Universitat de Lleida

Assessment of genomic and high-throughput phenotyping tools in a diversity panel of Mediterranean landraces and cultivars for wheat breeding under rainfed environments

Rubén Rufo Gómez

http://hdl.handle.net/10803/673165

Universitat de Lleida

IRTA
RESEARCH & TECHNOLOGY
FOOD & AGRICULTURE

Rubén Rufo Gómez

Assessment of genomic and high-throughput phenotyping tools in a diversity panel of Mediterranean landraces and cultivars for wheat breeding under rainfed environments

Universitat de Lleida

IRTA
RESEARCH & TECHNOLOGY
FOOD & AGRICULTURE

Assessment of genomic and high-throughput phenotyping tools in a diversity panel of Mediterranean landraces and cultivars for wheat breeding under rainfed environments

Rubén Rufo Gómez

Universitat de Lleida

TESI DOCTORAL

Assessment of genomic and high-throughput phenotyping tools in a diversity panel of Mediterranean landraces and cultivars for wheat breeding under rainfed environments

Rubén Rufo Gómez

Memòria presentada per optar al grau de Doctor per la Universitat de Lleida
Programa de Doctorat en Ciència i Tecnologia Agrària i Alimentària

Direcció:
Dr. Jose Miguel Soriano Soriano
Dr. Joaquim Bellvert Ríos
Programes de Cultius Extensius Sostenibles i Ús Eficient de l'Aigua
Institut de Recerca i Tecnologia Agroalimentàries (IRTA)

Tutoria acadèmica:
Dr. Ignacio Romagosa Clariana
Departament de Producció Vegetal i Ciència Forestal
Universitat de Lleida (UdL)

Lleida, Juliol 2021

Agraïments

La present memòria de Tesi Doctoral s'ha realitzat a l'àrea de Cultius Extensius Sostenibles de l'Institut de Recerca i Tecnologia Agroalimentària (IRTA) de Lleida sota la direcció del Dr. Jose Miguel Soriano i la co-direcció del Dr. Joaquim Bellvert. A tots dos els vull agraïr la seva enorme dedicació, voluntat i efectivitat en la conducció de totes i cadascuna de les etapes d'aquesta tesi.

A tot l'equip de professionals que forma aquesta àrea també els vull agraïr la seva gran ajuda en les activitats de camp i de laboratori així com també el seu recolzament davant le situacions més complicades. Andrea, Teresa, Maria, moltes gràcies per tots els bons moments que m'heu fet passar, sense vosaltres no hagués estat el mateix. També vull donar les gràcies a les Doctores Conxita Royo i Marta da Silva, per la seva experiència, guia i consells que tant m'han servit per aprendre a plantejar els problemes des de diversos enfocs.

Vull agraïr enormement al Dr Silvio Salvi i a tots els integrants del Dipartimento di Scienze e Tecnologie Agro-Alimentari (DISTAL) de la Universitat de Bolònia (Unibo) per haver-me convidat a fer l'estada de la Tesi formant part del seu grup i haver-me fet sentir com un més d'ells. Moltes gràcies per tot el que m'heu ensenyat i per tots aquells meravellosos moments que hem viscut plegats. A tots els alumnes i doctorands d'en Silvio que em veu ajudar tant i que des del primer moment em veu donar suport i amistat, moltíssimes gràcies Gianmarco, Bart, Sara, Adelaide, Giuseppe, Francesco, Laura, Raffaella, Richard, Win, Kai, Alberto i una gran llista de noms. En especial vull agraïr als meus amics Giuseppe Condorelli (ara ja Dr) i David Baldo, no només per la seva ajuda i suport en la feina feta, sinó per la vida que m'han donat durant els 5 mesos d'estada a Bolònia. Gràcies de tot cor.

També volia agraïr al Ministerio de Ciencia e Innovación (MINECO) per haver-me concedit la beca de Formació de Personal Investigador (FPI) BES-2016-078247 associada al projecte AGL2015-65351-R i als projectes RTA2015-00072-C03-01, PID2019-109089RB-C31 i RTI2018-099949-R-C21.

En últim lloc i no menys important, vull donar les gràcies a tota la gent que he fet a Lleida i que m'ha donat el suport necessari per seguir endavant, en especial a la *cuadrilla* (Juanjo, Fermín, Efren, Salomé, Martina i Asier). També a la meva familia i en especial a la Nadet, qui sempre ha estat al meu costat en els moments més durs i en els més gratificants, animant-me, recolzant-me i fent possible que em nantingués tranquil i animat per seguir endavant.

Per acabar, voldria tornar a agraïr al meu director i amic Jose Soriano per l'atenció que m'ha donat en tot moment. Per una feina ben feta, per fer-me sentir valorat, per la teva confiança i per haver-me fet sentir que erem un equip. Si això ha estat possible, en gran part ha estat gràcies a tu.

Index

List of Tables

Chapter 1:

From landraces to improved cultivars: Assessment of genetic diversity and population structure of Mediterranean wheat using SNP markers

Chapter 2:

Exploring the Genetic Architecture of Root-Related Traits in Mediterranean Bread Wheat Landraces by Genome-Wide Association Analysis

Chapter 3:

Using Unmanned Aerial Vehicle and Ground-Based RGB Indices to Assess Agronomic Performance of Wheat Landraces and Cultivars in a Mediterranean-Type Environment

Annexes

List of Figures

Introduction

Chapter 1:
From landraces to improved cultivars: Assessment of genetic diversity and population structure of Mediterranean wheat using SNP markers

Chapter 2:
Exploring the Genetic Architecture of Root-Related Traits in Mediterranean Bread Wheat Landraces by Genome-Wide Association Analysis

Chapter 3:
Using Unmanned Aerial Vehicle and Ground-Based RGB Indices to Assess Agronomic Performance of Wheat Landraces and Cultivars in a Mediterranean-Type Environment

Chapter 4:
Identification of QTL hotspots affecting agronomic traits and high-throughput vegetation indices in rainfed wheat

Discussion

Abstract

Bread wheat (*Triticum aestivum* L.) is the main crop cultivated around the world, but climate change will significantly affect its production, with special impact in the Mediterranean basin. The ultimate purpose of this PhD thesis is to provide scientific knowledge and useful tools for the development of the next generation of superior bread wheat varieties resilient to the increased drought expected in the next decades as consequence of climate change. To achieve this objective, the existence of genetic, phenotypic and/or geographic structures in the germplasm collections was explored and quantitative trait loci (QTLs) controlling traits related to terminal drought resistance using a genome wide association study (GWAS) were identified. The MED6WHEAT IRTA-panel, with 170 landraces from 24 Mediterranean countries and 184 to modern varieties cultivated in 19 countries in the region, was characterized with more than 10K single nucleotide polymorphism (SNP) markers. A clear geographical pattern was found for the landraces, with three subpopulations (SPs) representing the western, northern, and eastern Mediterranean, whereas the modern cultivars were structured according to the breeding programmes operating in the region: CIMMYT/ICARDA, France/Italy, and Balkan/eastern European countries. Landraces were used to investigate their seminal root system architecture (RSA) adapted to rainfed Mediterranean conditions. Those from northern Mediterranean countries showed the highest number of seminal roots with a root angle not statistically different from the western Mediterranean ones, whereas eastern Mediterranean landraces showed the lowest number of roots but the widest angle, the longest shoots, and the lowest seed weight. A GWAS detected marker-trait associations (MTAs) linked to root-related traits and 31 candidate genes related to RSA traits, seed size, root development and abiotic stress tolerance were found within 15 QTL hotspots. The whole panel was evaluated on a two-year field trial using high-throughput phenotyping (HTP) technologies with the aim to predict agronomic traits. The best estimation of LAI was achieved through the modified triangular vegetation index (MTVI2), and ground-based RGB vegetation indices (VIs) showed better predictions of agronomic traits. The predictive value of the models developed for modern genotypes increased when the data of more than one growing season were aggregated to build them. Results based on a three consecutive year study found significant differences for agronomic traits between subpopulations, pointing out the division of the whole set into landraces and modern cultivars. Modern SPs showed higher values of grain yield and components, harvest index and biomass and longer grain filling duration than landrace SPs, which were taller. The highest grain yield was observed for modern cultivars from France and Italy. A GWAS identified 2579 markers associated with agronomic and VIs–related traits that were simplified to 11 QTL hotspots. *In silico* analysis of candidate genes detected 12 differentially expressed genes (DEG) upregulated under abiotic stress within 6 QTL

hotspots. Among them, five genes were previously reported to be involved in abiotic stress tolerance. Overall, these results proved that Mediterranean wheat landraces are a valuable source of variability to introgress new alleles for desirable traits in the breeding programs in the Mediterranean Basin. The use of remote sensing technology is an efficient and rapid tool for the assessment of agronomic traits. Finally, GWAS have resulted a useful approach for the identification of genomic regions controlling important traits in the bread wheat Mediterranean germplasm.

Resum

El blat fariner (*Triticum aestivum* L.) és el principal cultiu arreu del món, però la seva producció es veurà afectada per l'impacte del canvi climàtic, sobretot a la conca Mediterrània. L'objectiu final d'aquesta tesi és oferir coneixement científic i les eines necessàries pel desenvolupament les noves varietats de blat fariner resilients a l'esperat increment de la sequera en les següents dècades com a conseqüència del canvi climàtic. Per aconseguir aquest objectiu, l'existència d'estructures genètiques, fenotípiques i/o geogràfiques en les col·leccions de germoplasma va ser explorada i els QTLs controlant caràcters relacionats amb la resistència terminal a la sequera usant l'anàlisi GWAS van ser identificats. La col·lecció MED6WHEAT del IRTA, amb 170 varietats tradicionals originàries de 24 països mediterranis i 184 varietats modernes cultivades a 19 països de la regió va ser caracteritzada amb més de deu mil marcadors SNP. Un clar patró geogràfic associat a l'estructura genètica es va trobar en les varietats tradicionals, amb 3 subpoblacions (SPs) representant l'oest, el nord i l'est del Mediterrani, mentre que les modernes van ser estructurades d'acord amb els programes de millora genètica de la regió: CIMMYT/ICARDA, França/Itàlia i països balcànics/est-europeus. Les varietats tradicionals van ser utilitzades per investigar l'arquitectura del sistema radicular seminal (RSA) adaptat a les condicions de secà de la regió mediterrània. Les varietats dels països Nord-Mediterranis van mostrar el màxim nombre d'arrels seminals amb un angle radical no significativament diferent al dels de l'oest del Mediterrani, mentre que les varietats tradicionals Est-Mediterrànies van mostrar el mínim nombre d'arrels, major angle radical, les tiges més llargues i el mínim pes de llavor. L'estudi GWAS va detectar MTAs relacionats amb arrels i es van identificar 31 gens candidats vinculats a RSA, mida de la llavor, i tolerància a estrès abiòtic en 15 regions genòmiques. La col·lecció completa va ser avaluada a un assaig a camp de dos anys mitjançant tecnologies de fenotipat d'alt rendiment (HTP) amb l'objectiu de predir caràcters agronòmics. La millor estimació de l'índex d'àrea foliar (LAI) es va assolir amb l'índex MTVI2, i els índexs de vegetació RGB mesurats manualment a camp van mostrar millors prediccions dels caràcters agronòmics. Els resultats basats en un estudi de tres anys consecutius van mostrar diferències significatives pels caràcters agronòmics entre SPs, ressaltant la divisió en varietats tradicionals i varietats modernes. Les SPs modernes van mostrar majors valors de rendiment i dels seus components, de l'índex de collita i de biomassa i una durada de l'ompliment del gra més llarga que les SPs de les varietats tradicionals, que van ser més altes. El major rendiment va ser observat en les varietats modernes de França i Itàlia. El GWAS va identificar 2579 marcadors associats a caràcters agronòmics i relacionats amb els VIs que, finalment, van quedar simplificats en 11 regions genòmiques. L'anàlisi in silico de gens candidats va detectar 12 gens sobre-expressats en condicions d'estrès abiòtic en 6 d'aquestes regions. Cinc d'aquests gens ja havien estat descrits prèviament per

estar relacionats amb la tolerància a l'estrès abiòtic. Globalment, aquests resultats van constatar que les varietats tradicionals Mediterrànies són una font important de variabilitat per introduir nous al·lels de caràcters desitjats en programes de millora genètica al voltant de la conca Mediterrània. L'ús de noves tecnologies de teledetecció és una eina eficient i ràpida per la mesura dels caràcters agronòmics. Finalment, el GWAS ha resultat ser un enfoc molt útil per la identificació de regions genòmiques que controlen caràcters importants del germoplasma del blat fariner Mediterrani.

Resumen

El trigo harinero (*Triticum aestivum* L.) es el principal cultivo en el mundo, pero su producción se verá seriamente afectada por el cambio climático, principalmente en la cuenca Mediterránea. El objetivo final de esta tesis es desarrollar el conocimiento científico y las herramientas necesarias para el desarrollo de las nuevas variedades de trigo harinero resilientes al incremento de la sequía en las próximas décadas como consecuencia del cambio climático. Para ello se ha explorado la existencia de las estructuras genética y fenotípica, así como geográfica en la colección de germoplasma utilizada y se han identificado QTLs que controlan caracteres relacionados con la resistencia a la sequía terminal mediante análisis GWAS. Se caracterizó la colección MED6WHEAT del IRTA, con 170 variedades tradicionales de 24 países mediterráneos y 184 variedades modernas cultivadas en 19 países de la región, con más de diez mil marcadores SNP. Se encontró un patrón geográfico asociado a la estructura genética con 3 subpoblaciones (SPs) de variedades tradicionales que representan el oeste, el norte y el este de la cuenca Mediterránea, mientras que las variedades modernas fueron estructuradas de acuerdo con los programas de mejora de la región: CIMMYT/ICARDA, Francia/Italia y países del este de Europa. Las variedades tradicionales se usaron para investigar la arquitectura del sistema radicular (RSA) adaptado a las condiciones de secano del Mediterráneo. Las variedades de los países del norte mostraron el mayor número de raíces seminales con un ángulo radical no significativamente diferente al de los del oeste, mientras que las variedades tradicionales del este mostraron el menor número de raíces, un mayor ángulo radical, unos tallos más largos y el menor peso de la semilla. El análisis GWAS detectó la presencia de marcadores asociados a los distintos caracteres y se identificaron 31 genes candidatos vinculados a RSA, el tamaño de la semilla, y la tolerancia a estrés abiótico en 15 regiones genómicas. La colección completa fue evaluada en un ensayo en campo de dos años mediante tecnologías de fenotipado masivo (HTP) con el objetivo de predecir caracteres agronómicos. La mejor estimación del LAI se logró con el índice MTVI2, y los índices de vegetación (VI) RGB medidos manualmente en campo mostraron mejores predicciones para los caracteres agronómicos. Los resultados basados en un estudio de tres años consecutivos mostraron diferencias significativas para los caracteres agronómicos entre SPs, resaltando la división en variedades tradicionales y variedades modernas. Las SPs modernas mostraron mayores valores de rendimiento y de sus componentes, mayor índice de cosecha, biomasa y duración de llenado de grano que las SPs de las variedades tradicionales, que fueron a su vez más altas. El mayor rendimiento fue observado en las variedades modernas de Francia e Italia. El análisis GWAS identificó 2579 marcadores asociados a caracteres agronómicos y relacionados con los VIs, que finalmente quedaron simplificados en 11 regiones genómicas. El análisis in silico de genes candidatos detectó 12 genes en condiciones de estrés abiótico en 6 de estas regiones. Cinco de estos genes ya

se han descrito previamente relacionados con la tolerancia a estrés abiótico. Estos resultados constataron que las variedades tradicionales Mediterráneas son una fuente importante de variabilidad para introducir nuevos alelos en programas de mejora genética en la cuenca Mediterránea. El uso de nuevas tecnologías de teledetección es una herramienta eficiente y rápida para estimar caracteres agronómicos. Finalmente, el análisis GWAS ha resultado ser un enfoque muy útil para la identificación de regiones genómicas que controlen caracteres importantes del germoplasma del trigo harinero Mediterráneo.

