozygotes e, whereas a dominant marker detects just the presence or absence of one particular allele (Kadirvel et al., 2015; Janni et al., 2017).

Markers may belong to either transcribed or non-transcribed regions of the genome; those developed without prior knowledge of their function are called ‘anonymous’ or ‘random markers’ whereas those derived from the functional variants present in a genic region, which causally govern the trait variation, are called Functional Markers (FMs). The FMs are diagnostic of the desired alleles and are the markers of choice for targeted MAS (marker-assisted selection); in plant breeding programs they can be used in the segregating population without the issue of recombination (Varshney et al., 2005). With the rapid progress in functional genomics together with whole genome and RNA sequencing, a lot of sequence data have become available for the development of FMs.

SNPs are single nucleotide base difference between DNA sequences of individuals, which occur within coding/non-coding regions of genes or in the intergenic regions. They are numerous and occur at different

frequencies in different regions of the genome. SNP is the most preferred marker system nowadays due to their abundance (two to three per kilo base), high level of polymorphism, high throughput capability and cost-effectiveness (Kadirvel et al., 2015). The use of SNPs as FMs is convenient as they are present in large number and distributed throughout the genome (Kage et al., 2016), and can be used in wheat improvement (Bagge and Lübberstedt, 2008).

The numerous FMs developed in wheat, rice, maize, sorghum, millets and other crops, and successfully used in breeding programs have been recently reviewed by Kage and co-workers (2016). Thanks to the recent increment of sequencing data available for several crop species, SNP markers have been developed and have largely replaced other type of molecular markers as SSR (Simple Sequence Repeats). Because of their low assay cost, high genomic abundance, locus specificity, codominant inheritance, simple documentation, potential for high-throughput analysis, and relatively low genotyping error rates (Schlötterer, 2004), SNPs have emerged as powerful tools for many genetic applications and are often the preferred marker system in MAS and marker-assisted breeding (Kage et al., 2016).

The NGS technologies have drastically increased the speed at which the DNA sequence can be generated, while reducing the costs and allowed the development of genotyping assays for SNPs (Kage et al., 2016). Numerous SNP genotyping platform are currently available combining a variety of chemistries or allele discrimination techniques (restriction enzyme digestion, sequencing, hybridization with allele-specific probes, single nucleotide primer extension, allele-specific amplification, oligonucleotide ligation), detection methods (colorimetry, spectrometry, fluorescence, fluorescence resonance energy transfer, fluorescent polarization, chemiluminescence), and reaction formats (solution-phase, solid-phase, and gel electrophoresis, NGS) (Semagn et al., 2014)

New approaches have been developed to use sequencing itself as a genotyping tool for genetic studies with discovery and breeding applications. This approach is called “genotyping by sequencing” (Chung et al., 2017). Currently, multiplexed chip-based technology is the highest throughput direct SNP genotyping platform generating till million SNPs per run (Semagn et al., 2014).

The most used techniques to detect SNPs rely basically on PCR through allele-specific hybridization, allele-specific nucleotide incorporation or the generation of length polymorphisms. The PCR-based methods to detect SNPs without relying on length polymorphisms include HRM, TaqMan® and KASP™ (Kompetitive Allele Specific PCR). To perform HRM, primers are usually designed to amplify products smaller than 150 bp, with the polymorphic site near the middle of the product (Dong et al., 2009). If the single SNP within the amplicon product is able to modify its melting temperature sufficiently to differentiate the alternative alleles the resulting marker may be codominant (with three distinct melting curves for the heterozygote and the two homozygotes), otherwise the resulting marker may be dominant (with the heterozygote distinguishable from one homozygote but not from the other) or it may distinguish the heterozygote from homozygotes but not the homozygotes from each other (Garcia and Mather, 2014).

TaqMan® assays use PCR primers in combination with a dual-labelled allele-specific probe. The probe contains a fluorophore at its 5’ end and a quencher at its 3’ end. When the probe is intact, the quencher is close enough to the fluorophore to reduce the emission of fluorescence by fluorescence resonance energy

transfer (FRET). During PCR, the probe binds specifically to the target site between the primers. During primer extension, the 5' exonuclease activity of the polymerase degrades the probe, releasing the fluorophore from the quencher. Fluorescence is emitted and can be detected in real-time with a proper thermal cycler. Probes and primers for Taqman® assays can be designed using software provided by commercial suppliers of the probes. The probes are usually between 20 and 30 bp long, providing specificity, with the target polymorphism as close as possible to the center of the probe. The primers are designed to closely flank the probe target sequence, to amplify a product no longer than 150 bp (Garcia and Mather, 2014).

KASP is a fluorescence-based genotyping technology. The KASP genotyping system (LGC Genomics, UK), initially developed by KBioscience, is based on the allele-specific primer extension and FRET for signal generation. Each assay involves three unlabelled primers: two allele-specific primers and one common primer. The allele-specific primers are designed with their 3' ends complementary to each of the SNP alleles and with a non-complementary tail in the 5' end. Their tail sequences differ from each other and are complementary to two FRET quenching reporter oligonucleotides (each labelled with a different fluorophore) that are present in the KASP Mastermix. The KASP Master mix (LGC Genomics, UK) is provided in a ready-to-use 2× format containing universal fluorescent reporting dyes FAM™ and HEX™ as well as Rhodamine X (ROX) dye as the passive reference; the mix contains the two 5' fluor-labelled primers with their complementary quencher-labelled primers. The KASP reaction consists of the KASP assay mix (assay specific, composed of the two allele-specific primers together with the common primer) and the KASP Mastermix (universal; used with any assay mix) which are combined with the DNA sample. In the initial stage of PCR, the appropriate allele-specific primer binds to its complementary region directly upstream of the SNP (with the 3' end of the primer positioned at the SNP nucleotide) (Figure 1.17). The common reverse primer also binds and PCR proceeds, with the allele-specific primer becoming incorporated into the template. During this phase, the fluor-labelled primers remain bound to their quencher-labelled complementary primers, and no fluorescent signal is generated. As PCR proceeds further, one of the fluor-labelled primers, corresponding to the amplified allele, is also incorporated into the template, and is hence no longer bound to its quencher-labelled complement. As the fluorophore is no longer quenched, the resulting fluorescence can be detected using a plate reader or a real-time PCR instrument. If the genotype at a given SNP is homozygous, only one or the other of the possible fluorescent signals will be generated. If the individual is heterozygous, the result will be a mixed fluorescent signal (Garcia and Mather, 2014).

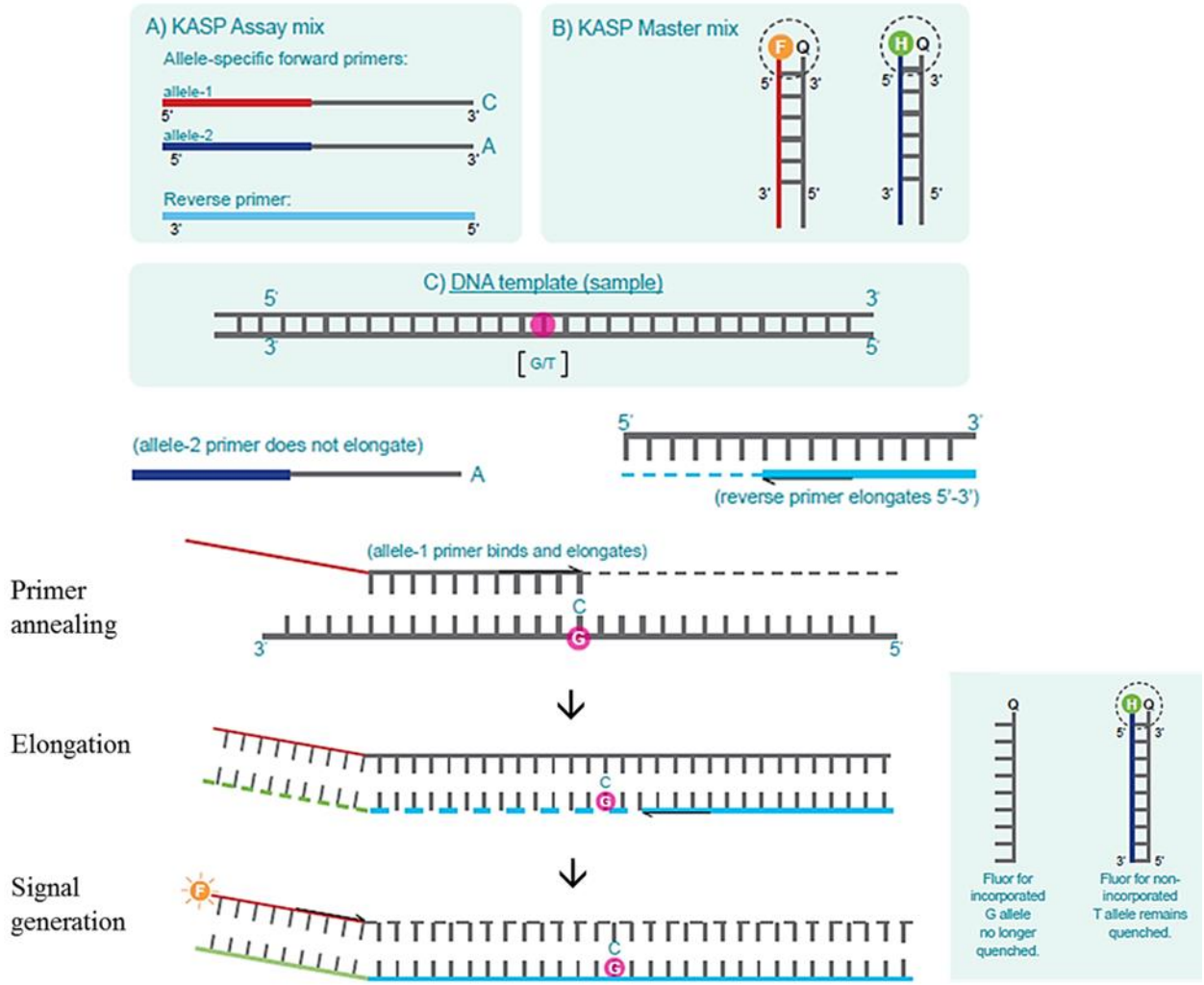


Figure 1.17 KASP mechanism of action. The KASP assay components are reported in A, B and C and include the KASP Assay mix (A), the KASP Mastermix (B) and the DNA template (C). The KASP Assay mix contains the two allele-specific primers with FAM and HEX complementary tails (colored in red and blue respectively) together with the common primer. The KASP Mastermix contains the two FRET cassette composed of two 5' fluor-labelled primers with their complementary quencher–labelled primers indicated as Q). The orange and green circles indicate the FAM and HEX dye. The pink circle indicates the target SNP (Modified from KASP genotyping chemistry User guide and manual).

1.5.5 KASP genotyping applied to molecular breeding

KASP is a convenient technology for the use in single-plex SNP genotyping platform that are suitable for the application requiring a moderate number of SNPs detected for a large number of samples (King et al., 2015; Semagn et al., 2014). Recently a KASP assay specifically developed for the genotyping of leaf rust resistance locus Lr21 has been successfully used in differentiating resistant and susceptible genotypes over a panel of 384 US wheat lines (Neelam et al., 2013).

KASP SNP-based markers can be also used to follow the mutation detected by TILLING through the back-crossing steps. To make the application of KASP marker easier, a bioinformatics pipeline called PolyMarker (http://polymarker.tgac.ac.uk;(Ramirez-Gonzalez et al., 2015a) has been recently implemented to automate the design of the three-primer needed for the KASP assay. PolyMarker

generates a multiple alignment between the target SNP sequence and the CSS scaffolds (IWGSC, 2014) for each of the three wheat genomes allowing developing genome specific primers taking advance of homoeologous SNPs with respect to the target genome. The common primer is selected to incorporate a homoeologue-specific or semi-specific base at the 3' end, whereas the competing diagnostic primers are designed to incorporate the alternative varietal bases at their 3' ends (Ramirez-Gonzalez et al., 2015b). This tool allows distinguishing between heterozygous and homozygous individuals in polyploid species and is therefore a powerful tool for TILLING in wheat. The pipeline has been successfully applied to validate putative SNPs from RNAseq data and successfully converted these SNPs into codominant genome-specific KASP assays that allow coping with the difficulties of the polyploid wheat genome and have been used for diagnostics across UK pedigree programmes.

1.6 Plant Phenomics, a necessary tool for prebreeding and functional genomics

The increasing availability of genetic resources for wheat provides invaluable opportunities for genetic improvement of crop resilience to abiotic stress arising from climate change. This requires analysis of hundreds of lines grown under diverse environmental scenarios (Tardieu et al., 2017).

On one hand genotyping has reached a high-throughput level at a relatively low cost through advances in DNA marker assays and sequencing technologies. On the other hand, equivalent improvements to generate high-throughput and valuable phenotypic information are urgently needed. This is the object of the field of plant phenomics.

Plant phenomics can be defined as the development and application of the suite of tools and methods used for three major goals — capturing information on structure, function and performance of large numbers of plants, together with their environment; analysing, organizing and storing the resulting datasets; and developing models able to disentangle and simulate plant behaviour in a range of scenarios (Tardieu et al 2017).

While advances in high-throughput genotyping have provided fast and inexpensive genomic information, the molecular breeding strategies, which place greater focus on selections based on genotypic information, still require as many as possible phenotypic data.

Lot of high-throughput phenotyping platforms (HTPPs) around the world have been developed (Araus and Cairns, 2014).

In Italy by ALSIA-Metapontum Agrobios (http://www.alsia.it/opencms/opencms/Ricerca/Agrobios/) is the reference infrastructure for plant phenomics in the Mediterrenean area (www.phen-italy.it). It exploits the *in vivo* imaging of a plant of interest as a source of high-throughput phenotyping. A semi-automated system allows constantly monitoring plants by the collection of several images with different light-sources: Near Infra-Red (NIR), Visible (VIS/RGB), and Fluorescence (FLU). Different light-sources can be used to retrieve several morphological and physiological parameters such as the

leaf water content and thickness (NIR data), the biovolume and morphology of plants (VIS data) or even the Photosystem II efficiency and chlorophyll content (FLU).

By HTPPs plant health status under both control and stress conditions (i.e. DS) can be monitored in a non-destructive manner. Vegetation indices derived from RGB images are used as an affordable means of assessing genotypic differences in grain yield in response to a wide range of stress conditions in wheat including water stress (Yousfi et al., 2016). RGB data can measure the plants biomass and thus health indexes under stress conditions, while NIR data can be used to measure water level in plant, since the reflection in the NIR increases with the decreasing of water content in leaves(Tardieu et al., 2017).

Nowadays, the availability of high-throughput genotyping and phenotyping methods, which provide a large set of data, together with the use of interconnected information network between different academic areas such as plant breeding and biology, could help breeders to search for useful accessions to be used in pre-breeding programs. It is a necessary first step in the use of diversity arising from wild relatives and other unimproved materials.

2 AIMS OF THE PROJECT

Wheat is one of the most important staple crops, of which durum wheat accounts for 5% of the total production, but still it represents a major local commodity in many countries of the Mediterranean area. Indeed the second producer worldwide is represented by Italy, for which durum wheat accounts for the 65% of the total wheat production.
Wheat growth and yield can be severely affected by DS at any developmental stage, and despite advances in our understanding about drought response in plants, many aspects of the molecular and physiological mechanism involved in drought tolerance remain unknown in wheat.
Moreover, the massive genetic improvement of the last century of this crop, lead to the high-yield varieties typically cropped for the main cultivation of wheat, with narrowed genetic basis.
This situation is worsened by the upcoming climate change that is forecast to reduce the wheat yield in the upcoming years.
In this contest, the identification of durum wheat genotypes with increased adaptability to drought is mandatory to address the ongoing challenges for both researchers and breeders.

The overall objective of this work is the identification of new target genes potentially involved in drought response in durum wheat, and to exploit the natural variation occurring in the genetic resources to identify new durum wheat genotypes ready for these challenges.

Intermediate goals of this thesis project can be summarized in the following principal pillars:

- The isolation and the characterization of new targets involved in the drought tolerance along different growing stages of the life-cycle of durum wheat;
- The characterization at molecular and transcriptional level of these genes, in particular regarding their role in the drought defence response;
- The identification of mutations in target genes in a durum wheat TILLING populations exploitable for functional genomic studies to dissect the role of these targets in the drought response in durum wheat;
- The identification of natural allele variants within genetic resources of durum wheat to be introduced in wheat breeding programs.

3 MATERIAL AND METHODS

3.1 Identification and isolation of durum wheat sequences of *TdBCAT* and *TdABA7* genes

3.1.1 Identification of wheat sequences and primer design

The IWGSC CSS (International Wheat Genome Sequencing Consortium Chromosome Survey Sequencing) assembly was investigated from the EnsemblPlants database (https://plants.ensembl.org/Triticum_aestivum/Info/Index) for wheat sequences matching that of *H.vulgare HvABA7* (AJ296273.1) and *HvBCAT-1* (AJ574850.2) using the BLASTn algorithm. The genomic scaffolds carrying putative *T.aestivum ABA7* and *BCAT* genes were aligned to *HvABA7* and *HvBCAT-1* to identify the putative sequences of *T.aestivum*. Bread wheat gene sequences were used for the design of gene-specific primers for the PCR-based isolation of *T.durum* sequences *TdABA7* and *TdBCAT*. All the primers were designed with Primer3 (Untergasser et al., 2012).

3.1.2 Plant material

Seeds of *T. durum* cv. Cham1 were germinated at room temperature in dark conditions. Ten-day-old seedling leaves were sampled for DNA extraction. Seeds were kindly provided by Prof. Mario Pagnotta (University of Tuscia, Viterbo, IT).

3.1.3 DNA extraction and DNA quality control

Genomic DNA was extracted with the GenElute™ Plant Genomic DNA Miniprep Kit (SIGMA-ALDRICH, St. Louis, MO) from 100 mg of frozen leaves. DNA quality and quantity was assessed with Cary 50 UV-Vis Spectrophotometer (Agilent Technologies, Santa Clara, CA). The DNA integrity and quantity were also checked on agarose gel (1% agarose w/v, 1X TAE buffer, GelRed™ as Nucleic Acid gel stain) by the comparison with λ genomic DNA standards.

3.1.4 Isolation of durum wheat sequences

The full *TdABA7-A*, *TdABA7-B*, *TdBCAT-A* and *TdBCAT-B* were PCR-amplified from cv. Cham1 template DNA using GoTaq Colorless Master Mix 2X (Promega, Madison, WI, USA) in 20 μL reactions each containing 10–20 ng template and 0.2 μM of each primer.

The primer pairs (Table 3.3) used for the complete isolation were:

- TdABA7_A1F-TdABA7_A1R for *TdABA7-A*,
- TdABA7_B1F-TdABA7_B1R for *TdABA7-B*,
- TaBCAT_A1F-TaBCAT_A1R, TaBCAT_A2F-TaBCAT_A2R, TaBCAT_U1F-TaBCAT_A3R for *TdBCAT-A*,
- TdBCAT_U2F-TdBCAT_B3R, TaBCAT_A2F-TaBCAT_B1R, TaBCAT_U1F-TaBCAT_B2R for *TdBCAT-B*.

PCR amplifications were performed as follows: 1) initial denaturation at 94° for 2 min 2) 35 cycles of 94 °C for 30 s, annealing for 30 s, and elongation at 72 °C for 60s/kb, and 3) 1 cycle at 72° for 5 min. The optimized annealing temperature for each primer pair are reported in Table 3.3.

The PCR products were separated through electrophoresis on a TAE agarose gel, extracted from the gel using a NZYGelpure kit (NZYtech, Lisbon, Portugal) and submitted for sequencing (Eurofins Genomics, Edersberg, Germany).

3.1.5 *In silico* prediction of gene model

The gene models and the coding sequences were predicted with FGENESH+ (http://www.softberry.com/berry.phtml?topic=fgenes_plus&group=programs&subgroup=gfs) (Solovyev, 2008) of the Softberry suites (Softberry Inc., NY, USA) using as similar protein *HvABA7* (CAC17774.1) for *TdABA7* gene model and *HvBCAT* (CAE00460.2) for *TdBCAT*. The protein sequences of TdABA7s and TdBCATs were deduced on the basis of the results of FGENESH+.

3.1.6 *In silico* analysis investigation of *T. durum* genome assemblies for chromosome localization and promoter analysis, phylogenetic analysis

Durum wheat genome assemblies of cv. Kronos (https://wheatis.tgac.ac.uk/grassroots-portal/blast) and cv. Svevo (https://d-data.interomics.eu/blast/nucleotide/nucleotide) were investigated for sequences matching that of *TdABA7* and *TdBCAT* previously isolated and sequenced in cv. Cham1, using BLASTn algorithm. All the multiple alignments of nucleotide sequences were performed using DNAMAN version

4.15 (Lynnon Biosoft) to confirm the correct isolation of targets. Proximal promoter region (1000 bp upstream start codon ATG) of cv. Svevo were analysed with Plant Pan 2.0 (https://d-data.interomics.eu/blast/nucleotide/nucleotide) using *Arabidopsis* database as reference (Chow et al., 2016) and PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) (Lescot et al., 2002).

Phylogenetic and molecular evolutionary analyses of BCATs were conducted using *MEGA* version 6.06 (Tamura et al., 2013) on multiple alignment performed with MUSCLE algorithm. The phylogenetic analysis was performed using the Neighbor-joining method, Phylogeny Test with the Bootstrap method (1000 Bootstrap Replications), using the Poisson model for the substitution model.

3.2 Gene Expression analysis

The expression of genes *TdABA7-A, TdABA7-B, TdBCAT-A, TdBCAT-B* was evaluated under water-limiting conditions on a number of *Triticum durum* cultivars during two different growing phases:

(1) Germination;

(2) Reproductive stage during anthesis and early post-anthesis events.

The experiments have been named accordingly.

3.2.1 Plant materials and growth conditions

The following durum wheat cultivars were used as experimental materials: Cappelli, kindly provided by the Institute of Biosciences and Bioresources (IBBR) of Bari (Italy), Colosseo, kindly provided by Council for Agricultural Research and Economics (CREA) of the Research centre for Genomics & Bioinformatics of Fiorenzuola d'Arda (Italy), Kronos, kindly provided by Prof. Cristobal Uauy (John Innes Center, Norwich, UK), and Primadur, Simeto, Messapia, Ciccio, Creso, Saragolla and Svevo, kindly provided by Center of Cereal Research (CREA-CER) of Foggia (Italy).
Seeds were surface-sterilized with 10% (v/v) sodium hypochlorite for 2 minutes with gentle agitation, rinsed in distilled sterile water, and grown on wet Whatman filter paper for 9 days at 6°C in dark conditions to germinate. Two experiments were done:
Experiment (1)

Ten durum wheat cultivars were used as experimental material: Kronos, Primadur, Simeto, Colosseo, Messapia, Ciccio, Creso, Saragolla, Svevo, and Cappelli. After germination, and vernalization, 8-days old coleoptiles (Z09, Zadoks scale) were dehydrated for 8 hours on dry-paper as stress condition, while controls were kept well-hydrated on wet paper. Ten coleoptiles from each cultivar and condition were

sampled and immediately frozen in liquid nitrogen and kept at -80°C until RNA extraction for the transcriptional analysis.

Experiment (2)

Two different durum wheat cultivars Colosseo (susceptible) and Cappelli (tolerant) were grown in a growth chamber under controlled conditions (L:D 16:8 hours, T=20˚C, RH=80%). Plants were grown under well-watered condition until the late tillering stage and then transplanted, with three plants per pot (15 x 15 x 20 cm). Each pot contained the same amount of soil (8:4:1 soil:peak:perlite). After transplantation the water status of the soil was monitored through the Relative Soil Water Content, RSWC until the end of the experiment. To calculate the RSWC, dry weight (DW) and field capacity (FC) of 700 gr of fresh weight (FW) of soil were calculated before transplantation. RSWC was calculated as $RSWC\ (\%) = \left(\frac{FW-DW}{FC-DW}\right) * 100$. The weight of pots and the plant biomass was evaluated during this estimation. Pots were weighted twice a day for all the experiment to measure the FW, watering pots when necessary to keep constant RSWC for all the duration of the experiment. The experimental design is illustrated in Figure 3.1.

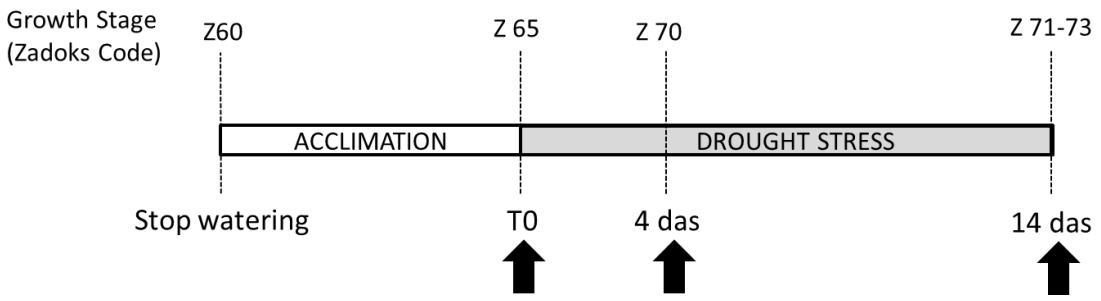


Figure 3.1 Experimental design of the drought experiment during the reproductive stage. Black arrows indicate the 3 time-points in which tissues (flag leaves and spikes) were harvested for transcriptional analysis and for the physiological evaluations.

All plants were kept at 80% RSWC until the end of heading phase (Z60, Zadoks scale), when for a plot of plants the water administration was stopped to reach a 30% of the RSWC which is considered a severe DS condition. The time since the plants reach RSWC 30% (7 days for both cultivars) was considered as the acclimation to DS and named accordingly. All the measurement and tissue sampling for both the physiologic and transcriptional analysis were performed after 2 hours of the beginning of the light-period at three time-points: 1) T0 or acclimation to DS; 2) 4 days after stress (4 das); 3) 14 days of DS (14 das) (Figure 3.1). For each time-point spikes and flag leaves were harvested and immediately frozen in liquid nitrogen, being kept at -80°C until RNA extraction. The leaf below the flag leaf was used for the physiologic evaluation. For each variety and time-point, from 3 to 4 plants were sampled.

3.2.2 Physiologic analysis

In experiment (2) physiological response to drought was investigated by measuring the RWC and the chlorophyll content (SPAD unit) of the leaf below the flag leaf sampled for the transcriptional analysis. The SPAD measurement and the excision of the leaf disks for the RWC were done immediately after the sampling of tissues for the transcriptional analysis, always at the same time of the day (2 hours after the start of the light period), for each time-point, condition and genotype. Each physiological analysis was performed on 3 different biological replicates. Student's t test was applied to compare data of control and stressed plants.

RWC

Leaf RWC (%) was measured in control and stressed leaves of both Cappelli and Colosseo cultivars. Five disks (Ø=1 cm) were clipped from each leaf and their fresh weight (FW) was immediately recorded. Then leaf-disks were soaked for 4 h in distilled water at room temperature under constant light, and the turgid weight (TW) was recorded. Tissues were dried for 24 h at 80°C to record the total dry weight (DW). RWC was calculated according to the formula $RWC\ (\%) = \left(\frac{FW-DW}{TW-DW}\right) * 100$.

Chlorophyll Content

Chlorophyll content was measured in control and stressed leaves of both Cappelli and Colosseo cultivars with a SPAD-502 Minolta (Spectrum Technologies Inc., Plainfield, IL, USA) according to the manufacture instruction. Ten measures for each leaf were recorded along all the leaf length.

3.2.3 RNA extraction and retrotranscription

For both experiments (1) and (2), RNA was extracted with RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) from all samples, and visualised on agarose gel (1% agarose w/v, 1X TAE buffer, GelRed™ as Nucleic Acid gel stain) to assess its integrity.

For each condition, 1 µg of total RNA samples extracted from 100 mg of frozen tissue were retro-transcripted into cDNA with QuantiTect Reverse Transcription Kit (Qiagen) according the manufacturer instructions.

3.2.4 Quality control of cDNA

Firstly cDNAs were tested by PCR to assess the absence of genomic DNA using primer pairs ACT-Fw/ACT-Rev of the *TaActin* gene (AB181991) (Table 3.1) designed on two different exons of the target sequence. PCR reactions were set up in 10µL containing 1µL of 1:10 dilution of cDNA, 0.2 µM of each

primer and and 1X GoTaq® Colorless Master Mix (Promega). PCR amplifications were performed on each cDNA as follows: 1) initial denaturation at 95° for 2 min, 2) 35 cycles at 95 °C for 30 s, 60°C for 45 s, and elongation at 72 °C for 20 s, and 3) 1 cycle at 72° for 5 min. PCR products were checked on agarose gel (1.5% w/v, 1X TAE buffer, GelRed™ as Nucleic Acid gel stain).

3.2.5 Real-time quantitative RT-PCR

For both the experiments (1) and (2), the analysis was performed as a two-step real-time quantitative RT-PCR (RT-qPCR) to quantify the transcripts level in the cDNAs obtained. 1 uL of 1:10 dilution of cDNA was used for Real Time (RT) qPCR analysis by using the ABI PRISM® 7000 Sequence Detection System (Applied Biosystems, Foster City, USA). The reactions were set up in 10 µL reaction volume with 2X PowerUp SybrGreen Master Mix (Life Technologies, Carlsbad, CA, USA) and 30 nM of gene-copy specific forward and reverse primers ABA7-A_Fw1/ABA7-A_Rev1, ABA7-B_Fw1/ABA7-B_Rev1, BCAT-A_Fw1/BCAT-A_Rev1, BCAT-B_Fw1/BCAT-B_Rev1, ACT-Fw/ACT-Rev, DHN15.3_For/DHN15.3_Rev (Table 3.1); manufacturer indications were followed for the cycling conditions. *TdActin* was used as house-keeping gene for normalization; the expression of durum wheat drought-related gene *TdDHN15.3* (AM180931.1) was measured to evaluate the stress-status of the tissues.

Table 3.1 List of primers used for the Real time qPCR

Target	Primer pair		5'-3' sequence		Amplicon size (bp)	
	Forward	Reverse	Forward	Reverse	gDNA (*)	cDNA
TdABA7-A	ABA7-A_Fw1	ABA7-A_Rev1	GACTGCCTCGTGAAATTGCTGT	CCATGTTCCTGCTCAGATCTGC	87	87
TdABA7-B	ABA7-B_Fw1	ABA7-B_Rev1	ACGAGACCCATGAGCAGGTT	TCTCTCCGATACCTTTCATCACCTC	132	132
TdBCAT-A	BCAT-A_Fw1	BCAT-A_Rev1	GACAAGGTACGAGTTCAGGACC	CGACCGTCCATCCCTTCTTG	110	110
TdBCAT-B	BCAT-B_Fw1	BCAT-B_Rev1	ACAAGGTACGAGTTCAGGGCT	TAACCGTCCATCCTTTCTTGTCCT	109	109
TdACT	ACT-Fw	ACT-Rev	CTTGTATGCCAGCGGTCGAA	TGAGGAAGCGTGTATCCCTCG	173	96
TdDHN15.3	DHN15.3_For	DHN15.3_Rev	GGAGGAGGAAGAAGGGCATCA	CATCCCTGCCGTATGACCTTG	116	116

(*) *genomic DNA*

All primers were designed with Primer3 (Untergasser et al., 2012) on the target sequences retrievable on the NCBI for *TdDHN15.3* (AM180931.1) and *TaActin* gene (AB181991), or according to the coding sequence of genes *TdABA7-A*, *TdABA7-B*, *TdBCAT-A*, *TdBCAT-B* after their *in silico* characterization.

The absence of non-specific PCR products and primer dimer artefacts was checked by melting curves to verify the presence of a single peak for all the sequence-specific primer pairs listed in Table 3.1. The

efficiency of each primer pair was evaluated according to the manual provided by ABI PRISM® 7000 Sequence Detection System (Applied Biosystems).

The relative quantitation analysis was determined by using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). We considered as relevant modulations of the gene expression only fold-change values (i.e. $2^{-\Delta\Delta Ct}$) greater than 2-fold ($FC \leq 0.5$ and $FC \geq 2$). The hierarchical clustering analysis (Eucliden distance matrix, average linkage) of log2-transformed fold-change values was carried out and plotted using MultiExperiment Viewer (MeV) software (version 4.9).

3.3 TILLING approach

3.3.1 SNP detection by *in silico* TILLING

The sequences of *TdABA7* and *TdBCAT* obtained as described in Paragraph 3.1 were used as query for searching the wheat TILLING database (www.wheat-tilling.com) to search allele variations in *TdABA7-A, TdABA7-B, TdBCAT-A* and *TdBCAT-B.* For each Kronos mutant line identified, the predicted mutations and the effects on both the nucleotide and protein sequences were manually verified on the sequence of all the target genes.

3.3.2 Plant material and growth conditions

The mutant lines selected in the wheat TILLING database and the parental wild type cv Kronos were provided by Dr. Cristobal Uauy (JIC, Norwich UK). From 15 to 35 seeds for each mutant line and 20 seeds for each parental wild type were sown. Seeds were surface sterilized by immersion in sodium hypochlorite (0.5% v/v) for 30 min, and then rinsed thoroughly in sterile water and sown in Petri dishes in imbibition conditions for 10 days at 5 °C. Seedlings were then transferred in peat and sand mix (85% fine peat, 15% Grit, 2.7 Kg/m^3 Osmocote, wetting agent, 4 kg/m^3 Maglime, 1 Kg PG mix) in three 96-wells trays and grown under controlled conditions (L:D 10:14, 6°C, RH=70%). At seedling stage (Z11-Z12 of Zadoks scale) the plants were transplanted in John Innes Cereal Mix (40% medium grade peat, 40% sterilized soil, 20% horticultural grit, 1,3 kg/m^3 PG Mix 14-16-18 + Te base fertilizer, 1 kg/m^3 Osmocote Mini 16-8-11 2 mg + Te 0.02% B, wetting agent, 3 kg/m^3 Maglime, 300 g/m^3 Exemptor) and grown till grain harvesting. For the DNA extraction leaves were sampled taking 2-4 cm of tissues at Z11-Z12. For the DNA extraction protocol see paragraph 3.6.1

3.3.3 SNP validation through KASP assay

The SNP assay (KASP assay 1-7 in

Table 3.3) was conducted for each genotype of the M4 cv. Kronos to confirm the mutation and the zygosity predicted from the wheat TILLING resource.
For details on the KASP assay see paragraph 3.6.2.
The plates were read on a Tecan Safire plate reader. Fluorescence was detected at room temperature. Data analysis was performed manually using Klustercaller software (version 2.22.0.5, LGC Genomics). Additional amplification cycles were conducted and the sample when necessary and then plates were read again.

3.3.4 Generation of double mutant lines and back-crosses

Double mutant lines were generated between TILLING mutant lines carrying mutation on the A or B copy of the same gene (i.e. *TdABA7-A* or *TdBCAT),* and then backcrossed with the recurrent wild-type parent cv Kronos when possible. The emasculation and the subsequent pollination have been carried out in the glasshouse and all the crossed spikes have been closed with a cellophane bag to avoid cross-contamination and seeds dispersion. At maturation, the F_1 spikes were smpled and dried at 30 °C for 1 month before seed harvesting.

KASP assays were performed on the F_1 plants to test the presence of the mutation on putative double mutant lines, by using the markers 1-7 reported in Table 3.3. For details on DNA extraction and KASP genotyping see paragraph 3.6.
The putative double-mutant lines are tested with two KASP markers, one for each mutation (A and B gene). The KASP endpoint genotyping analysis was conducted on Real Time PCR 7900HT (Life Technologies Foster City, CA, USA) for the data capture following the LGC guide “Guide to running KASP genotyping on the ABI7900 instrument”. The wild-type DNA (cv. Kronos) and the parental Hom mutant DNA were included in the analysis as reference.

3.4 Allele mining by Targeted-Resequencing to find natural variation of *TdBCAT* and High-throughput Phenotyping

3.4.1 Plant material

This part of the work was done in collaboration with CNR-IBBR and ALSIA Metapontum Agorbios, the National node of the Plant Phenotyping Infrastructure "Phen-Italy". Plant growth, stress and monitoring were done at the ALSIA station.