Introduction

Introduction

1. Origin and evolution of wheat

Wheat (*Triticum* sp) is the most important cereal of Old World agriculture (Zohary and Hopf 2000), belongs to the Poaceae family and its domestication occurred in the Fertile Crescent between 12,000 and 10,000 years before present (BP). Domestication refers to the selection process of plants to get an increased adaptation for cultivation by human purposes (Brown et al. 2009).

Wild wheat (*Triticum Urartu* L.) was diploid (2n=2x=14, genome AA) and crossed with goat grass (*Aegilops speltoides* L.) (2n=2x=14, genome BB) 300–500 thousand years BP to produce a wild emmer wheat (*Triticum dicoccoides* L.) (2n=4x=28, genome AABB), which was highly adapted to extreme drought periods and presented a wide range of morphological ecotypes (Mac Key 2005). Once humans began to cultivate this new species, the selection of a cultivated emmer wheat (*Triticum dicococcum* L.) occurred 10,000 years BP probably in southeast Turkey. Molecular studies suggested that emmer wheat was domesticated in the Diyarbakir region in south-eastern Turkey (Luo et al. 2007). A new hybridization event was produced 9,000 years BP between cultivated emmer wheat and another goat grass, *Aegilops tauschii* L., whose D genome was incorporated into the spelt cereal forming the hexaploidy wheat (*Triticum spelt* L., 2n=6x=42, genome AABBDD) that after domestication and centuries of cultivation and selection originated the modern bread wheat (*Triticum aestivum* L.) (Venske et al. 2019) in the south and west of the fertile crescent, where goat grass belonged (Giles and Brown 2006). However as discussed by other authors as Dvorak et al. (2010), the real origin of the hexaploid wheat is still a dilemma, and its diversity was originated by gene flow from the ancestors, thus the new centres of diversity would not reflect the geography of crop origin.

Subsequently, from the western border of the Fertile Crescent, wheat spread to south-east of Europe through the Caucasus and reached the Balkan Peninsula and Greece over 8,000 years BP. Farming cultures moved it northwards following the rivers Danube (7,000 years BP) and Rhine, reaching England and Scandinavia by 5,000 years BP and, therefore, causing the adaptation of wheat to harsher climates (Helbaek H. 1959). Primitive wheat also spread along the Mediterranean basin to reach Italy and Spain by 7,000 years BP (Feldman 2001).

This migration of wheat and both natural and human selection resulted in the development of local landraces adapted to the different climatic conditions of the site of origin. They were able to tolerate biotic and abiotic stresses with a high yield stability and maintaining an intermediate yield level under low input agricultural systems (Zeven 1998). Domestication caused a substantial genetic erosion from the wild relatives, which was reinforced during modern breeding processes increasing

susceptibility and vulnerability to environmental stresses, pests and diseases (Nevo 2009, 2011). A study conducted by Haudry et al. (2007) in wild and domesticated durum and bread wheats revealed that diversity was further reduced in cultivated forms during domestication by 69% and 84% in bread and durum wheat, respectively.

2. Importance of the crop

The history of wheat is closely associated with a changing relationship of humans to their environment, especially to the efforts to protect their population from hunger and to master food supply and use (Igrejas and Branlard 2020). According to FAO statistics, (http://www.fao.org/faostat/), China and India were the top wheat-producing countries in 2020, mainly because wheat culture has the advantage of requiring less water input for cultivation than other comparable crops while being the main ingredient of a variety of processed foods valued in modern life (Igrejas and Branlard 2020). Nowadays, wheat grain is the most important source of food on earth, containing 75–80% carbohydrates, fibre, many vitamins (especially B vitamins), calcium, iron, and many macro and micro-nutrients, and provides 18% of the total human intake of calories and 20% of protein intake (http://www.fao.org/faostat/). It is the staple food for 40% of the world's population, mainly in Europe, North America, and the western and northern parts of Asia (http://www.croptrust.org). Wheat production progressively increased in the twentieth century, especially after World War II, due to the high demand of a population growing exponentially. It was possible thanks to the agronomic and genetic advances occurred during the Green Revolution, which made wheat an essential crop to humankind. The adaptability of wheat has made possible its cultivation in almost all regions in the world, growing on 17% of all crop areas in the temperate, Mediterranean-type and subtropical parts of both hemispheres, from 67°N in Norway, Finland and Russia to 45°S in Argentina (Peng et al. 2011). It is important to mention the wheat yield potential achieved between 1961 and 2013. Whereas the total land area of wheat increased only by 6.8%, from 204 million hectares (Mha) to 218 Mha, the world production increased by 321% from 222 million tons (Mt) (a worldwide yield of 1 t/ha) to 713 Mt (3.2 t/ha) (Igrejas and Branlard 2020). Nowadays wheat production represents near the 40% of the total cereal production (International Grains Council, IGC, https://www.igc.int/en/default.aspx).

Nearly 95% of cultivated wheat is bread wheat, whereas only 5% is represented by durum wheat (*Triticum durum* L.). The large success of bread wheat has been attributed to the plasticity of the hexaploid genome, which allowed a wider adaptation capacity than tetraploid wheat (Peña-Bautista et al. 2017; Giunta et al. 2018). Despite the D genome present in hexaploid wheat, durum and bread wheat have traditionally been used for different end-use products due to their distinct quality and technological properties. Thus, durum wheat has been used to produce pasta, bulgur, or couscous,

while the final use of bread wheat has been to produce the different types of bread. Nevertheless, their roles can be exchanged, as durum wheat flat breads such as chapatti, tortilla, baladi, tanoori, and pita are mainly consumed in North Africa, West Asia and India, and pasta made with bread wheat flour also exist (Mastrangelo and Cattivelli 2021). Since 1980s, wheat quality improvement accelerated when genetic analysis of wheat storage proteins, the components of gluten, were deeply investigated (Biesiekierski JR 2017).

Despite wheat grain can be stored for long periods of time as food reserves, global wheat demand is predicted to increase by 60% by the year 2050, so there is an urgent need to raise wheat production by 1.7% per year until then (Leegood et al. 2010).

3. Impact of drought stress in the Mediterranean basin

Climate change may be the single unifying and chronic issue that will affect everyone and every aspect of the economy. Changes in weather patterns and variability, as well as differential combinations of effects in different parts of Europe and the Mediterranean region are expected. The Mediterranean Basin embraces countries between 27° and 47°N and between 10°W and 37°E extending over three continents and a coastline of 46,000 km (Royo et al. 2017). It is naturally exposed to several hazards, including earthquakes, volcano eruptions, floods, fires, and/or droughts. Recent accelerated climate change has exacerbated existing environmental problems in the Mediterranean Basin that are caused by the combination of changes in land use, increasing pollution and declining biodiversity. Several new challenges from climate change arise, including warming, more severe droughts, changing extreme events, sea level rise and ocean acidification. The North will see warmer and wetter weather, whereas the South will experience more frequent and severe droughts and heat waves, and in both cases, there will be a shifting pattern of incidence of pests and diseases. Average annual temperatures are now approximately 1.5°C higher than during the preindustrial period (1880 -1899) and above current global warming trends (+1.1°C). The expected change in the global climate will significantly affect wheat production, with special impact in the Mediterranean basin, where prediction models have projected a rise of temperatures by 3–5ºC and a decrease of annual rainfall by 25–30% in the next decades (Giorgi and Lionello 2008).

In the Mediterranean Basin, wheat is mainly cultivated under rainfed conditions with an irregular precipitation pattern across years and locations and along the plant growth cycle resulting in major yield variations (Soriano et al. 2018). Besides, wheat usually experiences terminal drought originated by high temperatures during the grain-filling period (Araus et al. 2008), causing a reduction on yield potential of about 50% (Altenbach 2012). The Intergovernmental Panel on Climate Change

(IPCC) models predict that average yields of cereals will fall due to drought, insect predation and diseases, whilst the demand for food will rise significantly due to population growth. An increasing frequency and severity of terminal drought stress will negatively influence wheat grain weight, quality, and yield (Araus et al. 2003; Slafer et al. 2005; Kulkarni et al. 2017). Therefore, there is a need to improve the selection of crops able to maintain acceptable levels of yield and stability in semi-arid environments, which have been identified as the regions most sensitive to the effects of climate change (Rufo et al. 2021). To achieve that, strategies to retain and increase the genetic diversity are being explored, since climate change is expected to constrain it (Iglesias and Garrote 2015). The adaptability and stability of new cultivars that can be successfully grown in dry areas will be the main concern in breeding programs.

4. Landraces as source of diversity

A landrace is defined as a traditional variety with a high capacity to tolerate biotic and abiotic stresses, resulting in high yield stability and an intermediate yield level under a low input agricultural system (Zeven 1998). Landraces were developed during the evolution of wheat along new territories by human selection after the advent of agriculture. Mediterranean landraces have a good adaptation to their environments, forming populations with different genetic constitutions and are the reservoir of the greatest genetic variability of the species (Royo et al. 2017). The pioneer Mediterranean farmers started selecting the plants with the most favourable characteristics in terms of vigour, phenological adaptation, spike length and yield with the aim to produce improved lines (Royo et al. 2017). Wheat landrace collections contain wider genetic diversity than most breeding programmes and this diversity includes adaptation to different conditions according to the place of origin (Lopes et al. 2015). Mediterranean gene banks have been collecting a large number of wheat germplasm accessions during the last century, creating core collections in order to maximize the genetic variation present in the whole collection with a minimum of repetitiveness (Ruiz et al. 2013). Hybridization is the main source to create variability and parents used in crosses can proceed from different genetic resources, including primitive cultivars, wild species, landraces, or modern accessions. The genetic pool of *Triticum* spp has become exceptionally wide due to the origin of the tribe *Triticeae*, which was originated by alloploidization through hybrid speciation (Zaharieva and Monneveux 2006).

The increasing reliance on relatively few varieties has led to the loss of genetic diversity and, therefore, the genetic vulnerability of wheat cultivars (Skovmand et al. 2005). International research centres such as CIMMYT (International Maize and Wheat Improvement Centre) and ICARDA (International Center for Agricultural Research in the Dry Areas), have largely helped to widen the genetic pool of current

cultivars identifying new sources of genetic variability to find new and useful candidate genes for their introgression into commercial cultivars. Knowledge of the genetic diversity and population structure of landraces is essential for their conservation and efficient use in breeding programmes (Soriano et al. 2016; Pascual et al. 2020) , especially with respect to the field of adaptation to climate change and the quality of the end-products (Lopes et al. 2015). To achieve that, several studies using molecular markers such as simple sequence repeats (SSRs) or single nucleotide polymorphism (SNP) are currently conducting since they have been proven to be very useful for evaluating the genetic diversity and population structure of Mediterranean wheat collections (Soriano et al. 2016; Rufo et al. 2019).

5. The MED6WHEAT diversity panel at IRTA

A set of accessions, including landraces and improved commercial varieties, representative of the variability existing in the species in the Mediterranean Basin, was selected, purified, and multiplied by the Sustainable Field Crops Programme at IRTA in Lleida (Spain). The landraces were selected from a larger collection, comprising 730 accessions of different origins based on phenology, spike characteristics and passport data. Landrace populations were provided by public gene banks from Germany (IPK, Gatersleben), Italy (ISC, S. Angelo Lodigiano), Romania (Suceava GenBank, Suceava), Russia (VIR, St. Petersburg), Spain (CRF-INIA, Madrid), the Netherlands (CGN-WUR, Wageningen) and the USA (NSGC-USDA, Aberdeen, ID). A selection among 300 modern varieties from different countries was based on agronomic and pedigree data. Modern cultivars were provided by public institutions (CIMMYT, ICARDA, INRA and the University of Belgrade), breeding companies and own germplasm from the IRTA breeding programme. To improve the reliability of the subsequent analyses to be carried out with the collection, care was taken to have enough number of genotypes representatives of different drought resistance capacities. Accessions were bulk-purified during two cropping cycles to select the dominant type and seed was increased in plots in the same field to ensure a common origin for all lines. The approach used was to ensure having a balanced number of genotypes with common constitutive traits between the four climatic zones identified in the Mediterranean Basin for wheat cultivation reported in Royo et al. (2014). To represent the past and current cultivated variability in the Mediterranean Basin, the MED6WHEAT IRTA-panel consisted of a germplasm collection of 354 bread wheat genotypes: 170 landraces from 24 Mediterranean countries and 184 modern varieties cultivated in 19 countries in the region. A description of this panel is shown in Annex (Table 1 and 2).

The genetic structure of the whole collection was carried out during this PhD thesis using 557 common SNPs markers between landraces and modern cultivars evenly distributed across the genome (Rufo et al. 2019). The analysis showed

a clear geographical pattern for the landraces, which were clustered into three subpopulations (SPs) representing the western (SP1), northern (SP2) and eastern (SP3) Mediterranean, whereas the modern cultivars were structured according to the breeding programmes that developed them: France/Italy (SP4), Balkan/eastern European countries (SP5) and CIMMYT/ICARDA (SP6). More detailed information is given in the **chapter 1** of this thesis.