A set of durum wheat landraces and durum wheat cultivars was selected for the allele mining strategy with the targeted-resequencing by NGS. The germplasm panel comprised 33 durum wheat genotypes, selected from a 452 genotypes core set, named SSD collection, produced by single seed descent from a worldwide germplasm collection (Pignone et al., 2015). In the recent years, the whole core set of this germplasm collection had been analysed in the ALSIA phenotyping centre in Metaponto. The NIR image-based indices acquisition which can measure the water content of plant tissues after the image convolution, led to a selection of the durum wheat genotypes used in the present study that are highly representative of the variation of response to DS of the entire SSD collection (Janni et al., personal communication). Four reference Italian varieties (Svevo, Saragolla, Cappelli, Colosseo) and the American variety (Kronos) were included with the 33 SSD durum wheat genotypes (Table 3.2) for the allele mining approach.

Only 35 out of the 38 durum wheat genotypes, due to seed availability problems, were furtherly investigated for the phenotype evaluation under DS condition on the high-throughput phenotyping platform at the Research Center Metapontum Agrobios (ALSIA).

3.4.2 DNA extraction

Genomic DNA of cv Cappelli, Colosseo, Saragolla, Svevo and Kronos was extracted with the GenElute™ Plant Genomic DNA Miniprep Kit (Sigma-Aldrich) from leaves tissues of durum wheat genotypes and quantified through electrophoresis on agarose gels 1.2% (w/v) in TAE buffer 1X using λ genomic DNA as reference.

3.4.3 Isolation of *TdBCAT-A* and *TdBCAT-B*

Target regions (total 6.8 Mb) are both promoter and coding sequences of *TdBCAT-A* and *TdBCAT-B*, which were PCR-amplified from all the durum wheat genotypes listed in Table 3.1 using TaqDNA Polymerase (New England BioLabs, Hitchin, UK) in 25 μL reactions each containing 10–20 ng template

DNA, 0.2 μM of each forward and reverse primers, 0.2 mM of each dNTPs, 1X Standard Taq Buffer (New England BioLabs). All the DNA of the SSD genotypes was kindly provided by IBBR-CNR (Bari, Italy).

PCR conditions were optimized for different groups of primers (Table 3.2) as follows:

- Primer pairs TdBCAT_U2F/TaBCAT_a2R and TdBCAT_U2F/TaBCAT_b1R
  PCR amplifications were performed as follows: 1) initial denaturation at 95° for 2 min 2) 35 cycles at 95 °C for 30 s, annealing for 60 s, elongation at 72 °C for 2 min, 3) 1 cycle at 72°C for 10 min. The list of optimized annealing temperatures for each primer pair are provided in Table 3.3.
- Primer pairs bcatA_PrF1/bcat_UR6, BCATB_PrF3/bcat_UR6 and bcat_UF2/bcat_UR5
  PCR amplifications were performed with a touch-down PCR as follows: 1) initial denaturation at 95° for 2 min 2) 10 cycles at 94 °C for 30 s, annealing at 65°C for 30 s with a drop -0.5°C/cycle, elongation at 72 °C for 2 min, 3) 35 cycles of 94 °C for 30 s, annealing for 30 s, elongation at 72 °C for 2 min, 4) 1 cycle at 72°C for 10 min.
  The list of optimized annealing temperatures of stage 3) of the touch-down PCR is provided in Table 3.3.

The PCR products were electrophoretically separated through a TAE agarose gel to check the correct size of amplicons. PCR product obtained with primer pairs TdBCAT_U2F/ TaBCAT_a2R, TdBCAT_U2F/TaBCAT_b1R, bcatA_PrF1/bcat_UR6 and BCATB_PrF3/bcat_UR6 were sequenced with BigDye™ Terminator v3.1 Cycle Sequencing Kit (Thermofisher Scientific, Waltham, MA) according to the manufacturer instructions. For primer pair bcat_UF2/ bcat_UR5 the presence of both targets was confirmed with restriction fragment length polymorphism by the digestion with BmgBI (New England BioLabs) according to the instructions.

3.4.4 DNA pooling and sequencing of targets

A total of 5 PCR reactions were performed in 96-well plate, according to the scheme in Figure 3.2. The 5 target regions were randomly pooled in 4 tubes by a two-step pooling. Final pools represent all the targets sequences from different genotypes as follows:

1. pool 1 (SSD 35, SSD 92, SSD 122, SSD 171, SSD 253, SSD 269, SSD 416, SSD 487. SSD 494 and cv. Kronos);
2. pool 2 (SSD 44, SSD 109, SSD 178, SSD 322, SSD 343, SSD 397, SSD 415, SSD 511, cv. Colosseo);
3. pool 3 (SSD 69, SSD 99, SSD 116, SSD 244, SSD 335, SSD 409, SSD 441, SSD 451, cv. Cappelli and cv. Saragolla);
4. pool 4 (SSD 64, SSD 112, SSD 135, SSD 195, SSD 278, SSD 325, SSD 459, SSD 499, Svevo).

	1	2	3	4	5	6	7	8	9	10	11	12
a	SSD 35	SSD 116	SSD 269	SSD 415	SSD 511							
b	SSD 44	SSD122	SSD 278	SSD 416	Saragolla							
c	SSD 64	SSD 135	SSD 322	SSD 441	Svevo							
d	SSD 69	SSD 171	SSD 325	SSD 451	Colosseo							
e	SSD 92	SSD 178	SSD 335	SSD 459	Cappelli							
f	SSD 99	SSD 195	SSD 343	SSD 487	Kronos							
g	SSD 109	SSD 244	SSD 397	SSD 494								
h	SSD 112	SSD 253	SSD 409	SSD 499								

pool 1
pool 2
pool 3
pool 4

Figure 3.2 Scheme of the 96-plate of each PCR reaction used for targeting *TdBCAT*. Colours indicate the pool in which the relative genotype (SSD or cultivar) is represented.

3.4.5 Bioinformatics analysis of the NGS data

The paired-end (PE) reads raw data from the NGS were checked for QC using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The two fastq files of the PE of each sample were aligned to Kronos reference genome v.1.1 (https://opendata.earlham.ac.uk/opendata/data/Triticum_turgidum/EI/) using BWA (version 0.7.17) with parameters as default to generate the SAM files. Samtools (version 1.7) was used to convert the SAM to sorted BAM files for each sample. The BAM files were directly visualized with the Integrative Genomic Viewer IGV v. 2.4 (Broad Institute) for the SNP identification.

3.4.6 SNP assay

SNPs with a frequency ≥5% for each pool were targeted with KASP molecular markers Kompetitive Allele Specific PCR (KASP) (LGC Genomics, Teddington, UK) designed with Primer3 (Untergasser et al., 2012) on the basis of the reference sequence of durum wheat cv. Kronos. A full list of primers is provided in Table 3.3 (KASP ASSAY 8-27).

The KASP endpoint genotyping analysis was conducted on Real Time PCR 7900HT (Life Technologies). The reference DNA (cv. Kronos) was included in the analysis as reference for all the assays. For details on the KASP genotyping see paragraph 3.6.2.

3.5 Phenotyping

Seeds of 33 genotypes of the SSD collection (Pignone et al., 2015) were germinated at room temperature for maximum of 4 days on moist filter paper in Petri dishes and then transplanted into polystyrene plateaus. The plateaus were then stored at 4 °C for two weeks in order to synchronize the plantlets growth. Individual plants were then transferred into two liter pots filled with a 1:1(v/v) mixture of river sand and peat moss until a total weight of 1200 g. Then plants were grown in glasshouse under natural light conditions, and environmental conditions were monitored every 30 minutes using a datalogger (Watchdog Model 450, Spectrum Tecnologies, Inc.).

Three treated and three control replicates of each accession were randomly distributed in greenhouse to minimize through spatial distribution the possible establishment of microclimatic variation in the greenhouse. To identify each single pot a barcode was applied in proper position to allow automatic reading of plant identifier.

Wheat plants were manually kept fully irrigated up to the booting stage (Z45, 104 days after sowing, DAS), then DS was imposed for 43 days (up to 147 DAS) by maintaining the amounts of water in the soil around 50% field capacity (FC) .

3.5.1 Phenotyping analysis

Images in the RGB domain (white light) for high-throughput phenotyping were captured every other day following Petrozza et al. (2014), and using a Scanalyzer 3D device (LemnaTec GmbH, Aachen, Germany).

Three imaging chambers with different levels of illumination, near-infra-red (NIR), white light (RGB), and fluorescence (UV), were used. For each chamber three images were taken, one from above the plant and two laterally at an orthogonal angle. The images collected in the NIR chamber were used to evaluate the plant water content. Images from the RGB chamber were used to evaluate the state of health of the plant via color classification (i.e. digital biovolume), as well as for morphometric measurements, such as digital biovolume and height.

The SSDs plants were monitored through image acquisition at two-day intervals from 55 DAS up to 147 DAS for a total of 92 days.

3.6 Genotyping with KASP molecular markers

3.6.1 DNA extraction

Genomic DNA was extracted from all samples obtained as described in 1.3.2. The extractions were performed in 96-wells 1.5 mL plates by adding 500 µl of extraction buffer (1M Tris-HCl pH 7.5, 0.5 M EDTA, 10% SDS) at 65 °C and grounding the samples in the 2010 Geno/Grinder (SPEX Sample Prep, Metuchen, NJ) at 1500 strokes min^{-1} for 90 s. After 1 h incubation at 65 °C followed by 15 mins at 4 °C, 250 µl of 6 M ammonium acetate were added and the plates were leaved 15 mins at 4 °C for protein precipitations. The plates were centrifuged at 4000 rpm for 15 mins and 600 µl of supernatant containing the DNA were recovered in new plates containing 360 µl of cold isopropanol. The plates were than mixed and leaved 5 mins at room temperature for DNA precipitation. After 15 mins centrifugation at 4000 rpm the supernatant was discarded and the DNA pellet was washed in 500 µl of 70% ethanol. The plates were centrifuged again at 4000 rpm for 15 mins, the supernatant was discarded and the DNA pellet was dried and dissolved in 100 µl of sterile water. The DNA was vortex, centrifuged at 4000 rpm for 5 mins and diluted 1:10 before using for the SNP assay.

3.6.2 SNP assay

Kompetitive Allele Specific PCR (KASP) (LGC Genomics, Teddington, UK) were designed with Primer3 (Untergasser et al., 2012) to target the selected SNP on the mutant lines. Oligonucleotides were purchased from Sigma-Aldrich (Gillingham, UK), with primers carrying standard FAM or HEX compatible tails (FAM tail: 5' GAAGGTGACCAAGTTCATGCT 3'; HEX tail: 5' GAAGGTCGGAGTCAACGGATT 3'). The full list of primers is provided in table 3.3). The KASP assay mix was prepared as recommended by LGC Genomics: 46 µl dH_2O, 30 µl of 100 µM common primer and 12 µl of each 100 µM allele-specific primer. The reactions were set up according as recommended by LGC Genormics in 4 µl or 10 µl of final volume, depending on the plate-format used (4 µl for the 384-well plate and 10 µl for the 96-well plate) with 10-20 ng of DNA to test

Two technical replicates were done for each genotype, using the DNA of the Kronos wild type as reference for the signal of the allele 1.

PCR cycling was performed following two protocols depending on the KASP assay:

1) KASP assay 1-22, 26 (table 3.3): 1 cycle at 94 °C for 15 mins, followed by 10 touchdown cycles (94 °C for 20 s; touchdown 65 °C, -0.8 °C per cycle, 60 s) then followed by 30 cycles of amplification (94 °C 20 s; 57 °C 60 s).

2) KASP assay 23-25, 27 (table 3.3): 1 cycle at 94 °C for 15 mins, followed by 10 touchdown cycles (94 °C for 20 s; touchdown 61 °C, -0.6 °C per cycle, 60 s) then followed by 30 cycles of amplification (94 °C 20 s; 55 °C 60 s).

Table 3.1 List of genotypes used for the allele mining. The genotypes that were phenotyped with the High-throughput phenotyping system are indicated

SSD Entry	Origin
35	Algeria
44	Tunisia
64	Morocco
69	Morocco
92	Usa-ND North Dakota
99	Ethiopia
109	Iraq
112	Iraq
116	Iraq
122	Usa-ND North Dakota
135	Turkey
171	Peru
178	France
195	Saudi Arabia
244	Ethiopia
253	Cyprus
269	Iran
278	Bulgaria
322	Turkey
325	Syria
335	Iraq
343	Iran
397	Crete
409	Grece
415	Crete
416	Grece
441	Crete
451	Iraq
459	USA
487	Grece
494	Grece
499	
511	Lybia
Colosseo	Italy
Kronos	USA
Cappelli	Italy
Saragolla	Italy
Svevo	Italy

Table 3.2 List of primer pairs used for the targeting of *TdABA7* and *TdBCAT* gene sequences. For each primer pair it is indicated the name and 5'-3' sequence, the optimized annealing temperature, the PCR product size, the targeted gene and for which application *(b)* the primer pair was used

Target gene	Forward primer (5'-3' sequence)	Reverse primer (5'-3' sequence)	T ann (°C) *(a)*	Amplicon lenght (bp)	Application *(b)*
TdABA7-A	TdABA7-A1F (TATCCAAGCATGTGTCAGC)	TdABA7-A1R (GTACAGACATGCACGTGTA)	56	805	*
TdABA7-B	TdABA7_B1F (CATCCAAGCATGCCTCAGA)	TdABA7_B1R (CTGTTGCTACAATAACTCGGC)	58	798	*
TdBCAT-A	TaBCAT_A1F (GGTCAATTGCCATCTCCGT)	TaBCAT_A1R (GACGTACATGTAGTCGGTCG)	58	1093	*
TdBCAT-A	TaBCAT_A2F (GTGGAAGTCGTCGCAGCC)	TaBCAT_A2R (GAGCTCGATGATGCTCCGA)	61	1202	*
TdBCAT-A	TaBCAT_U1F (CTCAAGGCGCAGATGGAC)	TaBCAT_A3R (TTACAAGCCTCACTCTCAGCA)	58	940	*
TdBCAT-B	TdBCAT_U2F (CTCCACGACCTATATAGAAACCTC)	TdBCAT_B3R (CGGCCTGATGTACAATGCT)	58	1289	*
TdBCAT-B	TaBCAT_A2F (GTGGAAGTCGTCGCAGCC)	TaBCAT_B1R (GAGCTCGATGATGCTCCTG)	62	1201	*
TdBCAT-B	TaBCAT_U1F (CTCAAGGCGCAGATGGAC)	TaBCAT_B2R (CTAGCTAGGCACCATAGAGAA)	58	888	*
TdBCAT-A	bcatA_PrF1 (GGTAGTGGCCGACGAAGG)	bcat_UR6 (CAGACGACAGCACAGCCAT)	56 *(c)*	1661	**
TdBCAT-B	BCATB_PrF3 (GCCTTGACACGACGACTTTC)	bcat_UR6 (CAGACGACAGCACAGCCAT)	56 *(c)*	899	**
TdBCAT-A	TdBCAT_U2F (CTCCACGACCTATATAGAAACCTC)	TaBCAT_a2R (GAGCTCGATGATGCTCCGA)	56	1722	**
TdBCAT-B	TdBCAT_U2F (CTCCACGACCTATATAGAAACCTC)	TaBCAT_b1R (GAGCTCGATGATGCTCCTG)	56	1760	**
TdBCAT-A	bcat_UF2 (GAGATCCACCGCGCCATG)	bcat_UR5 (CAGTAGCATCAGAGAAACCATC)	60 *(c)*	945	**
TdBCAT-B				925	**

(a) Annealing temperature,
*(b) Application: * gene characherization,see paragraph 3.1.1** target-enrichment for NGS, see paragraph 3.4.3;*
(c) Final annealing temperature (x35 cycles) of Touch-down PCR

Table 3.3 Full list of KASP assay used for the SNP assessment of TILLING lines and the allele mining by NGS. Each KASP assay is composed of three primers: a common primer and two allele-specific primers carrying at the 5' the FAM or VIC/HEX complementary sequence which is not reported in the table (FAM tail: 5' GAAGGTGACCAAGTTCATGCT 3', HEX tail: 5' GAAGGTCGGAGTCAACGGATT 3')

Target	KASP Assay	Target SNP *	5'-3' sequence			Use ****
			Common	Allele 1 **	Allele 2 ***	
TdABA7-A	1	8 *(a)*	ATGTCACTGCCAAGCTTGTTC	(VIC/HEX)-GCTTTGGGGAGCATAAGC	(FAM)-GGCTTTGGGGAGCATAAGT	A
TdABA7-A	2	10 *(a)*	CACCTCTTCATGATCCCTGC	(VIC/HEX)-AGCATAAGCAGCTCAACTGG	(FAM)-AGCATAAGCAGCTCAACTGA	A
TdABA7-B	3	20 *(b)*	AGATACGCGGAGCATTGGA	(VIC/HEX)-GGACGGTAGCAAGGATTTTG	(FAM)-TGGACGGTAGCAAGGATTTTA	A
TdABA7-B	4	14 *(b)*	GGCCACGAGACCCATGAG	(VIC/HEX)-CTGAGATAGGACATCCTCTGATC	(FAM)-CTGAGATAGGACATCCTCTGATT	A
TdBCAT-A	5	10 *(c)*	GGAACGATGAGCCTGCATGC	(VIC/HEX)-GGAAGTCGTCGCAGCCGC	(FAM)-TGGAAGTCGTCGCAGCCGT	A
TdBCAT-B	6	16 *(d)*	CAGTGCTGAGATATGTGCGTT	(VIC/HEX)-CCAGCCCTGAACTCGTACC	(FAM)-CCCAGCCCTGAACTCGTACT	A
TdBCAT-B	7	5 *(d)*	CGGCCATGCTTTGTACGTAG	(VIC/HEX)-GTCAGCGGCGGCCTCC	(FAM)-GTCAGCGGCGGCCTCT	A
TdBCAT-A	8	A3 *(e)*	TGGCAAATCCCCATCCCTC	(VIC/HEX)-ACAAAGGGCTCTGATGGGG	(FAM)-GACAAAGGGCTCTGATGGGT	B
TdBCAT-A	9	A5 *(e)*	CGACAAAGGGCTCTGATGGG	(VIC/HEX)-TTCCCATCGAAACGCGCC	(FAM)-TTCCCATCGAAACGCGCG	B
TdBCAT-A	10	A6 *(e)*	CGACAAAGGGCTCTGATGGG	(VIC/HEX)-TCCCTCGTTTCCCATCGAAA	(FAM)-TCCCTCGTTTCCCATCGAAG	B
TdBCAT-A	11	A8 *(e)*	AGCACAACCTCTTAGTCCCC	(VIC/HEX)-GGCGACGCCATTGACT	(FAM)-GGCGACGCCATTGACG	B
TdBCAT-A	12	A9 *(e)*	TTTTCTAGTGGTATATCGGTGTGA	(VIC/HEX)-GCTAAAACTTCCCTACGCCATTG	(FAM)-GCTAAAACTTCCCTACGCCATTA	B
TdBCAT-A	13	A10 *(e)*	TTTTCTAGTGGTATATCGGTGTGA	(VIC/HEX)-GACGGGGCTGCGTGGA	(FAM)-GACGGGGCTGCGTGGG	B
TdBCAT-A	14	A11 *(e)*	TTTTCTAGTGGTATATCGGTGTGA	(VIC/HEX)-ACAGTTCCTACCAAAATTACCATATC	(FAM)-ACAGTTCCTACCAAAATTACCATATT	B
TdBCAT-A	15	A13 *(e)*	ACACCATCCGACCATGCTAG	(VIC/HEX)-GAAGGTGAGACGACGGCA	(FAM)-AAGGTGAGACGACGGCG	B
TdBCAT-A	16	A21 *(e)*	GTCTCGGTGGTGCGTCTA	(VIC/HEX)-CCAGAGACAGACATCCACACA	(FAM)-CAGAGACAGACATCCACACG	B
TdBCAT-A	17	A22 *(e)*	GTCTCGGTGGTGCGTCTA	(VIC/HEX)-AGACAATCCCCGACCGAATT	(FAM)-AGACAATCCCCGACCGAATC	B
TdBCAT-A	18	A23 *(e)*	GTCTCGGTGGTGCGTCTA	(VIC/HEX)-GCCGCATGAACGTAGATGA	(FAM)-GCCGCATGAACGTAGATGG	B
TdBCAT-A	19	A24 *(e)*	CACCGTCCCTCCCTTGTC	(VIC/HEX)-TCTTCATCAGCGGTGTTTGC	(FAM)-CTCTTCATCAGCGGTGTTTGT	B
TdBCAT-A	20	A25 *(e)*	CACCGTCCCTCCCTTGTC	(VIC/HEX)-CTGGTGCGCTGGTCCTATAG	(FAM)-CTGGTGCGCTGGTCCTATAT	B
TdBCAT-A	21	A27 *(e)*	TCCCGAAAATACTTCCATTCTGT	(VIC/HEX)-CCTTCGGCTCACTTCAATTC	(FAM)-CCTTCGGCTCACTTCAATTT	B
TdBCAT-A	22	A28 *(e)*	GGAGGGACGGTGACAGT	(VIC/HEX)-ATAATAAAAGATACAGTCAAATCCAC	(FAM)-ATAATAAAAGATACAGTCAAATCCAT	B
TdBCAT-B	23	B1 *(e)*	GTTGGAGCTGGGCATAACTTA	(VIC/HEX)-ACGACGACTTTCCGACTATCTA	(FAM)-CGACGACTTTCCGACTATCTG	B
TdBCAT-B	24	B2 *(e)*	GCCCAGCTAATTTTTTCCAA	(VIC/HEX)-ATTTTTTATTATTTTAGGTGTTCATT	(FAM)-ATTTTTTATTATTTTAGGTGTTCATA	B
TdBCAT-B	25	B3 *(e)*	GCCCAGCTAATTTTTTCCAA	(VIC/HEX)-GGTGTTCATTGTACTGTTATAATATT	(FAM)-GGTGTTCATTGTACTGTTATAATATA	B

Table 3.3 continue

Target	KASP Assay	Target SNP *	5'-3' sequence			
			Common	**Allele 1 ****	**Allele 2 *****	**Use ******
TdBCAT-B	26	B4 *(e)*	CAGCGAGTTAAAATGGCAACC	(VIC/HEX)-GCCAAATTCTCCACGACCTAA	(FAM)-GCCAAATTCTCCACGACCTAT	B
TdBCAT-B	27	B5 *(e)*	CGAGTACCTAGCGGCTCC	(VIC/HEX)-CGAGCAATATAATGGCCAGCC	(FAM)-CGAGCAATATAATGGCCAGCT	B

* For relative details on each SNP see table (a)Table 4.9 (b) Table 4.10(c)Table 4.11 (d) Table 4.12 (e)Table 4.19

**Allele 1 is the wt allele for TILLING lines and the reference allele for the allele mining by NGS

*** Allele 2 refers to the mutant allele for the TILLING lines, and to the alternative allele for the allele mining by NGS

**** A = genotyping of TILLING mutants, B= SNP identification with the allele mining by NGS

4 RESULTS AND DISCUSSION

4.1 Identification and isolation of durum wheat sequences of *TdBCAT and TdABA7* genes

The isolation of genes *TdBCAT* and *TdABA7* from *T. durum* was pursued by preliminary *in silico* analysis of the sequences available in *T. aestivum* databases, followed by the physical isolation and sequencing of the *T. durum* target regions.

The sequences *HvBCAT-1* (AJ971373) and *HvABA7* (AJ296273), previously identified in barley were used to retrieve the orthologous sequences of bread wheat (Chinese Spring) from the Ensembl Plants database on the basis of the CSS assembly (IWGSC, 2014). In Table 4.1 the accession numbers of *TaBCAT* (A and B) and *TaABA7* (A and B) are reported in relation to the different wheat genome assemblies available in the Ensembl Plants database: the CSS assembly (IWGSC, 2014), the TGAC v.1 (Clavijo et al., 2017) and the RefSeq v.1 (IWGSC, 2018).

Table 4.1 Matches of the durum wheat genes *TdBCAT-A*, *TdBCAT-B*, *TdABA7-A* and *TdABA7*-B on the Ensembl Plants database for the three different assemblies of *T. aestivum*: CSS (IWGSC, 2014), TGAC v1 (Clavijo et al., 2017), and RefSeq v1 (IWGSC, 2018).

GENE	CSS	TGAC v.1	RefSeq v.1
TdBCAT-A	Traes_4AS_6C0AB5E18	TRIAE_CS42_4AS_TGACv1_307232_AA1018650	TraesCS4A01G059800
TdBCAT-B	Traes_4BL_FC8629A72	TRIAE_CS42_4BL_TGACv1_321464_AA1060900	TraesCS4B01G235400
TdABA7-A	N/A	TRIAE_CS42_6AS_TGACv1_485170_AA1538760	TraesCS6A01G008000
TdABA7-B	N/A	TRIAE_CS42_2BS_TGACv1_146862_AA0474150	TraesCS2B01G066200

The BLASTn analysis was performed using as query the coding sequence (CDS) of the barley genes, on the *T. aestivum* genomic database. The best matches were selected based on scores, chromosome position, and transcript annotation (Traes), if available. Top-scoring matches were then used to reconstruct *in silico* the bread wheat sequences of *BCAT* (Table 4.2) and *ABA7* genes (Table 4.3).

Table 4.2 Summary of the *in silico* analysis on Ensembl Plants database (IWGSC CSS assembly) for the gene *BCAT*. The genomic location, the identification (Traes) of the overlapping genes, the % of identity between the query *HvBCAT-1* and the target, as well as the corresponding match on the gene structure of *HvBCAT-1*, are reported.

Genomic location *	**Overlapping gene***	**%ID**	***HvBCAT-1* exon**
4A:24465877-24465914	Traes_4AS_6C0AB5E18	92.1	1
4D:22280222-22280253	Traes_4DL_9D7A18BBC	93.8	
4A:24466205-24466271	Traes_4AS_6C0AB5E18	98.5	2
4D:22279863-22279929	Traes_4DL_9D7A18BBC	98.5	
4A:24466438-24466803	Traes_4AS_6C0AB5E18	96.2	3
4D:22279324-22279689	Traes_4DL_9D7A18BBC	95.4	
4D:22278974-22279208	Traes_4DL_9D7A18BBC	97.0	4
4A:24466921-24467154	Traes_4AS_6C0AB5E18	95.7	
4A:24467240-24467384	Traes_4AS_6C0AB5E18	95.9	5
4D:22278746-22278890	Traes_4DL_9D7A18BBC	95.9	
IWGSC_CSS_4BL_scaff_7034480:3-147	Traes_4BL_FC8629A72	94.5	
4A:24467551-24467667	Traes_4AS_6C0AB5E18	96.6	6
IWGSC_CSS_4BL_scaff_7034480:310-426	Traes_4BL_FC8629A72	94.9	
4D:222784479-22278595	Traes_4DL_9D7A18BBC	94.9	
IWGSC_CSS_4BL_scaff_7034480:310-426	Traes_4BL_FC8629A72	89.4	7
4D:22278259-22278371	Traes_4DL_9D7A18BBC	89.4	
4A:24467775-24467887	Traes_4AS_6C0AB5E18	87.6	

*****_Genomic location and the identification numbers of the overlapping genes refer to the IWGSC CSS assembly of bread wheat (Chinese Spring)_

Table 4.3 Summary of the *in silico* analysis on EnsemblPlants database (IWGSC CSS assembly) for the gene *ABA7*. The genomic location on IWGSC CSS assembly, the % of identity between the query *HvABA7* and the target, as well as the corresponding match on the gene structure of *HvABA7* are reported.

Genomic location	**Overlapping gene**	**%ID**	**HvABA7 gene exon**
IWGSC_CSS_6AS_scaff_4359967:1187-1678	N/A	91.7	1
6D:2060809-2061306	N/A	92.4	1
IWGSC_CSS_2BS_scaff_5169659:2865-3362	N/A	93.4	1

For *TdBCAT* the best matching sequences are on group 4 chromosomes of *T. aestivum*, in agreement with the map position of *HvBCAT-1* on chromosome 4H of barley (Malatrasi et al., 2006). In particular, the best matches are on the short arm of chromosome 4A (4AS), and on the long arm of both chromosome 4B (4BL) and 4D (4DL). This finding can be explained by the evolutionary history of the wheat genome. Indeed, a pericentric inversion took place on chromosome 4A, which caused the partial transfer of 4AL to 4AS (Devos et al., 1995).

The investigation of Ensembl Plants database allowed to identify the complete sequence of *TaBCAT-A*. For *TaBCAT-B*, it was predicted a 3-exons structure due to a lack of genomic sequence upstream the first exon of Traes_4BL_FC8629A72, which corresponds to exon 5 in *HvBCAT-1* (Figure 4.1), perfectly matching in terms of gene structure with the last three exons of *TaBCAT-A*. The genomic scaffold in which Traes_4BL_FC8629A72 is annotated starts 3 bases upstream the beginning of the CDS of this

gene (i.e.Traes_4BL_FC8629A72). The predicted protein product of both Traes_4AS_6C0AB5E18 and Traes_4BL_FC8629A72 is a BCAT enzyme of the metabolic pathway of BCAAs.

The conserved regions of Traes_4AS_6C0AB5E18 and Traes_4DL_9D7A18BBC were used for the primer design strategy, to retrieve the complete sequence of *TdBCAT-B*.

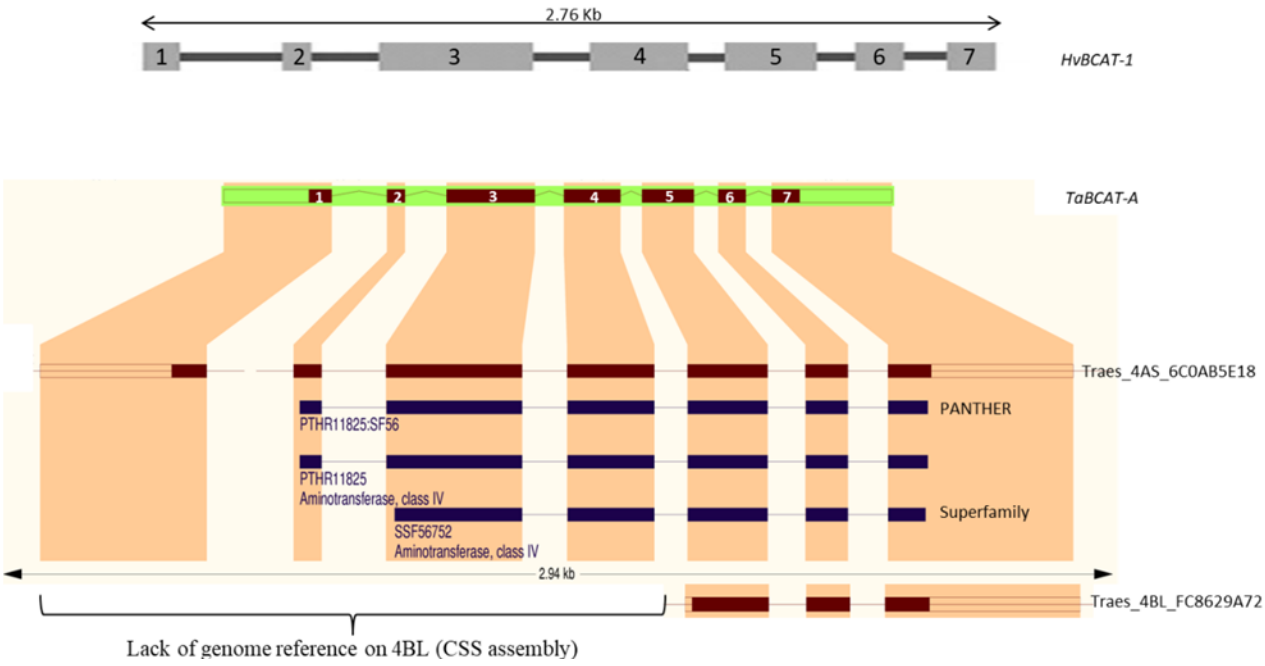


Figure 4.1 Comparison of the genes structures and the protein domains annotation (PANTHER and Superfamily repositories) for gene *TaBCAT-A* (Traes_4AS_6C0AB5E18) and *TaBCAT-B* (Traes_4BL_FC8629A72). On the top the gene model of gene *HvBCAT-1* is shown.

A similar analyses was done also for *TaABA7*, for which the best matching results were found on chromosomes 6AS, 6DS and 2BS of *T. aestivum*. These results are consistent with those of Gulli (1995) and co-workers, which localized *HvABA7* on chromosome 6H of barley, and on the distal ends of chromosome 6AS, 6DS and 2BS of wheat. The presence of the copy on chromosome 2BS is supported by the evidence of a 2BS-6BS reciprocal translocation of the distal ends of these chromosomes (Devos et al., 1993).

A 1-exon-structure was predicted in barley, the three best matches are summarized in Table 4.3. On the CSS assembly, no overlapping genes were found for these genomic regions, as these genes were not annotated at the time (December 2015). The more recent investigation on Ensembl Plants database, after the deposition of the TGACv1 (Clavijo et al., 2017) and RefSeqv1.1 (IWGSC et al., 2018) assemblies, allowed the identification of the annotated gene sequences for both the A and B copies of bread wheat (Table 4.1), while for the D-copy the annotation is still missing. Notwithstanding the big steps forward made with the sequencing projects, the function of this gene is still unknown.

4.2 Isolation and sequencing of *TdBCAT* and *TdABA7* from *T. durum*

The sequences of *TaBCAT* and *TaABA7* were used as reference for gene-specific primer design to isolate the genomic sequences of the orthologous genes *TdBCAT* (*A* and *B*) and *TdABA7* (*A* and *B*) from durum wheat cv. Cham1 as described in paragraph 3.1.

4.2.1 *TdBCAT*

In total, 2655 and 2407 bp were identified for *TdBCAT-A* and *TdBCAT-B* respectively. The gene structure, predicted with FGENESH+, consisted of 7 exons in both A and B sequences. The CDS is 1200 bp long encoding a protein of 399 amino acids, for both *TdBCAT-A* and *TdBCAT-B* (Table 4.4).

For *TdBCAT-A*, a 400 bp-long sequence upstream the predicted starting codon (ATG) was isolated, in which a transcription starting site (TSS) was predicted at -261 (considering +1 as the A of the starting codon), and a polyadenylation site was predicted 2085 bp downstream the ATG.

For *TdBCAT-B*, it was possible to isolate 151 bp upstream the ATG, in which a TSS was not predicted. A polyadenylation site was predicted 2066 bp downstream the ATG. The gene models are summarized in Table 4.4 and represented in Figure 4.2.

Table 4.4 Summary of *TdBCAT-A* and *TdBCAT-B* gene structures based on FGENESH+ analysis

	CDS (bp)	Protein (aa)	EXON 1	EXON 2	EXON 3	EXON 4	EXON 5	EXON 6	EXON 7
TdBCAT-A	1200	399	1-93	320-393	565-925	1046-1276	1365-1577	1678-1790	1897-2011
TdBCAT-B	1200	399	1-93	322-395	566-926	1048-1278	1366-1578	1675-1787	1878-1992

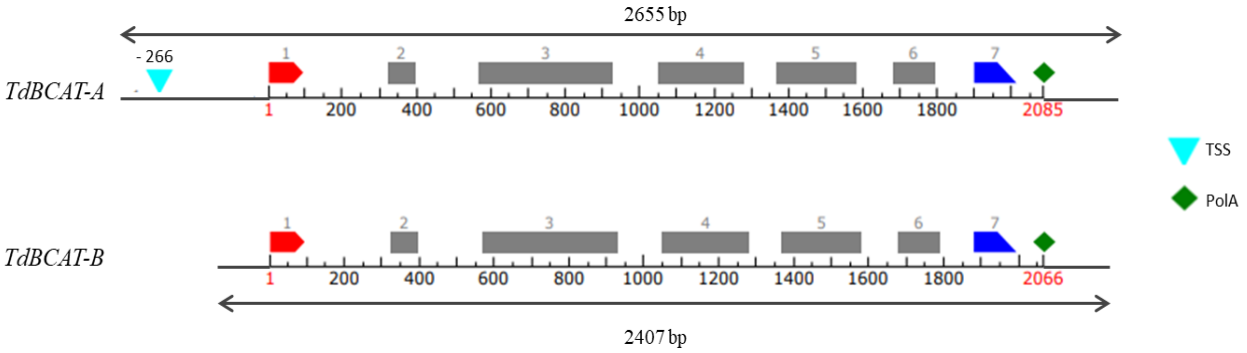


Figure 4.2 Representation of *TdBCAT-A* and *TdBCAT-B* gene models. The Transcription Starting Site (TSS) and the poly Adenilation site (PolA) are indicated when predicted by FGENESH+.

4.2.2 *TdABA7*

In total, 777 bp and 771 bp were isolated and sequenced for *TdABA7-A* and *TdABA7-B* respectively. The gene structure predicted with FGENESH+ allowed to identify a 1 exon structure for both the sequences. The CDS is 498 bp long encoding a protein of 165 amino acids, for both *TdABA7-A* and *TdABA7-B.* Upstream the starting codon, two regions of 126 bp and 133 bp, for *TdABA7-A* and *TdABA7-B* respectively, were isolated. For *TdABA7-A,* the TSS was predicted at - 89, while for *TdABA7-B* none was found. A poly-adenylation site was predicted 580 bp downstream the starting codon for both *TdABA7* genes. The gene structure results are summarized in Table 4.5 and represented in Figure 4.3.

Table 4.5 Summary of *TdABA7-A* and *TdABA7-B* gene structure based on FGENESH+ analysis

	CDS (bp)	**PROTEIN (aa)**	**EXON 1**
TdABA7-A	498	165	1-498
TdABA7-B	498	165	1-498

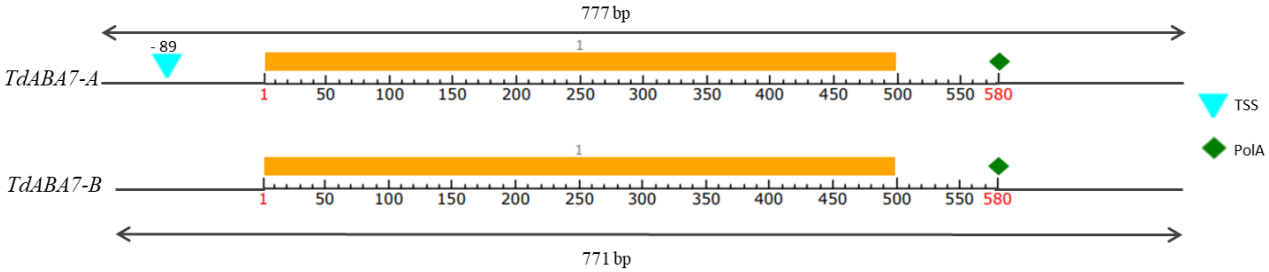


Figure 4.3 Representation of *TdABA7-A* and *TdABA7-B* gene model as predicted with FGENESH+. The Transcription Starting Site (TSS) and the poly Adenilation site (PolA) are indicated when predicted by FGENESH+.

Thanks to the recent advances in cereal genomics, the durum wheat genome of the Italian cultivar Svevo (https://d-data.interomics.eu/) and of the UK cultivar Kronos (https://wheatis.tgac.ac.uk/grassroots-portal/blast) are now available.

Both databases were blasted to verify the homology between bread and durum wheat sequences *TdBCAT-A, TdBCAT-B, TdABA7-A,* and *TdABA7-B* (Figure 4.4). The identification of a unique major match (High-scoring Segment Pair-HSP) in the predicted chromosomal localization was observed, even if for the *TdBCATs* some minor matches have been found (bit score < 200), which represent repetitive elements along the respective chromosomes.

The results of these BLASTn analyses using the genes isolated in cv. Cham1 present a 100% of identity for all the genes, with the exception of *TdABA7-A,* which shows a 99.1% identity in both Kronos and Svevo assemblies due to the presence of 7 SNPs (Figure 4.5) distributed along the sequence. Two SNPs are in the 5' untranslated region (5'-UTR), three in the coding region (AgA>AtA R17I; aAA>gAA K88E; Gag>Gat E131D), and two in the 3' untranslated region (3'-UTR).

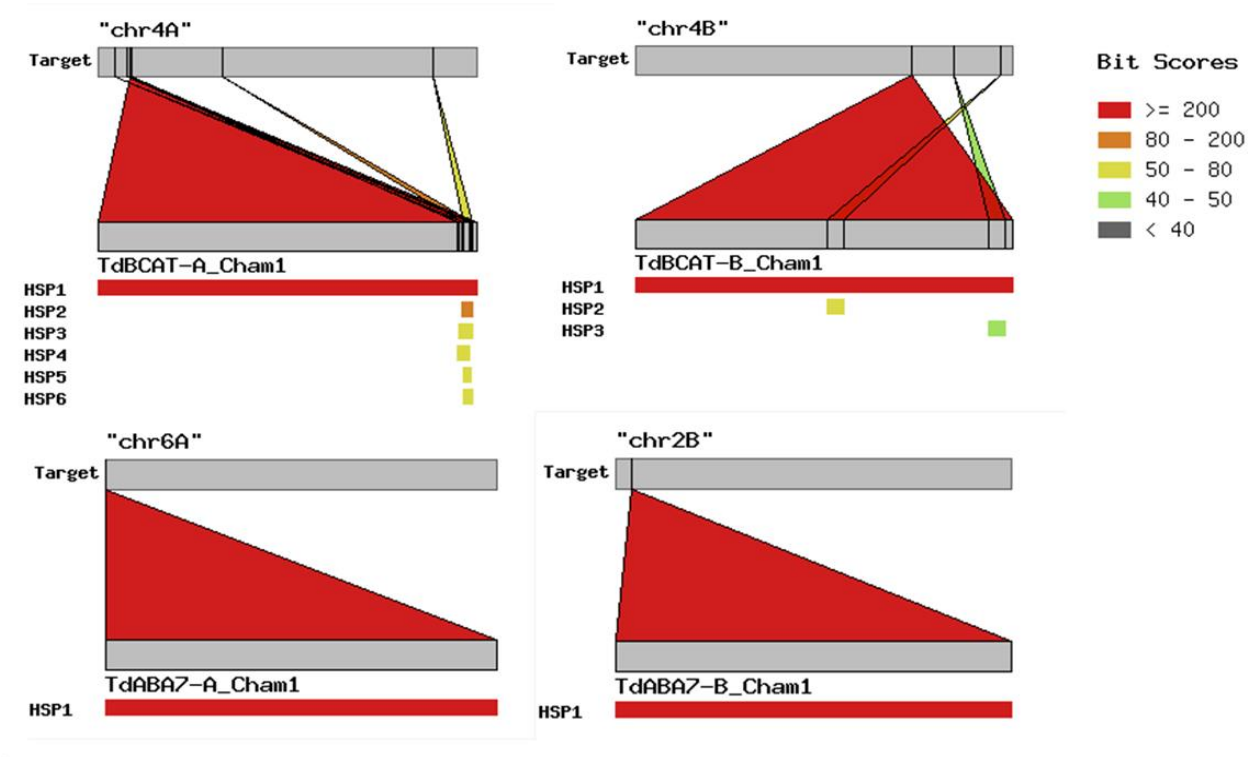


Figure 4.4 Results of the BLASTn analysis on the Svevo assembly (https://d-data.interomics.eu/) using as query the sequences isolated from cv. Cham1 of *TdABA7* and *TdBCAT* genes.

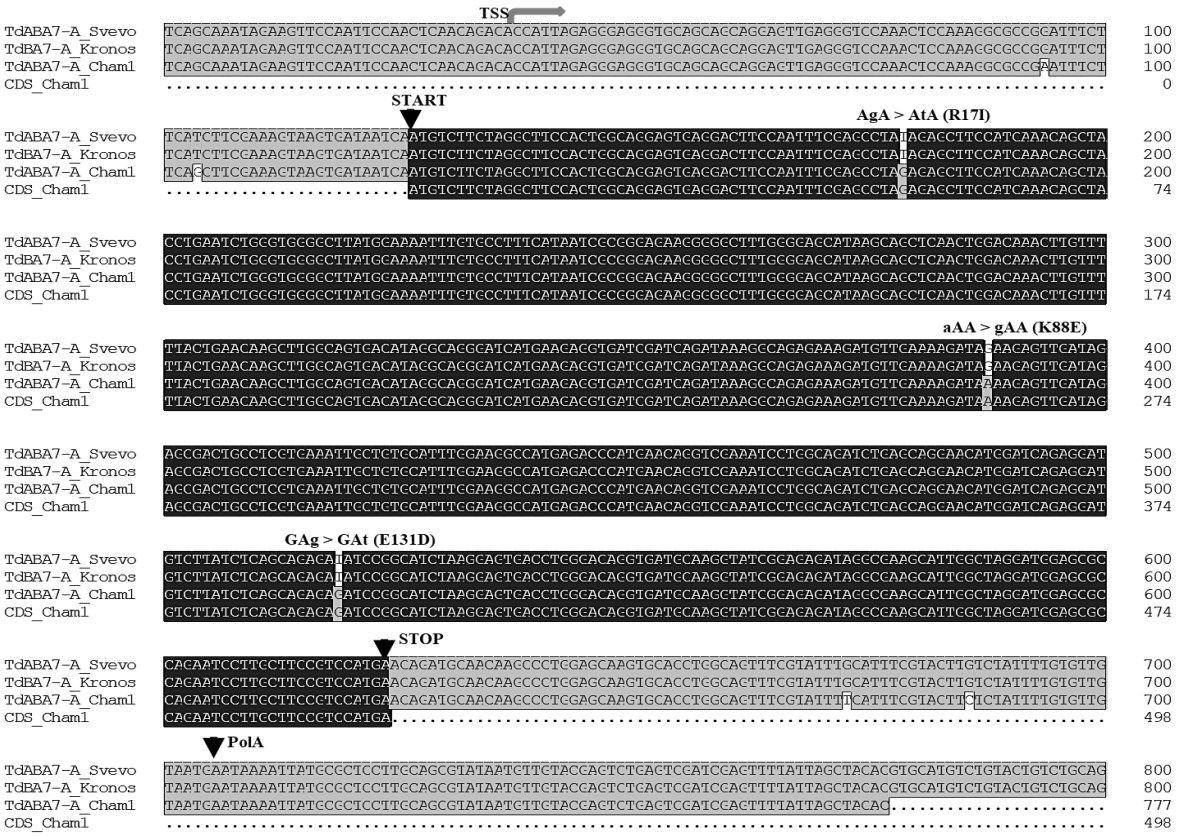


Figure 4.5 Multiple alignment of *TdABA7-A* of Cham1, Svevo and Kronos cultivars, and the CDS of Cham1 gene. The START and STOP codon, as well as the predicted TSS and polyadenylation signal (PolA), are indicated by black arrows. For each SNP, the position and the effect on the coding sequence is indicated.

4.2.3 *In silico* Analysis of Proximal Promoter Regions

Since the durum wheat genes isolated from the cv. Cham1 are nearly identical (>99%) to Svevo's sequences, the analysis of promoter regions was performed on Svevo sequences. Svevo sequences are 100% identical to Kronos's for all the genes except for *TdBCAT-B*, which presents a single SNP 191 bp upstream the starting codon (i.e. Svevo's allele is T, while Kronos' is A).
The proximal promoter sequence of 1000 bp upstream the ATG was analysed for each gene. Plant Pan 2.0 software identified several *cis*-acting regulatory elements targeted by several TFs. The *in silico* analysis was also implemented searching into PlantCARE database (Lescot et al., 2002), especially to check the presence of ABREs. Here we present in particular the results concerning the *cis*-acting elements involved in abiotic stress response, which had a matching score between 0.9 and 1 (1 is the maximum value).

4.2.3.1 Promoter of TdABA7 genes

The most interesting results of the *in silico* analysis of the promoter region are presented for *TdABA7-A* in Figure 4.6 and for *TdABA7-B* in Figure 4.7. In *TdABA7-A* promoter, a TATA-BOX was predicted 70 bp upstream the predicted TSS, while for *TdABA7-B* it is predicted around 200 bp upstream the starting codon of the CDS. In both genes, many ACGT-motifs were found, a general target site for bZIP TFs. These motifs overlap with other *cis*-acting elements targeted by TFs involved in abiotic stress response. Among these, a major regulatory element was found around 500 bp upstream the starting codon of both *TdABA7* genes, in which two different *cis*-acting elements (TFmatrixID_0184, TFmatrixID_0111) are present. The TFmatrixID_0184 are targeted by ABA-responsive TFs such as ABF1 (AT1G49720) and AREB2/ABF4 (AT3G19290) both of the bZIP family, while PIF3 (AT1G09530) targets TFmatrixID_0111. ABF1 is involved in the ABA regulated gene expression, and abiotic stress response such as heat and cold, and it is required for the establishment of *Arabidopsis* seedlings during winter (Sharma et al., 2011). More recently, a study in *O. sativa* showed that the overexpression of *OsABF1* significantly improved rice performance under drought conditions (Zhang et al., 2017). AREB2/ABF4 is, along with AREB1 and AREB3, a master regulator of the ABRE-dependent ABA signalling in DS tolerance, regulating many downstream genes, including several Late Embryogenesis Abundant (LEA) proteins (Yoshida et al., 2010). The bHLH TF called PIF3 in *Arabidopsis* is involved in the phytochrome signalling. Transgenic rice plants, which are induced to express the *Zea mays* PIF3, show an improved drought tolerance, as well as an enhanced water-saving behaviour by the decreased stomatal aperture and reduced transpiration (Gao et al., 2018).

An ABRE is represented by TFmatrixID_0277, bound by TFs RGL2 (AT3G03450) and RGL3 (AT5G17490) of the GRAS family subgroup of DELLA proteins, which are involved in many biological activities such as the ABA biosynthesis and regulation, especially in seed dormancy and germination, and

the acquisition of salt and oxidative stress tolerance. These two TFs can act as a functional interface of ABA-mediated abiotic stress response and the gibberellic acid (GA)-controlled development signalling (Golldack et al., 2014).

In a recent investigation of the transcriptional regulatory activity of differentially expressed genes under drought and salinity conditions, it was found that among all the *cis*-acting elements, the over-expressed genes share the abundance of GATA-motifs (Shariatipour and Heidari, 2018).

In *Arabidopsis*, GATA1, GATA2, and GATA3 TFs can bind both palindromic and single GATA sites ((A/T)GATAA) (Reyes et al., 2004). *TdABA7-B* promoter contains more GATA-motifs in comparison with *TdABA7-A*, thus it could represent a preferred target for GATA proteins compared to *TdABA7-A*.

The PlantCARE (Lescot et al., 2002) database confirmed the presence of several ABREs overlapping the regulatory elements highlighted with the *in silico* analysis with PlantPAN 2.0 (Chow et al., 2016).

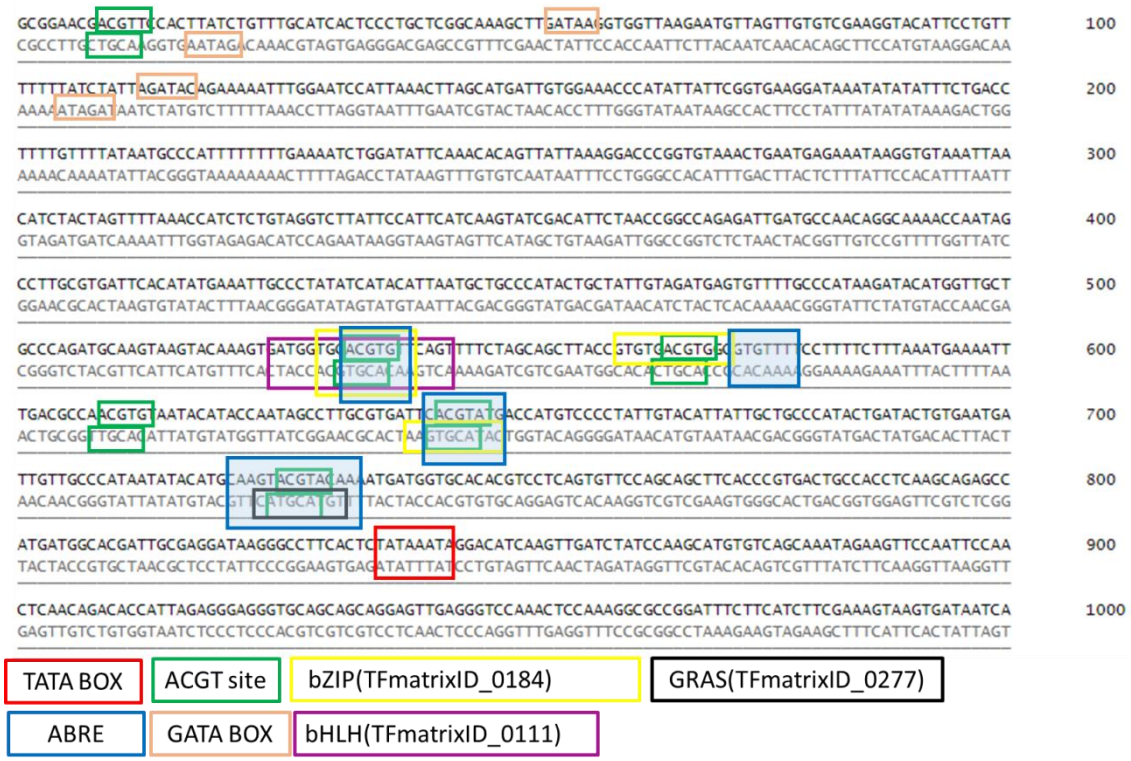


Figure 4.6 Scheme of the main abiotic responsive elements and the TATA BOX retrieved by the *in silico* analysis of *TdABA7-A* promoter.

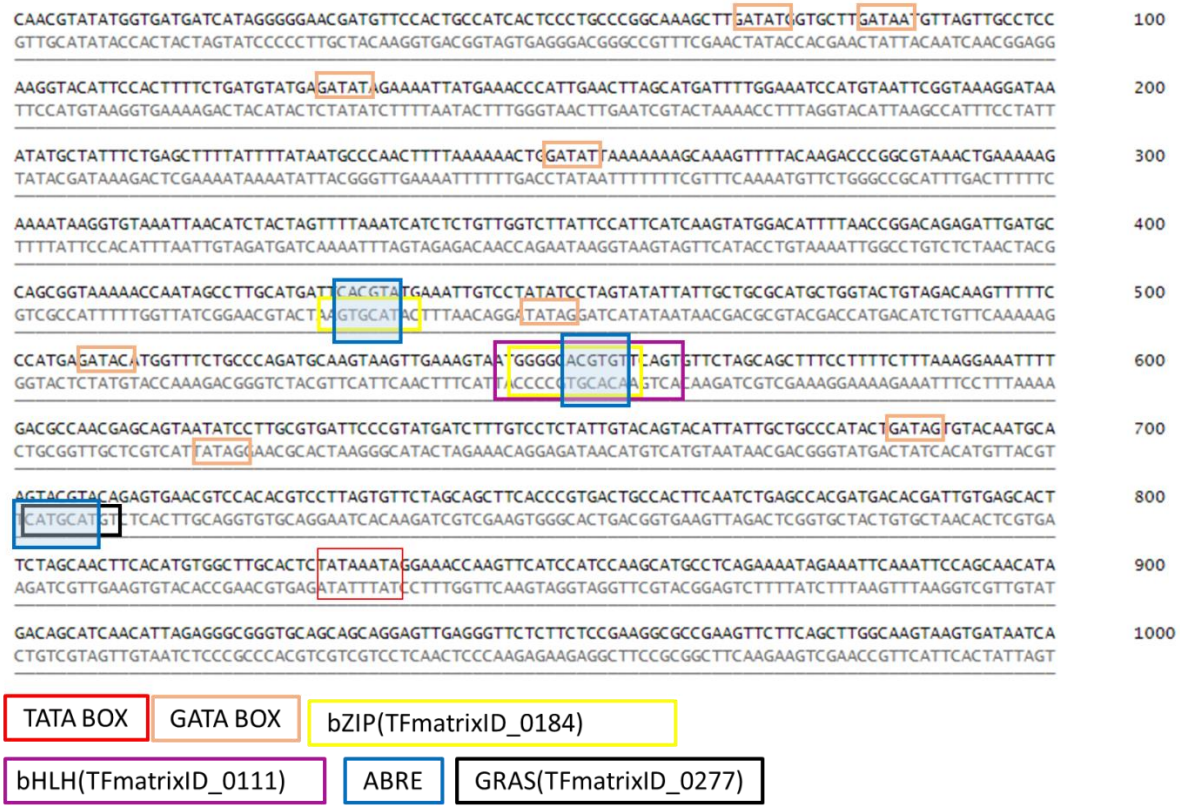


Figure 4.7 Scheme of the major abiotic responsive elements and the TATA BOX retrieved by the *in silico* analysis of *TdABA7-B* promoter

4.2.3.2 Promoter of *TdBCAT* genes

The most interesting results of the *in silico* analysis of the promoter region are presented for *TdBCAT-A* (Figure 4.8) and for *TdBCAT-B* (Figure 4.9). In *TdBCAT-A* promoter the TATA-BOX was predicted 80 bp upstream the predicted starting codon (ATG) by PlantPAN 2.0, while the analysis with PlantCARE identified a TATA BOX around 260 bp upstream the starting codon. For *TdBCAT-B* promoter the TATA BOX could be identified only with PlantCare around 76 bp upstream the starting codon (ATG).

Both promoters are particularly rich of *cis*-acting elements targeted by several TFs of the AP2/ERF family.

A DREB element is predicted around 750 bp upstream the starting codon in the *TdBCAT-A* promoter. This element is a target of several TFs involved in both abiotic and biotic stress regulation of gene expression, as well as in stress signal transduction: DEAR3 (AT2G23340), RAP2.10/DEAR4 (AT4G36900), RAP2.4 (AT1G78080), DREB2C (AT2G40340), DREB2/DREB2A (AT5G05410), ATABI4 (AT2G40220), CRF6 (AT3G61630), CBF1/DREB1B/ATCBF1 (AT4G25490), DDF1 (AT1G12610), ATERF38/ERF38 (AT2G35700), CBF4/DREB1D (AT5G51990), TINY2 (AT5G11590).

For *TdBCAT-B*, the AP2/ERF elements are targets of RAP2.6 (AT1G43160), DREB2C (AT2G40340), DREB2/DREB2A (AT5G05410), and ATABI4 (AT2G40220).

Both promoters are target of MYB TFs, especially the one related to the circadian clock-controlled genes. Members of this family are key factors in regulatory networks controlling development, metabolism and responses to biotic and abiotic stresses in plant genome (Roy, 2015).

Interestingly for both genes, a cis-acting element target of CG-1/CAMTA family was found. The predicted TFs are CAMTA3/SR1 (AT2G22300) and a calmodulin-binding transcription activator protein with CG-1 and Ankyrin domain (AT5G64220). These TFs can regulate transcriptional activity in response to calcium signals. AtCAMTA3 has a role in cold tolerance by inducing *CBF* genes in cooperation with AtCAMTA1 and AtCAMTA2 (Zeng et al., 2015).

Moreover, to testify the various involvements into abiotic stress response of *TdBCATs*, a CH2H element was found in both BCATs. This site is target of several repressors of the Zinc-finger protein (ZFP) family, which are involved in the salt-tolerance mechanism. The recognized TFs are CH2H/ZAT6/CZF2 (AT5G04340), STZ/ZAT10 (AT1G27730), AZF3 (AT5G43170), and AZF2 (AT3G19580), which can also act as a positive regulator of leaf senescence. In *Arabidopsis* CH2H ZFPs such as AZF2, AZF3 and ZAT10 have been reported to have a role in the drought and salt stress response (Sakamoto et al., 2000). In particular, Sakamoto and co-workers showed that AZF3 is induced by low-temperature, while AZF2 is induced by ABA treatment and high salinity. These results show how AZF2 and AZF3 are both involved in water-stress response.

Figure 4.8 Scheme of the major abiotic responsive elements and TATA BOX retrieved by the *in silico* analysis of *TdBCAT-A* promoter. In *TdBCAT-A* promoter the TATA-BOX was predicted 80 bp upstream the predicted starting codon (ATG) by PLANT PAN 2.0, while the analysis with Plant Care identified a TATA BOX around 260 bp upstream the starting codon.

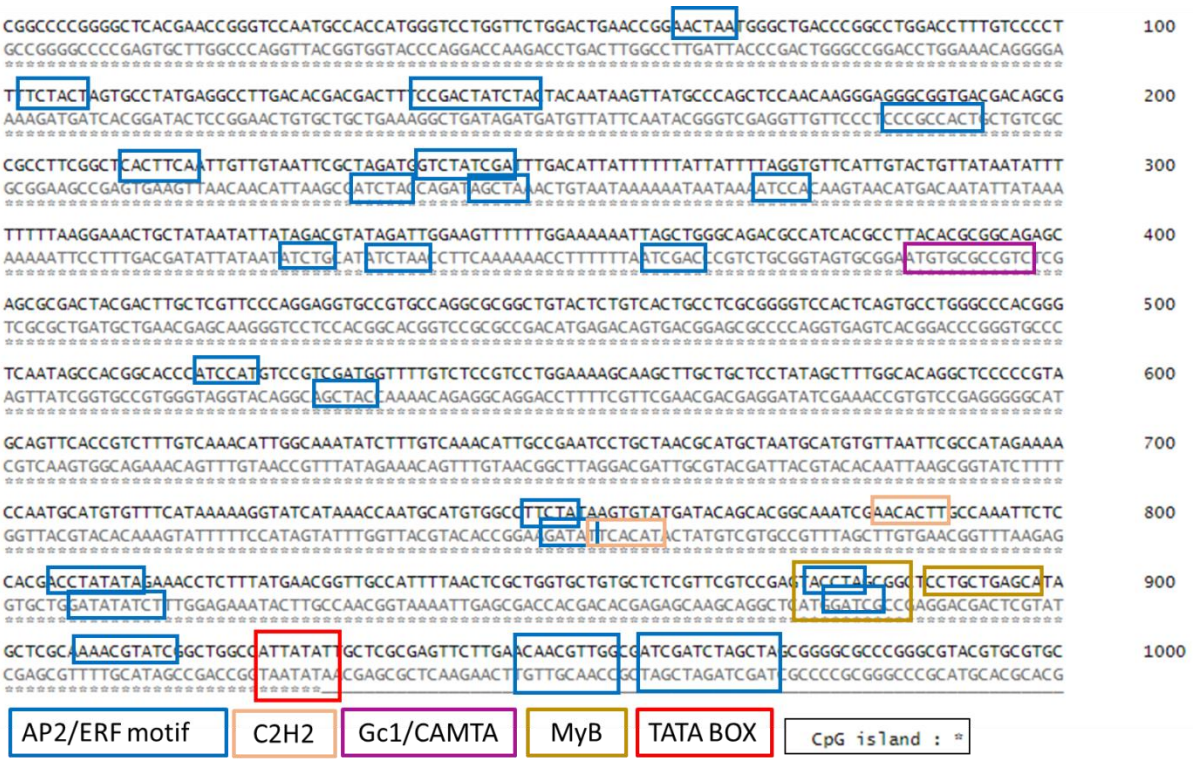


Figure 4.9 Scheme of the major abiotic responsive elements and TATA BOX retrieved by the *in silico* analysis of *TdBCAT-B* promoter

4.2.4 Phylogenetic analysis and subcellular localization of TdBCATs

On the basis of the results discussed so far, the sequences of *TdBCAT*s were translated (https://web.expasy.org/translate) and analysed *in silico* to predict their sub-cellular localization.

A phylogenetic analysis was also performed using several BCAT protein sequences. Durum wheat protein sequences were used as query for a BLASTp search on the NCBI database to retrieve the most similar BCAT proteins. Most proteins of the first 100 results are classified as chloroplastic or chloroplastic-like Branched-chain aminotransferase. The only three BCATs descripted as mitochondrial are BCAT1 of *Morus notabilis* (XP_024029919.1), BCAT1 of *Ziziphus jujuba* (XP_015885109.1) and BCAT1 of *Momordica charantia* (XP_022132131.1), with a 66% of identity to both TdBCAT-A and TdBCAT-B.

For the phylogenetic analysis the following proteins were included: i) BCATs of species evolutionary close to durum wheat; ii) three mitochondrial BCATs; iii) all the functional BCATs (1-6) known in *Arabidopsis*; iv) bread wheat orthologues of HvBCAT-1. Protein sequences of Arabidopsis and bread wheat were retrieved on Ensembl Plant database, including all the available isoforms of each gene.

Three major clusters are obtained in relation to the subcellular localization of the *Arabidopsis* BCATs (Figure 4.10). Both BCAT-A and BCAT-B of durum wheat are in the same sub-tree with all the orthologues bread wheat BCATs, and with the two chloroplastic BCATs of *Arabidopsis* (AtBCAT-3 and AtBCAT-5). The three proteins which are predicted as mitochondrial BCATs (XP_024029919.1,

XP_015885109.1 and XP_022132131.1) cluster with the mitochondrial AtBCATs (AtBCAT-1 and AtBCAT-2), and look more similar to AtBCAT-2 than AtBCAT-1. The two cytosolic AtBCATs (AtBCAT-4 and AtBCAT-6) cluster separately.

Moreover, we can notice that all the wheat (durum and bread) BCATs form as many clusters as the number of genomes (A, B and D), which contain also the related ancestor, where the sequence is available. Indeed the BCAT from the donor of A genome (*T. Urartu*) is in the same subtree of the BCAT-A of durum and bread wheat, as well as the BCAT of the donor of the D genome (*A.* tauschii) clusters with the BCAT-D of wheat.

Also, the evolutionary order is respected along the *Gramineae* species, as we can see the closest species to wheats and their relatives is barley which belongs to the *Triticeae* tribe, followed by *B. distachyon*, *O. sativa*, *S. italic*, *S. bicolor* and finally *Zea mays.*

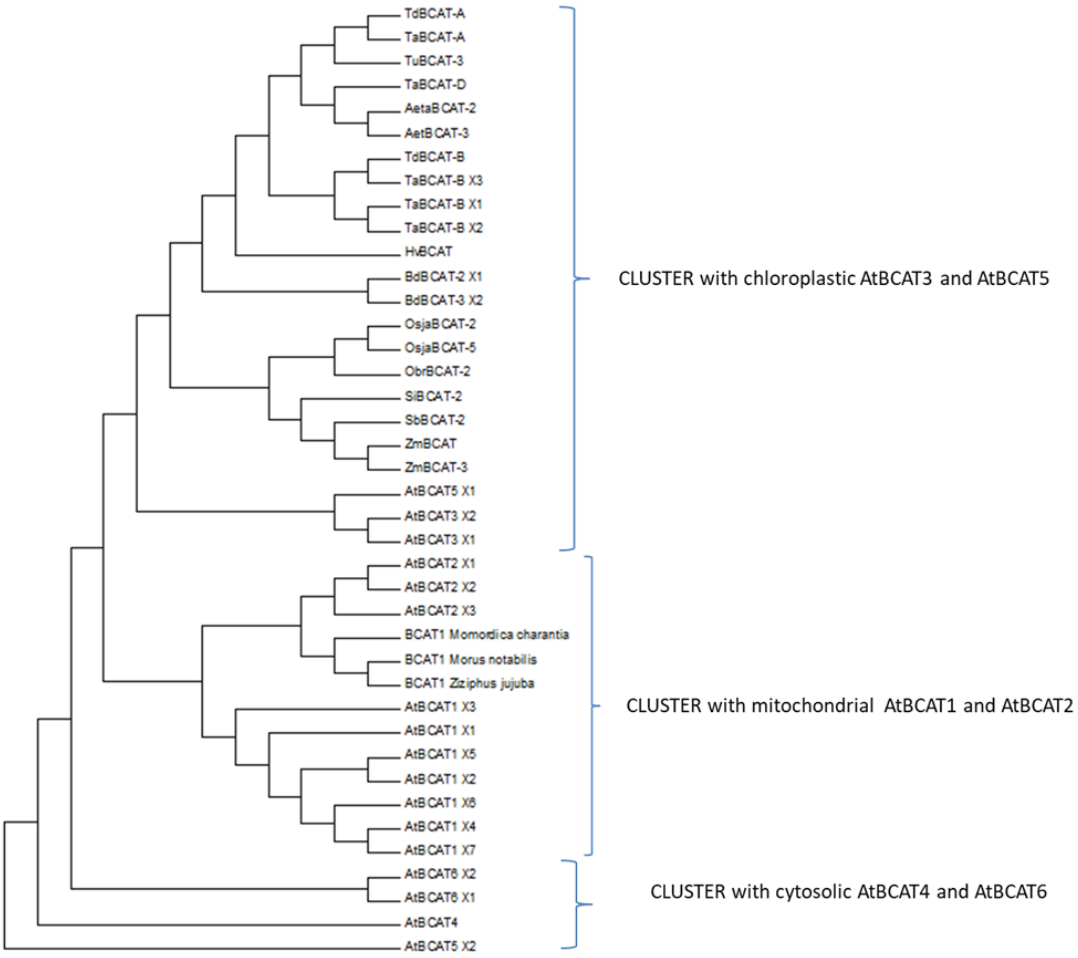


Figure 4.10 Phylogeny tree of BCATs. Sequences retrieved on the NCBI database are TuBCAT-3 (*Triticum urartu*, EMS62409.1), HvBCAT (*Hordeum vulgare* CAE00460.2), Aeta BCAT-2 (*Aegilops tauschii* XP_020159191.1), AetBCAT-3 (*Aegilops tauschii*, XP 020173856.1), BdBCAT-2 X1 (*Brachypodium distachyon*, XP_003558433.1), BdBCAT-3_X2 (*Brachypodium distachyon*,XP_010228932.1), OsjaBCAT-2 (*Oryza sativa* Japonica group, XP_015630708.1), OsjaBCAT-5 (*Oryza sativa* Japonica group, ABF94786.1), ObrBCAT-2 (*Oryza brachyantha*, XP_006651192.1), SiBCAT-2 (*Setaria italica*, XP_004985120.1), SbBCAT-2 (Sorghum bicolor, XP_021317757.1), ZmBCAT (Zea mays, XP_008674369.1), ZmBCAT-3 (*Zea mays*, PWZ58251.1). The only three BCATs descripted as mitochondrial are BCAT1 of *Morus notabilis* (XP_024029919.1), BCAT1 of *Ziziphus jujuba* (XP_015885109.1) and BCAT1 of *Momordica charantia* (XP_022132131.1). Sequences retrieved from Ensembl Plant database are the *Triticum aestivum* TaBCAT-A (TraesCS4A01G059800.1), TaBCAT-B_X1, X2 and X3 (TraesCS4B01G235400.1, .2, .3), TaBCAT-D (TraesCS4D01G236800.1), and the *Arabidopsis thaliana* AtBCAT1_X1, X2, X3, X4, X5, X6, X7 (AT1G10060.1, .2, .3, .4, .5, .6, .7), AtBCAT2_X1, X2, X3 (AT1G10070.1, .2, .3), AtBCAT3_X1, X2 (AT3G49680.1, .2), AtBCAT5_X1, X2 (AT5G65780.1, .2), AtBCAT6_X1, X2 (AT1G50110.1, .2), AtBCAT4 (AT3G19710)

4.3 Expression analysis under water-limiting conditions in different growth stages, tissue and cultivars of durum wheat

The modulation of the expression of genes *TdABA7* (A and B) and *TdBCAT* (A and B) under water-limiting conditions was monitored in several cultivars of durum wheat during two fundamental phases of its life cycle: the germination and the reproductive stages.

To monitor the response to DS we checked the expression of the GOI along with a durum wheat drought-related gene, *TdDHN15.3* (AM180931.1) that encodes for a low molecular weight dehydrin (DHN3). DHN3 is a LEA protein belonging to the YSK2 type of the dehydrin family which is strongly up-regulated in dehydrated durum wheat coleoptiles (Rampino et al., 2006) as well as in barley flag leaves under terminal drought (Karami et al., 2013). The expression rate of *TdDHN15.3* is used to verify the effectiveness of the stress imposed to wheat plants in both growth conditions.

4.3.1 Analysis of the differential expression of *TdABA7* and *TdBCAT* genes in 8-days old dehydrated coleoptiles

We analysed the expression of all GOIs in ten different durum wheat cultivars during the germination phase under water-limiting condition as described in section 3.2.

The most interesting results concern the expression of *TdABA7*, as both *TdABA7-A* and *TdABA7-B* alleles are up regulated in coleoptiles of all durum wheat cultivars in response to DS (Figure 4.11; Table 4.6).