6. Breeding under a climate change scenario

Although bread wheat has relatively high adaptability to drought environments than other cereals, its production is threatened by the impacts of climate change. Drought is one of the most severe factors that reduce wheat yield, with a loss of 20% if plants are grown with 40% less water than required to avoid the stress (Daryanto et al. 2016). In the Mediterranean Basin, grain yield and quality are generally constrained by dry growing-season precipitation, as well as by higher temperatures towards the end of the crop cycle (Rharrabti et al. 2003). Breeding can significantly contribute to the mitigation of climate change effects on production by developing drought-tolerant wheat germplasm with improved grain yield, resistance to main diseases and reaching the industry quality requirements. Thus, the main challenge faced by breeders is to identify genotypes able to tolerate multiple stresses that occur simultaneously (Reynolds et al. 2012). To achieve this goal, it is essential to better understand yield formation and the plant mechanisms involved in drought tolerance and their adaptation to the environment. Plants can cope with drought stress through an extensive root system to explore deeper layers of soil (Anjum et al. 2011), with an osmotic adjustment and the accumulation of solutes (Keyvan 2010), maintaining the balance between antioxidants and reactive oxygen species (ROS) levels (Anjum et al. 2011), and inducing leaf senescence as a survival mechanism to maintain favourable water status (Munné-Bosch and Alegre 2004). All this knowledge is critical for breeding programmes to produce and deliver adapted germplasm for suboptimal conditions of diverse wheat growing regions worldwide.

6.1. Phenology fitting

Growth or developmental stages of wheat are specific times at which recognizable physical changes can be seen on the plant. The need to identify such stages is very important because they are accompanied by various morphological and physiological changes from seed germination to plant maturity through different stages: tillering, spike differentiation, stem elongation, heading, anthesis and grain filling.

Under drought conditions, wheat productivity can vary depending on the phenological stage at which the water deficit occurs (Crespo-Herrera et al. 2018), being larger when water is limited at reproductive stages than if it occurs only

at the vegetative stage (Daryanto et al. 2016). Therefore, matching phenology to growing season length changing the cultivar day-length and temperature response could be a useful prospect of adaptation to climate change (Kumudini et al. 2014). The genetics of flowering time in wheat is complex due a strong genotype x environment (GxE) interaction (Mastrangelo et al. 2005). The genetics of wheat development is determined mainly by the allelic diversity within the loci regulating the vernalization requirement (*Vrn*) and photoperiod sensitivity (*Ppd*). A third group of genes controlling earliness when the vernalization and photoperiod requirements are accounted for are the earliness per se loci (*Eps*), characterized by a polygenic inheritance and lower effects than *Vrn* and *Ppd* loci (Royo et al. 2020). Vernalization is the induction of flowering by cold exposure, and plants can be grouped in winter types (vernalization requirement), spring types (no vernalization requirements) or facultative types (intermediate). Photoperiod sensitive cultivars can only flower under long-days, whereas no limitation of light hours is needed for photoperiod insensitive cultivars.

Flowering time (or anthesis) and grain filling period are considered the main target occasions to be focused by breeders since they are defined as sensitive stages of wheat development. Anthesis is physiologically one of the most critical stages highly affected by temperature and is considered a primary trait determining wheat adaptation to a particular set of growing conditions (Worland et al. 1998; Snape et al. 2001). This is the most critical factor to optimize adaptation in environments differing in water availability and distribution during the growing season (Richards 2006). The grain filling period involves milk and dough development, ending at the physiological maturity. It starts once the flowering is complete and corresponds with the early kernel formation stage. Kernel size then increases rapidly during this stage which occurs one to two weeks after pollination. At the end of the period, most of the kernel dry weight (starch and protein content) accumulates. In Mediterranean environments grain filling is limited mainly by rising temperatures and reducing water supply, which reduce photosynthesis rate after anthesis, increasing the contribution of remobilization of pre-anthesis assimilates, thus constraining yield potential (Álvaro et al. 2008). Then, controlling the time to reach a particular growth stage is of paramount importance to allow wheat plants to escape from terminal drought (Habash et al. 2009).

6.2. Grain Yield

Global wheat demand is predicted to increase by 60% by the year 2050, so there is an urgent need to raise wheat production by 1.7% per year until then, which is more than the improvement reached by the Green Revolution (Leegood et al. 2010). Yield improvement in rainfed conditions as Mediterranean-type environments is complex to achieve due to the abiotic stresses characteristic of this region, high

temperatures and drought during the grain filling period, and the low heritability due to the highly polygenic nature of the trait (Royo et al. 2009). Grain yield can be analysed in terms of three primary yield components: number of spikes per unit area (NSm^2), number of grains per spike (NGS) and grain weight (Moragues et al. 2006). Number of grains per unit area (NGm^2) is defined by the product of NSm2 and NGS. The number of grains and their weight are established sequentially during plant development, with the potential number of grains being determined before anthesis, and the grain weight after it (García Del Moral et al. 2005). That explains the negative correlation observed between grain number and grain weight (Sadras and Rodriguez 2007). The improvement of grain yield have been mostly based on increases in NGm^2 and harvest index (HI), defined as the grain yield expressed as the decimal fraction of the aboveground biomass production, thus indicating the allocation of biomass to grain and the partitioning between grain and straw production, thanks to increasing the NSm^2 and grain setting (Royo et al. 2007), while the weight of the grains has remained unchanged (Álvaro et al. 2008). Drought stress affects all the yield components, but particularly reduces the number of fertile spikes per unit area, thus reducing the NGS (Giunta et al. 1993). Grain weight is negatively influenced by high temperatures and drought during ripening (Chmielewski and Köhn 2000).

6.3. Crop traits related to drought tolerance

The crop traits to be considered as selection targets under drought conditions must be genetically correlated with yield and should have a greater heritability than yield itself (Royo and Villegas 2011). Among them, early vigour, leaf area duration, relative water content, radiation use efficiency and root architecture have been identified to be associated with yield under rainfed conditions (reviewed in Tuberosa 2012).

Early vigour is defined as a fast development of leaf area or crop biomass. It allows annual crops to optimize water use efficiency (WUE) under low evapotranspiration conditions and limit the loss of water due to direct evaporation from the soil surface, contributing to increase yield in drought-prone environments (Slafer et al. 2005; Rebetzke et al. 2007).

The **leaf area duration (LAD)** serves as an indicator of premature leaf senescence that may be caused by high temperature and water scarcity during and after flowering (Prasad et al. 2007; Álvaro et al. 2008) that result on a yield loss and poor grain quality. Therefore, delaying leaf senescence maintains transpiration and increases cumulative photosynthesis over the crop life cycle.

Yield potential (YP) can be expressed as a product of intercepted light (IL), **radiation use efficiency (RUE)** and harvest index (HI) (Reynolds et al. 2009).

The theoretical limit for HI has been estimated for wheat as 0.62 (Austin et al. 1980), but historically it has not surpassed 0.50 (Royo et al. 2007). Thus, further yield gains should arise from the traits improving IL (early vigour and leaf area duration), and RUE (Reynolds et al. 2009). Enhanced RUE will probably lead to greater crop growth rate (Fischer and Edmeades 2010), and it is worth to study ancient wheat genotypes since wild relatives of wheat have been found to have greater RUE (Evans 1993).

Root architecture. Roots exhibit a high level of morphological plasticity in response to soil conditions allowing plants a better adaptation, particularly under drought conditions. A deep root system helps the plant to avoid drought stress by extracting water store in deep soil layers (Narayanan and Vara Prasad 2014). Therefore, identifying and introgressing alleles from deeper rooting varieties in adapted phenotypes are desirable approaches for the breeding programs.

7. New approaches for plant phenotyping. From ground to air

One of the main challenges in the breeding programmes is to improve phenotyping for further yield enhances, as this is the bottleneck of the process (Araus and Cairns 2014). Remote sensing has experimented a growing interest since it can deliver more detailed information about the biophysical crop parameters in many situations (Bellvert et al. 2021). Advances in plant phenotyping are even more crucial in scenarios of high environmental variability occurring under climate change (Araus et al. 2018). In recent years, high-throughput phenotyping (HTP) based on the use spectral vegetation indices (VIs) acquired from Unmanned Aerial Vehicles (UAVs) have increasingly improved the evaluation of agronomic traits (Gracia-Romero et al. 2017; Xie and Yang 2020) in large germplasm collections in a rapid, cost-effective and high spatial resolution way (Duan et al. 2017). HTP platforms are capable of simultaneously taking multiple measurements of plant characteristics to capture and provide reliable estimates of trait phenotypes (De Vita and Taranto 2019) using non-intrusive and non-destructive technology (White et al. 2012; Rufo et al. 2021).

Indirect assessments of agronomic and physiological traits can be performed using digital photography and VIs which have a great potential for high-throughput phenotyping in wheat breeding programs in a non-destructive manner (Royo et al. 2005). Digital photography has also been used for quantifying plant traits such as leaf area (Baker et al. 1996; Lukina et al. 1999; Campillo et al. 2008; Casadesús and Villegas 2014) as well as leaf senescence (Adamsen et al. 1999; Ide and Oguma 2010), grassland coverage (Li et al. 2005) and biomass (Casadesús and Villegas 2014). RGB indices such as GA, GGA, a*, and u* have been proven to be suitable for predicting higher yield due to their capacity to calculate a combination of physiological components related to biomass (Casadesús et al. 2007; Casadesús and

Villegas 2014; Rufo et al. 2021).

Differences in the spectral response of vegetation among genotypes has been widely used by breeding programs in order to assess differences in agronomic traits (Royo et al. 2003; Haboudane et al. 2004). At the 70's, most of the studies of field phenotyping were conducted throughout spectroradiometer (Aparicio et al. 2000). However, spectroradiometers remain expensive for many breeding programs as well as measurements of large trials are time-consuming (Casadesús et al. 2007). More recently, with the advance of new remote sensing technologies and the improvement of 'low-cost' multispectral cameras with wavelengths mostly located in the visible/near-infrared (VIS-NIR), it has been a growing interest on using UAV platforms for high throughout phenotyping (Lobos et al. 2014). The main advantage is the capacity of precisely screening hundreds of plots in a short period of time (Araus and Cairns 2014). Development of novel sensor modules, imaging devices, automated systems, unmanned aerial vehicles (UAV), LED lightings and portable devices has been of paramount importance to achieve significant progress in plant phenotyping (Fiorani and Schurr 2013). These approaches are being used to estimate agronomic traits such as early vigour, plant height, biomass, nutritional status, and biotic and abiotic stress (De Vita and Taranto 2019). Most of the agronomic traits useful for breeding programs are estimated through the so-called vegetation indices (VIs). These indices have been proved useful with certain limitations (Royo and Villegas 2011), such as the environmental influence, the noise associated to the sensors, and the cost of the equipment. The most common VIs are based on calculations of different spectral wavelengths located in the visible (400-700nm) and near-infrared (700-1100nm). Among them, the most widely used VIs is the normalized difference vegetation index (NDVI) (Royo et al. 2003; Álvaro et al. 2008). Positive correlations between NDVI and green biomass have been reported in bread and durum wheat and barley (Bellairs et al. 1996; Peñuelas and Filella 1998; Royo et al. 2003; Casadesús et al. 2007; Darvishzadeh et al. 2008; Casadesús and Villegas 2014). Other indices have been developed to estimate crop's water status, as the water index (WI) (Peñuelas et al. 1993), the normalized water indices NWI1 to NWI4 (Babar et al. 2006; Prasad et al. 2007) or the crop water stress index (CWSI) (Gonzalez-Dugo et al. 2015). The photochemical reflectance index (PRI) has also been used to assess the RUE of the plants (Peñuelas et al. 1993; Garbulsky et al. 2011). Finally, there also exist some works of field phenotyping with hyperspectral imagery (Zarco-Tejada et al. 2012; Gonzalez-Dugo et al. 2015). These images, for instance, has the potential to provide much richer datasets by the collection of several narrow spectral bands sensitive to the absorption of specific photosynthetic pigments (Zarco-Tejada et al. 2012). Phenotyping through remote sensing technologies improves the collection of phenotypic data and facilitates the creation of repository databases useful for genetic analyses (Lippman et al. 2007; Welcker et al. 2011). In this regard several

authors have stressed out the interest of HTP, since the accurately measurements in agronomic traits is required in the field trials to gene mapping with high precision (Condorelli et al. 2018; Gizaw et al. 2018).

8. Genetics and genomics in wheat

Breeding from a solid scientific base began only after the rediscovery of Mendel's findings (Mendel 1866), at the beginning of the last century. Other advances took place gradually over the decades, until a major leap was made with the so-called "Green Revolution" of the mid-1960s, consisted in the development of semi-dwarf, photoperiod insensitive and high yielding modern cultivars (Venske et al. 2019). These new genotypes became widely adopted, especially in developing countries, and generated an impact on the reduction of hunger and poverty (Borlaug 2007). Classical breeding implies the selection of varieties carrying the desired characteristics for the target trait, usually morphological or visual features (Colasuonno et al. 2021). Even though genetic gains for yield and quality have been obtained in bread wheat (Sanchez-Garcia et al. 2015), several years are necessary to develop a commercial variety, and has some limitations, especially when target traits are highly dependent on the environment due to low heritability (Jackson et al. 1996). More recently, wheat breeders have substantially used the pedigree-based prediction of breeding values for the genetic improvement of complex traits (Crossa et al. 2006; Piepho et al. 2008).

The development of molecular biology, especially the invention of the polymerase chain reaction (PCR) (Mullis et al. 1986) allowed the use of markers based on polymorphisms in the DNA for the identification of traits enhancing the agronomic performance in the earlier stages of development. New efficient strategies to combine molecular markers with accelerated development of elite germplasm have been needed to fast-track development and delivery of improved germplasm. Marker-assisted selection (MAS) resulted in an important advantage integrated with traditional breeding methods to enhance the efficiency of cultivar development. This strategy depends on the genetic linkage of the trait of interest with molecular markers, and usually only major effect genetic loci are exploited in this way due to that only a limited proportion of genetic variance can be captured by the markers (Goddard and Hayes 2009). Since many of the agronomic traits present a multigenic quantitative nature and the effect of the environment on them needs to be assessed, MAS cannot replace traditional breeding methods for these traits, especially in later generation screening and cultivar evaluation.

To cover this gap the availability of high-density low-cost marker genotyping platforms has enabled a change in plant breeding, the genomic selection (GS). The GS refers to select genotypes using the genomic information on a genome wide

scale to make selection (Isidro et al. 2016). Genomic selection uses genome-wide markers to estimate the effects of all genes or chromosome positions simultaneously (Meuwissen et al. 2001) to predict the breeding values of progeny, which are used for selection of individuals without costly phenotyping, saving money and time, and increasing accuracy of selection.

8.1. Molecular markers, from gels to chips.

Molecular markers resolve many of the shortcomings of morphological and biochemical methods as they are not influenced by the environmental or developmental stage and can detect DNA-level variations (Ashraf et al. 2014). Molecular markers have been widely used for the construction of linkage maps, describing the position and relative genetic distance between markers and genes. The marker-based breeding started in the 1990s (Figure 1) with the first use of restriction fragment length polymorphism (RFLP) mainly for genetic mapping (Gale et al. 1990; Hart 1994) and genetic diversity analyses (Gupta and Varshney 2000). However, their use required DNA cleavage with restriction enzymes and hybridization with radioactive labelled probes making the process time consuming and reporting low genome coverage. Then, science focused on PCR-based markers like randomly amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) and microsatellites or simple sequence repeats (SSR). Among them, SSR markers, have been extensively used in wheat molecular breeding as they present a high level of polymorphism, have codominant inheritance, and are widely distributed in the genome (Rasheed and Xia 2019). In the last decade, with the rapid development of high-throughput genotyping platforms and the reduction of sequencing costs researchers changed the view to thousands of single nucleotide polymorphism (SNP) markers (Jaganathan et al. 2020). SNPs are based only in a difference one nucleotide in the DNA sequence and are the most extended polymorphisms in the genome.