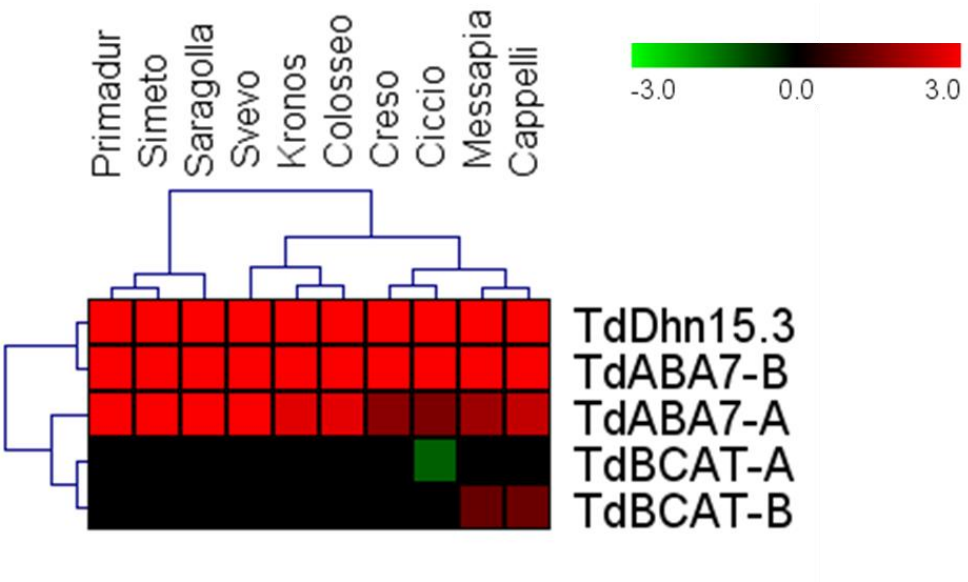


Figure 4.11 Heat-map representing the relative expression of GOIs and stress-gene (*TdDHN15.3*) in dehydrated coleoptiles of ten durum wheat cultivars. All the up- and down-regulation less than 2-fold are in black. Data are presented as log2(FC). Log2(FC) > 3 are of the same intensity

Table 4.6 Relative expression of five different genes in dehydrated coleoptiles of ten durum wheat cultivars. Values represent fold-change mean values.

Gene name	Creso	Kronos	Messapia	Primadur	Ciccio	Simeto	Saragolla	Svevo	Colosseo	Cappelli
TdDhn15.3	317.37	1,946.50	406.37	1,016.93	421.68	661.68	1,072.43	1,930.82	789.61	301.29
TdABA7-A	3.11	6.33	3.67	69.79	2.78	92.20	97.01	12.85	6.60	4.95
TdABA7-B	344.89	702.65	325.91	597.03	357.05	536.21	1,237.60	342.11	766.94	348.09
TdBCAT-A	0.84	0.92	0.84	1.09	0.46	0.87	1.02	0.69	1.78	1.33
TdBCAT-B	1.24	1.44	2.35	1.84	0.67	1.18	1.42	0.90	1.47	2.43

In particular, the up-regulation of *TdABA7-A* is the more diverse across the different genotypes. In most of the cultivars, the up-regulation is in the range between 2 and 3.5-fold (Creso, Messapia, Ciccio) and between 3.5 and 7-fold (Kronos, Colosseo and Cappelli) (Table 4.6). In Svevo *TdABA7-A* is up-regulated 12-fold, but the most intense up-regulation is observed in Primadur (70-fold), Simeto and Saragolla (90-fold) (Table 4.6). This cultivar-specific expression of *TdABA7-A* is related to a lower basal expression of this gene in Primadur, Simeto and Saragolla compared to all the other seven cultivars.

Remarkably, *TdABA7-B* shows the highest up-regulation (Table 4.6; Figure 4.11). Low expression levels are detected in control coleoptiles (Ct values t=30-33, data not shown) depending on the cultivar. After 8 hours of dehydration, the expression is up-regulated from 300-fold to almost 2000-fold depending on the cultivar similarly to *TdDhn15.3* (Table 4.6, Figure 4.11). This is consistent with the observation reported by Rampino et al (2006), which observed comparable expression levels of many dehydrin genes, including *TdDhn15.3,* after 8-hours of dehydration in wheat coleoptiles and furtherly supported at protein level, as the accumulation of DHNs was observed in wheat coleoptiles grown in water-limiting conditions (Shakirova et al., 2016).

Wheat and barley DHNs are induced by several abiotic stresses (low temperatures, drought, salinity) and have been proposed to play a role into the acquisition and maintenance of an enhanced frost tolerance during vernalisation (Kosová et al., 2014). Considering the comparable transcriptional profile of *TdDhn15.3* with *TdABA7-B*, we can speculate a possible involvement of these genes in the same pathway of response to water-limiting conditions during the very early germination phase of wheat. This would be also in accordance with previous results reported for barley gene *HvABA7* (Gulli et al. 1995), which was up-regulated in barley coleoptiles in response to dehydration, as well as cold and ABA treatment, an important regulator of the transitions from dormancy to germination and from germination to growth in cereal crops (Kermode, 2005).

The very evident homoeologs-specific modulation of *TdABA7* could be related to a different function during the stress and/or in the early phases of germination, as the basal expression of *TdABA7-A* is higher compared to *TdABA7-B* in most of the cultivars.

A homoeologs- and cultivar-specific expression profile is visible also for *TdBCAT* genes (Table 4.6, Figure 4.11), but with a less variable modulation compared to *TdABA7* genes. *TdBCAT-B* is up-regulated

in most of the cultivars, with the exception of Ciccio and Svevo in which the down-regulation is still below 2-fold compared to controls. In Messapia and Cappelli, *TdBCAT-B* expression increases for about 2-fold in dehydrated coleoptiles (Table 4.6); interestingly Cappelli is a drought resilient cultivar (Aprile et al., 2013). In Ciccio, *TdBCAT-A* is down-regulated in dehydrated coleoptiles (Table 4.6, Figure 4.11).

The results presented in this section allow appreciating the homoeologs-specific behaviour of each gene and secondly, the cultivar-specific expression profile of each GOI (Figure 4.11). The peculiar modulation of gene *TdABA7*, with a special regard to *TdABA7-B*, suggests an important role of this gene under water-limiting condition in the early phase of germination.

4.3.2 Drought stress at the reproductive stage in two cultivars with different drought tolerance

Two durum wheat cultivars (Figure 4.12A) were chosen on the basis of their tolerance to DS: Colosseo as susceptible (Rampino et al., 2006) and Cappelli as tolerant (Aprile et al., 2013). Plants were grown in controlled conditions as described in section 3.2.1, and they were maintained well-watered until the end of the heading when we stopped watering some of the plants to impose DS. The correct growth stage was evaluated by carefully checking the maturity of the reproductive organs (Figure 4.12.B) according to the Zadoks code (see *Annex*, Table 6.2). The relative expression and the physiological analysis for both the cultivars were evaluated at three time points: T0 (plants in Z65, Figure 4.12.C); 4das (Z70); 14das (plants in Z71-Z73). The flowering stage, seed formation and grain filling were chosen as the most crucial phases in which DS may severely hamper grain yield (Altenbach, 2012).

To confirm the occurrence of stress conditions, some physiological measurements were performed, in particular the RWC and the chlorophyll content were recorded.

Figure 4.12 (A) The two *cultivars* used for the experiment Colosseo (left) and Cappelli (right). (B) To check the exact developmental stage of the plants we opened the middle spikelet to verify the maturation level of the reproductive organs. The drought stress occurred during (C) the optimum flowering period or anthesis (Z65) and lasted for a total of 14 days during which plants were sampled at 0, 4 and 14 das.

4.3.2.1 Relative Water Content

For both Cappelli and Colosseo we evaluated the RWC on the leaf immediately below the flag leaf, at the same time point in which plant tissues (spikes and flag leaf) were sampled for the transcriptional analysis.

RWC values progressively decrease in stressed plants in relation to the duration of the water-limiting conditions along all the life cycle (Rampino et al., 2006; Slama et al., 2018).

The RWC of stressed plants reduces significantly only 4 das in both Colosseo ($p<0.01$) and Cappelli ($p<0.01$) (Figure 4.13), while it is not possible to appreciate any significant differences between controls and stressed plants during the acclimation to the stress treatment in both cultivars. The differences between the RWC values of control and stressed plants exacerbate along with the duration of the stress in both cultivars, but in the sensitive genotype Colosseo the time-dependant decrease is more consistent (Figure 4.13). Indeed, after 2 weeks of stress Colosseo's leaves reach an average RWC of about 50%, while Cappelli's leaves have still a RWC equal to 71%. These results are consistent to a similar study in common wheat during the anthesis in which RWC decreased along with the stress duration, and was kept higher in genotype resistant to drought, which had also an higher photosynthetic capacity (Ritchie et al., 1990). Eftekhari et al. (2017) saw comparable reductions of RWC between resistant and sensitive wheat cultivars under DS also in open-field.

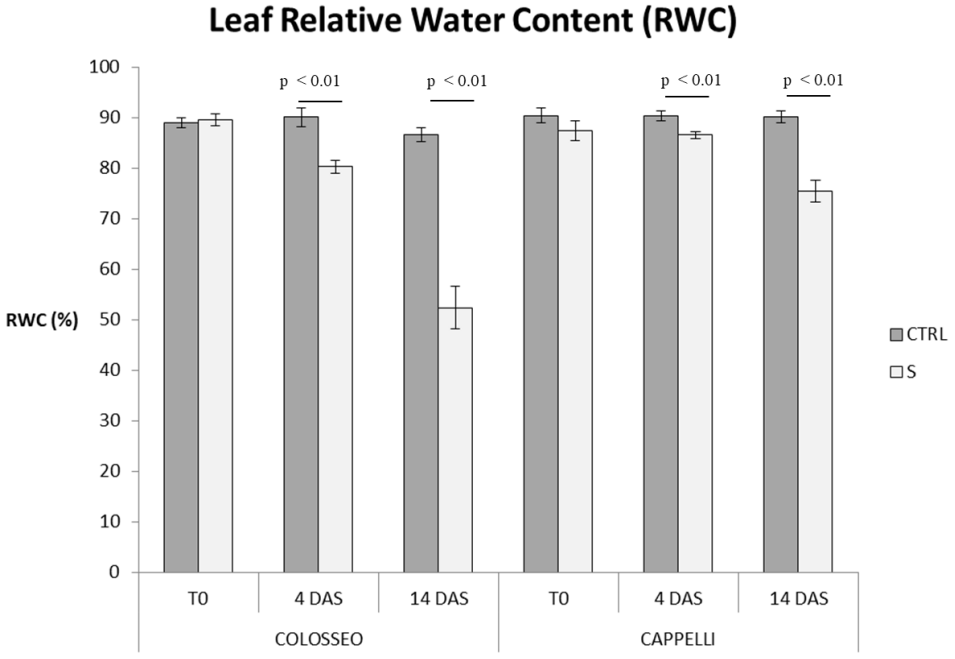


Figure 4.13 Relative Water Content recorded at three different time-points in Colosseo and Cappelli during anthesis and post-anthesis events. RWC (%) values are presented with the relative standard errors bars. The significant differences between control and stressed plants are calculated on the basis of Student's t-test. (DAS= days after stress; CTRL = control plants; S=stressed plants)

4.3.2.2 Chlorophyll content

Photosynthesis, together with cell growth, is among the primary processes to be affected by drought (Chaves et al., 2009). Chlorophyll is the major chloroplast component of the photosynthetic apparatus, and its content correlates positively with the photosynthetic capacity (Murchie and Horton, 1997). In this study, the chlorophyll content was measured with a SPAD 502 unit, which is broadly used for the measurement of leaf chlorophyll content. Lopes and Reynolds (2012) used the SPAD units as a measure of the leaf chlorophyll content to monitor the onset of senescence under abiotic stress condition during the anthesis and post-anthesis events in wheat.

The susceptible cultivar Colosseo shows a significant lower chlorophyll content ($p<0.01$ at T0 and 4 das, $p<0.05$ at 14 das) in stressed plants compared to the control plants in all time points, with the higher difference during the acclimation to DS (Figure 4.14). On the other hand, the resistant cultivar Cappelli keeps higher overall chlorophyll content, without any significant loss due to the DS treatments in all time points (Figure 4.14).

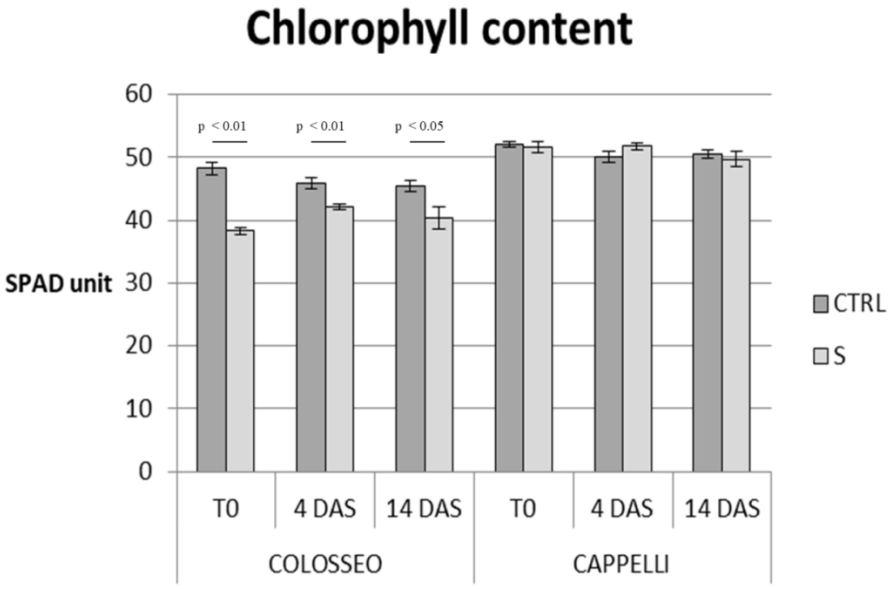


Figure 4.14 Effect of drought stress on chlorophyll content (reported as SPAD units) recorded in Colosseo and Cappelli at three different time-points during anthesis and post-anthesis events. SPAD values are presented with the relative standard errors bars. The significant differences between control and stressed plants are calculated on the basis of t-test. (DAS= days after stress; CTRL = control plants; S=stressed plants)

In addition, a stay green phenomena was observed in Cappelli under drought (while a rapid senescence was seen in Colosseo. The effect of water deprivation was visible on cultivar Colosseo starting from around 6 das, as plants turned to yellowish colour very soon during the experiments. On the contrary, Cappelli showed the so-called “stay green” ability for almost all the duration of the treatment, with some visible effect of turgor loss and yellow leaf in the last 2-4 days of stress. The stay green is a key indicator

to stress adaptation and can be evaluated by the use of many phenological trait, including the measurement of the chlorophyll content (Lopes and Reynolds, 2012). An higher chlorophyll content has been seen to improve yield and transpiration efficiency under water-limiting conditions in many crops, including bread wheat (Verma et al., 2004) and durum wheat (Hafsi et al., 2000). The stay green trait is genetic-related, and has been used in wheat to select genotypes more resistant to drought (Barakat et al., 2015), which is consistent with our knowledge about the different ability to resist to DS of the two cultivar Colosseo (susceptible) and Cappelli (resistant).

4.3.2.3 Expression analysis during the reproductive stage

The transcriptional analysis was performed on all the samples of Cappelli and Colosseo, as described in paragraph 3.2.

The overall results are shown in the heat-map in Figure 4.15, while the relative quantification data are presented in Table 4.7.

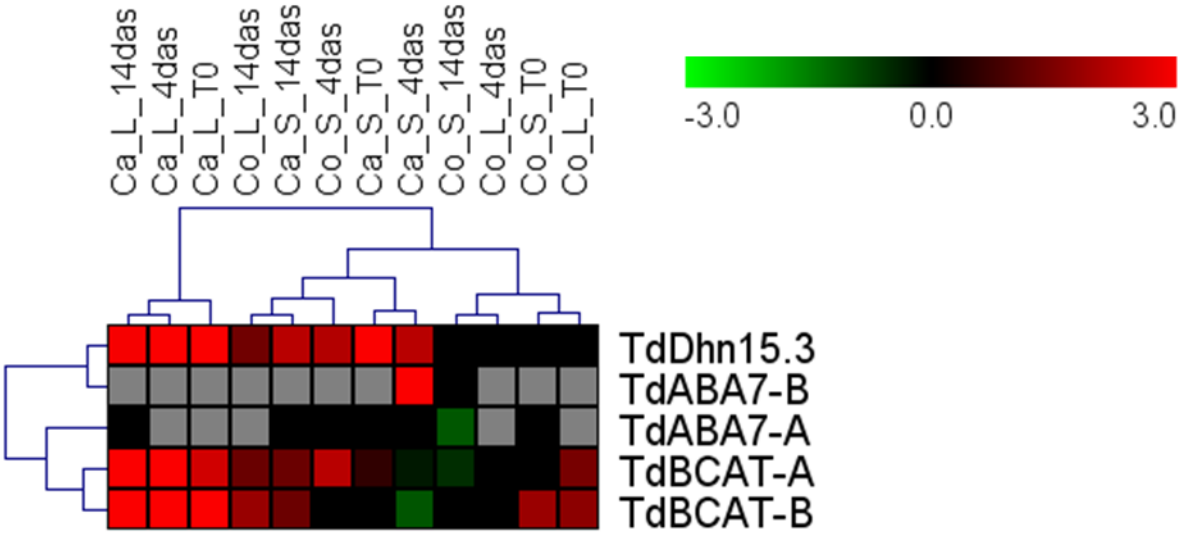


Figure 4.15 Heat-map showing the relative expression of GOIs and stress-genes (*TdDHN15.3*) in response to drought during the reproductive stage at three different time-points (T0, 4das, 14das). All the up- and down-regulation less than 2-fold are in black. Grey boxes represent the undetectable expression in both control and stressed conditions. Data are presented as log2 Fold-change (FC) (Ca =Cappelli; Co=Colosseo L= Flag Leaf; S = spike; das = days after stress). Log2 (FC) > 3 are of the same intensity.

Table 4.7 Fold-change values of 5 genes in spike and flag leaves of 2 durum wheat cultivar at three different time-point: T0 acclimation to DS during anthesis; 4 das and 14 das during the early post-anthesis events. (n/d: not detectable).

	COLOSSEO						**CAPPELLI**					
	FLAG LEAF			**SPIKE**			**FLAG LEAF**			**SPIKE**		
Gene	**T0**	**4das**	**14das**	**T0**	**4das**	**14das**	**T0**	**4das**	**14das**	**T0**	**4das**	**14das**
TdDhn15.3	0.63	1.87	2.55	1.39	4.40	1.17	19.84	9.52	7.73	8.63	4.63	4.63
TdABA7-A	n/d	n/d	n/d	0.57	1.70	0.49	n/d	n/d	1.69	0.96	1.02	1.04
TdABA7-B	n/d	n/d	n/d	n/d	n/d	0.52	n/d	n/d	n/d	n/d	27.95	n/d
TdBCAT-A	2.07	0.88	2.40	0.51	4.62	0.69	5.64	9.95	13.50	1.50	0.82	2.42
TdBCAT-B	3.21	1.24	3.54	3.61	0.75	1.47	8.53	8.77	16.37	1.65	0.50	2.44

The expression profile of all the target genes is strongly dependent on growth stage, tissue, and cultivar, showing remarkable differences between the two homoeologs especially for gene *TdABA7*.

The expression profile of the stress-related gene *TdDHN15.3* shows a clear up-regulation in most of the samples (Table 4.7). The up-regulation of *TdDHN15.3* is more consistent in the drought resistant cultivar Cappelli, in which the modulation within the same type of tissue (flag leaf or spike) shows the same decreasing time-dependant trend, with the highest value at time point T0, after the acclimation phase. Considering the enhanced tolerance to DS of Cappelli compared to Colosseo, a major up-regulation of *TdDhn15.3* in Cappelli is in accordance with the results of Karami and colleagues (2013) in barley. This gene was induced in barley flag leaf under terminal drought, with remarkable higher rate in drought resistant genotypes. This is the umpteenth proof of concept that testifies the enhanced drought tolerance of Cappelli compared to Colosseo.

TdBCAT-A and *B* genes are mainly up-regulated in the flag-leaf under DS, especially in the resistant genotype Cappelli. In Colosseo, the sensitive cultivar, we could measure an up-regulation in flag leaf at T0 and 14 das (2- and 3.5-fold respectively) lower than the lowest value measured in Cappelli (5.6-fold at T0). In Cappelli's spikes, the expression is mainly constant, and the only relevant modulation does not go over 2.5 fold (Table 4.7). This result indicates that the function of this gene under DS could be more relevant in leaves than in spike tissues.

Both *TdBCAT-A* and *TdBCAT-B* show the same expression profile within the same tissues in the three different time points, but *TdBCAT-B* had the greatest up-regulation compared to the homoeologs *TdBCAT-A*. Moreover, the up-regulation is time-dependant as it increased along with the duration of the stress. In Colosseo's leaves, after 4-days of stress we could not see any relevant modulation.

The *TdBCAT* gene encodes for a predicted chloroplast BCAT enzyme, as the phylogenetic analysis suggested (paragraph 4.2.4). The *Arabidopsis* chloroplastic BCATs catalyse the last step of the biosynthesis of BCAAs.

In *Triticae* species, the most significant changes in metabolites under DS can be observed in the amino acids, organic acids and sugars (Ullah et al., 2017). Although the role of BCAAs under such conditions is still unknown, several studies indicated the pool of these amino acids accumulates in leaves of barley (Templer et al., 2017) and wheat (Bowne et al., 2012) under water-limiting conditions. Moreover, even if is still not clear if the BCAAs accumulates under DS because of a *de novo* synthesis or protein degradation, Urano and colleagues (2009) observed an accumulation of BCAAs, regulated at the transcriptional level, under water-limiting conditions in *Arabidopsis*.

Considering the type of modulation of this gene, it is reasonable to assume a role for this gene in the maintenance of the homoeostasis of BCAAs under DS conditions. It is known that primary metabolism is one of the first metabolic target affected by DS. The primary response of the plants to avoid water loss is the stomatal closure, which indirectly decreases the CO_2 diffusion and thus the photosynthetic rate, until a state of metabolic constraint. It has been proposed that BCAAs under conditions of energy shortage are catabolized by the TCA cycle to generate energy (Galili et al., 2016), thus an up-regulation of the anabolic enzyme could contribute to keep at a constant level this alternative fuel of the primary metabolism under DS condition in wheat flag leaves.

A clear correlation between the accumulation of BCAAs under DS and their *de novo* synthesis was not possible to assume so far, but on the basis of the evidences in literature (Scarpeci et al., 2017) it can be argued that gene *bcat3* contributes to acquire an higher tolerance as well as a good yield under DS in *Arabidopsis*. Considering all these results, the durum wheat *TdBCAT* looks a good candidate for the drought tolerance during the reproductive season, which is one of the most important in terms of yield in wheat.

TdABA7 gene showed a very peculiar grain-specific expression. Indeed, it was not expressed in leaf samples with the only exception in flag leaves of cv. Cappelli after 14 days of stress (Table 4.7). In this sample *TdABA7-A* is expressed in both control and treated tissues, but with no relevant modulation, as its up-regulation in stressed flag leaves is less than 2-fold. The expression of *TdABA7-A* is mainly constant across the three different points in spike tissues of both cultivars, but with no significant modulations in response to the stress treatment.

TdABA7-B shows the most interesting behaviour: it is detectable only in spikes at 14 das for the sensitive Colosseo and at 4 das for in Cappelli. In Colosseo, we can see a 2-fold down-regulation in stressed plants compared to controls, while in Cappelli the expression increases more than 25-fold in stressed spikes. *TdABA7-B* expression completely switched-on in response to DS in flag leaves of cv. Cappelli, as no transcription was detected in CTRL plants of cv. Cappelli. Nevertheless this transcript was expressed at low level in response to drought in stressed conditions, as the average Ct we could record was equal to 33, and we could not find any other later expression.

As we have already said, DS deeply influenced the maturation of Colosseo's kernel. It is likely that the differences in the timing of kernel maturation are influenced especially by the regulation of genes

involved into the seed development in which *TdABA7* could play an active role. Checking transcriptomic data of *Triticum aestivum* on expVIP browser (www.wheat-expression.com) (Ramírez-González et al., 2018), we could assess that the expression of *TaABA7* is highly related to grain development, as both A and B copies are expressed in the grain only after 15 days-after-anthesis in non-stress conditions. *TaABA7-B* is extremely grain-specific compared to the A-gene which is expressed also in various moment of the plant development in leaf tissues. Particularly, data from a study focused on the developing grain of Chinese Spring, showed that the expression is specific of the transfer cells of the caryopsis (Pfeifer et al., 2014). Endosperm transfer cells are involved into nutrients absorption and transport from the maternal tissues into the embryo and starch endosperm (Zheng et al., 2014).

Searching in databases (i.e. NCBI, EnsemblPlants) we could identify orthologues of *TdABA7* only in *Triticum aestivum*, *Triticum urartu, Aegilops tauschii, Hordeum vulgare*, and *Oryza punctate*, which are all cereal crops of the *Poaceae* family. The peculiar grain-specific expression along with the presence of this gene only in some cereals, suggests that *Aba7* could have co-evolved along with some species within the Gramineae taxon, with a possible function in the development of the caryopsis, which in wheat has some peculiar aspects respect to maize and sorghum (Zheng et al., 2014).

Further investigation will be necessary to unravel the function of this gene. A useful tool has been considered the analysis of tilling lines carrying mutation, which can determine a functional knock out, which results are presented in next paragraph.

4.4 The wheat TILLING approach

4.4.1 SNP detection by *in-silico* TILLING

The recently developed wheat TILLING resource (www.wheat-tilling.com) was used to identify mutations in *TdABA7-A*, *TdABA7-B*, *TdBCAT-A* and *TdBCAT-B* using the previously isolated gene sequence as query to gain TILLING data.

The database enquiry allowed the identification of 229 Kronos lines (Table **4.8**), which are predicted to carry 216 SNPs in target sequences.

Table 4.8 Summary of the Kronos mutant lines detected by the *in silico* TILLING. Amount of Kronos mutant lines are indicated for each variant category: nonsense, splice variant, missense, synonymous and intron variant.

[a] Gene	Nonsense variant	Splice variant	Missense variant	Synonymous variant	Intron variant	Total
TdABA7-A	2	0	16	6	0	24
TdABA7-B	1	0	20	16	0	37
TdBCAT-A	1	1	50	37	35	124
TdBCAT-B	1	1	16	8	17	44

The identification of mutations is hampered sometimes by the correspondence retrieved ahead between the gene sequence and the genomic scaffolds on which the wheat TILLING pipeline is based. In fact, for the *TdBCAT* genes a full length match was retrieved only for *TdBCAT-A*, allowing the identification of a number of mutation spread in the entire gene sequence, while for the *TdBCAT-B* a portion of this gene was uncovered, between a portion of exon 3 till the beginning of exon 5. Thus for this gene, no SNPs could be found for the corresponding region. The full lists of Kronos mutant lines detected by the *in-silico* TILLING database for *TdABA7-A*, *TdABA7-B*, *TdBCAT-A* and *TdBCAT-B* genes are provided in Table **4.9**, Table **4.10**, Table **4.11** and Table **4.12** respectively.

A total of 23, 33, 118 and 42 mutations were retrieved for *TdABA7-A*, *TdABA7-B*, *TdBCAT-A* and *TdBCAT-B* respectively. SNPs on the coding sequence, with different level of predicted zygosity (heterozygous or homozygous), were grouped in nonsense, missense, synonymous, splice-variant and intron variant, on the basis of the possible effect on the gene product.

The SNPs positions predicted in the database were manually aligned on the *TdABA7* and *TdBCAT* protein sequences.

Table 4.9 Kronos lines carrying mutation on the *TdABA7-A* gene retrieved in the wheat TILLING database.

Kronos Line	Position [a]	Mutation Confidence [a]	Zygosity [a]	Nucleotide Change [a] [b]	Consequence [b]	Amino Acid change [b]	SNP [c]
3721	16053	Medium	Het	G>A	Missense	P5S	1
0573	16004	High	Het	G>A	Missense	S22L	2
1242	15980	High	Het	C>T	Missense	G30E	3
4522	15971	High	Hom	C>T	Missense	G33E	4
3205	15945	High	Hom	C>T	Missense	G42R	5
3908	15929	Medium	Het	C>T	Missense	G47E	6
2513	15927	High	Het	C>T	Missense	E48K	7
1426	15915	High	Hom	G>A	Missense	L52F	8
2477	15915	High	Het	G>A	Missense	L52F	
2472	15882	High	Het	G>A	Missense	L63F	9
2590	15822	High	Het	C>T	Missense	D83N	10
2398	15716	High	Het	C>T	Missense	S118N	11
0358	15692	High	Het	G>A	Missense	S126F	12
2433	15662	High	Het	C>T	Missense	R136K	13
2221	15620	High	Het	C>T	Missense	R150H	14
2051	15579	High	Het	G>A	Missense	R164C	15
3781	15918	High	Het	G>A	Non-sense	Q51*	16
0323	15907	High	Het	C>T	Non-sense	W54*	17
2825	15979	High	Hom	C>T	Synonymous	G30=	18
0865	15949	High	Het	G>A	Synonymous	I40=	19
3931	15889	Medium	Het	C>A	Synonymous	L60=	20
2630	15883	High	Het	C>T	Synonymous	K62=	21
3720	15850	High	Hom	C>T	Synonymous	E73=	22
0890	15729	High	Het	G>A	Synonymous	L114=	23

[a] Position on CSS scaffold, confidence, SNP zygosity, nucleotide change have been predicted by the database. The reference CSS scaffold is the IWGSC_CSS_6AS_scaff_4428898 mapped on chromosome 6A.

[b] The nucleotide change, consequence and the amino acid change have been verified manually for all the mutations on the reference genomic sequence of durum wheat cv. Kronos retrievable at https://wheatis.tgac.ac.uk/grassroots-portal/blast.

[c] SNPs 1-17 are reported in Figure 4.16.A

Table 4.10 Kronos lines carrying mutation on the *TdABA7-B* gene retrieved in the online wheat TILLING database.

Kronos Line	Position (a)	Mutation Confidence (a)	Zygosity (a)	Nucleotide Change (a) (b)	Consequence (b)	Amino Acid change (b)	SNP (c)
3560	3331	medium	het	G>A	Missense	R11K	1
0375	3310	High	het	G>A	Missense	R18K	2
0278	3306	medium	het	G>T	Missense	E19D	3
2125	3292	High	het	G>A	Missense	S24N	4
3091	3284	High	het	G>A	Missense	D27N	5
4090	3278	High	het	G>A	Missense	G29S	6
2140	3265	High	het	G>A	Missense	G33E	7
4473	3239	High	het	G>A	Missense	G42R	8
3825	3122	High	hom	G>A	Missense	E81K	9
0456	3079	High	het	C>T	Missense	P95L	10
2237	3047	High	hom	G>A	Missense	E106K	11
3179	3019	High	hom	G>A	Missense	R115K	12
3960	3007	High	hom	G>A	Missense	R119K	13
3950	2999	High	hom	G>A	Missense	D122N	14
3502	2966	High	het	C>T	Missense	R133W	15
0095	2966	High	hom	C>T	Missense	R133W	
1164	2965	High	het	G>A	Missense	R133Q	16
2701	2923	High	het	G>A	Missense	G147E	17
3216	2914	High	het	G>A	Missense	R150H	18
3126	2899	High	het	G>A	Missense	R155K	19
4027	2888	High	het	C>T	Non-sense	Q159*	20
3935	3330	High	het	G>A	Synonymous	R11=	21
1181	3138	High	het	C>T	Synonymous	I75=	22
4247	3129	High	hom	C>T	Synonymous	I78=	23
1023	3105	medium	het	G>A	Synonymous	K86=	24
3591	3045	High	het	G>A	Synonymous	E106=	25
2225	3024	High	het	C>T	Synonymous	I113=	26
3240	3024	High	hom	C>T	Synonymous	I113=	
3840	2976	High	het	C>T	Synonymous	S129=	27
4670	2967	High	het	C>T	Synonymous	I132=	28
2327	2964	High	hom	G>A	Synonymous	R133=	29

table 4.10 continue

Kronos Line	Position [a]	Mutation Confidence [a]	Zygosity [a]	Nucleotide Change [a] [b]	Consequence [b]	Amino Acid change [b]	SNP [c]
4479	2955	High	hom	G>A	Synonymous	R136=	30
3974	2946	High	het	G>A	Synonymous	L139=	31
3091	2925	High	het	C>T	Synonymous	I146=	
3213	2925	High	het	C>T	Synonymous	I146=	32
2348	2898	High	het	G>A	Synonymous	R155=	
3487	2898	High	het	G>A	Synonymous	R155=	33

[a] Position on CSS scaffold, confidence, SNP zygosity, nucleotide change have been predicted by the database. The reference CSS scaffold is the IWGSC_CSS_2BS_scaff_5169659 mapped on chromosome 2B.

[b] The nucleotide change, consequence and the amino acid change have been verified manually for all the mutations on the reference genomic sequence of durum wheat cv. Kronos retrievable at https://wheatis.tgac.ac.uk/grassroots-portal/blast.

[c] SNPs 1-20 are reported in Figure 4.16.B

Table 4.11 Kronos lines carrying mutation on the *TdBCAT-A* gene retrieved in the wheat TILLING database.