Different technologies are nowadays available for high-throughput genotyping as the diversity array technology (DArT) (Jaccoud et al. 2001), genotyping by sequencing (GBS) (Li et al. 2015) and SNP arrays (Cabral et al. 2014). Several SNP arrays has been developed until now for wheat genotyping: the Illumina iSelect SNP arrays: 9K, 15K, 25K, 55K and 90K; the Axiom® arrays: Wheat Breeders' 35K, Wheat 660K, Wheat 820K SNP array; and the Wheat 50K Triticum TraitBreed array (reviewed in Colasuonno et al. 2021). SNPs have been used for the development of highly saturated linkage maps and consensus maps (Wang et al. 2014; Colasuonno et al. 2014), genetic diversity (Marcotuli et al. 2016), and structure analysis (Rufo et al. 2019). Because of the next generation sequencing (NGS) technologies, GBS maps containing 20 to 450 K loci have already been generated for wheat (Poland et al. 2012; Saintenac et al. 2013).

To individualise single markers from these arrays, a new technique appears, the kompetitive allele specific PCR or KASP markers (Ramirez-Gonzalez et al. 2015). This technique allowed the development of functional markers in wheat for several traits like plant height, grain weight and length, pre-harvest sprouting, biotic and abiotic stresses (Chandra et al. 2017).

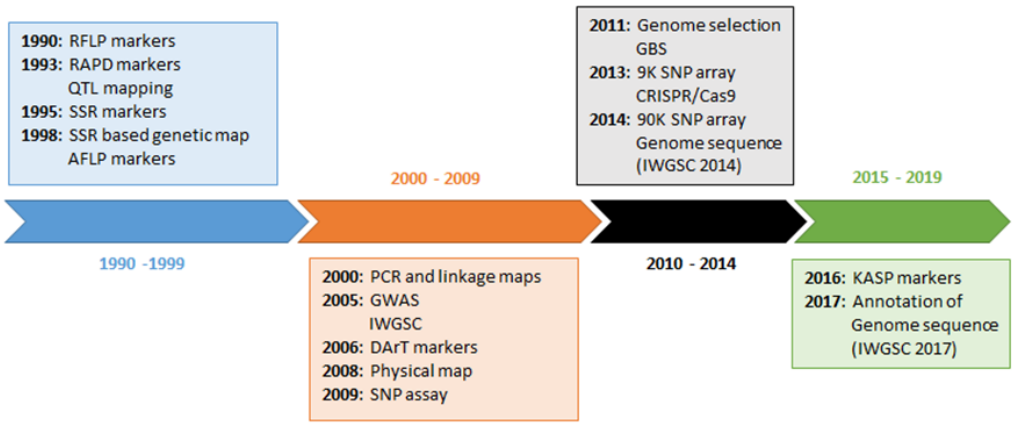

Figure 1. Timeline of markers and molecular techniques development in wheat. Adapted from Alotaibi et al. (2021).

8.2. Mapping quantitative traits

Quantitative trait loci (QTL) mapping is an approach used to determine the genetic architecture of complex traits, i.e. controlled by multiple genes. QTL mapping has been performed predominantly in populations derived from two known ancestors (biparental) with simple structure and different phenotypic performances. The most used populations in QTL analysis are F2, recombinant inbreed lines (RIL), double haploids (DH) and backcrosses (BC). The success in detecting QTL depends on the marker density, population size, and the heritability of the trait. Unfortunately, the identification of QTLs related to complex characters such as heat and drought is strongly affected by the low proportion of variance explained by most of the identified QTLs, and that are specific to a particular phenotyping environment or genetic background (De Vita and Taranto 2019). The genetic resolution of QTL mapping often remains in a range of 5-10 cM due to the limited number of recombination events derived from a biparental population and great efforts to develop new lines are needed to reduce mapping precision.

In order to provide broader allelic coverage and higher mapping resolution a complementary approach to QTL mapping was developed, the genome-wide association studies (GWAS) or association mapping (AM), based on linkage disequilibrium (LD). LD is defined as the non-random association of alleles at

different loci (Flint-Garcia et al. 2003). It is important to differentiate the LD due to physical linkage from LD due to population structure and genetic relatedness among genotypes. Thus, prior to carry out a GWAS, a population structure and diversity analysis must be performed to avoid spurious associations. The main differences among QTL mapping and GWAS are: 1) the type of germplasm used as GWAS is based in germplasm collections with many recombination events; 2) no need to construct a linkage map, as it can be used a consensus or reference map; 3) LD is not only due to linkage, but also to mutation, selection, inbreeding, genetic drift, and migration.

In the last year GWAS has been used to identify genomic regions related to numerous traits in wheat: disease resistance (Tomar et al. 2021; El Hanafi et al. 2021; Alemu et al. 2021), grain yield (Akram et al. 2021; Gao et al. 2021a; Hu et al. 2021), root traits (Rufo et al. 2020; Xu et al. 2021; Fatima et al. 2021), abiotic stress (Maulana et al. 2020; Abou-Elwafa and Shehzad 2021; Quan et al. 2021), kernel traits (Muhammad et al. 2020; Gao et al. 2021b), plant architecture (Muhammad et al. 2021), micronutrient concentration (Liu et al. 2021) and herbicide resistance (Shi et al. 2020), among others.

8.3. The release of the wheat genome sequence

The establishment of the International Wheat Genome Sequencing Consortium (IWGSC) took place in 2005 through the efforts to create a reference genome of wheat for the scientific community. The vision of the IWGSC was to provide a high-quality genome sequence of bread wheat to accelerate the development of improved varieties and to empower all aspects of basic and applied wheat research (https://www.wheatgenome.org). The first version of this sequence was published in 2014 for the cultivar Chinese Spring (IWGSC, 2014) and, since 2017, it was available and in continuous improvement (IWGSC, 2018a and 2018b) (http://wheat-urgi.versailles.inra.fr/). Nowadays a new sequencing project is ongoing, the 10+Wheat Genomes Project (http://www.10wheatgenomes.com/). It aims to characterize the wheat pan genome, developing strategies and resources to compare multiple wheat genome sequences and to identify genes that can be used for improving wheat production and quality (Venske et al. 2019).

The annotation of wheat genes opens the possibility for the development other bioinformatic tools for assisting wheat research, as the open databases of RNAseq experiments http://www.wheat-expression.com/ (Ramírez-González et al. 2018) that have made possible to carry out a candidate gene (CG) approach within QTL intervals without performing new functional studies.

9. The future of genetic improvement of bread wheat

The expected effects of climate change and the declining availability of water will require the release of cultivars with an enhanced genetic capacity to maintain acceptable yield levels and yield stability under harmful environmental conditions. The new high-throughput genotyping and phenotyping technologies will facilitate the characterization and utilization of underexploited germplasm as wild relatives or landraces (Wang et al. 2017). The success of hybrid wheat breeding depends on the clustering of suitable germplasm into heterotic groups and on the identification of a high-yielding heterotic pattern identification (Kempe et al. 2014). Heterotic groups are sets of genotypes displaying similar hybrid performance when crossed with individuals from another, genetically distinct germplasm group, and genetic variation within them improves the heterotic pattern simplifying the identification of superior single crosses (Zhao et al. 2015).The development of **speed breeding** has helped to accelerate the plant improvement, saving time for breeders, reducing the length of the breeding cycles, and selecting the best genotypes in combination with MAS. It is based on photoperiod and temperature manipulation in growth chambers and glasshouses, and as reported by Ghosh et al. (2018) it is possible to achieve up to six generations per year for spring wheat. Genome or **gene editing** can accurately target segments of the genome for modification, either by deletion, insertion or substitution of nucleotides (Sander and Joung 2014). Despite the complexity of the wheat genome and the difficulty in genetic transformation, recent progress in plant genome editing using the CRISPR/Cas9 system has been reported for bread wheat (Wang et al. 2014).

Another genome-based approach that is taking relevance in the last years is the **genome selection (GS)** or prediction. The availability of high-density and low-cost marker genotyping systems made possible a change in plant breeding. The selection of genotypes is done using genomic information on a genome-wide scale (Isidro et al. 2016). The effects of all genes for a trait are estimated simultaneously to predict the breeding value of the progenies, thus they are selected without phenotyping costs.

Overall, efforts in the development and implementation of improved strategies must continue in wheat breeding programmes to face the increasing demand for food by an ever-growing world population in a constant scenario of climate change.

10. References

Abou-Elwafa SF, Shehzad T (2021) Genetic diversity, GWAS and prediction for drought and terminal heat stress tolerance in bread wheat (Triticum aestivum L.). Genet Resour Crop Evol 68:711–728. https://doi.org/10.1007/s10722-020-01018-y

Adamsen FJ, Pinter PJ, Barnes EM, et al (1999) Measuring Wheat Senescence with a Digital Camera. Crop Sci 39:719–724. https://doi.org/10.2135/cropsci1999.0011183X003900030019x

Akram S, Arif MAR, Hameed A (2021) A GBS-based GWAS analysis of adaptability and yield traits in bread wheat (Triticum aestivum L.). J Appl Genet 62:27–41. https://doi.org/10.1007/s13353-020-00593-1

Alemu A, Brazauskas G, Gaikpa DS, et al (2021) Genome-Wide Association Analysis and Genomic Prediction for Adult-Plant Resistance to Septoria Tritici Blotch and Powdery Mildew in Winter Wheat. Front Genet 12:661742. https://doi.org/10.3389/fgene.2021.661742

Alotaibi F, Alharbi S, Alotaibi M, et al (2021) Wheat omics: Classical breeding to new breeding technologies. Saudi J Biol Sci 28:1433–1444. https://doi.org/10.1016/j.sjbs.2020.11.083

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

Álvaro F, Royo C, García Del Moral LF, Villegas D (2008) Grain filling and dry matter translocation responses to source-sink modifications in a historical series of durum wheat. Crop Sci 48:1523–1531. https://doi.org/10.2135/cropsci2007.10.0545

Anjum SA, Xie X-Y, Wang L-C, et al (2011) Morphological, physiological and biochemical responses of plants to drought stress. African J Agric Res 6:2026–2032. https://doi.org/10.5897/AJAR10.027

Aparicio N, Villegas D, Casadesus J, et al (2000) Spectral vegetation indices as nondestructive tools for determining durum wheat yield. Agron J 92:83–91. https://doi.org/10.2134/agronj2000.92183x

Araus JL, Bort J, Steduto P, et al (2003) Breeding cereals for Mediterranean conditions: ecophysiological clues for biotechnology application. Ann Appl Biol 142:129–141. https://doi.org/10.1111/j.1744-7348.2003.tb00238.x

Araus JL, Cairns JE (2014) Field high-throughput phenotyping: The new crop breeding frontier. Trends Plant Sci. 19:52–61

Araus JL, Kefauver SC, Zaman-Allah M, et al (2018) Translating High-Throughput Phenotyping into Genetic Gain. Trends Plant Sci. 23:451–466

Araus JL, Slafer GA, Royo C, Serret MD (2008) Breeding for yield potential and stress adaptation in cereals. CRC. Crit. Rev. Plant Sci. 27:377–412

Ashraf K, Ahmad A, Chaudhary A, et al (2014) Genetic diversity analysis of Zingiber officinale Roscoe by RAPD collected from subcontinent of India. Saudi J Biol Sci 21:159–165. https://doi.org/10.1016/j.sjbs.2013.09.005

Austin RB, Morgan CL, Ford MA, Blackwell RD (1980) Contributions to grain yield from pre-anthesis assimilation in tall and dwarf barley phenotypes in two contrasting seasons. Ann Bot 45:309–319. https://doi.org/10.1093/oxfordjournals.aob.a085826

Babar MA, Reynolds MP, Van Ginkel M, et al (2006) Spectral reflectance indices as a potential indirect selection criteria for wheat yield under irrigation. Crop Sci 46:578–588. https://doi.org/10.2135/cropsci2005.0059

Baker B, Olszyk DM, Tingey D (1996) Digital image analysis to estimate leaf area. J Plant Physiol 148:530–535. https://doi.org/10.1016/S0176-1617(96)80072-1

Bellairs SM, Turner NC, Hick PT, Smith RCG (1996) Plant and soil influences on estimating biomass of wheat in plant breeding plots using field spectral radiometers. Aust J Agric Res 47:1017–1034. https://doi.org/10.1071/AR9961017

Bellvert J, Nieto H, Pelechá A, et al (2021) Remote Sensing Energy Balance Model for the Assessment of Crop Evapotranspiration and Water Status in an Almond Rootstock Collection. Front Plant Sci 12:288. https://doi.org/10.3389/fpls.2021.608967

Biesiekierski JR (2017) No TitleWhat is gluten? J Gastroenterol Hepatol 32:78–81

Borlaug N (2007) Feeding a Hungry world. Science (80-.). 318:359

Brown TA, Jones MK, Powell W, Allaby RG (2009) The complex origins of domesticated crops in the Fertile Crescent. Trends Ecol. Evol. 24:103–109

Cabral AL, Jordan MC, McCartney CA, et al (2014) Identification of candidate genes, regions and markers for pre-harvest sprouting resistance in wheat (Triticum aestivum L.). BMC Plant Biol 14:1–12. https://doi.org/10.1186/s12870-014-0340-1

Campillo C, Prieto MH, Daza C, et al (2008) Using digital images to characterize canopy coverage and light interception in a processing tomato crop. HortScience 43:1780–1786. https://doi.org/10.21273/hortsci.43.6.1780

Casadesús J, Kaya Y, Bort J, et al (2007) Using vegetation indices derived from conventional digital cameras as selection criteria for wheat breeding in water-limited environments. Ann Appl Biol 150:227–236. https://doi.org/10.1111/j.1744-7348.2007.00116.x

Casadesús J, Villegas D (2014) Conventional digital cameras as a tool for assessing leaf area index and biomass for cereal breeding. J Integr Plant Biol 56:7–14. https://doi.org/10.1111/jipb.12117

Chandra S, Singh D, Pathak J, et al (2017) SNP discovery from next-generation transcriptome sequencing data and their validation using KASP assay in wheat (Triticum aestivum L.). Mol Breed 37:1–15. https://doi.org/10.1007/s11032-017-0696-7

Chmielewski FM, Köhn W (2000) Impact of weather on yield components of winter rye over 30 years. Agric For Meteorol 102:253–261. https://doi.org/10.1016/S0168-1923(00)00125-8