Kronos Line	Position [a]	Mutation Confidence [a]	Zygosity [a]	Nucleotide Change [a] [b]	Consequence [b]	Amino Acid change [b]	SNP [c]
4486	1732	low	het	G>A	Missense_variant	A2T	1
1051	1745	medium	het	C>T	Missense_variant	S6F	2
4479	1748	medium	het	C>T	Missense_variant	A7V	3
2619	1775	high	het	C>T	Missense_variant	A16V	4
2928	1775	high	hom	C>T	Missense_variant	A16V	
0184	1817	low	het	G>A	Missense_variant	G30E	5
2738	2058	high	het	G>A	Missense_variant	G35D	6
2236	2067	high	het	C>T	Missense_variant	S38F	7
2172	2069	high	hom	C>T	Missense_variant	L39F	8
0557	2075	high	het	C>T	Missense_variant	P41S	9
2898	2102	high	het	C>T	stop_gained	Q50*	10
0338	2303	high	het	G>A	Missense_variant	E60K	11
2615	2309	medium	het	G>A	Missense_variant	G62S	12
4268	2337	high	het	G>A	Missense_variant	G71D	13
3825	2351	high	het	C>T	Missense_variant	P76S	14
2397	2355	high	het	C>T	Missense_variant	T77I	15
2547	2355	low	het	C>T	Missense_variant	T77I	
3868	2355	high	het	C>T	Missense_variant	T77I	
2275	2374	high	het	G>A	Missense_variant	M83I	16
2380	2409	high	het	G>A	Missense_variant	G95D	17
3716	2442	high	het	G>A	Missense_variant	S106N	18
2557	2474	high	het	G>A	Missense_variant	G117R	19
2321	2502	low	het	G>A	Missense_variant	R126Q	20
3326	2514	high	het	G>A	Missense_variant	G130E	21
4271	2517	high	het	C>T	Missense_variant	P131L	22
4601	2540	high	het	G>A	Missense_variant	E139K	23

Table 4.11 continue

Kronos Line	Position [a]	Mutation Confidence [a]	Zygosity [a]	Nucleotide Change [a] [b]	Consequence [b]	Amino Acid change [b]	SNP [c]
2734	2573	high	het	G>A	Missense_variant	G150R	24
0282	2576	high	het	C>T	Missense_variant	R151C	25
2369	2618	high	het	G>A	Missense_variant	A165T	26
0255	2621	high	het	G>A	Missense_variant	V166I	27
3691	2648	high	het	C>T	Missense_variant	R175C	28
2348	2774	medium	het	G>A	Missense_variant&splice_region	V177M	29
1222	2780	high	het	C>T	Missense_variant	P179S	30
3932	2823	high	het	G>A	Missense_variant	G193E	31
2376	2831	high	hom	G>A	Missense_variant	A196T	32
0497	2865	high	het	C>T	Missense_variant	T207I	33
2885	2955	high	het	C>T	Missense_variant	A237V	34
4319	2963	high	het	G>A	Missense_variant	G240S	35
4209	3102	high	het	G>A	Missense_variant	A257T	36
3321	3156	high	het	G>A	Missense_variant	V275I	37
1166	3211	high	het	G>A	Missense_variant	G293D	38
4289	3211	high	het	G>A	Missense_variant	G293D	39
0323	3213	high	het	G>A	Missense_variant	G294S	40
3128	3237	high	het	G>A	Missense_variant	E302K	41
4059	3241	high	het	G>A	Missense_variant	G303E	42
1242	3436	high	hom	C>T	Missense_variant	L335F	43
2824	3478	high	hom	G>A	Missense_variant	G349S	44
0467	3481	high	het	G>A	Missense_variant	V350I	45
4210	3511	high	hom	G>A	Missense_variant	G360R	46
2450	3515	high	hom	C>T	Missense_variant	T361I	47
4346	3624	high	het	G>A	splice_acceptor	Splice Junction	48
3719	3717	high	het	G>A	Missense_variant	G393E	49

Table 4.11 continue

Kronos Line	Position [a]	Mutation Confidence [a]	Zygosity [a]	Nucleotide Change [a] [b]	Consequence [b]	Amino Acid change [b]	SNP [c]
3511	1743	high	hom	G>A	Synonymous_variant	S5=	50
4289	1797	high	het	C>T	Synonymous_variant	G23=	51
1234	2083	high	het	G>A	Synonymous_variant	R43=	52
0548	2116	high	het	C>T	Synonymous_variant	V54=	53
4409	2299	high	het	C>T	Synonymous_variant	D58=	54
2323	2305	medium	het	G>A	Synonymous_variant	E60=	55
3188	2329	medium	het	C>T	Synonymous_variant	D68=	56
2937	2350	medium	het	C>T	Synonymous_variant	T75=	57
2557	2389	low	het	G>A	Synonymous_variant	E88=	58
2825	2398	high	het	C>T	Synonymous_variant	G91=	59
2235	2404	high	het	C>T	Synonymous_variant	S93=	60
0537	2419	high	het	C>T	Synonymous_variant	A98=	61
3598	2446	high	hom	C>T	Synonymous_variant	P107=	62
4539	2449	high	het	C>T	Synonymous_variant	S108=	63
2991	2473	high	het	G>A	Synonymous_variant	Q116=	64
2018	2476	high	het	G>A	Synonymous_variant	G117=	65
4389	2485	high	het	G>A	Synonymous_variant	E120=	66
3900	2521	high	het	G>A	Synonymous_variant	G132=	67
2694	2542	high	het	G>A	Synonymous_variant	E139=	68
2997	2545	high	het	G>A	Synonymous_variant	E140=	69
3025	2548	high	het	C>T	Synonymous_variant	N141=	70
0118	2557	high	het	G>A	Synonymous_variant	R144=	71
2692	2611	high	het	C>T	Synonymous_variant	F162=	72
2554	2836	high	het	C>T	Synonymous_variant	I197=	73
1139	2953	high	het	C>T	Synonymous_variant	R236=	74
4219	2983	high	hom	G>A	Synonymous_variant	K246=	75
3100	3146	high	het	C>T	Synonymous_variant	Y271=	76

Table 4.11 continue

Kronos Line	Position [a]	Mutation Confidence [a]	Zygosity [a]	Nucleotide Change [a] [b]	Consequence [b]	Amino Acid change [b]	SNP [c]
4043	3164	high	het	G>A	Synonymous_variant	K277=	77
3233	3191	high	het	C>T	Synonymous_variant	C286=	78
3954	3197	high	hom	C>T	Synonymous_variant	L288=	79
3581	3242	high	hom	G>A	Synonymous_variant	G303=	80
3622	3254	high	het	G>A	Synonymous_variant	P307=	81
0447	3456	high	het	G>A	Synonymous_variant	V341=	82
1027	3480	high	hom	C>T	Synonymous_variant	G349=	83
3335	3688	high	hom	C>T	Synonymous_variant	I383=	84
3714	3688	high	het	C>T	Synonymous_variant	I383=	
4364	3712	high	het	G>A	Synonymous_variant	K391=	85
2054	2766	high	het	G>A	splice_region&Intron_variant		86
2573	1830	high	het	C>T	Intron_variant		87
0776	1837	high	hom	C>T	Intron_variant		88
0478	1864	high	hom	C>T	Intron_variant		89
0216	1884	high	het	C>T	Intron_variant		90
2731	1942	high	hom	C>T	Intron_variant		91
2670	1952	high	hom	G>A	Intron_variant		92
2117	1959	high	het	C>T	Intron_variant		93
2101	1995	high	hom	G>A	Intron_variant		94
3538	2025	high	het	G>A	Intron_variant		95
2224	2036	high	het	C>T	Intron_variant		96
3353	2036	high	het	C>T	Intron_variant		
0640	2152	high	het	C>T	Intron_variant		97
3284	2157	high	het	C>T	Intron_variant		98
0110	2186	high	het	G>A	Intron_variant		99
2657	2189	high	het	C>T	Intron_variant		100

Table 4.11 continue

Kronos Line	Position [a]	Mutation Confidence [a]	Zygosity [a]	Nucleotide Change [a] [b]	Consequence [b]	Amino Acid change [b]	SNP [c]
3726	2201	high	het	G>A	Intron_variant		101
0909	2209	high	het	C>T	Intron_variant		102
2514	2236	low	het	C>T	Intron_variant		103
3842	2241	high	het	C>T	Intron_variant		104
3078	2248	high	het	G>A	Intron_variant		105
4525	2261	high	hom	G>A	Intron_variant		106
3221	2267	high	het	C>T	Intron_variant		107
3510	2267	high	het	C>T	Intron_variant		
4660	2268	high	het	C>T	Intron_variant		108
0119	2695	high	het	C>T	Intron_variant		109
1186	2705	high	het	C>T	Intron_variant		110
4595	2712	high	het	C>T	Intron_variant		111
2592	2739	high	het	G>A	Intron_variant		112
0214	2754	high	het	C>A	Intron_variant		113
0337	3360	high	hom	C>T	Intron_variant		114
0596	3379	high	het	G>A	Intron_variant		115
2879	3540	high	hom	C>T	Intron_variant		116
4096	3570	high	het	G>A	Intron_variant		117
0118	3589	high	hom	G>A	Intron_variant		118

[a] Position on CSS scaffold, confidence, SNP zygosity, nucleotide change, have been predicted by the database. The reference CSS scaffold is the IWGSC_CSS_4AS_scaff_5887798 mapped on chromosome 4A.

[b] The nucleotide change, consequence and the amino acid change have been verified manually for all the mutations on on the reference genomic sequence of durum wheat cv. Kronos retrievable at https://wheatis.tgac.ac.uk/grassroots-portal/blast

[c] SNPs 1-49 are reported in Figure 4.17.A

Table 4.12 Kronos lines carrying mutation on the *TdBCAT-B* gene retrieved in the wheat TILLING database. For gene *TdBCAT-B* two reference scaffolds were identified, the CSS IWGSC_CSS_4BL_scaff_7034480 scaffold and the UCW_Kronos scaffold UCW_Kronos_U_jcf7180000453710

Kronos Line	Scaffold (a)	Position (b)	Mutation Confidence (b)	Zygosity (b)	Nucleotide Change (b) (c)	Consequence (c)	Amino Acid change (c)	SNP (d)
2569	U	194	high	hom	C>T	Missense_variant	A7V	1
3214	U	215	high	hom	G>A	Missense_variant	G14D	2
2329	U	230	high	hom	G>A	Missense_variant	G19D	3
2155	U	239	high	hom	G>A	Missense_variant	S22N	4
0852	U	250	medium	het	C>T	STOP gained	R26*	5
4065	U	503	medium	het	G>T	Missense_variant	G34V	6
2365	U	556	high	het	G>A	Missense_variant	D52N	7
0685	U	560	high	het	C>T	Missense_variant	P53L	8
3723	4B	61	high	hom	G>A	missense_variant	D273N	9
3321	4B	67	high	hom	G>A	missense_variant	V275M	10
3954	4B	125	high	het	G>A	missense_variant	G294D	11
0411	4B	128	high	het	C>T	missense_variant	A295V	12
0382	4B	182	high	het	G>A	missense_variant	S313N	13
0257	4B	355	high	hom	G>A	missense_variant	D339N	14
2087	4B	355	high	hom	G>A	missense_variant	D339N	
4672	4B	425	high	het	G>A	missense_variant	R362K	15
0860	4B	515	high	het	G>A	splice_acceptor_variant	Splice Junction	16
0112	4B	592	high	hom	G>A	missense_variant	A388T	17
0846	4B	608	high	het	G>A	missense_variant	G393E	18
3196	U	540	high	hom	G>A	Synonymous_variant	S46=	19
3490	4B	12	medium	het	G>A	Synonimous_variant	K256=	20
2631	4B	52	high	het	C>T	synonymous_variant	L270=	21
4568	4B	105	high	het	C>T	synonymous_variant	N287=	22
2691	4B	120	high	het	G>A	synonymous_variant	K292=	23

Table 4.12 continue

Kronos Line	Scaffold [a]	Position [b]	Mutation Confidence [b]	Zygosity [b]	Nucleotide Change [b] [c]	Consequence [c]	Amino Acid change [c]	SNP [d]
4327	4B	168	high	hom	G>A	synonymous_variant	G308=	24
2701	4B	180	high	het	G>A	synonymous_variant	R312=	25
1348	4B	321	high	het	G>A	synonymous_variant	E327=	26
4027	U	271	high	hom	C>T	Intron_variant		27
1107	U	276	high	het	C>T	Intron_variant		28
3901	U	286	low	het	C>T	Intron_variant		29
2694	U	382	high	het	G>A	Intron_variant		30
2281	U	400	low	het	G>A	Intron_variant		31
2631	U	444	high	het	G>A	Intron_variant		32
3900	U	477	high	hom	G>A	Intron_variant		33
0377	U	484	high	het	C>T	Intron_variant		34
3273	U	605	low	hom	C>T	Intron_variant		35
2338	U	643	high	hom	C>T	Intron_variant		36
4623	U	643	high	hom	C>T	Intron_variant		
4590	U	666	high	hom	C>T	Intron_variant		37
4320	4B	236	high	het	C>T	intron_variant		38
2024	4B	447	medium	het	C>T	intron_variant		39
2928	4B	448	high	het	G>A	intron_variant		40
2520	4B	470	high	het	G>A	intron_variant		41
4066	4B	501	high	het	G>A	intron_variant		42

[a] Scaffolds were retrieved on the database. U referes to UCW_Kronos_U_jcf7180000453710 and 4B refers to CSS IWGSC_CSS_4BL_scaff_7034480 .

b) Position on scaffold, confidence, SNP zygosity, nucleotide change, PSSM Difference and SIFT score have been predicted by the database.

[c] The nucleotide change, consequence and the amino acid change have been verified manually for all the mutations on the reference genomic sequence of durum wheat cv. Kronos retrievable at https://wheatis.tgac.ac.uk/grassroots-portal/blast

[d] SNPs 1-18 are reported in Figure 4.17.B

For *TdABA7-A*, 23 mutations were detected with the *in silico* TILLING (Figure 4.16.a). Six were synonymous mutations, whereas 15 were classified as missense mutations. Importantly, 2 nonsense mutation (Q51*, and Q54*) were detected.

For *TdABA7-B*, 33 mutations were detected (Figure 4.16.b), of which 13 were synonymous mutations, 19 were classified as missense mutations and 1 lead to a premature STOP gained (Q159*).

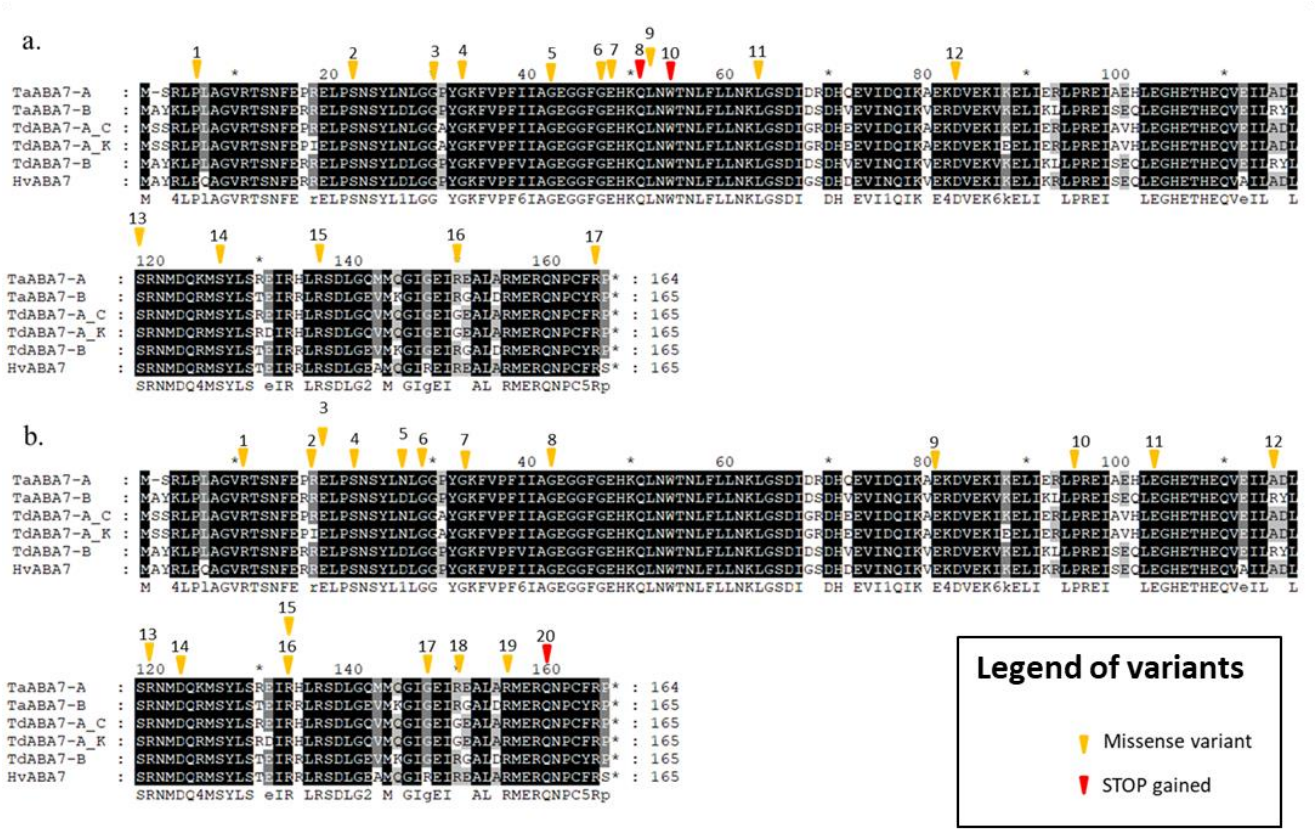


Figure 4.16 Multiple alignment of ABA7 proteins of *T. aestivum*, *T. durum* and *H. vulgare* with missense and nonsense mutations detected by searching the online *in silico* TILLING database using *TdABA7* sequence as query. The missense mutation are indicated with the orange arrows, while the nonsense mutations with the red arrows (a.) 15 missense mutations and 2 nonsense were found for TdABA7-A and (b.) 19 missense and 1 nonsense for TdABA7-B. For protein TdABA7-A both Kronos (TdABA7-A_K) and Cham1 (TdABA7-A_C) sequences are displayed in the multiple alignment.

For *TdBCAT-A*, 118 mutations were detected (**Table 4.11**), of which 33 intron variants, 36 synonymous mutations. All the other variants are reported in Figure 4.17.a, of which 47 missense mutations, 1 splice variant 1 splice variant and 1 lead to a premature STOP gained (Q50*).

For *TdBCAT-B* 42 mutations were detected (**Table 4.12**), of which 16 are intron variants and 8 are synonymous mutations. In figure 4.17.b are reported the 15 missense mutations, 1 splice variant and 1 non-sense which lead to a premature STOP gained (R26*).

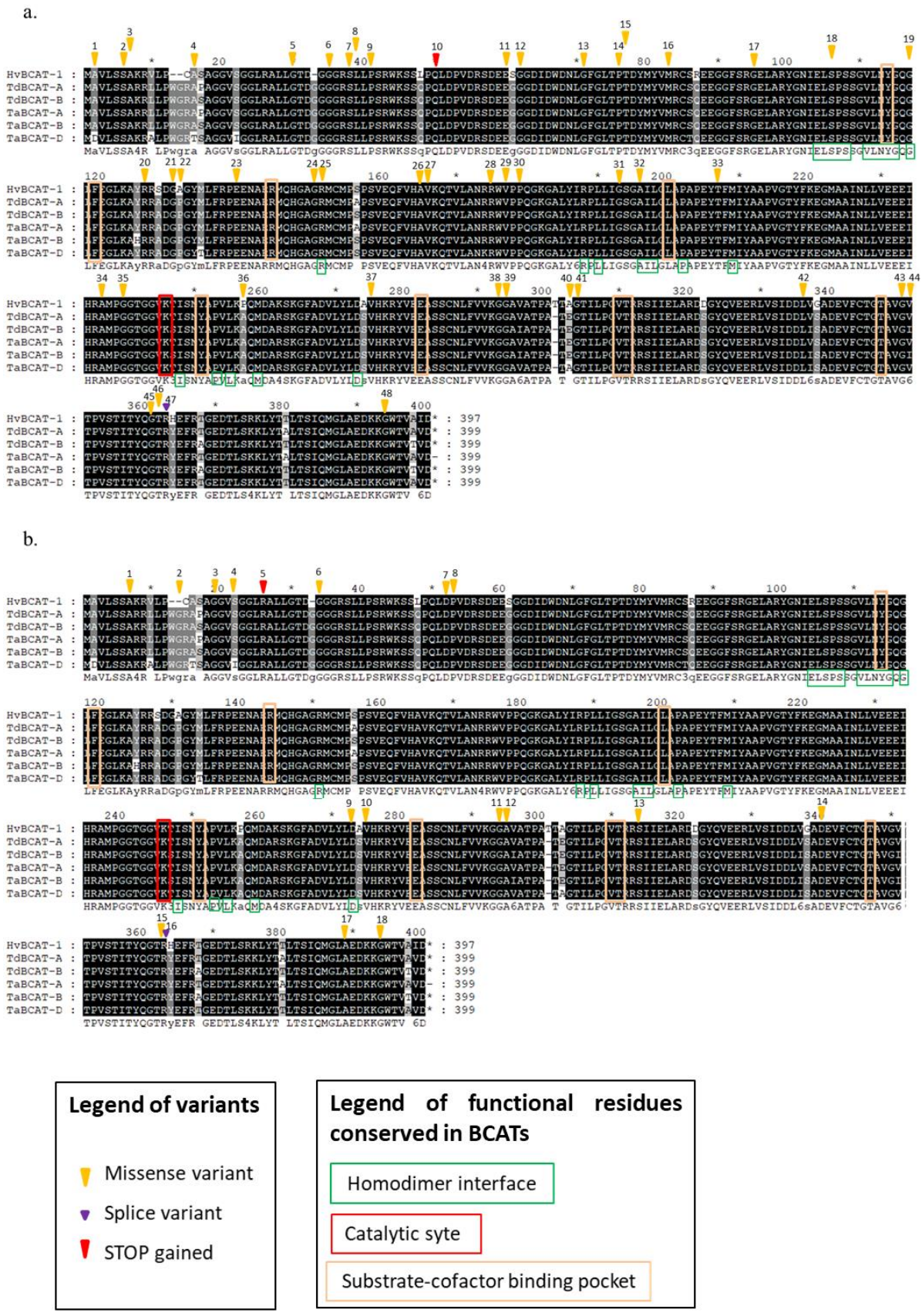


Figure 4.17 Multiple alignment of BCAT proteins of *T. aestivum*, *T. durum* and *H. vulgare* with missense and nonsense mutations, and splice variants detected by searching the online *in silico* TILLING database using TdBCAT sequence as query. The missense mutation are indicated with the orange arrows, the splice variant with purple arrow and the nonsense mutation with the red arrow. In a) the mutations for TdBCAT-A, and in b) the mutations for TdBCAT-B. Functionally important residues are reported on the basis of the annotation of the Superfamily Protein Domain PRK13357 retrieved on the NCBI-Conserved Domains database at https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml).

To demonstrate the involvement of the target genes in the mechanism of the DS resilience a set of mutant lines where further selected with the final goal to link a specific mutation to a phenotype under stress conditions. For this purpose, we focused on mutations that lead to a deleterious effect on the protein function: nonsense, splice-variant and missense mutation in conserved protein domains.

A top down approach was followed in the selection of the mutant lines: we first selected truncation mutations independently of their position then splice variants (only for *TdBCAT-A* and *TdBCAT-B*) and missense mutations in conserved positions. Seven mutants were selected for the *in vivo* validation (Table 4.13).

Table 4.13 Mutant lines selected for *in vivo* validation. For each line, the target gene which carries the mutation is indicated along with the predicted effect and the expected zygosity

Line	Target Gene	Effect	Predicted zygosity
Kronos 3781	*TdABA7-A*	stop gained (Q51*)	Het
Kronos 323	*TdABA7-A*	stop gained (W54*)	Het
Kronos 4027	*TdABA7-B*	stop gained (Q159*)	Het
Kronos 3950	*TdABA7-B*	missense variant (D122N)	Hom
Kronos 2898	*TdBCAT-A*	stop gained (Q50*)	Het
Kronos 860	*TdBCAT-B*	splice acceptor variant	Het
Kronos 852	*TdBCAT-B*	stop gained (R26*)	Het

For TdABA7-A two lines were found to carry missense mutations, while for TdABA7-B only one (Q159*) was retrieved. Since this gene has a 1 exon-structure, the second choice was selected among the possible missense mutations. For TdABA7 there are no information about conservative protein domains. The mutant line Kronos 3950 was chosen, it carries an allele that leads to an amino acid change (D122N) of a conserved residue in all the putative protein orthologues of ABA7 (*T. urartu, A. tauschii, H. vulgare, O. punctata*) available on NCBI and EnsemblPlants database. The amino acid change is between a negative-charged amino acid and the positive-charged asparagine, thus it is likely that this substitution can alter the protein structure/function.

For *TdBCAT-A* a very premature nonsense mutation (Q50*) is promising for the knock-out of protein function. On the other hand, for *TdBCAT-B* the detection of an unanchored Kronos scaffold (UCW_Kronos_U_jcf7180000453710) on the *in-silico* TILLING database was fundamental to identify a STOP-gained variant (R26*). Another interesting allele was chosen, a splice acceptor variant between exon 6 and exon 7. This mutation should cut out the last 37 amino acids of the protein, which are highly conserved among all the BCAT proteins.

4.4.2 SNP validation through KASP assay

To confirm *in vivo* the presence of the detected SNPs, KASP homeologs-specific markers were designed and tested on the DNA extracted from the selected Kronos lines. All the SNPs were confirmed by KASP analysis by evaluating the FAM and HEX fluorescence signal on a single genotype. The distribution of the fluorescence signals within each line was analyzed to confirm the predicted level of zygosity of the mutants.

An example of the results obtained is reported in Figure 4.18 for lines Kronos 3781, Kronos 4027, Kronos 860 and Kronos 3950 where the KASP markers were applied. Kronos 3950 was confirmed as homozygous for the mutant allele as predicted by the TILLING database with 100% of homozygous individuals. For all the remaining lines, the TILLING *in silico* predicted a heterozygous state. The KASP analyses on an average of 20 individuals for each Kronos mutant line showed a slight distortion of the expected 1:2:1 ratio (Hom wt: Het: Hom mut) for the predicted heterozygous, which could be related to the low number of plants tested (Figure 4.19).

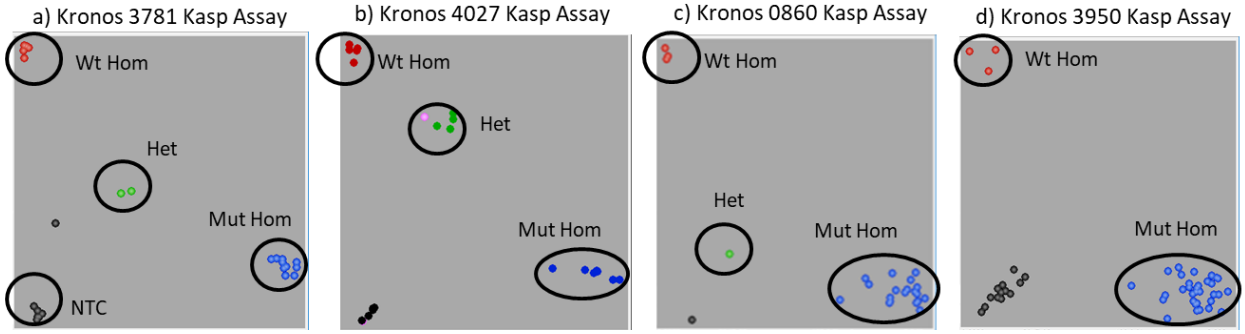


Figure 4.18 Examples of the KASP assay output (KBiosciences Klustercaller software) for Kronos TILLING mutant lines a) 3781, b) 4027, c) 0860, d) 3950. Analyses were performed on Kronos lines carrying mutations in the target genes. Each KASP assay was used to test the Kronos mutants and the Kronos wild type. (wt =wild-type, Mut=mutant, Hom=homozygous, Het=heterozygous, NTC=non-template control).

The KASP assay confirmed for all the selected lines the presence of the mutant allele and the predicted zygosity, allowing the planning of the following required backcrosses for the background cleaning.

Table 4.14 Results of the assay on the mutant plants.

Line	Target Gene	Effect	Predicted zygosity	Total plants	Hom wt (a)	Het (b)	Hom mut (c)
Kronos 3781	*TdABA7-A*	stop gained (Q51*)	Het	15	3	2	10
Kronos 0323	*TdABA7-A*	stop gained (W54*)	Het	34	15	9	10
Kronos 4027	*TdABA7-B*	stop gained (Q159*)	Het	16	3	5	8
Kronos 3950	*TdABA7-B*	missense variant (D122N)	Hom	35	0	0	35
Kronos 2898	*TdBCAT-A*	stop gained (Q50*)	Het	20	9	3	8
Kronos 0860	*TdBCAT-B*	splice variant	Het	21	1	1	19
Kronos 0852	*TdBCAT-B*	stop gained (R26*)	Het	21	6	4	11

(*a) Number of individuals homozygous for the wild-type allele; (b) number of heterozygous plants; (c) number of individuals homozygous for the mutant allele.*

In total, out of 162 individuals of 7 mutant lines tested, 62% were Hom mut , 23% Het and 15% Hom wt.

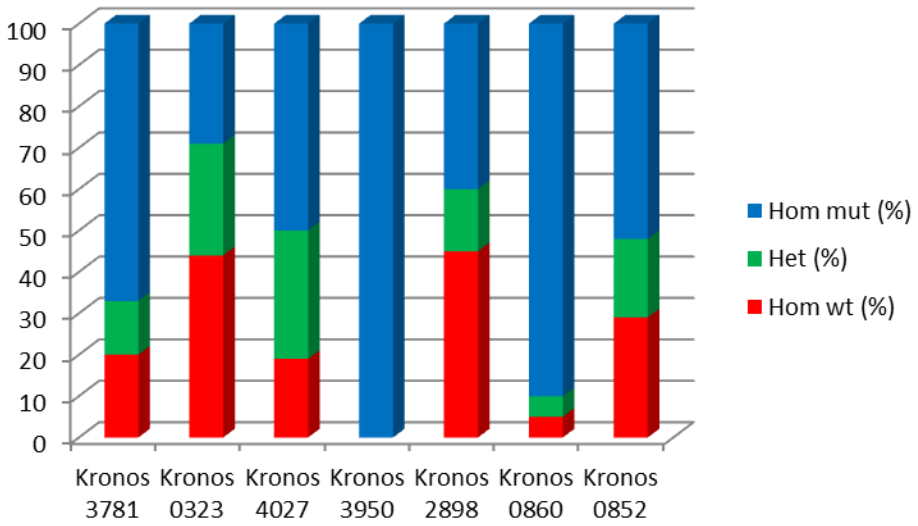


Figure 4.19 Histogram representing % of the three diverse allele combinations (y-axis) obtained with the KASP assay for each line tested.

4.4.3 Generation of double mutant lines and back-crosses

The TILLING approach involves a random mutation spanning the entire genome and thus requires several rounds of backcrosses to clean the high-load of background mutations (Adamski et al., 2018). Moreover, due to the polyploidy nature of the wheat genome a functional redundancy is known (Borrill et al., 2015), thus the generation of double mutants lines carrying knockout mutations in both homoeologs is required to appreciate a phenotype related to a knock-out of the target gene.

The double mutant lines for *TdBCAT* and *TdABA7* selected lines were generated. In addition, single mutants were generated when possible, by back-crossing with wild type individuals of cv. Kronos.

A total of 11 and 15 lines were obtained through crossing of A and B mutant lines for *TdBCAT* (Table 4.15) and *TdABA7* (Table 4.16).respectively.

Table 4.15 Full list of the successful crosses events for *TdBCAT* genes. For each target gene the event id, the parental lines, the type of mutation and the number of seeds obtained are listed.

Target gene	event id	Parental lines: E=emasculated; P=pollen donor		mutation type	number of seed
TdBCAT-B	VB33	Parental line 1 (E)	Kronos 0852-18	stop gained (R26*)	13
		Parental line 2 (P)	Kronos -wt	wt allele	
TdBCAT-B	VB66	Parental line 1 (E)	Kronos 0852-6	stop gained (R26*)	7
		Parental line 2 (P)	Kronos -wt	wt allele	
TdBCAT-A	VB73	Parental line 1 (E)	Kronos 2898-19	stop gained (Q50*)	3
		Parental line 2 (P)	Kronos -wt	wt allele	
TdBCAT-B	VB55	Parental line 1 (E)	Kronos 0860-18	splice acceptor variant	1
TdBCAT-A		Parental line 2 (P)	Kronos 2898-10	stop gained (Q50*)	
TdBCAT-A	VB89	Parental line 1 (E)	Kronos 2898-2	stop gained (Q50*)	2
TdBCAT-B		Parental line 2 (P)	Kronos 0860-10/17	splice acceptor variant	
TdBCAT-B	VB92	Parental line 1 (E)	Kronos 0860-8	splice acceptor variant	9
TdBCAT-A		Parental line 2 (P)	Kronos 2898-10	stop gained (Q50*)	
TdBCAT-A	VB93	Parental line 1 (E)	Kronos 2898-5	stop gained (Q50*)	11
TdBCAT-B		Parental line 2 (P)	Kronos 0860-21	splice acceptor variant	
TdBCAT-A	VB95	Parental line 1 (E)	Kronos 2898-6	stop gained (Q50*)	7
TdBCAT-B		Parental line 2 (P)	Kronos 0860-19	splice acceptor variant	
TdBCAT-A	VB96	Parental line 1 (E)	Kronos 2898-6	stop gained (Q50*)	19
TdBCAT-B		Parental line 2 (P)	Kronos 0860-10	splice acceptor variant	
TdBCAT-B	VB97	Parental line 1 (E)	Kronos 0860-19	splice acceptor variant	11
TdBCAT-A		Parental line 2 (P)	Kronos 2898-2	stop gained (Q50*)	
TdBCAT-B	VB47	Parental line 1 (E)	Kronos 0852-5	stop gained (R26*)	4
TdBCAT-A		Parental line 2 (P)	Kronos 2898-2	stop gained (Q50*)	
TdBCAT-B	VB54	Parental line 1 (E)	Kronos 0852-13	stop gained (R26*)	4
TdBCAT-A		Parental line 2 (P)	Kronos 2898-2	stop gained (Q50*)	
TdBCAT-B	VB94	Parental line 1 (E)	Kronos 0852-16	stop gained (R26*)	3
TdBCAT-A		Parental line 2 (P)	Kronos 2898-20	stop gained (Q50*)	
TdBCAT-B	VB98	Parental line 1 (E)	Kronos 0852-12	stop gained (R26*)	15
TdBCAT-A		Parental line 2 (P)	Kronos 2898-6	stop gained (Q50*)	

Table 4.16 Full list of the successful crosses events for *TdBCAT* genes. For each target gene the event id, the parental lines, the type of mutation and the number of seeds obtained are listed

Target gene	Cross event	Parental lines: E=emasculated; P=pollen donor		mutation type	number of seed
TdABA7-A	VB1	Parental line 1 (E)	Kronos 0323-23	stop gained (W54*)	9
		Parental line 2 (P)	Kronos -wt	Wt-allele	
TdABA7-A	VB2	Parental line 1 (E)	Kronos 0323-9	stop gained (W54*)	8
		Parental line 2 (P)	Kronos -wt	Wt-allele	
TdABA7-B	VB65	Parental line 1 (E)	Kronos 3950-20	missense variant (D122N)	3
		Parental line 2 (P)	Kronos -wt	Wt-allele	
TdABA7-B	VB88	Parental line 1 (E)	Kronos 4027-3	stop gained (Q159*)	1
		Parental line 2 (P)	Kronos -wt	Wt-allele	
TdABA7-A	VB104	Parental line 1 (E)	Kronos 3781-8	stop gained (Q51*)	15
		Parental line 2 (P)	Kronos -wt	Wt-allele	
TdABA7-A	VB4	Parental line 1 (E)	Kronos 0323-1	stop gained (W54*)	5
TdABA7-B		Parental line 2 (P)	Kronos 4027-13	stop gained (Q159*)	
TdABA7-A	VB5	Parental line 1 (E)	Kronos 0323-13	stop gained (W54*)	2
TdABA7-B		Parental line 2 (P)	Kronos 4027-13	stop gained (Q159*)	
TdABA7-A	VB13	Parental line 1 (E)	Kronos 0323-27	stop gained (W54*)	8
TdABA7-B		Parental line 2 (P)	Kronos 4027-3	stop gained (Q159*)	
TdABA7-A	VB14	Parental line 1 (E)	Kronos 0323-22	stop gained (W54*)	8
TdABA7-B		Parental line 2 (P)	Kronos 4027-13	stop gained (Q159*)	
TdABA7-B	VB19	Parental line 1 (E)	Kronos 3950-27	missense variant (D122N)	2
TdABA7-A		Parental line 2 (P)	Kronos 0323-10	stop gained (W54*)	
TdABA7-B	VB20	Parental line 1 (E)	Kronos 3950-2	missense variant (D122N)	4
TdABA7-A		Parental line 2 (P)	Kronos 0323-10	stop gained (W54*)	
TdABA7-B	VB34	Parental line 1 (E)	Kronos 4027-11	stop gained (Q159*)	6
TdABA7-A		Parental line 2 (P)	Kronos 0323-3	stop gained (W54*)	
TdABA7-B	VB44	Parental line 1 (E)	Kronos 3950-13	missense variant (D122N)	13
TdABA7-A		Parental line 2 (P)	Kronos 0323-26	stop gained (W54*)	
TdABA7-B	VB60	Parental line 1 (E)	Kronos 3950-26	missense variant (D122N)	3
TdABA7-A		Parental line 2 (P)	Kronos 0323-5	stop gained (W54*)	
TdABA7-B	VB62	Parental line 1 (E)	Kronos 4027-9	stop gained (Q159*)	7
TdABA7-A		Parental line 2 (P)	Kronos 0323-6	stop gained (W54*)	
TdABA7-A	VB77	Parental line 1 (E)	Kronos 0323-23	stop gained (W54*)	1
TdABA7-B		Parental line 2 (P)	Kronos 3950-15	missense variant (D122N)	
TdABA7-A	VB99	Parental line 1 (E)	Kronos 3781-4	stop gained (Q51*)	14
TdABA7-B		Parental line 2 (P)	Kronos 4027-1	stop gained (Q159*)	

Table 4.16 continue

Target gene	Cross event	Parental lines: E=emasculated; P=pollen donor		mutation type	number of seed
TdABA7-A	VB101	Parental line 1 (E)	Kronos 3781-6	stop gained (Q51*)	7
TdABA7-B		Parental line 2 (P)	Kronos 4027-4	stop gained (Q159*)	
TdABA7-B	VB102	Parental line 1 (E)	Kronos 3950-31	missense variant (D122N)	15
TdABA7-A		Parental line 2 (P)	Kronos 3781-14	stop gained (Q51*)	
TdABA7-A	VB103	Parental line 1 (E)	Kronos 3781-15	stop gained (Q51*)	2
TdABA7-B		Parental line 2 (P)	Kronos 4027-4	stop gained (Q159*)	

Three double mutant lines for both *TdBCAT* and *TdABA7* were selected for the back-crosses with the recurrent parent cv. Kronos.