Colasuonno P, Gadaleta A, Giancaspro A, et al (2014) Development of a high-density SNP-based linkage map and detection of yellow pigment content QTLs in durum wheat. Mol Breed 34:1563–1578. https://doi.org/10.1007/s11032-014-0183-3

Colasuonno P, Marcotuli I, Gadaleta A, Soriano JM (2021) From genetic maps to qtl cloning: An overview for durum wheat. Plants 10:1–25. https://doi.org/10.3390/plants10020315

Condorelli GE, Maccaferri M, Newcomb M, et al (2018) Comparative Aerial and Ground Based High Throughput Phenotyping for the Genetic Dissection of NDVI as a Proxy for Drought Adaptive Traits in Durum Wheat. Front Plant Sci 9:893. https://doi.org/10.3389/fpls.2018.00893

Crespo-Herrera LA, Crossa J, Huerta-Espino J, et al (2018) Genetic Gains for Grain Yield in CIMMYT's Semi-Arid Wheat Yield Trials Grown in Suboptimal Environments. Crop Sci 58:1890–1898. https://doi.org/10.2135/cropsci2018.01.0017

Crossa J, Burgueño J, Cornelius PL, et al (2006) Modeling genotype × environment interaction using additive genetic covariances of relatives for predicting breeding values of wheat genotypes. Crop Sci 46:1722–1733. https://doi.org/10.2135/cropsci2005.11-0427

Darvishzadeh R, Skidmore A, Schlerf M, Atzberger C (2008) Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in a heterogeneous grassland. Remote Sens Environ 112:2592–2604. https://doi.org/10.1016/j.rse.2007.12.003

Daryanto S, Wang L, Jacinthe PA (2016) Global synthesis of drought effects on maize and wheat production. PLoS One 11:e0156362. https://doi.org/10.1371/journal.pone.0156362

De Vita P, Taranto F (2019) Durum wheat (Triticum turgidum ssp. durum) breeding to meet the challenge of climate change. In: Advances in Plant Breeding Strategies: Cereals. Springer International Publishing, pp 471–524

Duan T, Chapman SC, Guo Y, Zheng B (2017) Dynamic monitoring of NDVI in wheat agronomy and breeding trials using an unmanned aerial vehicle. F Crop Res 210:71–80. https://doi.org/10.1016/j.fcr.2017.05.025

Dvorak J, Luo M-C, Akhunov ED (2010) N.I. Vavilov's Theory of Centres of Diversity in the Light of Current Understanding of Wheat Diversity, Domestication and Evolution. Czech J Genet Plant Breed 47:20–27

El Hanafi S, Backhaus A, Bendaou N, et al (2021) Genome-wide association study for adult plant resistance to yellow rust in spring bread wheat (Triticum aestivum L.). Euphytica 217:1–14. https://doi.org/10.1007/s10681-021-02803-1

Evans LT (1993) Crop Evolution, Adaptation and Yield. Cambridge Univ Press

Fatima I, Gao Y, Xu X, et al (2021) Genome-Wide Association Mapping of Seedling Biomass and Root Traits Under Different Water Conditions in Wheat. Front Genet 12:663557. https://doi.org/10.3389/fgene.2021.663557

Feldman M (2001) Origin of cultivated wheat. World Wheat Book, a Hist Wheat Breed 3–56

Fiorani F, Schurr U (2013) Future scenarios for plant phenotyping. Annu Rev Plant Biol 64:267–291. https://doi.org/10.1146/annurev- arplant-050312-120137

Fischer RAT, Edmeades GO (2010) Breeding and Cereal Yield Progress. Crop Sci 50:S-85-S-98. https://doi.org/10.2135/cropsci2009.10.0564

Flint-Garcia SA, Thornsberry JM, Edward IV SB (2003) Structure of Linkage Disequilibrium in Plants. Annu. Rev. Plant Biol. 54:357–374

Gale MD, Chao S, Sharp PJ (1990) RFLP Mapping in Wheat — Progress and Problems. Gene Manip Plant Improv II 353–363. https://doi.org/10.1007/978-1-4684-7047-5_19

Gao L, Meng C, Yi T, et al (2021a) Genome-wide association study reveals the genetic basis of yield- and quality-related traits in wheat. BMC Plant Biol 21:1–11. https://doi.org/10.1186/s12870-021-02925-7

Gao Y, Xu X, Jin J, et al (2021b) Dissecting the genetic basis of grain morphology traits in Chinese wheat by genome wide association study. Euphytica 217:1–12. https://doi.org/10.1007/s10681-021-02795-y

Garbulsky MF, Peñuelas J, Gamon J, et al (2011) The photochemical reflectance index (PRI) and the remote sensing of leaf, canopy and ecosystem radiation use efficiencies. A review and meta-analysis. Remote Sens. Environ. 115:281–297

García Del Moral LF, Rharrabti Y, Elhani S, et al (2005) Yield formation in Mediterranean durum wheats under two contrasting water regimes based on path-coefficient analysis. Euphytica 146:203–212. https://doi.org/10.1007/s10681-005-9006-2

Ghosh S, Watson A, Gonzalez-Navarro OE, et al (2018) Speed breeding in growth chambers and glasshouses for crop breeding and model plant research. Nat Protoc 13:2944–2963. https://doi.org/10.1038/s41596-018-0072-z

Giles RJ, Brown TA (2006) GluDy allele variations in Aegilops tauschii and Triticum aestivum: implications for the origins of hexaploid wheats. Theor Appl Genet 112:1563–1572. https://doi.org/10.1007/s00122-006-0259-5

Giorgi F, Lionello P (2008) Climate change projections for the Mediterranean region. Glob Planet Change 63:90–104. https://doi.org/10.1016/j.gloplacha.2007.09.005

Giunta F, De Vita P, Mastrangelo AM, et al (2018) Environmental and genetic variation for yield-related traits of durum wheat as affected by development. Front Plant Sci 9:8. https://doi.org/10.3389/fpls.2018.00008

Giunta F, Motzo R, Deidda M (1993) Effect of drought on yield and yield components of durum wheat and triticale in a Mediterranean environment. F Crop Res 33:399–409. https://doi.org/10.1016/0378-4290(93)90161-F

Gizaw SA, Godoy JG V., Pumphrey MO, Carter AH (2018) Spectral Reflectance for Indirect Selection and Genome-Wide Association Analyses of Grain Yield and Drought Tolerance in North American Spring Wheat. Crop Sci 58:2289–2301. https://doi.org/10.2135/cropsci2017.11.0690

Goddard ME, Hayes BJ (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat. Rev. Genet. 10:381–391

Gonzalez-Dugo V, Hernandez P, Solis I, Zarco-Tejada PJ (2015) Using high-resolution hyperspectral and thermal airborne imagery to assess physiological condition in the context of wheat phenotyping. Remote Sens 7:13586–13605. https://doi.org/10.3390/rs71013586

Gracia-Romero A, Kefauver SC, Vergara-Díaz O, et al (2017) Comparative Performance of Ground vs. Aerially Assessed RGB and Multispectral Indices for Early-Growth Evaluation of Maize Performance under Phosphorus Fertilization. Front Plant Sci 8:2004. https://doi.org/10.3389/fpls.2017.02004

Gupta PK, Varshney RK (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica 113:163–185. https://doi.org/10.1023/A:1003910819967

Habash DZ, Kehel Z, Nachit M (2009) Genomic approaches for designing durum wheat ready for climate change with a focus on drought. In: Journal of Experimental Botany. Oxford University Press, pp 2805–2815

Haboudane D, Miller JR, Pattey E, et al (2004) Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies : Modeling and validation in the context of precision agriculture. Remote Sens 90:337–352. https://doi.org/10.1016/j.rse.2003.12.013

Hart GE (1994) RFLP maps of bread wheat. DNA-based markers in plants 327–358. https://doi.org/10.1007/978-94-011-1104-1_21

Haudry A, Cenci A, Ravel C, et al (2007) Grinding up wheat: A massive loss of nucleotide diversity since domestication. Mol Biol Evol 24:1506–1517. https://doi.org/10.1093/molbev/msm077

Helbaek H. (1959) Domestication of food plants in the old world. Domest food plants old world 130:365–372. https://doi.org/10.1126/science.130.3372.365

Hu P, Zheng Q, Luo Q, et al (2021) Genome-wide association study of yield and related traits in common wheat under salt-stress conditions. BMC Plant Biol 21:1–20. https://doi.org/10.1186/s12870-020-02799-1

Ide R, Oguma H (2010) Use of digital cameras for phenological observations. Ecol Inform 5:339–347. https://doi.org/10.1016/j.ecoinf.2010.07.002

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

Igrejas G, Branlard G (2020) The importance of wheat. In: Wheat Quality For Improving Processing And Human Health. Springer International Publishing, pp 1–7

Isidro J, Akdemir D, Burke J (2016) Genomic selection. A Hist Wheat Breed 3:1001–1023

Jaccoud D, Peng K, Feinstein D, Kilian A (2001) Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Res 29:25. https://doi.org/10.1093/nar/29.4.e25

Jackson P, Robertson M, Cooper M, Hammer G (1996) The role of physiological understanding in plant breeding; From a breeding perspective. F Crop Res 49:11–37. https://doi.org/10.1016/S0378-4290(96)01012-X

Jaganathan D, Bohra A, Thudi M, Varshney RK (2020) Fine mapping and gene cloning in the post-NGS era: advances and prospects. Theor. Appl. Genet. 133:1791–1810

Kempe K, Rubtsova M, Gils M (2014) Split-gene system for hybrid wheat seed production. Proc Natl Acad Sci U S A 111:9097–9102. https://doi.org/10.1073/pnas.1402836111

Keyvan S (2010) The effects of drought stress on yield, relative water content, proline, soluble carbohydrates and chlorophyll of bread wheat cultivars

Kulkarni M, Soolanayakanahally R, Ogawa S, et al (2017) Drought response in wheat: Key genes and regulatory mechanisms controlling root system architecture and transpiration efficiency. Front. Chem. 5:106

Kumudini S, Andrade FH, Boote KJ, et al (2014) Predicting maize phenology: Intercomparison of functions for developmental response to temperature. Agron J 106:2087–2097. https://doi.org/10.2134/agronj14.0200

Leegood RC, Evans JR, Furbank RT (2010) Food security requires genetic advances to increase farm yields. Nature 464:831

Li G, Wang Y, Chen MS, et al (2015) Precisely mapping a major gene conferring resistance to Hessian fly in bread wheat using genotyping-by-sequencing. BMC Genomics 16:1–10. https://doi.org/10.1186/s12864-015-1297-7

Li XB, Chen YH, Yang H, Zhang YX (2005) Improvement, comparison, and application of field measurement methods for grassland vegetation fractional coverage. J Integr Plant Biol 47:1074–1083. https://doi.org/10.1111/j.1744-7909.2005.00134.x

Lippman ZB, Semel Y, Zamir D (2007) An integrated view of quantitative trait variation using tomato interspecific introgression lines. Curr. Opin. Genet. Dev. 17:545–552

Liu J, Huang L, Li T, et al (2021) Genome-Wide Association Study for Grain Micronutrient Concentrations in Wheat Advanced Lines Derived From Wild Emmer. Front Plant Sci 12:651283. https://doi.org/10.3389/fpls.2021.651283

Lobos GA, Matus I, Rodriguez A, et al (2014) Wheat genotypic variability in grain yield and carbon isotope discrimination under Mediterranean conditions assessed by spectral reflectance. J Integr Plant Biol 56:470–479. https://doi.org/10.1111/jipb.12114

Lopes MS, El-Basyoni I, Baenziger PS, et al (2015) Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. J Exp Bot 66:3477–3486. https://doi.org/10.1093/jxb/erv122

Lukina E V., Stone ML, Raun WR (1999) Estimating vegetation coverage in wheat using digital images. J Plant Nutr 22:341–350. https://doi.org/10.1080/01904169909365631

Luo MC, Yang ZL, You FM, et al (2007) The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication. Theor Appl Genet 114:947–959. https://doi.org/10.1007/s00122-006-0474-0

Mac Key J (2005) Wheat: Its Concept, Evolution, and Taxonomy. Durum Wheat Breed 35–94. https://doi.org/10.1201/9781482277883-9

Marcotuli I, Houston K, Schwerdt JG, et al (2016) Genetic diversity and genome wide association study of β-glucan content in tetraploid wheat grains. PLoS One 11:e0152590. https://doi.org/10.1371/journal.pone.0152590

Mastrangelo AM, Belloni S, Barilli S, et al (2005) Low temperature promotes intron retention in two e-cor genes of durum wheat. Planta 221:705–715. https://doi.org/10.1007/s00425-004-1475-3

Mastrangelo AM, Cattivelli L (2021) What Makes Bread and Durum Wheat Different? Trends Plant Sci 1–8. https://doi.org/10.1016/j.tplants.2021.01.004

Maulana F, Huang W, Anderson J, Ma X (2020) Genome-Wide Association Mapping of Seedling Drought Tolerance in Winter Wheat. Front Plant Sci 11:573786. https://doi.org/10.3389/fpls.2020.573786

Mendel G (1866) Versuche über pflanzen-hybriden. Verh Naturforsch Ver Brünn 4:3–47

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Moragues M, García Del Moral LF, Moralejo M, Royo C (2006) Yield formation strategies of durum wheat landraces with distinct pattern of dispersal within the Mediterranean basin: II. Biomass production and allocation. F Crop Res 95:182–193. https://doi.org/10.1016/j.fcr.2005.02.008

Muhammad A, Hu W, Li Z, et al (2020) Appraising the genetic architecture of Kernel traits in hexaploid wheat using GWAS. Int J Mol Sci 21:1–21. https://doi.org/10.3390/ijms21165649

Muhammad A, Li J, Hu W, et al (2021) Uncovering genomic regions controlling plant architectural traits in hexaploid wheat using different GWAS models. Sci Rep 11:6767. https://doi.org/10.1038/s41598-021-86127-z

Mullis K, Faloona F, Scharf S, et al (1986) Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. Cold Spring Harb Symp Quant Biol 51:263–273. https://doi.org/10.1101/sqb.1986.051.01.032

Munné-Bosch S, Alegre L (2004) Die and let live: Leaf senescence contributes to plant survival under drought stress. Funct. Plant Biol. 31:203–216

Narayanan S, Vara Prasad P V. (2014) Characterization of a spring wheat association mapping panel for root traits. Agron J 106:1593–1604. https://doi.org/10.2134/agronj14.0015

Nevo E (2009) Ecological genomics of natural plant populations: The Israeli perspective. Methods Mol Biol 513:321–344. https://doi.org/10.1007/978-1-59745-427-8_17

Nevo E (2011) Triticum. In: Wild Crop Relatives: Genomic and Breeding Resources. Springer Berlin Heidelberg, pp 407–456

Pascual L, Ruiz M, López-Fernández M, et al (2020) Genomic analysis of Spanish wheat landraces reveals their variability and potential for breeding. BMC Genomics 21:1–17. https://doi.org/10.1186/s12864-020-6536-x

Peña-Bautista RJ, Hernández-Espinosa N, Jones JN, et al (2017) Wheat-Based Foods: Their Global and Regional Importance in the Food Supply, Nutrition, and Health. Cereal Foods World 62:231–249. https://doi.org/10.1094/CFW-62-5-0231

Peng JH, Sun D, Nevo E (2011) Domestication evolution, genetics and genomics in wheat. Mol Breed 28:281–301. https://doi.org/10.1007/s11032-011-9608-4

Peñuelas J, Filella I, Biel C, et al (1993) The reflectance at the 950-970 nm region as an indicator of plant water status. Int J Remote Sens 14:1887–1905. https://doi.org/10.1080/01431169308954010

Peñuelas J, Filella L (1998) Technical focus: Visible and near-infrared reflectance techniques for diagnosing plant physiological status. Trends Plant Sci 3:151–156. https://doi.org/10.1016/S1360-1385(98)01213-8

Piepho HP, Möhring J, Melchinger AE, Büchse A (2008) BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161:209–228

Poland JA, Brown PJ, Sorrells ME, Jannink JL (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One 7:32253. https://doi.org/10.1371/journal.pone.0032253

Prasad B, Carver BF, Stone ML, et al (2007) Genetic analysis of indirect selection for winter wheat grain yield using spectral reflectance indices. Crop Sci 47:1416–1425. https://doi.org/10.2135/cropsci2006.08.0546

Quan X, Liu J, Zhang N, et al (2021) Genome-Wide Association Study Uncover the Genetic Architecture of Salt Tolerance-Related Traits in Common Wheat (Triticum aestivum L.). Front Genet 12:563. https://doi.org/10.3389/fgene.2021.663941

Ramírez-González RH, Borrill P, Lang D, et al (2018) The transcriptional landscape of polyploid wheat. Science (80-) 361:. https://doi.org/10.1126/science.aar6089

Ramirez-Gonzalez RH, Uauy C, Caccamo M (2015) PolyMarker: A fast polyploid primer design pipeline. Bioinformatics 31:2038–2039. https://doi.org/10.1093/bioinformatics/btv069

Rasheed A, Xia X (2019) From markers to genome-based breeding in wheat. Theor Appl Genet 132:767–784. https://doi.org/10.1007/s00122-019-03286-4

Rebetzke GJ, Richards RA, Fettell NA, et al (2007) Genotypic increases in coleoptile length improves stand establishment, vigour and grain yield of deep-sown wheat. F Crop Res 100:10–23. https://doi.org/10.1016/j.fcr.2006.05.001

Reynolds M, Foulkes J, Furbank R, et al (2012) Achieving yield gains in wheat. Plant, Cell Environ 35:1799–1823. https://doi.org/10.1111/j.1365-3040.2012.02588.x

Reynolds M, Manes Y, Izanloo A, Langridge P (2009) Phenotyping approaches for physiological breeding and gene discovery in wheat. Ann. Appl. Biol. 155:309–320

Rharrabti Y, Villegas D, Royo C, et al (2003) Durum wheat quality in Mediterranean environments II. Influence of climatic variables and relationships between quality parameters. F Crop Res 80:133–140. https://doi.org/10.1016/S0378-4290(02)00177-6

Richards RA (2006) Physiological traits used in the breeding of new cultivars for water-scarce environments. In: Agricultural Water Management. Elsevier, pp 197–211

Royo C, Álvaro F, Martos V, et al (2007) Genetic changes in durum wheat yield components and associated traits in Italian and Spanish varieties during the 20th century. Euphytica 155:259–270. https://doi.org/10.1007/s10681-006-9327-9

Royo C, Aparicio N, Villegas D, et al (2003) Usefulness of spectral reflectance indices as durum wheat yield predictors under contrasting Mediterranean conditions. Int J Remote Sens 24:4403–4419. https://doi.org/10.1080/0143116031000150059

Royo C, Dreisigacker S, Ammar K, Villegas D (2020) Agronomic performance of durum wheat landraces and modern cultivars and its association with genotypic variation in vernalization response (Vrn-1) and photoperiod sensitivity (Ppd-1) genes. Eur J Agron 120:126129. https://doi.org/10.1016/j.eja.2020.126129

Royo C, Elias EM, Manthey FA (2009) Durum Wheat Breeding. In: Cereals. Springer US, pp 199–226

Royo C, García del Moral LF, Slafer G, et al (2005) Selection tools for improving yield-associated physiological traits, In: Durum Wheat Breeding: Current Approaches and Future Strategies. In: Food Products Press. pp 563–598

Royo C, Nazco R, Villegas D (2014) The climate of the zone of origin of Mediterranean durum wheat (Triticum durum Desf.) landraces affects their agronomic performance. Genet Resour Crop Evol 61:1345–1358. https://doi.org/10.1007/s10722-014-0116-3

Royo C, Soriano JM, Alvaro F (2017) Wheat: A Crop in the Bottom of the Mediterranean Diet Pyramid. Intech 16:381–399

1. Royo C, Villegas D (2011) Field measurements of canopy spectra for biomass assessment of small-grain cereals. Biomass-Detection, Prod Usage. https://doi.org/10.5772/17745

Rufo R, Alvaro F, Royo C, Soriano JM (2019) From landraces to improved cultivars: Assessment of genetic diversity and population structure of Mediterranean wheat using SNP markers. PLoS One 14:. https://doi.org/10.1371/journal.pone.0219867

Rufo R, Salvi S, Royo C, Soriano J (2020) Exploring the Genetic Architecture of Root-Related Traits in Mediterranean Bread Wheat Landraces by Genome-Wide Association Analysis. Agronomy 10:613. https://doi.org/10.3390/agronomy10050613

Rufo R, Soriano JM, Villegas D, et al (2021) Using Unmanned Aerial Vehicle and Ground-Based RGB Indices to Assess Agronomic Performance of Wheat Landraces and Cultivars in a Mediterranean-Type Environment. Remote Sens 13:1–18

Ruiz M, Giraldo P, Royo C, Carrillo JM (2013) Creation and Validation of the Spanish Durum Wheat Core Collection. Crop Sci 53:2530–2537. https://doi.org/10.2135/cropsci2013.04.0238

Sadras VO, Rodriguez D (2007) The limit to wheat water-use efficiency in eastern Australia. II. Influence of rainfall patterns. Aust J Agric Res 58:657–669. https://doi.org/10.1071/AR06376

Saintenac C, Jiang D, Wang S, Akhunov E (2013) Sequence-based mapping of the polyploid wheat genome. G3 Genes, Genomes, Genet 3:1105–1114. https://doi.org/10.1534/g3.113.005819

Sanchez-Garcia M, Álvaro F, Peremarti A, et al (2015) Breeding effects on dry matter accumulation and partitioning in Spanish bread wheat during the 20th century. Euphytica 203:321–336. https://doi.org/10.1007/s10681-014-1268-0

Sander JD, Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. 32:347–350

Shi C, Zheng Y, Geng J, et al (2020) Identification of herbicide resistance loci using a genome-wide association study and linkage mapping in Chinese common wheat. Crop J 8:666–675. https://doi.org/10.1016/j.cj.2020.02.004

Skovmand B, Warburton M, Sullivan S, Lage J (2005) Managing and collecting genetic resources. New York Haworth Press 142–163

Slafer GA, Araus JL, Royo C, García Del Moral LF (2005) Promising eco-physiological traits for genetic improvement of cereal yields in Mediterranean environments. Ann Appl Biol 146:61–70. https://doi.org/10.1111/j.1744-7348.2005.04048.x

Snape J, Butterworth K, Whitechurch E, Worland AJ (2001) Waiting for Fine Times: Genetics of Flowering Time in Wheat. Wheat a Glob Environ 67–74. https://doi.org/10.1007/978-94-017-3674-9_7

Soriano JM, Villegas D, Aranzana MJ, et al (2016) Genetic structure of modern durum wheat cultivars and mediterranean landraces matches with their agronomic performance. PLoS One 11:e0160983. https://doi.org/10.1371/journal.pone.0160983

Soriano JM, Villegas D, Sorrells ME, Royo C (2018) Durum wheat landraces from east and west regions of the mediterranean basin are genetically distinct for yield components and phenology. Front Plant Sci 9:1–9. https://doi.org/10.3389/fpls.2018.00080

Tomar V, Singh D, Dhillon GS, et al (2021) New QTLs for Spot Blotch Disease Resistance in Wheat (Triticum aestivum L.) Using Genome-Wide Association Mapping. Front Genet 11:1740. https://doi.org/10.3389/fgene.2020.613217

Tuberosa R (2012) Phenotyping for drought tolerance of crops in the genomics era. Front Physiol 3 SEP:1–26. https://doi.org/10.3389/fphys.2012.00347

Venske E, Dos Santos RS, Busanello C, et al (2019) Bread wheat: a role model for plant domestication and breeding. Hereditas 156:16

Wang S-X, Zhu Y-L, Zhang D-X, et al (2017) Genome-wide association study for grain yield and related traits in elite wheat varieties and advanced lines using SNP markers. PLoS One 12:e0188662. https://doi.org/10.1371/journal.pone.0188662

Wang S, Wong D, Forrest K, et al (2014) Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. Plant Biotechnol J 12:787–796. https://doi.org/10.1111/pbi.12183

Welcker C, Sadok W, Dignat G, et al (2011) A common genetic determinism for sensitivities to soil water deficit and evaporative demand: Meta-analysis of quantitative trait loci and introgression lines of maize. Plant Physiol 157:718–729. https://doi.org/10.1104/pp.111.176479

White JW, Andrade-Sanchez P, Gore MA, et al (2012) Field-based phenomics for plant genetics research. F Crop Res 133:101–112. https://doi.org/10.1016/j.fcr.2012.04.003

Worland A, Korzun W, Li M, et al (1998) The influence of photoperiod genes on the adaptability of European winter wheats. Euphytica 100:385–394. https://doi.org/https://doi.org/10.1023/A:1018327700985

Xie C, Yang C (2020) A review on plant high-throughput phenotyping traits using UAV-based sensors. Comput. Electron. Agric. 178:105731

Xu F, Chen S, Yang X, et al (2021) Genome-Wide Association Study on Seminal and Nodal Roots of Wheat Under Different Growth Environments. Front Plant Sci 11:602399. https://doi.org/10.3389/fpls.2020.602399

Zaharieva M, Monneveux P (2006) Spontaneous Hybridization between Bread Wheat (Triticum aestivum L.) and Its Wild Relatives in Europe. Crop Sci 46:512–527. https://doi.org/10.2135/cropsci2005.0023

Zarco-Tejada PJ, González-Dugo V, Berni JAJ (2012) Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. Remote Sens Environ 117:322–337. https://doi.org/10.1016/j.rse.2011.10.007

Zeven AC (1998) Landraces: A review of definitions and classifications. Euphytica 104:127–139. https://doi.org/10.1023/A:1018683119237

Zhao Y, Li Z, Liu G, et al (2015) Genome-based establishment of a high-yielding heterotic pattern for hybrid wheat breeding. Proc Natl Acad Sci U S A 112:15624–15629. https://doi.org/10.1073/pnas.1514547112

Zohary D, Hopf M (2000) Domestication of plants in the Old World: the origin and spread of cultivated plants in West Asia, Europe and the Nile Valley. Domest plants Old World Orig spread Cultiv plants West Asia, Eur Nile Val

Objectives

Objectives

The main objective of this PhD Thesis was to provide scientific knowledge for the wheat research community and useful tools for the breeding programs to help the development of new varieties with enhanced yield to be grown in a Mediterranean environment in a climate change scenario. To achieve this general objective, specific aims were defined in this PhD thesis:

1. To explore the existence of genetic, phenotypic and/or geographic structures in the germplasm collections used, exploring the gene exchange among subpopulations.

2. To assess the feasibility of using spectral vegetation indices (VIs) from multispectral images installed on-board unmanned air vehicles and from ground-based RGB images to evaluate agronomic traits and to estimate grain yield under rainfed environments.

3. To study the genetic architecture of seminal root traits in the bread wheat Mediterranean landraces and to identify associated molecular markers for wheat improvement.

4. To explore the genetic variability of the wheat QTLome among Mediterranean landraces and modern cultivars for yield related traits and vegetation indices to identify the genomic regions most involved in trait variation for breeding purposes.

The plant material used was the MED6WHEAT IRTA-panel, consisted of a germplasm collection of 354 bread wheat genotypes, of which 170 were landraces from 24 Mediterranean countries and 184 modern varieties cultivated in 19 countries in the region. This PhD thesis is structured in four chapters written as scientific articles. **Chapter 1** is published in PLoS ONE (2019), **Chapter 2** in Agronomy (2020), **Chapter 3** in Remote Sensing (2021) and **Chapter 4** in Frontiers in Plant Science (2021). The multidisciplinary expertise needed to achieve the objective in **Chapter 3** was obtained through a five months' research stay in the Department of Agricultural and Food Sciences at Bologna University in Italy.

Chapter 1:

From landraces to improved cultivars: Assessment of genetic diversity and population structure of Mediterranean wheat using SNP markers

Rubén Rufo, Fanny Alvaro, Conxita Royo, Jose Miguel Soriano

Sustainable Field Crops Programme, Institute for Food and Agricultural Research and Technology (IRTA), Lleida, Spain

Published in PLoS ONE (2019) 14(7): e0219867

doi: org/10.1371/journal.pone.0219867

From landraces to improved cultivars: Assessment of genetic diversity and population structure of Mediterranean wheat using SNP markers

1. Abstract

Assessment of genetic diversity and population structure in crops is essential for breeding and germplasm conservation. A collection of 354 bread wheat genotypes, including Mediterranean landraces and modern cultivars representative of the ones most widely grown in the Mediterranean Basin, were characterized with 11196 single nucleotide polymorphism (SNP) markers. Total genetic diversity (HT) and polymorphic information content (PIC) were 0.36 and 0.30 respectively for both landraces and modern cultivars. Linkage disequilibrium for the modern cultivars was higher than for the landraces (0.18 and 0.12, respectively). Analysis of the genetic structure showed a clear geographical pattern for the landraces, which were clustered into three subpopulations (SPs) representing the western, northern and eastern Mediterranean, whereas the modern cultivars were structured according to the breeding programmes that developed them: CIMMYT/ICARDA, France/Italy, and Balkan/eastern European countries. The modern cultivars showed higher genetic differentiation (GST) and lower gene flow (0.1673 and 2.49, respectively) than the landraces (0.1198 and 3.67, respectively), indicating a better distinction between subpopulations. The maximum gene flow was observed between landraces from the northern Mediterranean SPs and the modern cultivars released mainly by French and Italian breeding programmes.