VB92, VB93, VB98 lines were selected for *TdBCAT*. The observed germination rate (i.e. percentage of seeds germinated to the total seed sowned) were 44% for VB92 LINE and 90% for the others. VB34, VB5, VB99 were selected for *TdABA7* with a germination rate of 33% for VB34 and 100% of seeds germinated for the others.

Each plant tested was verified for the presence of both mutant alleles with KASP assays, confirming the presence of 100% of heterozygous individuals in both A and B *loci*.

In April 2018 a back-cross pipeline was established for the background cleaning and several double mutants lines have been already backcrossed reaching already the 50% of the background cleaning (BC1). The work presented in this section is still ongoing; however it represents a key step to gain a phenotype clearly associable to a deleterious mutation in TILLING mutants of wheat, which can have a really widespread background of mutations along the genome.

Taking all together these results, highlight the efficacy of the TILLING approach in polyploidy species, since a high number of mutant lines have been recovered. The feasibility of the system is also supported by the availability of the KASP assay to validate and follow the mutation through generation and represent a milestone of the project since the development of a number of KASP markers might be applied MAS also in field.

The ongoing process of backcrossing although necessary is time consuming, but the exploitation of emerging tool like the "speed breeding method" (Watson et al., 2018) will shorten the time required to obtain stable double mutants for the phenotype evaluations. The further phenotyping characterization required to link the mutation to a specific plant response is ongoing, even though preliminary experiments have been planned.

4.5 Allele mining by Targeted-Resequencing to find natural variation of *TdBCAT*

To identify the natural allele variants in the *TdBCAT* genes in durum wheat, a new targeted-resequencing strategy was used as an allele mining approach in a set of durum wheat landraces and five reference varieties, selected on the basis of their behaviour under DS.

The targeted-resequencing is a PCR-based approach that allows sequencing a panel of GOIs by NGS. NGS technologies nowadays can still be cost-limiting, especially for the screening of germplasm collection, which are frequently populated by a high number of accessions.

This approach could be particularly relevant for the identification of SNPs, into polyploid species like wheat, which are characterized by a big-size and complex genome, which thus did not make affordable to everyone whole-genome sequencing for variant detection.

A method coupling targeted NGS, pooling of target in different individuals and the application of KASP assay for the SNPs detections is here proposed as valuable method for the allele mining in durum wheat.

The *TdBCAT-A* and *–B* genes were selected for this allele mining approach due to its involvement in drought tolerance in *Arabidopsis* (Scarpeci et al., 2017), which were furtherly enforced by our results that showed a high rate of expression under water-limiting conditions at the reproductive stages, particularly in a cultivar more tolerant to DS paragraph 4.3.2.3.

A scheme of the overall procedure for allele mining approach by targeted resequencing coupled with KASP assay is presented in Figure 4.20.

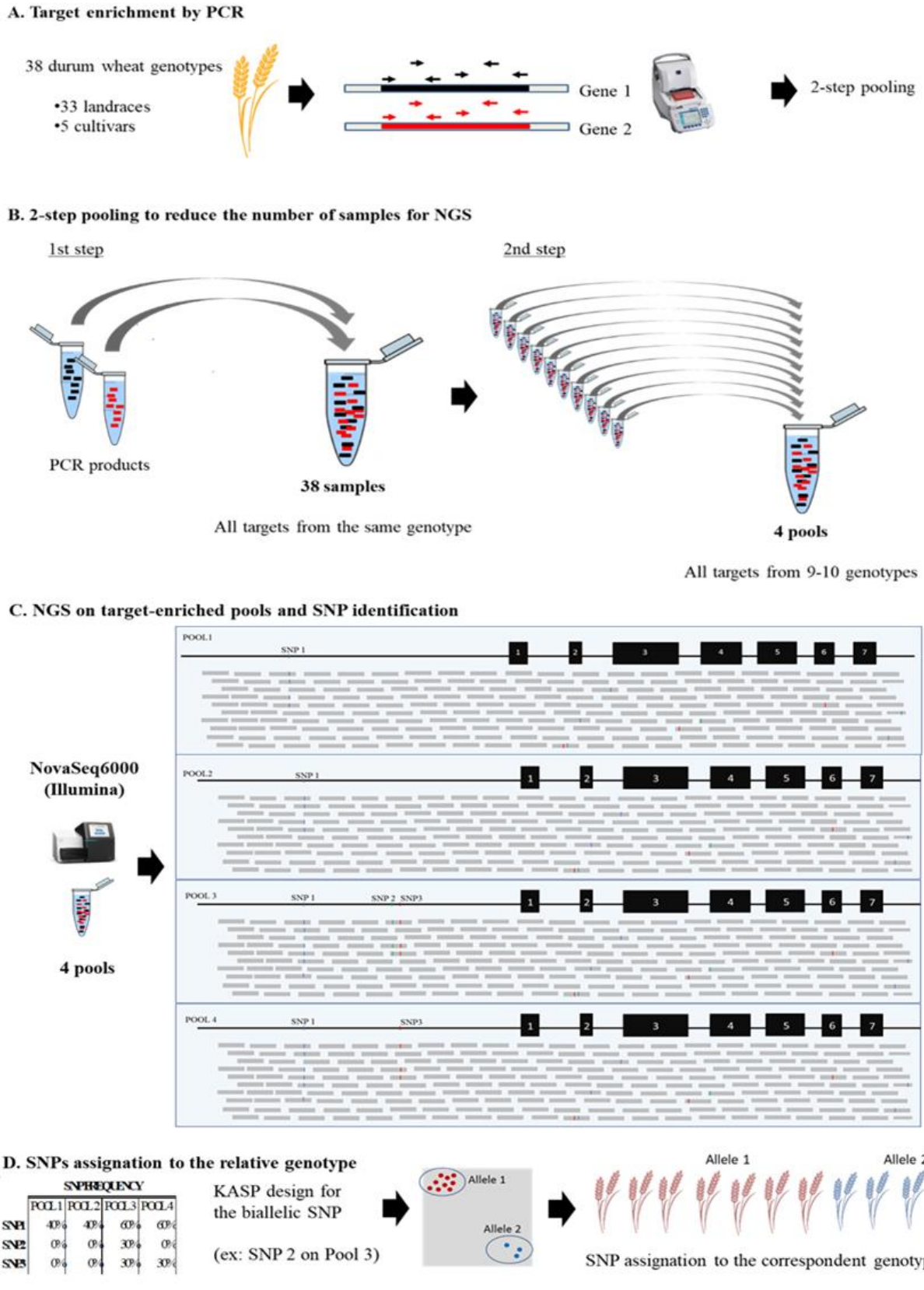


Figure 4.20 Scheme of the method used for the allele mining. (A) Targeted enrichment by PCR is followed by (B) a 2-step pooling to reduce the number of samples for the (C) sequencing by NGS. After the alignments of all the reads to the reference sequence, the identified SNPs are finally (D) assigned to the corresponding genotype with a KASP assay**.**

4.5.1 Targeted enrichment by PCR

The isolation of target sequences is mandatory for the application of this method (Figure 4.20A). The generation of multiple overlapping PCR-products, which cover the whole target sequences was chosen for the *TdBCAT-A* and *-B* genes (Figure 4.21). Five primer pairs were used to tag the gene region previously identified (Table 3.2), of which four homoeologs-specific primer pairs to isolate both A and B gene from the promoter region to exon 5. The last part of the gene was targeted with a non-specific primer pair, designed on conserved genomic region between the two homoeologs, to target the same region of both A- and B- genes in a single reaction.

In details, 3800 bp for *TdBCAT-A* including 1642 bp of the promoter sequence and 3010 bp for *TdBCAT-B*, of which only 880bp belongs to the promoter region, were considered for the analyses. The smaller portion of the *TdBCAT-B* gene is due to the presence of a transposon of the TIR family (732 bp) 888 bp upstream the ATG of this gene. This repetitive element hampered the isolation by PCR of part of the *TdBCAT-B* promoter region, for which a total of 880 bp were isolated and sequenced.

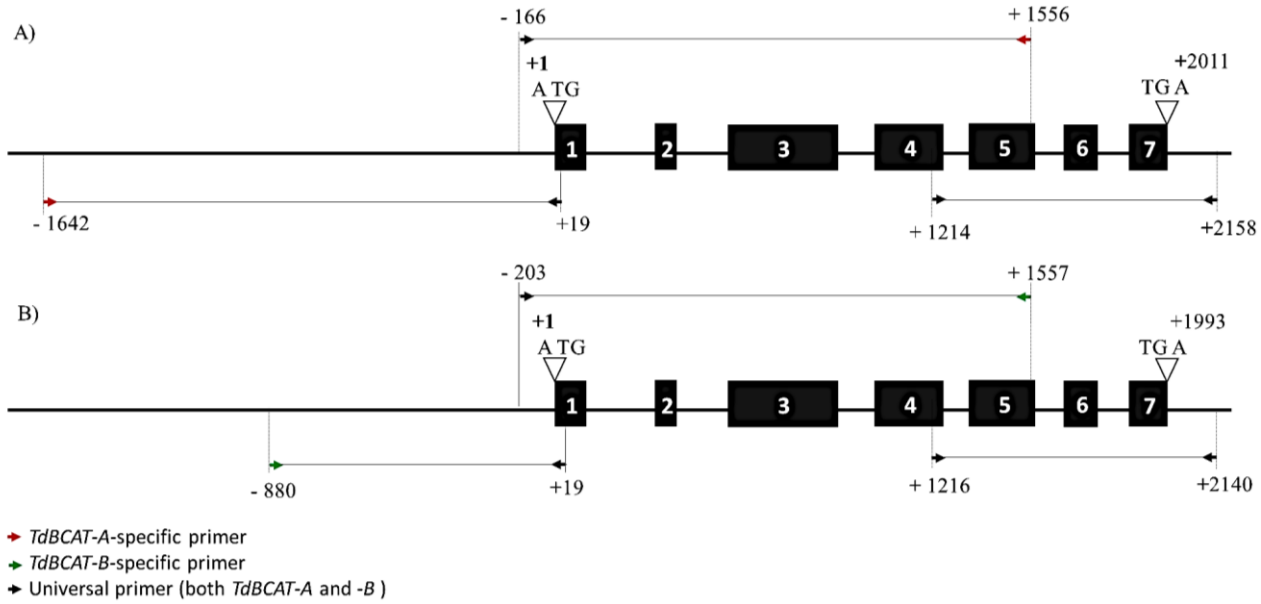


Figure 4.21 Gene models of A) *TdBCAT-A* and B) *TdBCAT-B* genes and primer designed for the target enrichment by PCR. The black boxes correspond to exons, which are numbered from 1 to 7. Each targeted region is shown with a continuous line and arrows, which represent primer pairs used for the PCR-based isolation of the corresponding target.

To reduce the number of samples and thus the cost of NGS, the PCR-products of the targeted sequences were pooled to constitute 4 pools (Figure 4.20.B). Each pool consists of the PCR amplicons obtained for *TdBCAT-A* and *–B* from 9 or 10 genotypes of durum wheat including both SSDs and Italian cultivars as references, depending on the pool.

The obtained PCR fragments were sequenced with NovaSeq6000 (Illumina) (Figure 4.20.C), a paired-end read sequencing and for the relative reads obtained, a FastQC report was generated confirming the high quality of the sequencing for all the fastq files received. Then, the reads were aligned to the whole genome sequence of durum wheat cv. Kronos, obtaining the relative position on the Kronos assembly (Table 4.17).

Table 4.17 Summary table on the results obtained by this first step of pooling.

Gene	**Kronos Scaffold ***	**Position ****	**Strand**	**Number of mapped reads**			
				POOL 1	**POOL 2**	**POOL 3**	**POOL 4**
TdBCAT-A	Triticum_turgidum_Kronos_Elv1.1_scaffold_059708	14825-11061	minus	1,385,670	964,745	1,070,464	1,254,931
TdBCAT-B	Triticum_turgidum_Kronos_Elv1.1_scaffold_014525	98379-95360	minus	1,746,314	1,320,628	1,584,427	1,409,427

** Indicates the location of the hit on the Kronos scaffold;*

*** Indicates the position on the relative Kronos scaffold*

4.5.2 SNPs identification

An analysis to discover possible SNPs, and thus allele variants within the genes of interest, was attempted. We firstly considered all the SNPs with a frequency ≥ 1% allowing for the identification of 93 variants, 62 for TdBCAT-A, and 31 for TdBCAT-B.

The identified SNP have to be assigned to a specific genotype, and considering that the genetic background of the SSD genotypes and of the selected cultivars should be homozygous (99%) for all the *loci*, a 9-10% of SNP frequency is expected for a single genotype variant.

The KASP assay allowed at the same time to test the validity of the SNPs detected but also to assign each SNP to a specific genotype with the exception of SNP A7 on the *TdBCAT-A* promoter that was not possible to assign due to the complexity of the gene region carrying the SNP

To assess the presence/absence of variants on a single genotype KASP assays were performed (Figure 4.20.D) for a set of SNPs identified. The SNP frequency for a single genotype varied from 5% to 16%, depending on the SNP, which can be related to an imperfect relative balance of the PCR-products during the 2-step pooling. SNPs with a frequency < 5 % revealed always false positive, prompting us to increase the cut-off ratio to 5%.

A total of 25 SNPs with a frequency ≥ 5% were identified. Interestingly only one of the SNPs identified falls in the CDS, suggesting a high conservation of the genes during the wheat domestication. 21 were located in the promoter region in both A and B genes (Table 4.18) and 3 SNPs were found in the 3'-UTR. The only missense variant on TdBCAT-B protein (amino acid variant A367T), was not considered as an interesting allele, as the same variant amino acid is present in the respective position of homoeologs protein TdBCAT-A. A total of 16 SNPs were found *on TdBCAT-A* promoter, and 5 SNPs on *TdBCAT-B* promoter.

These results reinforce the hypothesis of a strong conservation of TdBCATs during wheat evolution. The presence of the highest rate of variability in the regulatory elements leads us to hypothesise a role of these mutations in the adaptation to challenging agro-climate conditions.

Table 4.18 List of the 21 SNP variant discovered in *TdBCAT* promoters. Each variant is indicated with an alphanumeric identification with the A or B indicating the homoeologs. For each variant the nucleotide change is indicated with the position on the isolated 5'-3' sequence of the relative gene. Variants frequency (%) are presented for each pool, when detected

Gene	Variant	Nucleotide change	Variant Frequency			
			POOL 1	POOL 2	POOL 3	POOL 4
TdBCAT-A	A3	G100T		9%	9%	
TdBCAT-A	A5	G118C		11%		
TdBCAT-A	A6	T124C		10%	10%	
TdBCAT-A	A7	A139G			7%	
TdBCAT-A	A8	T283G		10%	10%	
TdBCAT-A	A9	G333A		10%	10%	
TdBCAT-A	A10	A348G		10%	10%	
TdBCAT-A	A11	C399T		10%	10%	
TdBCAT-A	A13	A592G		10%	10%	
TdBCAT-A	A21	T659C		10%	10%	
TdBCAT-A	A22	A692G		10%	10%	
TdBCAT-A	A23	T742C		10%	9%	
TdBCAT-A	A24	C810T			5%	37%
TdBCAT-A	A25	G834T	53%	48%	35%	38%
TdBCAT-A	A27	C938T	53%	48%	38%	35%
TdBCAT-A	A28	G962A	53%	48%	39%	35%
TdBCAT-B	B1	A30G			7%	10%
TdBCAT-B	B2	T163A				16%
TdBCAT-B	B3	T179A		16%		
TdBCAT-B	B4	A690T	75%	79%	81%	82%
TdBCAT-B	B5	G797A		11%		

In *TdBCAT-B* promoter a variant (SNP B4 in Table 4.18) with a particularly high frequency was detected in all the 4 pools (75-82%). The KASP assay confirmed this variant as a true-positive SNP, but at the same time it assessed that the observed frequency of each pool was not representative of the expected frequency, especially regarding pool 2 and pool 3 in which a lower number of genotypes carrying this allele were found. We could find 6, 4, 5 and 7 genotypes carrying this variant in pool 1, 2, 3 and 4 respectively. Considering the number of genotypes actually carrying the variant allele, the theoretical SNP frequency for such a number of genotype-variant should have been expected around 60%, 40%, 50% and 70% for poo1, pool2, pool3, and pool4 respectively.

A summary of the results of the genotyping with KASP is presented in table Table 4.19

Table 4.19 Summary of the genotyping with KASP assays for the identification of genotypes carrying the SNPs (variant allele) detected by NGS.

Gene	Variant	Position on target (*)	Reference allele	Variant allele	SSD/genotype with the variant allele
TdBCAT-A	A3	-1543	G	T	397, 409, 441
TdBCAT-A	A5	-1525	G	C	343
TdBCAT-A	A6	-1519	T	C	397, 409, 441
TdBCAT-A	A8	-1360	T	G	397, 409, 441
TdBCAT-A	A9	-1310	G	A	397, 409, 441
TdBCAT-A	A10	-1295	A	G	397, 409, 441
TdBCAT-A	A11	-1244	C	T	397, 409, 441
TdBCAT-A	A13	-1051	A	G	397, 409, 441
TdBCAT-A	A21	-984	T	C	397, 409, 441
TdBCAT-A	A22	-951	A	G	397, 409, 441
TdBCAT-A	A23	-901	T	C	397, 409, 441
TdBCAT-A	A24	-833	C	T	335
TdBCAT-A	A25	-809	G	T	69, 92, 109, 178, 195, 253, 269, 278, 322, 325, 343, 451, 494, Cappelli
TdBCAT-A	A27	-705	C	T	69, 92, 109, 178, 195, 253, 269, 278, 322, 325, 343, 451, 494, Cappelli
TdBCAT-A	A28	-681	G	A	69, 92, 109, 178, 195, 253, 269, 278, 322, 325, 343, 451, 494, Cappelli
TdBCAT-B	B1	-851	A	G	335, 459
TdBCAT-B	B2	-718	T	A	112
TdBCAT-B	B3	-702	T	A	322
TdBCAT-B	B4	-191	A	T	35, 44, 69, 99, 112, 116, 122, 135, 171, 195, 244, 253, 269, 278, 322, 325, 397, 487, 494, 499, 511, Colosseo, Cappelli, Svevo
TdBCAT-B	B5	-84	G	A	322

(*) refers to the position upstream the ATG (starting codon), which has been considered as +1

The KASP assay revealed that the number of genotypes carrying the allele variant B4 (Table 4.18) were not equally represented in all the 4 pools as would suggest the frequency of this variant. This alteration between the observed and expected variant frequency for each pool is related to the primer sequence used for the targeted-enrichment by PCR. Indeed SNP B4 (Table 4.18) falls in the region of the primer used for the targeted-enrichment, thus it has been introduced by the PCR reaction in the amplicons and altering the correct frequency. This should be considered as a proof of concept of the robustness and validity of KASP as molecular markers for the correct genotyping of biallelic SNPs.

Moreover even if in our case the design of the KASP markers was not a waste, the possibility of introducing false-positive should be considered for future primer design before applying this method, especially targeting region in which a higher variation is expected (i.e. regulatory elements, introns or gene-sequences coding out of conserved protein domain).

With the applied approach 4 alternative haplotypes for both A and B promoters have been identify (Figure 4.22, Figure 4.23).

Ten out of the 21 SNPs on the *TdBCAT-A* promoter belongs to the same haplotype, which is shared by 3 SSD (SSD 409, SSD 441 and SSD 397). For *TdBCAT-B* the most variable haplotype has 3 SNP and was identified in the SSD 322.

	Haplotype identification	Total SNPs	A3	A5	A6	A8	A9	A10	A11	A13	A21	A22	A23	A24	A25	A27	A28	SSD *	Varieties*
Haplotype 1 (Reference)	HA1	-	G	G	T	T	G	A	C	A	T	A	T	C	G	C	G	13	4
Haplotype 2	HA2	3	G	G	T	T	G	A	C	A	T	A	T	C	T	T	A	12	1
Haplotype 3	HA3	4	G	C	T	T	G	A	C	A	T	A	T	C	T	T	A	1	0
Haplotype 4	HA4	10	T	G	C	G	A	G	T	G	C	G	C	C	G	C	G	3	0
Haplotype 5	HA5	2	T	G	T	T	G	A	C	A	T	A	T	T	G	C	G	4	0

* number of SSD/varieties carrying the correspondent haplotype

Figure 4.22 Haplotypes detected in *TdBCAT-A* promoter.

	Haplotype identification	Total SNPs	B1	B2	B3	B4	B5	SSD *	Varieties*
Haplotype 1 (Reference)	HB1	-	A	T	T	A	G	12	2
Haplotype 2	HB2	1	G	T	T	A	G	2	0
Haplotype 3	HB3	2	A	A	T	T	G	1	0
Haplotype 4	HB4	3	A	T	A	T	A	1	0
Haplotype 5	HB5	1	A	T	T	T	G	17	3

* number of SSD/varieties carrying the correspondent haplotype

Figure 4.23 Haplotypes detected in *TdBCAT-B* promoter.

Considering the functional redundancy typical of wheat homoeologs, it is interesting to observe the relative combination of the haplotypes found for both A and B genes in each single SSD/variety. A total of 10 haplotypes were identified in promoters region, (5 in gene A and 5 in gene B) and correspond to 12 different haplotypes combinations considering both *TdBCAT* genes, of which 6 are unique and represented in a single SSD (Table **4.20**).

Table 4.20 Twelve combinations of haplotypes for *TdBCAT-A* and *TdBCAT-B* promoters in the 38 genotypes analysed with the targeted resequencing for the allele mining.

Haplotype Combination	**Haplotype of *TdBCAT-A* promoter** ***(a) (b)***	**Haplotype of *TdBCAT-B* promoter** ***(a) (c)***	**Genotypes**
1	HA1	HB1	64, 415, 416, Saragolla, Kronos
2	HA2	HB1	92, 109, 178, 451
3	HA3	HB1	343
4	HA4	HB1	397, 409, 441
5	HA1	HB2	459
6	HA2	HB2	253
7	HA5	HB2	335
8	HA5	HB3	112
9	HA2	HB4	322
10	HA1	HB5	35, 44, 99, 116, 122, 171, 244, 487, 511, Colosseo, Svevo
11	HA2	HB5	69, 195, 269, 278, 325, 494, Cappelli
12	HA5	HB5	135, 499

(a) Identifications refer to haplotypes in figure 4-22 and 4-23 for TdBCAT-A and TdBCAT-B respectively;
(b) 1642 bp of promoter region upstream the ATG;
(c) 880 bp of promoter region upstream the ATG

The results reported here confirmed the high genetic diversity of the 33 SSD as representative of a wider germplasm collection (Pignone et al., 2015), compared to all the five reference cultivars.

Among the cultivars, Cappelli was the most variable (4 SNPs), Svevo and Colosseo shared a single variant, while Saragolla was 100% identical to Kronos. Among all the analysed cultivars, Cappelli is the oldest cultivar, the only variety that originated before the intensive breeding programs, which led to the more productive and genetically uniform semi-dwarf cultivar in the middle of the last century.

The Green Revolution from the 1940s to 1960s worsened the phenomenon of genetic erosion of domesticated wheat, narrowing the genetic diversity of the modern elite cultivars. It is well understood that the modern breeding affected the genetic diversity in wheat (Reif et al., 2005), favouring the high-yield traits at the expense of other traits like tolerance to environmental constraints. In the last decades breeders and researchers started to care more about this decreasing trend of genetic diversity, and focused their attention to wheat natural genetic resources such as wild relatives and landraces, which have helped to invert the trend of genetic diversity loss (van de Wouw et al., 2010).

Lopes and colleagues (2015) underlined the importance of wheat landraces as a unique genetic tool for wheat breeding that can help to overcome the impending climate changes. As a consequence of the avowal of the importance of landraces, many recent works focused on the description of the genetic structure and variability panel of wheat landraces (Kabbaj et al., 2017; Soriano et al., 2016), which demonstrated the high rate of allelic variation within these wheat genetic resources. The availability of a low cost technique to discover and exploit the high reservoir of genetic variability contained in the genetic resources is desirable to help wheat breeding in the contest of the upcoming challenges posed by the climate change.

Many strategies for allele mining purpose have been used in crops to find new alleles in several genes involved into the abiotic stress response, like the enzymatic polymorphism detection with *CelI* for the EcoTILLING in rice (Yu et al., 2012), and the High Resolution Melting (HRM) coupled to the TILLING strategy in durum wheat (Comastri et al., 2018). Notwithstanding the efficiency and the utility of both these techniques for the detection of new variants in crops species, we know that both *CelI* and HRM can be quite tedious and time-consuming approaches.

NGS technologies costs significantly decreased in the recent years, but their affordability can still be a limit for many. Sequencing pools of individuals by NGS (pool-seq) is a recognised cost-effective technique compared to the sequencing of separate individuals for the genome-wide polymorphism (Schlötterer, 2004), but in species like wheat, which are characterized by big and complex genome, genome-wide sequencing is still a major limitation.

The method here applied, allows targeting any region of interest on a genome, including the regulatory elements which are extremely important for the regulation of gene expression. The method, is reliable for allele mining and cost-effective since allows a rapid detection of allele variants by coupling NGS and KASP markers for the SNP validation on a pool of individuals. All these elements, candidate this approach as an innovative and cost-effective method for the allele mining, since the final cost for the analysis can be reduced up to 10-fold compared to a classic NGS sequencing of separate individuals. Moreover, KASPs developed for the SNP identification are cheap PCR-based codominant markers, directly usable for the introgression of the candidate alleles in wheat cultivar by MAS. This real important feature offsets the time spent for the KASP design for the simple SNP validation.

This approach allowed us to identify a total of 21 out of 25 SNPs in the proximal promoter region of *TdBCAT* genes in this subset of 33 SSD from a wider germplasm collection of more than 400 accessions (Pignone et al., 2015), which could be further investigated by applying the same procedure.

Since only variant found in the coding sequence of TdBCAT-B lead to the same amino acid present on the respective position of the A- homoeologs, testifying the high-rate of conservation even in durum wheat landraces, which are recognised to be extremely heterogeneous in their genetic background. This is a proof of the importance and the conservativeness of the biological function of this gene.

Moreover *bcat3*, the *Arabidopsis* orthologues of *TdBCAT*, has been demonstrated to be involved in drought tolerance (Scarpeci et al., 2017). This study highlighted also how *bcat3* mutant plants yielded significantly less seeds under water-limiting conditions, compared to control plants. These results, along with our finding in the transcriptional profiling during the reproductive stages of durum wheat, lead us to consider *TdBCAT* as a very interesting candidate for drought tolerance and possibly, a high yield-stability under water-limiting conditions.

For all these reasons, a phenotypic evaluation of the genotyped SSD is desirable to assess their response under DS conditions and to find valuable genotypes for further breeding programs.

4.6 Phenotypical evaluation by high-throughput phenotyping

Once the natural variation was identified, the set of SSD lines and the varieties Cappelli, Saragolla and Svevo were subjected to a preliminary phenotypical evaluation with a non-invasive HTP system, to evaluate the effect of DS.

The phenotyping experiment was conducted in controlled environment through the acquisition and processing of images (3D Scanalyzer System, Lemnatec) in the visible spectrum (RGB) and NIR thanks to the ongoing collaboration with the national infrastructure of plant phenotyping ALSIA located in Metaponto (MT, Italy).

Two indices have been acquired from the image analyses as related to the DS: the digital biovolume and the dry index (DI).

4.6.1 Phenotyping evaluation of the RGB-based data for the digital biovolume ratio

The RGB images are an excellent source of information regarding the morphometric traits of a growing plant and have been used to efficiently and can be expressed as "digital biovolume". To assess the effect of stress on different genotypes, the digital biovolume ratio (DBR) was extrapolated for all the 35 genotypes under analysis. The DBR is a parameter useful to assess the growth performance of a genotype and thus its healthiness during a stress event like drought, which can affect plant growth and thus reduce its biomass under stress conditions (Poiré et al., 2014). Figure 4.24 shows the biovolume ratio change over the 43 days monitoring of the plants during the DS. Higher DBR indicates a higher biomass under stress conditions thus a higher adaptability of the plant to drought.

In general, in most of the landraces a decrease in the DBR during the stress that occurs 104 days after sowing (DAS) was maintained till the end of the experiment, a behaviour previously observed in also in tomato (Petrozza et al., 2014). The DBR reached a minimum at 147 DAS where both the effect of the DS and the natural senescence are overlapping (Figure 4.24).

The varieties tested in the experiment showed the expected behaviour, in particular Cappelli and Svevo showed a stay green behaviour until the end of the trial and Saragolla already at 8 days of stress showed a strong reduction in the plant biomass (Figure 4.24).

On the basis of the mean DBR from 104 to 147 DAS, 3 major categories of genotypes were identified, as high, medium and low resilient to drought. Eight genotypes had a high resilience (DBR $\geq$ 0.85), with Svevo, Cappelli, SSD 409 and SSD 195, with the highest and most stable biovolume ratio (mean > 0.9). Eleven genotypes belong to the medium category (0.85 < DBR $\leq$ 0.8) which are characterized by an

intermediate behaviour between the high and low resilience. Last, all the genotypes with a DBR < 0.8 were considered as low resilient under drought conditions.

4.6.1.1 Considerations on the haplotype combinations of TdBCAT

A number of SSDs (SSD 409, SSD 441, SSD 397) that share the haplotype combination 3 (Table **4.20**) showed a high (SSD 409) or a medium (SSD 441, SSD 397) resilience to drought. Between them, SSD441 had the highest DBR. These three genotypes are the only that actually carry the haplotype HA4 of *TdBCAT-A* which presented the highest number of SNPs. It is interesting to note that these three SSDs are originated from the same geographic area: Greece (SSD 409) and Crete (SSD 397, SSD 441), suggesting a possible role of the Greek islands.

Among all the SSDs that carry a unique combination of haplotypes (i.e. SSD 459. SSD 253, SSD, 335, SSD112, SSD 135, SDD 322 and SSD 343;Table **4.20**), only SSD 322 and SSD 343 are classified as high resilient. SSD 343 is the only SSD that carries the haplotype HA3 of *TdBCAT-A* (Table **4.20**) and it does not show any variation in the DBR slope. Moreover the DBR rose during the DS reaching a maximum at 124 DAS. All the other SSDs which carry a unique combination of haplotypes (i.e. SSD 459. SSD 253, SSD, 335, SSD 112, SSD 135) are all categorized as low resilient to DS. In more details, all the genotypes with haplotype HA5 (Table **4.20**, Figure 4.24) in *TdBCAT-A* are classified in the low resilience class, which the worst performer SSD 135 is the only genotype with the haplotype combination 12 (Table **4.20**, Figure 4.24). Another haplotype present only in low resilient SSDs of is HB2 in *TdBCAT-B* (Table **4.20**, Figure 4.24).

The two cultivar Svevo and Cappelli identified as high resilient have the same haplotype HB5 (Figure 4.24, Table **4.20**) of *TdBCAT-B*, but a different haplotype for *TdBCAT-A.* Saragolla performed as a low resilient, but its haplotype combination 1 is shared within many other genotypes which behaved heterogeneously considering the DBR.

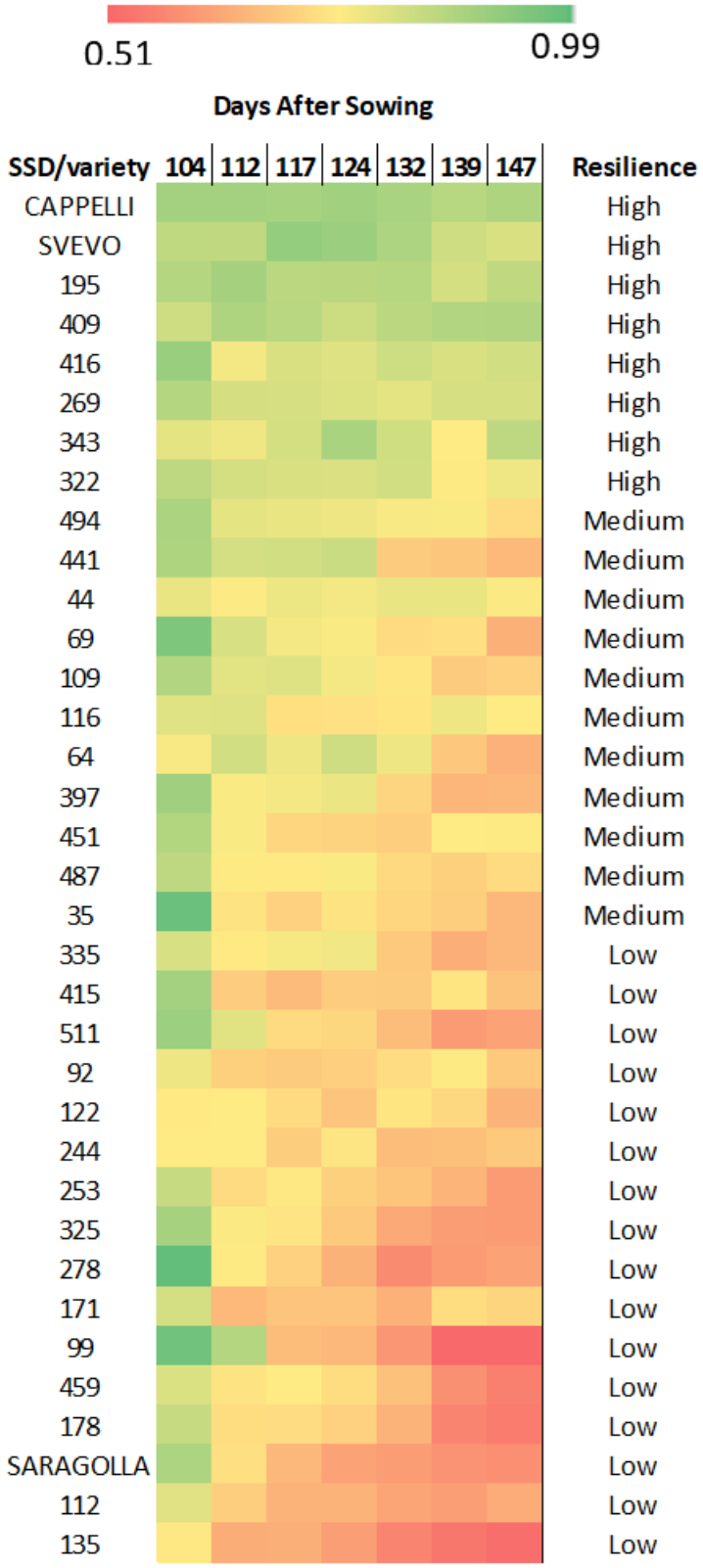


Figure 4.24 Colorimetric representation of the digital biovolume ratio (DBR). The picture represents the biovolume ratio of 35 durum wheat genotypes at 6 different time-points during the drought experiment in the high-throughput phenotyping platform from 104 days after sowing (DAS) when the drought stress was imposed, until 139 DAS. For each genotype the corresponding performance category is indicated.

4.6.2 Phenotyping evaluation of the NIR-based water content for the dry index

The plant water content was estimated from the NIR imaging. Plants showing a lower DI have an higher water content, while plants with a higher DI have a lower water content. A full list of the dry-index values for both control and stressed plants of the 35 analysed genotypes is provided in Table 6.1.

An evaluation as for the DBR of the most resilient SSD genotypes was done on the basis of the acquired DI. To evaluate the trend of the DI during DS, for each genotype the relative DI (RDI %) was calculated as $\frac{DIS-DIN}{DIN} * 100$, where DIS is the DI of stressed plants, while DIN is the DI of non-stressed plants. The evaluation of the mean RDI (%) from 104 to 147 DAS was calculated to rank the SSD for their drought resilience as follow: high 0.5 < RDI ≤3, medium 3 <RDI≤ 5 and low 5<RDI≤7.5. The full list of RDI is provided in Table 4.31.

The radar chart displayed in Figure 4.25 shows the highest value of RDI of the genotypes considered as low, compared to the one in both high and medium class. In the same chart the highest values of RDI was measured at 139 DAS for the majority of genotypes, with very few exceptions.