2. Introduction

Wheat is one of the most widely cultivated crops in the world, covering an area of 219 million ha with a production of nearly 772 million t in 2017 (http://www.fao.org/faostat). Of the daily intake of humans, wheat provides 19% of the calories and 21% of the protein (http://www.fao. org/faostat). It is generally accepted that to match the global population demand wheat production will need to increase by 1.7% per year by 2050 [1]. Additionally, the increasing unpredictability of the weather conditions imposed by climate change will require the release of cultivars with very high yield potential that are able to maintain acceptable yield levels and stability under a broad range of environmental conditions. In the Mediterranean Basin, the climatic regions vary greatly, including both favourable lands and drylands that are subject to frequent drought episodes and high temperature stress, particularly during grain filling [2].

Wheat is estimated to have originated around 10000 years BP in the Fertile Crescent. From there it spread to the west of the Mediterranean Basin and reached the Iberian Peninsula around 7000 years BP [3]. During this migration, domestication and selection by humans resulted in the development of landraces that were very well adapted to local environments [4]. From the middle of the 20th century, the cultivation of local landraces was progressively abandoned, as a consequence of the Green Revolution, and they were replaced by the improved and more productive semi-dwarf cultivars. Mediterranean landraces are an important group of genetic resources because of their documented resilience to abiotic stresses, their resistance to pests and diseases, and their genetic diversity [4, 5].

Knowledge of genetic diversity provides valuable information for understanding the relationships between cultivars and facilitates their characterization and classification, determination of population structure, etc., thus enriching breeding strategies for crop improvement, for example helping breeders to develop new cultivars reducing pre-breeding activities. In the last few decades, several markers have been used for genetic studies [6]. However, the high-density genome coverage provided at low cost in recent years by new high-throughput genotyping technologies such as single nucleotide polymorphism (SNP) arrays or genotyping-by-sequencing (GBS) have made them the procedures of choice for wheat genetic analysis [7–11].

The aim of this study was to explore the existence of genetic and/or geographic structures and genetic diversity in collections of wheat landraces from the Mediterranean Basin and modern cultivars representative of the ones most widely cultivated in the region.

3. Material and methods

3.1. Plant material

With the aim of representing the past and current cultivated variability in the Mediterranean Basin, the plant material consisted of a germplasm collection of 354 bread wheat (*Triticum aestivum* L.) genotypes, the MED6WHEAT IRTA-panel, of which 170 correspond to landraces from 24 Mediterranean countries and 184 to modern varieties cultivated in 19 countries in the region (S1 File). The landraces were selected from a larger collection comprising 730 accessions of different origins on the basis of phenology, spike characteristics and passport data.

Care was taken to include enough genotypes to represent the different climatic regions existing within the Mediterranean Basin [12]. Landrace populations were provided by public gene banks from Germany (IPK, Gatersleben), Italy (ISC, S. Angelo Lodigiano), Romania (Suceava GenBank, Suceava), Russia (VIR, St. Petersburg), Spain (CRF-INIA, Madrid), the Netherlands (CGN-WUR,

Wageningen) and the USA (NSGC-USDA, Aberdeen, ID). Modern cultivars were provided by public institutions (CIMMYT, ICARDA, INRA and the University of Belgrade), breeding companies and the germplasm collection of the IRTA breeding programme. Accessions were bulk-purified during two cropping cycles to select the dominant type and seed was increased in plots in the same field to ensure a common origin for all lines.

3.2. Molecular characterization

DNA isolation was performed from lyophilized leaf samples at Trait Genetics GmbH (Gate- rsleben, Germany). Accessions were genotyped with 13177 SNPs from the Illumina Infinium 15K Wheat SNP Chip at Trait Genetics GmbH (Gatersleben, Germany). Markers were ordered according to the SNP map developed by Wang et al. [13].

3.3. Data analysis

Polymorphic information content (PIC) values were calculated following the formula described by Botstein et al. [14] using the Cervus software v3.0.7 [15] (available at http://www. fieldgenetics.com). Genetic diversity was estimated as total diversity (HT) [16] using POP-GENE v1.32 [17]. The coefficient of genetic differentiation, i.e. the proportion of total variation that is distributed between populations (GST), was calculated as GST=DST/HT, where DST is the genetic diversity between populations. DST was calculated as DST=HT-HS, where HS is the mean genetic diversity within populations. Gene flow was estimated as Nm=0.5(1-GST)/GST according to McDonald and McDermott [18].

Linkage disequilibrium (LD) was estimated as the square of marker correlations (r^2) for markers with known map positions using TASSEL 5.0 [19] at a significance level of $P<0.001$ with a sliding window of 50 cM. The intra-chromosomal r^2 values were plotted against the genetic distance and a LOESS curve was fitted to determine the distance at which the curve intercepts the line of a critical value of r^2 in order to estimate how fast the LD decay occurs for each chromosome. The critical value of r^2 was determined as the mean r^2 for each genome.

The genetic structure of the collection was estimated using the Bayesian clustering algorithm implemented in the STRUCTURE software v2.3.4 [20] using an admixture model with burn-in and Monte Carlo Markov chain for 10000 and 100000 cycles, respectively. A continuous series of K were tested from 1 to 10 in seven independent runs. The Evanno method [21] was used to calculate the most likely number of subpopulations with STRUCTURE HARVESTER software [22]. Principal coordinates analysis (PCoA) based on genetic distance was performed using GenAlEx 6.5 [23].

Diversity analysis between accessions was determined by simple matching coefficient [24] implemented in DARwin software v.6 [25]. The un-rooted tree was calculated using the neighbour-joining clustering method [26]. The tree is divided in:

- Clusters: Main divisions of the tree
- Branches: Division within clusters
- Groups: Genotypes from the same subpopulation within a branch

4. Results

4.1. SNP polymorphism and diversity

A total of 11196 polymorphic markers were located in the map developed by Wang et al. [13]. In order to reduce the risk of errors in further analyses, markers and accessions were analysed for the presence of duplicated patterns and missing values. For the landrace collection, 8 markers with more than 25% of missing values as well as 730 markers with minor allele frequency (MAF) lower than 5% were excluded from the analysis, leaving a total of 10458 SNPs (Table 1). For the modern collection, 3 markers with more than 25% of missing values and 487 markers with an MAF lower than 5% were excluded, leaving 10706 polymorphic markers. The total number of polymorphic markers was 11074, of which 10090 (91%) were polymorphic in both collections, 368 only in landraces and 616 only in modern cultivars.

The D genome had the lowest number of markers, whereas the B genome showed the best coverage (Table 1). Genetic diversity (HT) and PIC were estimated for each chromosome. For the landraces, HT ranged from 0.39 (1A, 4B) to 0.25 (4D) (mean 0.36) and PIC ranged from 0.33 (1A) to 0.21 (4D) (mean 0.30). For the modern collection, HT ranged from 0.40 (2D, 6D) to 0.29 (7D) (mean 0.36) and PIC ranged from 0.33 (1D, 2D) to 0.25 (7D) (mean 0.30)

4.2. Linkage disequilibrium

LD was determined (r^2) for single chromosomes. In the landraces, LD ranged from 0.07 in chromosome 7D to 0.34 in chromosome 1D, with a mean among chromosomes of 0.14. The percentage of locus pairs showing a significant LD at $P<0.001$ ranged from 16% to 59% for chromosomes 7D and 2D, respectively, with a mean of 33% (Table 1). In modern cultivars the mean value of r^2 was 0.18, with 45% of the locus pairs in LD. Chromosome 7D, as reported for landraces, showed the lowest LD (0.08), with 21% of the locus pairs in LD at $P<0.001$ (Table 1). The maximum LD was found in chromosome 1D (0.41), showing 60% of the locus pairs with a significant LD at $P<0.001$. However, as reported for the landraces, the chromosome with the most locus pairs with a significant LD was 2D (68%, $r^2=0.4$).

Table 1. Number of SNP markers (N), gene diversity (HT), polymorphic information content (PIC) and linkage disequilibrium (LD) (r^2, % of markers in LD at P<0.001, and LD decay in cM) for all of the chromosomes and genomes in both types of germplasm.

	Landraces						Modern cultivars					
	N	H_T	**PIC**	**LD (r^2)**	**LD (%)**	**LD decay**	**N**	H_T	**PIC**	**LD (r^2)**	**LD (%)**	**LD decay**
1A	595	0.39	0.33	0.16	45	2	604	0.37	0.28	0.18	44	1
1B	812	0.36	0.30	0.11	35	2	834	0.33	0.28	0.17	45	2
1D	293	0.36	0.30	0.34	49	1	295	0.39	0.33	0.41	60	3
2A	576	0.38	0.30	0.13	30	1	590	0.36	0.30	0.16	42	3
2B	925	0.36	0.30	0.10	28	1	999	0.36	0.29	0.18	48	2
2D	376	0.37	0.30	0.30	59	2	388	0.40	0.33	0.40	68	2
3A	522	0.37	0.30	0.09	28	1	522	0.35	0.31	0.17	46	1
3B	770	0.35	0.30	0.10	31	1	781	0.34	0.30	0.14	45	2
3D	134	0.35	0.28	0.13	23	1	137	0.36	0.29	0.14	23	1
4A	435	0.35	0.30	0.10	26	1	433	0.35	0.29	0.16	43	1
4B	375	0.39	0.31	0.14	38	3	382	0.36	0.32	0.17	51	2
4D	39	0.25	0.21	0.09	23	2	51	0.39	0.30	0.12	32	2
5A	628	0.36	0.31	0.12	29	1	655	0.36	0.31	0.18	46	3
5B	860	0.38	0.31	0.11	31	2	867	0.37	0.31	0.18	53	2
5D	93	0.33	0.27	0.12	29	10	143	0.38	0.30	0.12	29	10
6A	648	0.36	0.30	0.14	34	7	674	0.36	0.31	0.18	52	9
6B	774	0.36	0.31	0.13	33	2	787	0.32	0.30	0.26	63	4
6D	135	0.35	0.28	0.18	45	1	146	0.40	0.32	0.15	38	4
7A	703	0.37	0.31	0.09	29	2	665	0.35	0.31	0.11	37	4
7B	652	0.37	0.30	0.11	34	2	637	0.36	0.31	0.17	56	2
7D	113	0.32	0.26	0.07	16	4	116	0.29	0.25	0.08	21	6
Genome A	4107	0.37	0.31	0.12	31	-	4143	0.36	0.30	0.16	43	-
Genome B	5168	0.37	0.31	0.11	31	-	5287	0.35	0.30	0.18	51	-
Genome D	1183	0.33	0.27	0.22	39	-	1276	0.37	0.30	0.20	44	-
Total	10458	0.36	0.30	0.14	32	-	10706	0.36	0.30	0.18	46	-

The extent of LD was also investigated for the three genomes. The highest r^2 was found in the D genome for both landraces and modern cultivars (0.22 and 0.20, respectively). The D genome had the most markers with significant LD at $P<0.001$ (39%) in the landraces, whereas in the modern cultivars the B genome had the most, with 51% of the locus pairs showing a significant LD at $P<0.001$ (Table 1). The decay of LD varied for each chromosome. For the A and B genomes, LD decay ranged from 1 to 9 cM, whereas for the D genome it ranged from 1 to 10 cM (Table 1, S3 File).

4.3. Population structure

The first analysis of the population structure of the 354 accessions was carried out using 557 common SNPs markers between landraces and modern cultivars evenly distributed across the genome. The markers were selected according to the distance of the LD decay for each chromosome in order to avoid the use of markers with a significant LD. The Bayesian clustering method using the Evanno test [21] to infer the most likely number of structured subpopulations (ΔK) revealed the presence of two distinct subpopulations, one including landraces and the other including modern cultivars (data not shown). On the basis of this result, a subsequent analysis of population structure was performed independently for the landraces and modern cultivars.

The highest value of ΔK for the landraces was observed for K=2 (2861), followed by K=3 (1031) (Fig 1A). Population structure for both K=2 and K=3 showed a geographical pattern. For K=2 the landraces were separated following an east-west pattern within the Mediterranean Basin, whereas for K=3 landraces mainly from the Balkan Peninsula were separated either from eastern Mediterranean ones or from French and Italian ones included in the western Mediterranean. The subpopulations (SPs) were classified as western (SP1), northern (SP2) and eastern Mediterranean (SP3) (Fig 1B) according to the geographical region of the countries most represented in the SP. The inferred population structure for K=3 showed that 45% of the accessions (77 out of 170) showed a strong membership coefficient (q-value) to one of the SPs ($q>0.7$), whereas using a moderate q-value (>0.5) the number of accessions within an SP increased to 144 (85%), leaving 26 as admixed (S1 File). The western Mediterranean SP (SP1) included 43 accessions, of which 28 corresponded to western countries (Spain, Portugal, France, Morocco, Algeria and Tunisia), with 23 of them having a mean q-value greater than 0.8 (Table 2). SP1 also included 10 landraces (23%) from eastern countries and 6 (14%) from northern countries. The northern Mediterranean SP (SP2) included 59 accessions, with 68%, 24% and 8% corresponding to northern (Albania, Bosnia and Herzegovina, Croatia, France, Greece, Italy, Macedonia, Romania and Serbia), western and eastern Mediterranean countries, respectively (Table 2 and Fig 1B and 1C). Finally, the eastern Mediterranean

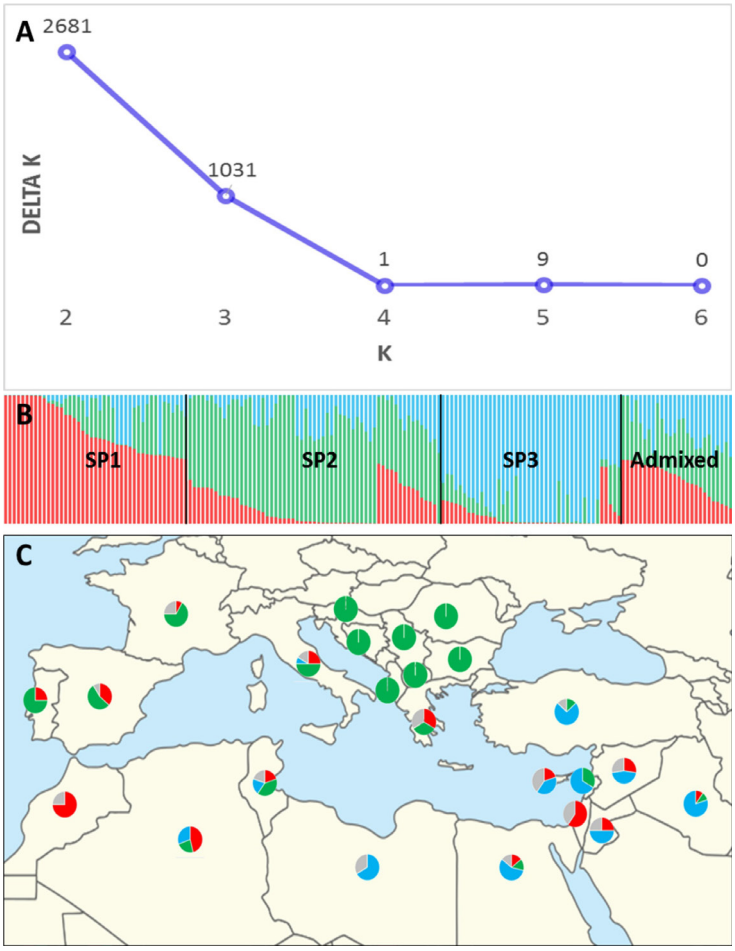

Figure 1. Genetic structure of Mediterranean landraces. (A) Estimation of the number of subpopulations (SPs) according to the Evanno test. (B) Inferred structure of the landrace collection based on 170 genotypes. Each individual is represented by a coloured bar with length proportional to the estimated membership to each of the three SPs. (C) Geographic distribution of the wheat subpopulations within the Mediterranean Basin. Circles indicate the proportion of each SP in the country. Red, SP1 (western Mediterranean); green, SP2 (northern Mediterranean); blue, SP3 (eastern Mediterranean); grey, admixture. Map source: https://commons.wikimedia.org/wiki/File:Middle_East_location_map.svg.