Ten genotypes (SSD 441, 343, 416, 64, 109, 269, 278, 511, 409 and cultivar Svevo) showed mean RDI < 3 as a consequence of their high water content in the tissues during the entire length of the stress. A match with the DBR selection is clearly evidenced since a number of these genotypes (Svevo, 343, 416 and 269) were identified as high-performing also for the DB evaluation. Among these 10 genotype, SSD 441, SSD 343 and SSD 416 had the lowest mean RDI (0.5 < RDI < 1.5), which of SSD 441 is the only one with a mean RDI<1, showing also an increased water content in stressed plants following the DS imposition from 104 DAS till 112 DAS (Figure 4.25). Thirteen genotypes belong to the medium category which is characterized by an intermediate resilience. Last, twelve genotypes were considered as low resilient under drought conditions.

The average trend of the DI of all the 35 genotypes is presented in Figure 4.26. The DI does not vary significantly in both the control and treated up to the stress imposition at 104 DAS, where in both cases the DI increases accordingly to the increasing of biomass. All the genotypes increased the RDI starting from 104 DAS, with a general increasing trend with the lasting of the stress. After the occurrence of the stress, the DI varied significantly in the stressed genotypes, reaching its maximum at 139 DAS, when the stressed plants despite the decrease in the overall DI due to senescence, showed almost an RDI equal to 7%.

When the varieties were considered, even if Cappelli and Svevo were both considered as high-performer for the biovolume change under DS, it was not possible to observe a similar trend for what concern the DI (Figure 4.27.A and B). Svevo is confirmed to be the highest resilience variety with a maximum DI increase of about 4% at 117 DAS, while Cappelli despite the expectation, in this experiment did not show comparable level of RDI as for the DBR.

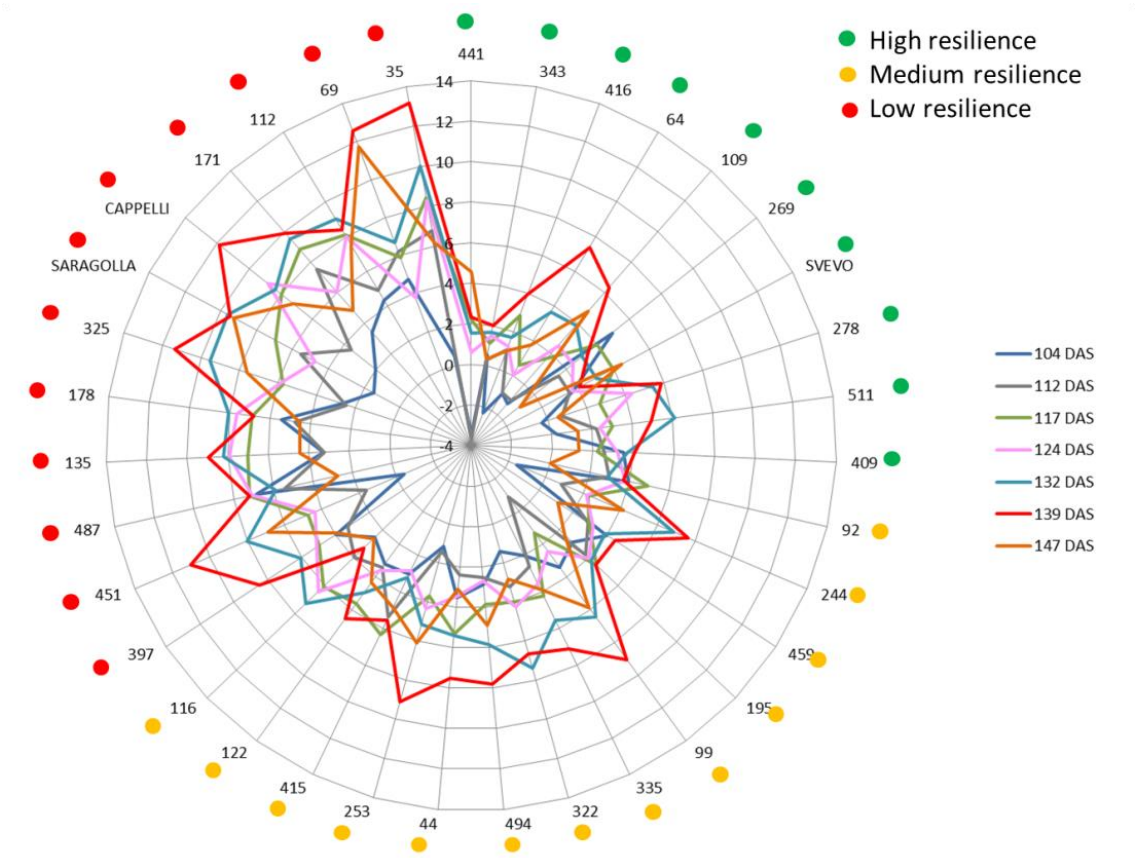


Figure 4.25 Radar chart representing the RDI (%) for all the genotypes from the imposition of the drought stress (104 days after sowing –DAS-) until 147 DAS.

Comparing the results of the biovolume ratio and the RDI among all the three cultivars Svevo is definitely, as expected, the best performer and more stable during the stress (Figure 4.27).

Within the high resilient genotypes, SSD 343 and SSD 441 were of particular interest, as the only genotypes having a RDI < 2% for all the time points during the stress until 139 DAS. In particular, SSD 441 is the one with the lower RDI which is an evidence of its major ability to limit water loss under water-limiting conditions. SSD 441 was also classified as medium resilient considering the DBR, but it showed a digital biovolume comparable to the high resilient genotypes until 124 DAS (20 days of DS) which started to decrease at a DBR < 0.8 only after 132 DAS. SSD 343 and SSD 441 showed highly similar DI slope between stressed and controls plants (Figure 4.27 C, D), making these lines good candidates for future breeding programs.

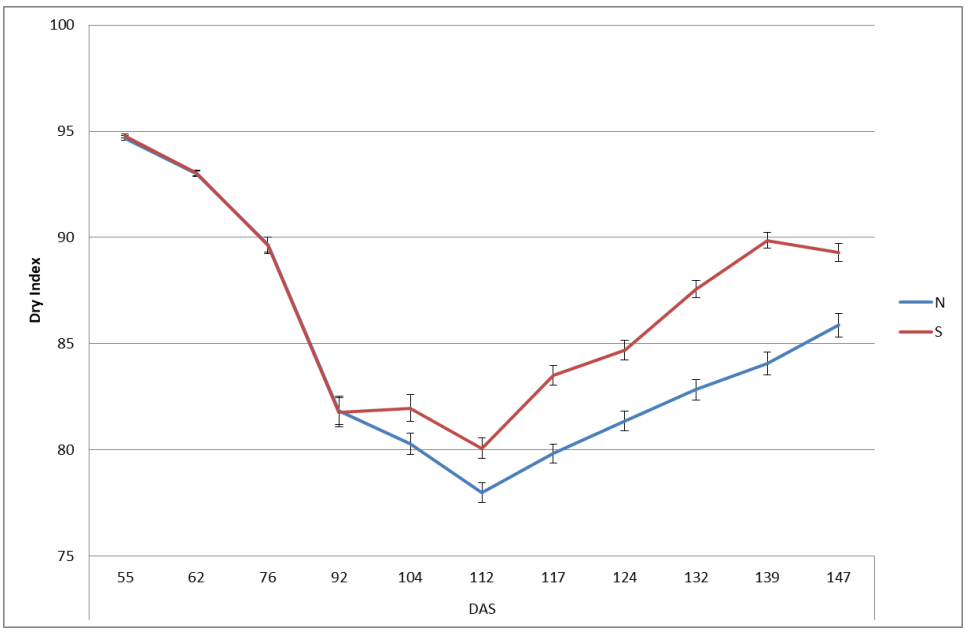


Figure 4.26 Means and standard errors of the NIR based Dry Index (DI) in (N) well-watered and (S) conditions**.**

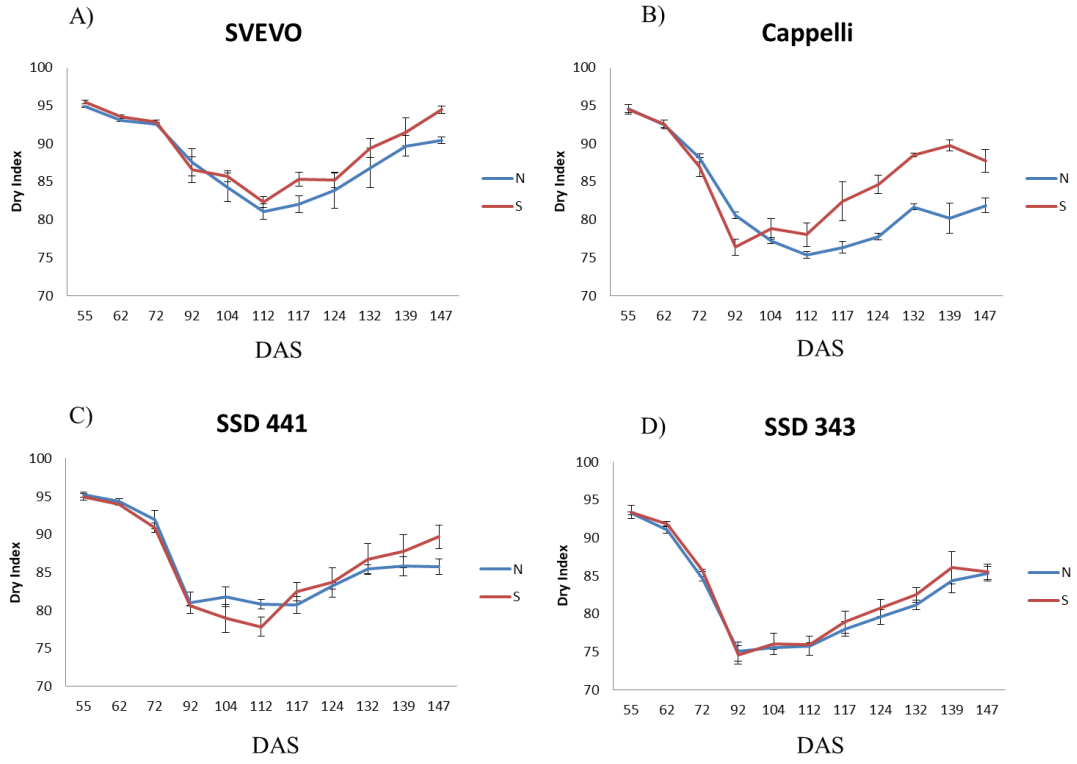


Figure 4.27 Means dry-indexes for genotypes A) Svevo, B) Cappelli, C) SSD 441, D) SSD 343. DAS=days after sowing.

Some interesting considerations referring to the haplotypes of *TdBCAT* genes can be done on the basis of the RDI results. First, we can see that the SSD 441, which has been considered has the highest resilient for the RDI, is characterized by one of the most peculiar combinations of haplotype for *TdBCAT-A*, (haplotype combination 3, Table **4.20**). This same allelic combinations is in common with SSD 409 which

is another high resilient genotype, and SSD 397 which is in the low resilience category. Moreover the DI trend of stressed plants showed an higher water content in stressed plants at 104 DAS and 124 DAS (Figure 4.27C), which can be related to an adaptive trait of this genotype to avoid water loss during water limiting conditions.

As we have already said, 7 SSDs carry a unique combination of haplotypes, of which only SSD 343 were classified as high resilient. SSD 343 is the only SSD that carries the haplotype HA3 of *TdBCAT-A* (Table **4.20**) and it showed a very similar DI slope between stressed and controls plants (Figure 4.27D).

Among all the other phenotyped SSDs that carry a unique combination of haplotypes (i.e. SSD 459. SSD 253, SSD 335, SSD112, SSD 322 and SSD 135), they are mostly medium resilient, with the exception of SSD 112 and SSD 135, which confirmed their low resilience to drought stress as observed measuring the DB.

Differently to what observed for the DBR, haplotype HB2 of *TdBCAT-B* is represented only in medium resilient genotype.

Taking all together these results reinforce the hypothesis that despite the existence of common regulatory mechanisms across species, some pathways/mechanisms are dependent upon genotype and the duration, intensity, and type of stress applied (Fleury et al., 2010).

This work provides another level of knowledge about the defence response of three common cultivated varieties: Svevo, Cappelli and Saragolla, confirming the high capability of Svevo to overcome the DS by preserving a high digital biovolume and high water content in the tissues during the DS.

Saragolla does not show high resilience in DS conditions, while in this experiment, Cappelli shows, at phenotyping level, contrasting results. The digital biovolume remains constant during the stress although the dry index decreases as in the control plants.

From our results, the exploitation of genetic resources is confirmed to be a valuable approach for the identification of useful traits for wheat breeding, in fact out of 33 SSDs, three to six SSDs on the basis of the phenotyping analyses have been identified as good candidates for wheat breeding programs.

Trying to demonstrate the direct involvement of TdBCAT in the increased resilience observed at least looking at two parameters (DBR and RDI), some comments about the haplotype composition can be made.

A number of SSDs (SSD 416, SSD 343, SSD 409, SSD 269) carrying specific haplotype combinations for *TdBCAT* genes, showed under DS a high maintenance of water balance in the tissues and do not showed a decrease in the biomass expressed as DB.

SSD 416 and SSD 269 have haplotypes combinations which are shared with SSD that have an heterogeneous behaviour under DS, considering either DB and DI, thus their allelic variants for *TdBCAT* genes are not probably contributing to enhance their drought tolerance under DS.

Actually, the two most interesting candidates to us are SSD 343 and SSD 409. SSD 343 has a unique haplotype combination, as it is the only one carrying the haplotype HA3 on *TdBCAT-A* promoter. On the other hand, SSD 409 has a combination of haplotypes for *TdBCAT-A* with the highest number of SNPs. The same haplotype combination is in common with SSD 441, which has the highest water content in DS conditions. SSD 441 could be considered thus another very interesting candidate, nevertheless it is not in the top rows for the DB evaluations, for which still has a good DBR. These findings allowed hypothesising a direct involvement of TdBCAT variants in the regulation of the DS response.

Table 4.21 Summary of the results of the classification of the drought resilience attitude considering both the Digital Biovolume Ratio (DBR), the Relative Dry Index RDI (%) and the combination of haplotypes on *TdBCAT-A* and *TdBCAT-B*

SSD/variety	DBR	RDI (%)	Haplotype combination *(a)*
416	High	High	1
64	Medium	High	1
SARAGOLLA	Low	Low	1
415	Low	Medium	1
CAPPELLI	High	Low	2
195	High	Medium	2
109	Medium	High	2
178	Low	Low	2
451	Medium	Low	2
92	Low	Medium	2
343	High	High	3
409	High	High	4
441	Medium	High	4
397	Medium	Low	4
459	Low	Medium	5
253	Low	Medium	6
335	Low	Medium	7
112	Low	Low	8
322	High	Medium	9
SVEVO	High	High	10
511	Low	High	10
35	Medium	Low	10
171	Low	Low	10
487	Medium	Low	10
99	Low	Medium	10
122	Low	Medium	10
244	Low	Medium	10
44	Medium	Medium	10
116	Medium	Medium	10
269	High	High	11
278	Low	High	11
325	Low	Low	11
69	Medium	Low	11
494	Medium	Medium	11
135	Low	Low	12

a) For the details of the different haplotypes in the corresponding genotype see Table 4.20

Table 4.22 Relative dry index (RDI %) of SSD genotypes and commercial varieties in response to drought

	Days After Sowing (DAS)											
Genotype	**55**	**62**	**72**	**92**	**104**	**112**	**117**	**124**	**132**	**139**	**147**	**167**
35	0.22	0.04	0.21	2.1	0.51	6.78	8.4	8.3	10.02	13.18	6.17	2.83
44	0.9	0.44	0.83	1.95	3.54	2.4	5.31	3.46	5.4	7.5	3.13	4.91
64	1.12	-0.08	-0.54	-4.62	-1	-0.96	0.62	0.03	3.66	7.38	1.76	3.33
69	-0.36	-0.14	-0.19	0.95	4.75	6.21	5.9	3.79	6.68	12.57	11.73	6.57
92	1.00	-0.64	1.17	3.98	3.95	2.96	4.91	3.77	2.87	3.70	0.01	2.56
99	-0.55	0.27	0.89	-0.21	3.4	-0.88	1.30	2.43	6.43	9.05	5.88	4.4
109	-0.26	0.28	2.20	0.10	-1.28	-1.05	1.70	2.55	3.88	6.34	4.79	2.15
112	1.08	1.18	-0.89	0.95	4.34	4.91	8.11	8.00	8.99	8.38	7.48	3.5
116	-0.22	0.47	-0.08	0.59	2.55	3.97	6.08	6.41	7.27	3.34	2.65	3.89
122	0.63	0.4	1.25	0.81	3.22	3.44	5.64	3.62	4.95	6.53	4.35	2.08
135	0.22	0.64	1.54	1.15	3.29	3.22	7.02	7.93	8.23	8.96	4.44	4.22
171	2.23	1.7	-1.46	1.14	3.4	7.55	8.84	6.05	9.54	9.91	4.83	5.93
178	0.51	-0.18	-0.2	1.51	5.41	4.75	6.89	7.63	8.05	6.8	4.5	2.31
195	-0.10	-0.24	-3.25	2.09	2.90	3.82	4.54	4.04	4.18	4.5	2.37	2.36
244	0.08	-1.25	-0.96	-5.49	-1.57	0.84	2.28	2.19	6.87	7.63	4.15	2.11
253	-1.33	0.14	-1.3	2.27	1.14	1.4	3.68	4.35	5.14	9.11	6.11	1.84
269	-0.26	-0.06	2.25	-0.86	4.9	1.52	3.94	2.43	3.01	3.36	-0.92	2.81
278	-0.31	-0.41	-1.05	-3.77	-0.34	0.73	2.67	4.33	5.4	5.85	0.51	2.45
322	-0.57	-1.85	-1.36	0.42	1.4	3.21	4.01	4.23	7.38	6.64	2.82	1.82
325	-0.92	0.2	-0.35	0.62	2.71	2.46	5.74	5.9	9.51	11.33	7.59	4.45
335	-0.97	0.32	-0.38	2.93	2	2.62	4.21	3.59	5.58	7.14	3.85	2.3
343	0.18	0.88	1.34	-0.6	0.54	0.27	1.13	1.52	1.68	2.03	0.32	1.8
397	1.19	0.24	-1.26	-2.28	3.98	3.27	4.95	5.1	6.06	8.47	3.85	2.8
409	0.21	0.23	0.34	-2.73	3.5	2.56	2.24	3.23	3.73	4.07	1.3	4.08
415	0.31	0.01	0.89	4.9	3.04	5.39	6.32	2.83	3.2	5.55	4.89	5.94
416	-0.15	-0.89	0.06	-2.91	-2.28	1.04	2.84	1.23	1.7	4	0.99	4.14
441	-0.35	-0.36	-1.11	-0.5	-3.45	-3.68	2.19	0.58	1.55	2.31	4.57	3.11
451	-0.22	-1.01	-0.98	-0.68	-0.39	1.63	4.7	4.37	8.01	11.01	6.88	2.63
459	-0.07	0.64	1.81	2.26	4.02	2.9	2.95	3.23	3.92	4.52	1.14	1.2
487	-0.39	-0.19	-0.25	1.18	6.87	5.43	7.2	7.08	5.89	7.2	2.74	3.55
494	-0.39	0.02	1.01	0.81	2.86	2.52	3.84	2.68	5.85	7.79	4.89	3.15
511	0.44	0.64	1.18	2.44	0.24	2.23	3.04	2.39	6.13	4.96	1.31	2.57
CAPPELLI	-0.1	0.09	-1.37	-5.24	2.04	3.55	7.96	8.8	8.34	11.89	7.19	2.27
SARAGOLLA	0.27	-0.3	-1.37	-6.61	1.46	5.52	6.94	4.73	9.64	9.48	9.28	4.76
SVEVO	0.65	0.5	0.29	-1.09	1.68	1.5	3.97	1.65	3.01	2.08	4.43	3.63

5 Conclusions

Four main results have been reached with this thesis work:

1. Two genes have been identified and then characterised at genomic and transcriptional level under DS. This characterization interested both the A and B homoeologs, which suggested a major involvement of *TdBCAT* under DS during the reproductive stage, and in particular at the anthesis and early post-anthesis events, while the role of *TdABA7* under water-limiting conditions is related to the germination phase.
2. All the isolated genes were used as targets in TILLING analysis to mine new alleles in a collection of durum wheat mutants obtaining a large number of SNPs, spanning the entire gene sequences, which could potentially affect the protein function. The *in silico* TILLING allowed to select 3 mutant lines for *TdBCAT* and 4 mutant lines for *TdABA7* that will undergo through the backcross process to obtain near isogenic lines. So far, 11 and 15 double mutant lines for *TdBCAT* and *TdABA7* respectively were generated to test for drought related traits.
3. A targeted-resequencing approach was successfully applied to exploit the natural variation available in a set of genetic resources (called SSD genotypes for the method applied for the selection). A set of 38 durum wheat landraces were analysed for *TdBCAT*s natural variation revealing 10 haplotypes for *TdBCATs.*
4. All the selected SSD genotypes have been grown in a HTPP to observe the phenotypes and measure variations under drought stress. In this way, five SSD lines were selected as highly resilient to drought considering the image based indices gained from the platform, three of them carrying unique haplotype of *TdBCAT-A* gene sequence.

Taken all together, these results represent a big step forward for the comprehension of the role played by *TdBCAT* in the durum wheat adaptation to DS and an initial step for the comprehension of the role played by *TdABA7* gene in wheat growth, development and defence response. The wide number of mutants identified open new perspectives for functional genomics studies on drought stress.

Moreover, the identification of a number of SSD lines carrying allele variants represent the first step for the development of "non GM" NILs to be introduced in wheat breeding programs.

The development of new wheat varieties with increased drought resilience is one of the main goals of the future 'smart' agriculture that will have to face the increase of global food demand in terms of agricultural commodities for food and feed. Furthermore, it will be necessary to preserve the water reservoir severely hampered by agriculture withdrawing.

6 Annex

Table 6.1 DI values for all the genotypes evaluated from 55 DAS until 147 DAS.

Genotype	S/N	Days After Sowing (DAS)										
		55	62	76	92	104	112	117	124	132	139	147
35	N	94.70	93.46	90.73	77.42	78.22	75.37	77.55	80.01	80.81	79.93	82.28
35	S	94.91	93.50	90.92	79.05	78.62	80.48	84.07	86.66	88.9	90.46	87.36
44	N	94.69	93.16	89.05	77.81	76.52	76.54	79.36	81.69	82.66	82.69	83.73
44	S	95.55	93.57	89.79	79.33	79.23	78.38	83.57	84.52	87.13	88.89	86.35
64	N	93.91	92.52	89.17	80.16	81.83	78.25	80.36	81.89	82.10	81.55	83.63
64	S	94.96	92.44	88.69	76.46	81.01	77.50	80.85	81.91	85.10	87.57	85.11
69	N	95.16	93.55	92.40	82.01	81.37	78.04	81.81	84.29	83.46	81.15	81.36
69	S	94.81	93.42	92.22	82.79	85.23	82.89	86.64	87.48	89.03	91.35	90.9
92	N	94.22	94.22	92.24	82.35	84.17	82.39	86.78	88.70	90.28	90.40	91.02
92	S	95.16	93.62	93.32	85.62	87.49	84.83	91.04	92.04	92.86	93.75	91.03
99	N	95.72	93.02	90.06	85.05	80.46	78.27	81.36	82.28	83.85	83.40	85.13
99	S	95.19	93.26	90.85	84.87	83.2	77.58	82.42	84.28	89.24	90.95	90.14
109	N	94.43	92.54	86.75	81.07	80.57	78.41	79.96	79.94	80.48	81.07	83.8
109	S	94.19	92.80	88.66	81.16	79.54	77.58	81.32	81.97	83.60	86.21	87.81
112	N	95.31	93.50	94.01	87.25	81.57	78.43	79.3	80.63	82.05	85.00	85.57
112	S	96.34	94.60	93.17	88.09	85.11	82.28	85.73	87.08	89.43	92.13	91.98
116	N	94.94	92.13	88.37	83.17	79.28	77.03	78.95	80.27	83.31	86.88	87.26
116	S	94.73	92.57	88.30	83.66	81.30	80.09	83.76	85.41	89.37	89.78	89.57
122	N	94.96	92.94	91.67	86.30	84.22	82.09	83.63	85.46	86.28	87.36	88.77
122	S	95.56	93.32	92.82	87.00	86.93	84.92	88.35	88.55	90.55	93.07	92.63
135	N	94.11	91.40	86.12	78.21	77.08	76.27	77.85	80.12	82.27	83.41	86.75
135	S	94.32	91.99	87.45	79.10	79.61	78.73	83.32	86.48	89.03	90.89	90.6
171	N	93.10	91.15	86.54	76.50	75.66	72.87	74.31	76.81	76.67	79.05	81.31
171	S	95.18	92.70	85.27	77.37	78.23	78.37	80.88	81.45	83.99	86.88	85.24
178	N	94.00	93.31	90.06	80.77	79.71	78.02	79.53	80.90	81.88	83.21	85.66
178	S	94.47	93.14	89.88	82.00	84.03	81.73	85.01	87.08	88.47	88.87	89.51
195	N	94.77	93.14	90.23	79.57	80.66	77.89	80.10	81.91	84.61	87.27	88.97
195	S	94.68	92.91	87.30	81.23	83.00	80.87	83.74	85.22	88.15	91.2	91.08
244	N	94.73	94.06	91.73	87.63	82.64	78.27	79.67	81.24	81.15	82.22	84.01
244	S	94.80	92.89	90.84	82.81	81.34	78.92	81.49	83.02	86.73	88.5	87.5
253	N	95.64	93.44	92.11	80.51	80.14	77.07	78.66	80.20	82.40	84.13	86.11
253	S	94.36	93.56	90.91	82.34	81.06	78.15	81.55	83.69	86.63	91.79	91.37
269	N	95.86	93.78	88.41	86.04	82.72	81.38	81.41	83.2	84.95	88.9	91.9
269	S	95.62	93.72	90.40	85.30	86.77	82.61	84.62	85.23	87.51	91.88	91.06
278	N	94.72	93.17	89.92	80.30	79.81	77.08	79.72	80.19	82.39	83.76	87.14
278	S	94.42	92.79	88.97	77.27	79.54	77.64	81.85	83.67	86.83	88.66	87.59
322	N	95.37	93.50	86.77	77.56	78.39	76.99	79.03	79.93	80.54	83.37	85.93
322	S	94.83	91.77	85.59	77.89	79.49	79.46	82.20	83.31	86.48	88.91	88.36
325	N	95.43	93.01	88.54	78.08	75.43	73.20	74.11	75.02	76.50	77.57	79.33
325	S	94.55	93.19	88.23	78.56	77.47	75.00	78.37	79.44	83.77	86.36	85.35
335	N	95.43	92.07	88.22	79.15	79.21	76.84	79.58	81.16	84.40	83.87	85.39

Table 6.1 continue

		Days After Sowing (DAS)										
Genotype	**S/N**	**55**	**62**	**76**	**92**	**104**	**112**	**117**	**124**	**132**	**139**	**147**
335	S	94.51	92.36	87.88	81.46	80.79	78.85	82.93	84.07	89.11	89.86	88.68
343	N	93.26	91.07	84.65	75.07	75.64	75.8	78.06	79.61	81.19	84.38	85.33
343	S	93.42	91.87	85.78	74.62	76.05	76.01	78.93	80.83	82.55	86.09	85.6
397	N	93.99	93.56	91.99	89.10	85.87	82.38	82.91	83.87	85.86	87.2	89.97
397	S	95.11	93.79	90.83	87.07	89.29	85.08	87.02	88.15	91.06	94.58	93.44
409	N	94.55	92.67	90.10	89.47	87.49	83.62	85.71	86.48	87.51	88.42	90.62
409	S	94.75	92.89	90.40	87.03	90.56	85.77	87.62	89.27	90.77	92.02	91.8
415	N	94.29	93.19	90.34	82.07	80.11	78.11	79.61	82.53	83.65	84.36	84.27
415	S	94.58	93.20	91.15	86.09	82.55	82.32	84.64	84.87	86.33	89.04	88.39
416	N	94.16	92.23	87.52	80.10	79.08	77.15	78.64	80.30	82.11	83.86	86.89
416	S	94.02	91.41	87.57	77.77	77.28	77.95	80.87	81.29	83.51	87.22	87.74
441	N	95.25	94.31	91.91	81.00	81.78	80.84	80.72	83.21	85.42	85.8	85.75
441	S	94.91	93.97	90.90	80.59	78.96	77.86	82.49	83.70	86.74	87.79	89.67
451	N	94.18	92.71	87.99	76.46	77.89	74.75	76.78	78.02	78.65	79.5	81.93
451	S	93.97	91.77	87.13	75.94	77.59	75.97	80.38	81.42	84.95	88.25	87.57
459	N	95.26	93.15	91.28	88.38	84.99	82.48	83.68	83.71	86.39	89.03	92.02
459	S	95.20	93.75	92.93	90.37	88.41	84.87	86.15	86.42	89.77	93.06	93.07
487	N	94.39	92.56	89.14	85.43	78.24	75.84	78.15	78.64	81.59	83.93	87.89
487	S	94.02	92.38	88.92	86.44	83.61	79.96	83.78	84.21	86.40	89.98	90.29
494	N	94.81	93.35	89.40	81.37	78.74	76.02	77.77	79.25	79.98	81.99	85.56
494	S	94.44	93.37	90.30	82.03	80.98	77.94	80.76	81.37	84.66	88.37	89.74
511	N	94.96	93.93	90.79	78.95	79.43	78.58	80.95	83.45	84.34	85.69	86.57
511	S	95.38	94.53	91.87	80.88	79.62	80.33	83.42	85.44	89.52	89.94	87.7
CAPPELLI	N	94.65	92.56	88.17	80.64	77.26	75.39	76.40	77.84	81.72	80.27	81.9
CAPPELLI	S	94.56	92.64	86.96	76.42	78.84	78.07	82.48	84.69	88.54	89.81	87.79
SARAGOLLA	N	93.63	93.52	88.99	82.28	79.24	76.79	79.44	80.88	80.69	81.48	81.21
SARAGOLLA	S	93.88	93.24	87.77	76.85	80.39	81.03	84.96	84.71	88.47	89.2	88.75
SVEVO	N	94.91	93.14	92.65	87.55	84.29	81.13	82.09	83.88	86.84	89.73	90.52
SVEVO	S	95.53	93.60	92.92	86.60	85.70	82.35	85.35	85.26	89.46	91.6	94.53

Table 6.2 Crop growth stages for cereals. Adapted Zadok Decimal Code (From Stapper 2007). The precise stage of a wheat plant can be designed with the Zadoks decimal code as Zxx (also DCxx, Dxx or GSxx).

0 Emergence
00 Dry seed sown
01 Seed absorbs water
03 Germination, seed swollen
05 Radicle emerged from seed
07 Coleoptile emerged from seed
09 Leaf at coleoptile tip
10 First leaf through coleoptile and tip visible
1 Seedling growth
11 1st leaf more then half visible
12 2nd leaf more then half visible
13 3rd leaf more then half visible
14 4th leaf more then half visible
19 6th leaf more then half visible
17 7th leaf more then half visible
18 8th leaf more then half visible
19 9 or more leaves visible and stem not elongating
2 Tillering
20 Main shoot only
21 Main shoot and 1 tiller
22 Main shoot and 2 tillers
23 Main shoot and 3 tillers
24 Main shoot and 4 tillers
25 Main shoot and 5 tillers
26 Main shoot and 6 tillers
27 Main shoot and 7 tillers
28 Main shoot and 8 tillers
29 Main shoot and 9 or more tillers
3 Stem elongation
30 stem starts to elongate, 'spike at 1cm'
31 swelling 1st node detectable
32 swelling 2nd node detectable
33 swelling 3rd node detectable
34 swelling 4th node detectable
35 swelling 5th node detectable
36 swelling 6th node detectable
4 Flag leaf to Booting
37 Flag leaf tip visible
38 Flag leaf half visible
39 Flag leaf ligule just visible
41 Early boot, flag sheath extending
43 Mid-boot, boot opposite ligule of 2nd last leaf
45 Full-boot, boot above ligule of 2nd last leaf
47 Flag leaf sheath opening
49 First awns visible

5 Heading
51 10% of spikes visible (ear peep)
52
53 30% of spikes visible
54
55 50% of spikes visible
56
57 70% of spikes visible
58
59 90% of spikes visible
60 Whole spike visible, no yellow anthers
6 Flowering (anthesis)
61 Early– 20% of spikes with anthers
62
63 30% of spikes with yellow anthers
64
65 Mid– half of spikes with anthers
66
67 70% of spikes with anthers
68
69 Late– 90% of spikes with anthers
7 Kernel and Milk development
70.2 Kernels middle spike extended 20%
70.5 Kernels middle spike half formed
70.8 Kernels middle spike extended 80%
71 Watery ripe, clear liquid
73 Early milk, liquid off-white
75 Medium milk, contents milky liquid
77 Late milk, more solids in milk
79 Very-late milk, half solids in milk
8 Dough development
81-85 spikes turn colour from light-green to yellow-green to yellow
81 Very early dough, more solids and slides when crushed
83 Early dough, soft, elastic and almost dry, shiny
85 Soft dough, firm, crumbles but fingernail impression not held
87 Hard dough, fingernail impression held, spike yellow-brown
89 Late hard-dough, difficult to dent
9 Ripening
91 Kernels hard, difficult to divide by thumb-nail
92 Harvest ripe, kernels can no longer be divided by thumb-nail and straw still firm
93 Kernels loosening in daytime
94 Over-ripe, straw brittle

7 ACKNOWLEGMENT

I thank the Environmental Biotechnology division of the Department of Chemistry, Life Sciences and Environmental Sustainability of the University of Parma for providing financial support, all the equipment and facilities which have supported all the three years of research behind this work.

I thank the Coordinator of the Ph.D course in Biotechnology and Life Sciences, Prof. Simone Ottonello.

I would like to express my sincere gratitude to my advisor Prof. Nelson Marmiroli, for the continuous support of my Ph.D study and related research. His guidance helped me in all the time of this research and I am really thankful for the opportunity he gave me in developing this entire work.

Besides my advisor, I would like to thank Prof. Mariolina Gullì of the University of Parma, and Dr. Michela Jannì or the CNR-IBBR of Bari, for their insightful comments and encouragement, but also for the incentive to widen my research from various perspectives.

My sincere thanks go to Prof. Cristobal Uauy who provided me the opportunity to join his team at the John Innes Centre (Norwich, UK), and who gave access to his laboratory and the research facilities at the JIC. Prof. Uauy's contribute has been fundamental for the development of TILLING mutant lines, and the development of the allele mining strategy. Among all the people of Prof. Uauy's lab, I would like to thank Dr. James Simmonds who taught me how to cross wheat plants, and Dr Ricardo Ramirez-Gonzalez who really helped me with the NGS data management.

All the phenotypic data and analysis related to the SSD collection, and the DNA of the SSDs used in this work, were kindly provided by the IBBR of Bari in collaboration with the ALSIA Metapontum Agrobios, of which I thank all the people that let me to have data access. In particular I'm really thankful to Dr. Donatella Danzi and Prof. Domenico Pignone from the IBBR of Bari, Dr. Francesco Cellini director of the ALSIA Metapontum Agrobios research center, Dr. Angelo Petrozza and Dr Stephan Summerer.

A final special thank goes to all the people of my laboratory team which supported any part of this research. I would like to thank Dr. Giacomo Lencioni for helping me during the hard work in the glasshouse, Dr. Davide Imperiale and Dr. Laura Paesano for their patience during the analysis at the Tecnopolo Infrastracture of the University of Parma.

I really want to thank both my office-mates, which have shared with me not only the physical space of our office, but also laughs, worries and friendship. So a big "thank-you" is needed to Dr Graziella Pira which was always there for me, nevertheless the time of the day, teaching me how to calm down, and how to see colours everywhere in the world around, and Dr Roberta Ruotolo which supported me in many moments during these last three years, not only as a colleague, but also as a role model of researcher, good person and true friend.

And finally, I have a real special thanks to the person that gave me the major strength to go on, to believe, to learn and understand. I know that I put strain on your patience many times, but I’m sure that you also know how much I loved everything you did for me.

8 References

Adamski, N.M., Borrill, P., Brinton, J., Harrington, S., Marchal, C., and Uauy, C. (2018). A roadmap for gene functional characterisation in wheat (PeerJ Inc.).

Altenbach, S.B. (2012). New insights into the effects of high temperature, drought and post-anthesis fertilizer on wheat grain development. J. Cereal Sci. 56, 39–50.

Aprile, A., Havlickova, L., Panna, R., Marè, C., Borrelli, G.M., Marone, D., Perrotta, C., Rampino, P., De Bellis, L., Curn, V., et al. (2013). Different stress responsive strategies to drought and heat in two durum wheat cultivars with contrasting water use efficiency. BMC Genomics 14, 821.

Araus, J.L., and Cairns, J.E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. Trends Plant Sci. 19, 52–61.

Arioca (2012). Plant Responses to Drought Stress: From Morphological to Molecular Features (Berlin Heidelberg: Springer-Verlag).

Bagge, M., and Lübberstedt, T. (2008). Functional markers in wheat: technical and economic aspects. Mol. Breed. 22, 319–328.

Barakat, M., Saleh, M., Al-Doss, A., Moustafa, K., Elshafei, A., and Al-Qurainy, F. (2015). Identification of new SSR markers linked to leaf chlorophyll content, flag leaf senescence and cell membrane stability traits in wheat under water stressed condition. Acta Biol. Hung. 66, 93–102.

Basu, S., Ramegowda, V., Kumar, A., and Pereira, A. (2016). Plant adaptation to drought stress. F1000Research 5.

Berger, B., Parent, B., and Tester, M. (2010). High-throughput shoot imaging to study drought responses. J. Exp. Bot. 61, 3519–3528.

Binder, S. (2010). Branched-Chain Amino Acid Metabolism in Arabidopsis thaliana. Arab. Book Am. Soc. Plant Biol. 8.

Blum, A. (1996). Crop responses to drought and the interpretation of adaptation. Plant Growth Regul. 20, 135–148.

Borrill, P., Adamski, N., and Uauy, C. (2015). Genomics as the key to unlocking the polyploid potential of wheat. New Phytol. 208, 1008–1022.

Botticella, E., Sestili, F., Hernandez-Lopez, A., Phillips, A., and Lafiandra, D. (2011). High resolution melting analysis for the detection of EMS induced mutations in wheat SBEIIa genes. BMC Plant Biol. 11, 156.

Bovina, R., Brunazzi, A., Gasparini, G., Sestili, F., Palombieri, S., Botticella, E., Lafiandra, D., Mantovani, P., and Massi, A. (2014). Development of a TILLING resource in durum wheat for reverse- and forward-genetic analyses. Crop Pasture Sci. 65, 112–124.

Bowne, J.B., Erwin, T.A., Juttner, J., Schnurbusch, T., Langridge, P., Bacic, A., and Roessner, U. (2012). Drought Responses of Leaf Tissues from Wheat Cultivars of Differing Drought Tolerance at the Metabolite Level. Mol. Plant 5, 418–429.

Brown, A.H.D. (2010). Variation under domestication in plants: 1859 and today. Philos. Trans. R. Soc. Lond. B Biol. Sci. 365, 2523–2530.

Budak, H., Hussain, B., Khan, Z., Ozturk, N.Z., and Ullah, N. (2015). From Genetics to Functional Genomics: Improvement in Drought Signaling and Tolerance in Wheat. Front. Plant Sci. 6.

Caicedo, A.L., and Purugganan, M.D. (2005). Comparative Plant Genomics. Frontiers and Prospects. Plant Physiol. 138, 545–547.

Callaway, E. (2017). Small group scoops international effort to sequence huge wheat genome. Nat. News.

Callaway, E. (2018). CRISPR plants now subject to tough GM laws in European Union.

Casañas, F., Simó, J., Casals, J., and Prohens, J. (2017). Toward an Evolved Concept of Landrace. Front. Plant Sci. 8.

Ceoloni, C., Kuzmanović, L., Forte, P., Gennaro, A., and Bitti, A. (2014). Targeted exploitation of gene pools of alien Triticeae species for sustainable and multi-faceted improvement of the durum wheat crop. Crop Pasture Sci. 65, 96–111.

Chaves, M.M., Flexas, J., and Pinheiro, C. (2009). Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. Ann. Bot. 103, 551–560.

Cheng, L., Wang, Y., He, Q., Li, H., Zhang, X., and Zhang, F. (2016). Comparative proteomics illustrates the complexity of drought resistance mechanisms in two wheat (Triticum aestivum L.) cultivars under dehydration and rehydration. BMC Plant Biol. 16, 188.

Chow, C.-N., Zheng, H.-Q., Wu, N.-Y., Chien, C.-H., Huang, H.-D., Lee, T.-Y., Chiang-Hsieh, Y.-F., Hou, P.-F., Yang, T.-Y., and Chang, W.-C. (2016). PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. Nucleic Acids Res. 44, D1154-1160.

Chung, Y.S., Choi, S.C., Jun, T.-H., and Kim, C. (2017). Genotyping-by-sequencing: a promising tool for plant genetics research and breeding. Hortic. Environ. Biotechnol. 58, 425–431.

Clavijo, B.J., Venturini, L., Schudoma, C., Accinelli, G.G., Kaithakottil, G., Wright, J., Borrill, P., Kettleborough, G., Heavens, D., Chapman, H., et al. (2017). An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations. Genome Res. 27, 885–896.

Comastri, A., Janni, M., Simmonds, J., Uauy, C., Pignone, D., Nguyen, H.T., and Marmiroli, N. (2018). Heat in Wheat: Exploit Reverse Genetic Techniques to Discover New Alleles Within the Triticum durum sHsp26 Family. Front. Plant Sci. 9.

Consortium (IWGSC), T.I.W.G.S., Investigators, I.R. principal, Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., Investigators, I. whole-genome assembly principal, Pozniak, C.J., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science 361, eaar7191.

Cook, J., Nuccitelli, D., Green, S.A., Richardson, M., Winkler, B., Painting, R., Way, R., Jacobs, P., and Skuce, A. (2013). Quantifying the consensus on anthropogenic global warming in the scientific literature. Environ. Res. Lett. 8, 024024.

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

Daryanto, S., Wang, L., and Jacinthe, P.-A. (2017). Global synthesis of drought effects on cereal, legume, tuber and root crops production: A review. Agric. Water Manag. 179, 18–33.

De Vries, A.P. (1971). Flowering biology of wheat, particularly in view of hybrid seed production — A review. Euphytica 20, 152–170.

Devos, K.M., Millan, T., and Gale, M.D. (1993). Comparative RFLP maps of the homoeologous group-2 chromosomes of wheat, rye and barley. Theor. Appl. Genet. 85, 784–792.

Devos, K.M., Dubcovsky, J., Dvořák, J., Chinoy, C.N., and Gale, M.D. (1995). Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. 91, 282–288.

Diebold, R., Schuster, J., Däschner, K., and Binder, S. (2002). The Branched-Chain Amino Acid Transaminase Gene Family in Arabidopsis Encodes Plastid and Mitochondrial Proteins. Plant Physiol. 129, 540–550.

Döll, P., Jiménez-Cisneros, B., Oki, T., Arnell, N.W., Benito, G., Cogley, J.G., Jiang, T., Kundzewicz, Z.W., Mwakalila, S., and Nishijima, A. (2015). Integrating risks of climate change into water management. Hydrol. Sci. J. 60, 4–13.

Dong, C., Vincent, K., and Sharp, P. (2009). Simultaneous mutation detection of three homoeologous genes in wheat by High Resolution Melting analysis and Mutation Surveyor®. BMC Plant Biol. 9, 143.

Dubcovsky, J., and Dvorak, J. (2007). Genome Plasticity a Key Factor in the Success of Polyploid Wheat Under Domestication. Science 316, 1862–1866.

Dwivedi, S.L., Ceccarelli, S., Blair, M.W., Upadhyaya, H.D., Are, A.K., and Ortiz, R. (2016). Landrace Germplasm for Improving Yield and Abiotic Stress Adaptation. Trends Plant Sci. 21, 31–42.

Eftekhari, A., Baghizadeh, A., Yaghoobi, M.M., and Abdolshahi, R. (2017). Differences in the drought stress response of DREB2 and CAT1 genes and evaluation of related physiological parameters in some bread wheat cultivars. Biotechnol. Biotechnol. Equip. 31, 709–716.

El Baidouri, M., Murat, F., Veyssiere, M., Molinier, M., Flores, R., Burlot, L., Alaux, M., Quesneville, H., Pont, C., and Salse, J. (2017). Reconciling the evolutionary origin of bread wheat (Triticum aestivum). New Phytol. 213, 1477–1486.

Eulalia, C.G., and Dessislava, Y. (2015). Water use in the EU - Think Tank.

Fahad, S., Bajwa, A.A., Nazir, U., Anjum, S.A., Farooq, A., Zohaib, A., Sadia, S., Nasim, W., Adkins, S., Saud, S., et al. (2017). Crop Production under Drought and Heat Stress: Plant Responses and Management Options. Front. Plant Sci. 8.

FAO (2015). AQUASTAT. FAO, Rome, Italy

Farooq, M., Hussain, M., and Siddique, K.H.M. (2014). Drought Stress in Wheat during Flowering and Grain-filling Periods. Crit. Rev. Plant Sci. 33, 331–349.

Fleury, D., Jefferies, S., Kuchel, H., and Langridge, P. (2010). Genetic and genomic tools to improve drought tolerance in wheat. J. Exp. Bot. 61, 3211–3222.

Galili, G., Amir, R., and Fernie, A.R. (2016). The Regulation of Essential Amino Acid Synthesis and Accumulation in Plants. Annu. Rev. Plant Biol. 67, 153–178.

Gao, Y., Wu, M., Zhang, M., Jiang, W., Liang, E., Zhang, D., Zhang, C., Xiao, N., and Chen, J. (2018). Roles of a maize phytochrome-interacting factors protein ZmPIF3 in regulation of drought stress responses by controlling stomatal closure in transgenic rice without yield penalty. Plant Mol. Biol. 97, 311–323.

Garcia, M., and Mather, D.E. (2014). From genes to markers: exploiting gene sequence information to develop tools for plant breeding. Methods Mol. Biol. Clifton NJ 1145, 21–36.

Giraldo, P., Royo, C., González, M., Carrillo, J.M., and Ruiz, M. (2016). Genetic Diversity and Association Mapping for Agromorphological and Grain Quality Traits of a Structured Collection of Durum Wheat Landraces Including subsp. durum, turgidum and diccocon. PLOS ONE 11, e0166577.

Golldack, D., Li, C., Mohan, H., and Probst, N. (2014). Tolerance to drought and salt stress in plants: Unraveling the signaling networks. Front. Plant Sci. 5.

Govindaraj, M., Vetriventhan, M., and Srinivasan, M. (2015). Importance of Genetic Diversity Assessment in Crop Plants and Its Recent Advances: An Overview of Its Analytical Perspectives. Genet. Res. Int. 2015.

Gulli, M., Mestri, E., Hartings, H., Raho, G., Perrotta, C., Devos, K., M., Marmiroli, N. (1995). Isolation and characterization of abscisic acid inducible genes in barley seedlings and their responsivness to environmental stress. Life Sci Adv – Plant Physiol 14: 89-96

Haddeland, I., Heinke, J., Biemans, H., Eisner, S., Flörke, M., Hanasaki, N., Konzmann, M., Ludwig, F., Masaki, Y., Schewe, J., et al. (2014). Global water resources affected by human interventions and climate change. Proc. Natl. Acad. Sci. 111, 3251–3256.

Hafsi, M., Mechmeche, W., Bouamama, L., Djekoune, A., Zaharieva, M., and Monneveux, P. (2000). Flag Leaf Senescence, as Evaluated by Numerical Image Analysis, and its Relationship with Yield under Drought in Durum Wheat. J. Agron. Crop Sci. 185, 275–280.

Hazard, B., Zhang, X., Colasuonno, P., Uauy, C., Beckles, D.M., and Dubcovsky, J. (2012). Induced mutations in the starch branching enzyme II (SBEII) genes increase amylose and resistant starch content in durum wheat. Crop Sci. 52, 1754–1766.

Henikoff, S., Till, B.J., and Comai, L. (2004). TILLING. Traditional mutagenesis meets functional genomics. Plant Physiol. 135, 630–636.

Henry, I.M., Nagalakshmi, U., Lieberman, M.C., Ngo, K.J., Krasileva, K.V., Vasquez-Gross, H., Akhunova, A., Akhunov, E., Dubcovsky, J., Tai, T.H., et al. (2014). Efficient Genome-Wide Detection and Cataloging of EMS-Induced Mutations Using Exome Capture and Next-Generation Sequencing. Plant Cell 26, 1382–1397.

Hou, H., Atlihan, N., and Lu, Z.-X. (2014). New biotechnology enhances the application of cisgenesis in plant breeding. Front. Plant Sci. 5.

Iglesias, A., and Garrote, L. (2015). Adaptation strategies for agricultural water management under climate change in Europe. Agric. Water Manag. 155, 113–124.

International Wheat Genome Sequencing Consortium (IWGSC) (2014). A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. Science 345, 1251788.

IPCC (2007): Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp.

IPCC (2014): Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Janni, M., Cadonici, S., Pignone, D., and Marmiroli, N. (2017). Survey and new insights in the application of PCR-based molecular markers for identification of HMW-GS at the Glu-B1 locus in durum and bread wheat. Plant Breed. 136, 467–473.

Janni, M., Cadonici, S., Bonas, U., Grasso, A., Dahab, A. A. D., Visioli, G., Pignone, D., Ceriotti, A., Marmiroli, N.(2018). Gene-ecology of durum wheat HMW glutenin reflects their diffusion from the center of origin. Scientific Reports. In press

Jaradat, A.A. (2013). WHEAT LANDRACES: A MINI REVIEW. Emir. J. Food Agric. 20–29.

Jia, M., Guan, J., Zhai, Z., Geng, S., Zhang, X., Mao, L., and Li, A. (2018). Wheat functional genomics in the era of next generation sequencing: An update. Crop J. 6, 7–14.

Jiménez Cisneros, B.E., T. Oki, N.W. Arnell, G. Benito, J.G. Cogley, P. Döll, T. Jiang, and S.S. Mwakalila, (2014): Freshwater resources. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L.White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 229-269.

Joshi, R., Wani, S.H., Singh, B., Bohra, A., Dar, Z.A., Lone, A.A., Pareek, A., and Singla-Pareek, S.L. (2016). Transcription Factors and Plants Response to Drought Stress: Current Understanding and Future Directions. Front. Plant Sci. 7.

Joshi, V., Joung, J.-G., Fei, Z., and Jander, G. (2010). Interdependence of threonine, methionine and isoleucine metabolism in plants: accumulation and transcriptional regulation under abiotic stress. Amino Acids 39, 933–947.

Kabbaj, H., Sall, A.T., Al-Abdallat, A., Geleta, M., Amri, A., Filali-Maltouf, A., Belkadi, B., Ortiz, R., and Bassi, F.M. (2017). Genetic Diversity within a Global Panel of Durum Wheat (Triticum durum) Landraces and Modern Germplasm Reveals the History of Alleles Exchange. Front. Plant Sci. 8.

Kadirvel, P., Senthilvel, S., Geethanjali, S., Sujatha, M., and Varaprasad, K.S. (2015). Genetic Markers, Trait Mapping and Marker-Assisted Selection in Plant Breeding. In Plant Biology and Biotechnology: Volume II: Plant Genomics and Biotechnology, B. Bahadur, M. Venkat Rajam, L. Sahijram, and K.V. Krishnamurthy, eds. (New Delhi: Springer India), pp. 65–88.

Kage, U., Kumar, A., Dhokane, D., Karre, S., and Kushalappa, A.C. (2016). Functional molecular markers for crop improvement. Crit. Rev. Biotechnol. 36, 917–930.

Karami, A., Shahbazi, M., Niknam, V., Shobbar, Z.S., Tafreshi, R.S., Abedini, R., and Mabood, H.E. (2013). Expression analysis of dehydrin multigene family across tolerant and susceptible barley (Hordeum vulgare L.) genotypes in response to terminal drought stress. Acta Physiol. Plant. 35, 2289–2297.

Kaur, G., and Asthir, B. (2017). Molecular responses to drought stress in plants. Biol. Plant. 61, 201–209.

Kermode, A.R. (2005). Role of Abscisic Acid in Seed Dormancy. J. Plant Growth Regul. 24, 319–344.

Kilian, B., Martin, W., and Salamini, F. (2010). Genetic Diversity, Evolution and Domestication of Wheat and Barley in the Fertile Crescent. In Evolution in Action, M. Glaubrecht, ed. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 137–166.

King, R., Bird, N., Ramirez-Gonzalez, R., Coghill, J.A., Patil, A., Hassani-Pak, K., Uauy, C., and Phillips, A.L. (2015). Mutation Scanning in Wheat by Exon Capture and Next-Generation Sequencing. PloS One 10, e0137549.

Knill, T., Schuster, J., Reichelt, M., Gershenzon, J., and Binder, S. (2008). Arabidopsis Branched-Chain Aminotransferase 3 Functions in Both Amino Acid and Glucosinolate Biosynthesis. Plant Physiol. 146, 1028–1039.

Kooyers, N.J. (2015). The evolution of drought escape and avoidance in natural herbaceous populations. Plant Sci. 234, 155–162.

Kosová, K., Vítámvás, P., and Prášil, I.T. (2014). Wheat and barley dehydrins under cold, drought, and salinity – what can LEA-II proteins tell us about plant stress response? Front. Plant Sci. 5.

Krasileva, K.V., Vasquez-Gross, H.A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., Simmonds, J., Ramirez-Gonzalez, R.H., Wang, X., Borrill, P., et al. (2017). Uncovering hidden variation in polyploid wheat. Proc. Natl. Acad. Sci. 114, E913–E921.

Lal, R. (2015). World Water Resources and Achieving Water Security. Agron. J. 107, 1526–1532.

Lata, C., and Prasad, A.Y. and M. (2011). Role of Plant Transcription Factors in Abiotic Stress Tolerance. Abiotic Stress Response Plants - Physiol. Biochem. Genet. Perspect.

Leiserowitz, A., Maibach, E.W., Roser-Renouf, C., Feinberg, G., and Howe, P. (2013). Climate Change in the American Mind: Americans' Global Warming Beliefs and Attitudes in April 2013. SSRN Electron. J.

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouzé, P., and Rombauts, S. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 30, 325–327.

Lesk, C., Rowhani, P., and Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. Nature 529, 84–87.

Licausi, F., Ohme- Takagi, M., and Perata, P. (2013). APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors: mediators of stress responses and developmental programs. New Phytol. 199, 639–649.

Lopes, M.S., and Reynolds, M.P. (2012). Stay-green in spring wheat can be determined by spectral reflectance measurements (normalized difference vegetation index) independently from phenology. J. Exp. Bot. 63, 3789–3798.

Lopes, M.S., El-Basyoni, I., Baenziger, P.S., Singh, S., Royo, C., Ozbek, K., Aktas, H., Ozer, E., Ozdemir, F., Manickavelu, A., et al. (2015). Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. J. Exp. Bot. 66, 3477–3486.

Malatrasi, M., Corradi, M., Svensson, J.T., Close, T.J., Gulli, M., and Marmiroli, N. (2006). A branched-chain amino acid aminotransferase gene isolated from Hordeum vulgare is differentially regulated by drought stress. Theor. Appl. Genet. 113, 965–976.

Mayer, K.F.X., Martis, M., Hedley, P.E., Šimková, H., Liu, H., Morris, J.A., Steuernagel, B., Taudien, S., Roessner, S., Gundlach, H., et al. (2011). Unlocking the Barley Genome by Chromosomal and Comparative Genomics. Plant Cell 23, 1249–1263.

McCallum, C.M., Comai, L., Greene, E.A., and Henikoff, S. (2000). Targeting Induced LocalLesions IN Genomes (TILLING) for Plant Functional Genomics. Plant Physiol. 123, 439–442.

Migliorini, P., Spagnolo, S., Torri, L., Arnoulet, M., Lazzerini, G., and Ceccarelli, S. (2016). Agronomic and quality characteristics of old, modern and mixture wheat varieties and landraces for organic bread chain in diverse environments of northern Italy. Eur. J. Agron. 79, 131–141.

Murchie, E.H., and Horton, P. (1997). Acclimation of photosynthesis to irradiance and spectral quality in British plant species: chlorophyll content, photosynthetic capacity and habitat preference. Plant Cell Environ. 20, 438–448.

Murtaza, G., Rasool, F., Habib, R., Javed, T., Sardar, K., Ayub, M.M., Ayub, M.A., and Rasool, A. (2016). A Review of Morphological, Physiological and Biochemical Responses of Plants under Drought Stress Conditions. 2, 7.

Nakashima, K., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2014). The transcriptional regulatory network in the drought response and its crosstalk in abiotic stress responses including drought, cold, and heat. Front. Plant Sci. 5.

Neelam, K., Brown-Guedira, G., and Huang, L. (2013). Development and validation of a breeder-friendly KASPar marker for wheat leaf rust resistance locus Lr21. Mol. Breed. 31, 233–237.

Obata, T., and Fernie, A.R. (2012). The use of metabolomics to dissect plant responses to abiotic stresses. Cell. Mol. Life Sci. 69, 3225–3243.

Oliveira, H.R., Campana, M.G., Jones, H., Hunt, H.V., Leigh, F., Redhouse, D.I., Lister, D.L., and Jones, M.K. (2012). Tetraploid Wheat Landraces in the Mediterranean Basin: Taxonomy, Evolution and Genetic Diversity. PLOS ONE 7, e37063.

Parry, M.A.J., Madgwick, P.J., Bayon, C., Tearall, K., Hernandez-Lopez, A., Baudo, M., Rakszegi, M., Hamada, W., Al-Yassin, A., Ouabbou, H., et al. (2009). Mutation discovery for crop improvement. J. Exp. Bot. 60, 2817–2825.

Peña-Bautista, R.J., Hernandez-Espinosa, N., Jones, J.M., Guzmán, C., and Braun, H.J. (2017). CIMMYT Series on Carbohydrates, Wheat, Grains, and Health: Wheat-Based Foods: Their Global and Regional Importance in the Food Supply, Nutrition, and Health ,. Cereal Foods World 62, 231–249.

Petrozza, A., Santaniello, A., Summerer, S., Di Tommaso, G., Di Tommaso, D., Paparelli, E., Piaggesi, A., Perata, P., and Cellini, F. (2014). Physiological responses to Megafol® treatments in tomato plants under drought stress: A phenomic and molecular approach. Sci. Hortic. 174, 185–192.

Pfeifer, M., Kugler, K.G., Sandve, S.R., Zhan, B., Rudi, H., Hvidsten, T.R., Consortium, I.W.G.S., Mayer, K.F.X., and Olsen, O.-A. (2014). Genome interplay in the grain transcriptome of hexaploid bread wheat. Science 345, 1250091.

Pignone, D., De Paola, D., Rapanà, N., and Janni, M. (2015). Single seed descent: a tool to exploit durum wheat (Triticum durum Desf.) genetic resources. Genet. Resour. Crop Evol. *62*, 1029–1035.

Pignone, D., & Hammer, K. (2013). In: Kole, C. (Ed), Genomics and breeding for climate-resilient crops. Springer, Berlin Heidelberg 1:9–26.

Poiré, R., Chochois, V., Sirault, X.R.R., Vogel, J.P., Watt, M., and Furbank, R.T. (2014). Digital imaging approaches for phenotyping whole plant nitrogen and phosphorus response in Brachypodium distachyon. J. Integr. Plant Biol. 56, 781–796.

Ramankutty, N., Mehrabi, Z., Waha, K., Jarvis, L., Kremen, C., Herrero, M., and Rieseberg, L.H. (2018). Trends in Global Agricultural Land Use: Implications for Environmental Health and Food Security. Annu. Rev. Plant Biol. 69, 789–815.

Ramirez-Gonzalez, R.H., Uauy, C., and Caccamo, M. (2015a). PolyMarker: A fast polyploid primer design pipeline. Bioinformatics 31, 2038–2039.

Ramirez-Gonzalez, R.H., Segovia, V., Bird, N., Fenwick, P., Holdgate, S., Berry, S., Jack, P., Caccamo, M., and Uauy, C. (2015b). RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. Plant Biotechnol. J. 13, 613–624.

Ramírez-González, R.H., Borrill, P., Lang, D., Harrington, S.A., Brinton, J., Venturini, L., Davey, M., Jacobs, J., Ex, F. van, Pasha, A., et al. (2018). The transcriptional landscape of polyploid wheat. Science 361, eaar6089.

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

Reif, J.C., Zhang, P., Dreisigacker, S., Warburton, M.L., van Ginkel, M., Hoisington, D., Bohn, M., and Melchinger, A.E. (2005). Wheat genetic diversity trends during domestication and breeding. TAG Theor. Appl. Genet. Theor. Angew. Genet. 110, 859–864.

Reyes, J.C., Muro-Pastor, M.I., and Florencio, F.J. (2004). The GATA Family of Transcription Factors in Arabidopsis and Rice. Plant Physiol. 134, 1718–1732.

Ripple, W.J., Wolf, C., Newsome, T.M., Galetti, M., Alamgir, M., Crist, E., Mahmoud, M.I., Laurance, W.F., and 15,364 scientist signatories from 184 countries (2017). World Scientists' Warning to Humanity: A Second Notice. BioScience 67, 1026–1028.

Ritchie, S.W., Nguyen, H.T., and Holaday, A.S. (1990). Leaf Water Content and Gas-Exchange Parameters of Two Wheat Genotypes Differing in Drought Resistance. Crop Sci. 30, 105–111.

Rizhsky, L., Liang, H., Shuman, J., Shulaev, V., Davletova, S., and Mittler, R. (2004). When Defense Pathways Collide. The Response of Arabidopsis to a Combination of Drought and Heat Stress. Plant Physiol. 134, 1683–1696.

Roy, S. (2015). Function of MYB domain transcription factors in abiotic stress and epigenetic control of stress response in plant genome. Plant Signal. Behav. 11.

Ruosteenoja, K., Markkanen, T., Venäläinen, A., Räisänen, P., and Peltola, H. (2018). Seasonal soil moisture and drought occurrence in Europe in CMIP5 projections for the 21st century. Clim. Dyn. 50, 1177–1192.

Sakamoto, H., Araki, T., Meshi, T., and Iwabuchi, M. (2000). Expression of a subset of the Arabidopsis Cys2/His2-type zinc-finger protein gene family under water stress. Gene 248, 23–32.

Savadi, S., Prasad, P., Kashyap, P.L., and Bhardwaj, S.C. (2018). Molecular breeding technologies and strategies for rust resistance in wheat (Triticum aestivum) for sustained food security. Plant Pathol. 67, 771–791.

Scarpeci, T.E., Frea, V.S., Zanor, M.I., and Valle, E.M. (2017). Overexpression of AtERF019 delays plant growth and senescence, and improves drought tolerance in Arabidopsis. J. Exp. Bot. 68, 673–685.

Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N.W., Clark, D.B., Dankers, R., Eisner, S., Fekete, B.M., Colón-González, F.J., et al. (2014). Multimodel assessment of water scarcity under climate change. Proc. Natl. Acad. Sci. 111, 3245–3250.

Schlötterer, C. (2004). The evolution of molecular markers--just a matter of fashion? Nat. Rev. Genet. 5, 63–69.

Semagn, K., Babu, R., Hearne, S., and Olsen, M. (2014). Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. Mol. Breed. 33, 1–14.

Sestili, F., Palombieri, S., Botticella, E., Mantovani, P., Bovina, R., and Lafiandra, D. (2015). TILLING mutants of durum wheat result in a high amylose phenotype and provide information on alternative splicing mechanisms. Plant Sci. Int. J. Exp. Plant Biol. 233, 127–133.

Shakirova, F., Allagulova, C., Maslennikova, D., Fedorova, K., Yuldashev, R., Lubyanova, A., Bezrukova, M., and Avalbaev, A. (2016). Involvement of dehydrins in 24-epibrassinolide-induced protection of wheat plants against drought stress. Plant Physiol. Biochem. 108, 539–548.

Shariatipour, N., and Heidari, B. (2018). Investigation of Drought and Salinity Tolerance Related Genes and their Regulatory Mechanisms in Arabidopsis (Arabidopsis thaliana). Open Bioinforma. J. 11.

Sharma, P.D., Singh, N., Ahuja, P.S., and Reddy, T.V. (2011). Abscisic acid response element binding factor 1 is required for establishment of Arabidopsis seedlings during winter. Mol. Biol. Rep. 38, 5147–5159.

Shewry, P.R. (2009). Wheat. J. Exp. Bot. 60, 1537–1553.

Shewry, P.R., and Hey, S.J. (2015). The contribution of wheat to human diet and health. Food Energy Secur. 4, 178–202.

Shimelis, H., and Laing, M. (2012). Timelines in conventional crop improvement: Pre-breeding and breeding procedures. Aust. J. Crop Sci. 6, 1542.

Slade, A.J., McGuire, C., Loeffler, D., Mullenberg, J., Skinner, W., Fazio, G., Holm, A., Brandt, K.M., Steine, M.N., Goodstal, J.F., et al. (2012). Development of high amylose wheat through TILLING. BMC Plant Biol. 12, 69.

Slama, A., Mallek-Maalej, E., Mohamed, H.B., Rhim, T., and Radhouane, L. (2018). A return to the genetic heritage of durum wheat to cope with drought heightened by climate change. PLOS ONE 13, e0196873.

Solovyev, V. (2008). Statistical Approaches in Eukaryotic Gene Prediction. In Handbook of Statistical Genetics, (Wiley-Blackwell), pp. 97–159

Soreng, R.J., Peterson, P.M., Romaschenko, K., Davidse, G., Zuloaga, F.O., Judziewicz, E.J., Filgueiras, T.S., Davis, J.I., and Morrone, O. (2015). A worldwide phylogenetic classification of the Poaceae (Gramineae). J. Syst. Evol. 53, 117–137.

Soriano, J.M., Villegas, D., Aranzana, M.J., García Del Moral, L.F., and Royo, C. (2016). Genetic Structure of Modern Durum Wheat Cultivars and Mediterranean Landraces Matches with Their Agronomic Performance. PloS One 11, e0160983.

Stapper M. (2007) Crop monitoring and Zadoks growth stages for wheat. CSIRO Plant Industry. Canberra, ACT. Available at: http://www.biologicagfood.com.au/wp-content/uploads/STAPPER-Crop-Monitoring-v21.pdf

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. *30*, 2725–2729

Tardieu, F., Cabrera-Bosquet, L., Pridmore, T., and Bennett, M. (2017). Plant Phenomics, From Sensors to Knowledge. Curr. Biol. CB 27, R770–R783.

Templer, S.E., Ammon, A., Pscheidt, D., Ciobotea, O., Schuy, C., McCollum, C., Sonnewald, U., Hanemann, A., Förster, J., Ordon, F., et al. (2017). Metabolite profiling of barley flag leaves under drought and combined heat and drought stress reveals metabolic QTLs for metabolites associated with antioxidant defense. J. Exp. Bot. 68, 1697–1713.

Trnka, M., Rötter, R.P., Ruiz-Ramos, M., Kersebaum, K.C., Olesen, J.E., Žalud, Z., and Semenov, M.A. (2014). Adverse weather conditions for European wheat production will become more frequent with climate change. Nat. Clim. Change 4, 637–643.

Tsai, H., Howell, T., Nitcher, R., Missirian, V., Watson, B., Ngo, K.J., Lieberman, M., Fass, J., Uauy, C., Tran, R.K., et al. (2011). Discovery of rare mutations in populations: TILLING by sequencing. Plant Physiol. 156, 1257–1268.

Uauy, C. (2017a). Wheat genomics comes of age. Curr. Opin. Plant Biol. 36, 142–148.

Uauy, C. (2017b). Plant Genomics: Unlocking the Genome of Wheat's Progenitor. Curr. Biol. 27, R1122–R1124.

Uauy, C., Paraiso, F., Colasuonno, P., Tran, R.K., Tsai, H., Berardi, S., Comai, L., and Dubcovsky, J. (2009). A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. BMC Plant Biol. 9, 115.

Ullah, N., Yüce, M., Neslihan Öztürk Gökçe, Z., and Budak, H. (2017). Comparative metabolite profiling of drought stress in roots and leaves of seven Triticeae species. BMC Genomics 18.

UNESCO, 2016. The United Nations World Water Development Report 2016. United Nations World Water Assessment Programme Programme Office for Global Water Assessment Division of Water Sciences, UNESCO 06134 Colombella, Perugia, Italy

United Nations, Department of Economic and Social Affairs, Population Division (2017). World Population Prospects: The 2017 Revision, Volume II: Demographic Profiles (ST/ESA/SER.A/400).

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3—new capabilities and interfaces. Nucleic Acids Res. *40*, e115.

Urano, K., Maruyama, K., Ogata, Y., Morishita, Y., Takeda, M., Sakurai, N., Suzuki, H., Saito, K., Shibata, D., Kobayashi, M., et al. (2009). Characterization of the ABA-regulated global responses to dehydration in Arabidopsis by metabolomics. Plant J. 57, 1065–1078.

Varshney, R.K., Graner, A., and Sorrells, M.E. (2005). Genomics-assisted breeding for crop improvement. Trends Plant Sci. 10, 621–630.

Verma, V., Foulkes, M.J., Worland, A.J., Sylvester-Bradley, R., Caligari, P.D.S., and Snape, J.W. (2004). Mapping quantitative trait loci for flag leaf senescence as a yield determinant in winter wheat under optimal and drought-stressed environments. Euphytica 135, 255–263.

Wang, C., Hu, S., Gardner, C., and Lübberstedt, T. (2017). Emerging Avenues for Utilization of Exotic Germplasm. Trends Plant Sci. 22, 624–637.

Wang, T.L., Uauy, C., Till, B., and Liu, C.-M. (2010). TILLING and associated technologies. J. Integr. Plant Biol. 52, 1027–1030.

Wang, T.L., Uauy, C., Robson, F., and Till, B. (2012). TILLING in extremis. Plant Biotechnol. J. 10, 761–772.

Watson, A., Ghosh, S., Williams, M.J., Cuddy, W.S., Simmonds, J., Rey, M.-D., Hatta, M.A.M., Hinchliffe, A., Steed, A., Reynolds, D., et al. (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. Nat. Plants 4, 23–29.

van de Wouw, M., van Hintum, T., Kik, C., van Treuren, R., and Visser, B. (2010). Genetic diversity trends in twentieth century crop cultivars: a meta analysis. TAG Theor. Appl. Genet. Theor. Angew. Genet. 120, 1241–1252.

Yoshida, T., Fujita, Y., Sayama, H., Kidokoro, S., Maruyama, K., Mizoi, J., Shinozaki, K., and Yamaguchi- Shinozaki, K. (2010). AREB1, AREB2, and ABF3 are master transcription factors that cooperatively regulate ABRE-dependent ABA signaling involved in drought stress tolerance and require ABA for full activation. Plant J. 61, 672–685.

Yousfi, S., Kellas, N., Saidi, L., Benlakehal, Z., Chaou, L., Siad, D., Herda, F., Karrou, M., Vergara, O., Gracia, A., et al. (2016). Comparative performance of remote sensing methods in assessing wheat performance under Mediterranean conditions. Agric. Water Manag. 164, 137–147.

Yu, S., Liao, F., Wang, F., Wen, W., Li, J., Mei, H., and Luo, L. (2012). Identification of Rice Transcription Factors Associated with Drought Tolerance Using the Ecotilling Method. PLOS ONE 7, e30765.

Zadoks, J.C., Chang, T.T., and Konzak, C.F. (1974). A decimal code for the growth stages of cereals. Weed Res. 14, 415–421.

Zeng, H., Xu, L., Singh, A., Wang, H., Du, L., and Poovaiah, B.W. (2015). Involvement of calmodulin and calmodulin-like proteins in plant responses to abiotic stresses. Front. Plant Sci. 6.

Zeven, A.C. (1998). Landraces: A review of definitions and classifications. Euphytica 104, 127–139.

Zhang, C., Li, C., Liu, J., Lv, Y., Yu, C., Li, H., Zhao, T., and Liu, B. (2017). The OsABF1 transcription factor improves drought tolerance by activating the transcription of COR413-TM1 in rice. J. Exp. Bot. 68, 4695–4707.

Zheng, Y., Wang, Z., and Gu, Y. (2014). Development and function of caryopsis transport tissues in maize, sorghum and wheat. Plant Cell Rep. 33, 1023–1031.