SP (SP3) included 42 accessions, of which 86% corresponded to eastern countries (Cyprus, Egypt, Jordan, Iraq, Lebanon, Libya, Syria and Turkey) and 12% and 2% to western and northern countries, respectively (Table 2). As shown in Table 2 and Fig 1, the northern Mediterranean SP included accessions from Balkan countries, but also French, Italian and Spanish landraces. On the other hand, three accessions from Israel belonged to the western Mediterranean SP, but their q-values (average q=0.605) were lower than those of the accessions from western Mediterranean countries. Also, 40% of the accessions from Tunisia belonged to SP2. Accessions

Table 2. Mean membership coefficients across the landraces from each country included in each SP. Only structured genotypes (q>0.5) are shown in the table. q-values>0.7 are indicated in bold.

	SP1		SP2		SP3	
Country	***q* mean**	**N**	***q* mean**	**N**	***q* mean**	**N**
Albania			**0.755**	3		
Algeria	**0.862**	6	**0.762**	3	0.601	4
Bosnia and Herzegovina			0.691	3		
Bulgaria			0.657	3		
Croatia			**0.700**	3		
Cyprus	0.661	1			0.501	2
Egypt	**0.992**	1	0.680	1	**0.706**	4
France	0.502	1	**0.707**	8		
Greece	0.565	2	0.540	2		
Iraq	0.517	1	0.546	1	**0.915**	8
Israel	0.605	3				
Italy	0.615	3	**0.727**	6	**0.986**	1
Jordan	0.614	1			**0.704**	2
Lebanon			0.676	1	**0.818**	2
Libya					**0.726**	2
Macedonia			**0.889**	4		
Morocco	**0.845**	15				
Portugal	**0.822**	1	0.592	3		
Romania			**0.774**	4		
Serbia			**0.761**	4		
Spain	0.526	4	0.544	6		
Syria	0.591	3			**0.966**	5
Tunisia	**0.973**	1	**0.738**	3	0.539	1
Turkey			0.552	1	**0.910**	11

SP1, western Mediterranean; SP2, northern Mediterranean; SP3, eastern Mediterranean.

from Portugal and Spain were distributed in SP1 and SP2. As shown in Table 2, the inconsistencies observed for these countries could be related to the low q-values for the most represented SP. The three Portuguese accessions included in SP2 had a mean q-value of 0.592, and the Spanish and Greek accessions showed q-values slightly higher than 0.5 for the two SPs (Table 2), denoting a high level of admixture in all of them.

Following the Evanno test, as previously reported for landraces, the most likely number of structured SPs (ΔK) for modern cultivars was K=2 (728) (Fig 2A). The first cluster grouped mainly cultivars developed by French and Italian breeding programmes, and the second one grouped mainly cultivars from North Africa, the Middle East and Spain, with evident CIMMYT and ICARDA genetic background (S1 File). A set of cultivars mainly from Serbia remained admixed. If the second highest value for the structured subpopulations was chosen (K=3) (105), a third SP clustered all Serbian accessions, plus 2 from Macedonia and 1 from Hungary (Fig 2B and 2C, S1 File). According to the classification into three structured subpopulations, 79% of the accessions (145 out of 184) showed a q>0.7. The first SP (SP4) included 73 cultivars from France, 10 from Italy and 1 from Serbia. The second SP (SP5) included cultivars from eastern Europe, mainly Serbia (21 out of 24). The third SP (SP6) included 33 cultivars from Spain, 15 cultivars from North African countries, 12 from the Middle East and Central Asia (Afghanistan), and finally 4 from northern Mediterranean countries (France and Italy). Finally, 12 cultivars remained as admixed.

The relationships between landraces and modern cultivars were also analysed by PCoA as a complementary way to visualize their clusters. In agreement with the results shown by STRUCTURE, the first two coordinates of the PCoA clearly separated the landraces from the modern cultivars, and within each group accessions were clustered matching the results of STRUCTURE (Fig 3). Landraces from the Balkan Peninsula 'TRI 1667' and 'TRI 1671' (Albania), 'Moriborska' (Bosnia and Herzegovina), '41-II/4-B' and 'Legan Bezosja' (Serbia), and 'Solonetu Nou' (Romania) were positioned on the positive side of the PCoA1 close to the origin of the axes, together with the modern Serbian cultivars. Additionally, two landraces from Italy ('TRI 16900' and 'TRI 16516') and three from France ('Mounton a Epi Rouge', 'TRI 14046' and 'TRI 17938') were located within the modern cultivars from SP4 that grouped modern French and Italian cultivars.

The six SPs showed a total genetic diversity (HT) ranging from 0.2651 for modern cultivars from the Balkans and eastern Europe (SP5) to 0.3690 for northern Mediterranean landraces (SP2) (Table 3). The genetic diversity among SPs (DST) was low (0.0706), resulting in a genetic differentiation (GST) among SPs of 0.1838. This means that only about 18% of the variability observed was due to differences

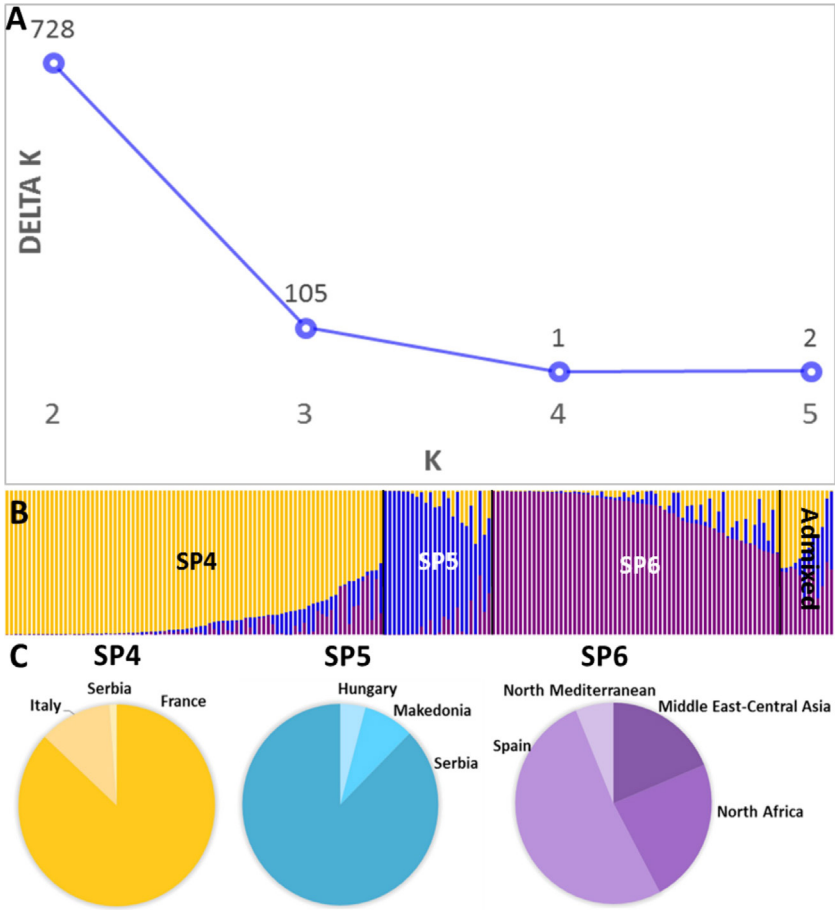

Figure 2. Genetic structure of the modern cultivars. (A) Estimation of the number of subpopulations (SPs) according to the Evanno test. (B) Inferred structure of the collection based on 184 genotypes. Each individual is represented by a coloured bar with length proportional to the estimated membership to each of the three subpopulations. (C) Proportion of cultivars from the different countries/regions within each SP. Yellow, SP4; blue, SP5; violet, SP6.

between SPs. The estimation of GST also allowed us to estimate the gene flow (Nm) among SPs. The value of this estimate (2.22) indicates a high level of gene exchange, which denotes a low genetic differentiation among the SPs. When analysed by type of accessions, the landraces showed a GST=0.1198 and an Nm=3.67, whereas the modern cultivars showed a GST=0.1673 and an Nm=2.49, indicating higher gene exchange between the landrace SPs (Table 3). When comparisons were made between two SPs, gene flow ranged from 2.53 between western Mediterranean landraces (SP1) and modern Balkan cultivars (SP5) to 9.39 between northern Mediterranean landraces (SP2) and cultivars developed mainly by French and Italian breeding programmes (SP4) (Table 3).

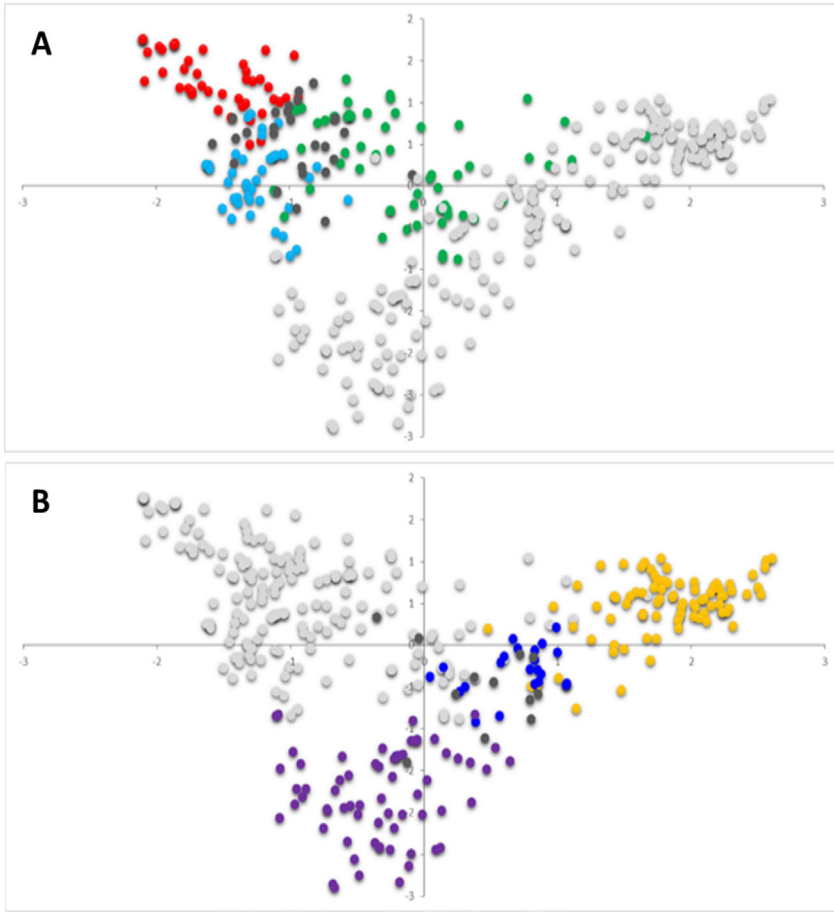

Figure 3. Principal coordinates analyses based on genetic distance. A) Landraces: red, SP1; green, SP2; blue, SP3; dark grey, admixed; light grey, modern cultivars. B) Modern cultivars: yellow, SP4; dark blue, SP5; violet, SP6; dark grey, admixed; light grey, landraces.

4.4. Cluster analysis

To better detail the kinship among accessions, a neighbour-joining tree was built using the common SNPs markers between the landraces and modern cultivars. The dendrogram showed two main clusters with a robust separation between them (Fig 4). Within each of these clusters, accessions were mainly grouped in agreement with the groups obtained previously by STRUCTURE and PCoA analysis (Fig 4).

The first cluster of modern cultivars (M_Q1) included cultivars from France and Italy (SP4), in addition to three landraces: 'Bahatane' and 'TRI17938' from France and 'TRI16516' from Italy. The second cluster of modern cultivars (M_Q2) represented mainly the accessions carrying CIMMYT and ICARDA genetic background (SP6), including cultivars developed in Turkey, Spain, Egypt, Syria, Morocco, Algeria,

Table 3. Genetic diversity and variation between the MED6WHEAT subpopulations (SPs).

	N	H_T	H_S	D_{ST}	G_{ST}	Nm
SP1	43	0.2873	-	-	-	-
SP2	59	0.3690	-	-	-	-
SP3	42	0.3132	-	-	-	-
SP4	82	0.2995	-	-	-	-
SP5	62	0.2651	-	-	-	-
SP6	24	0.3476	-	-	-	-
Total	312	0.3842	0.3136	0.0706	0.1838	2.22
Landraces	144	0.3672	0.3232	0.0440	0.1198	3.67
Modern	168	0.3652	0.3041	0.0611	0.1673	2.49
SP1-SP2	102	0.3594	0.3282	0.0312	0.0868	5.26
SP1-SP3	85	0.3363	0.3003	0.0360	0.1070	4.17
SP1-SP4	125	0.3452	0.2934	0.0518	0.1501	2.83
SP1-SP5	67	0.3307	0.2762	0.0545	0.1648	2.53
SP1-SP6	105	0.3690	0.3174	0.0516	0.1398	3.08
SP2-SP3	101	0.3737	0.3411	0.0326	0.0872	5.23
SP2-SP4	141	0.3521	0.3343	0.0178	0.0506	9.39
SP2-SP5	83	0.3628	0.3171	0.0457	0.1260	3.47
SP2-SP6	121	0.3844	0.3583	0.0261	0.0679	6.86
SP3-SP4	124	0.3549	0.3064	0.0485	0.1367	3.16
SP3-SP5	66	0.3445	0.2892	0.0553	0.1605	2.61
SP3-SP6	104	0.3304	0.2892	0.0412	0.1247	3.51
SP4-SP5	106	0.3176	0.2823	0.0353	0.1111	4.00
SP4-SP6	144	0.3658	0.3236	0.0422	0.1154	3.83
SP5-SP6	86	0.3594	0.3063	0.0531	0.1477	2.88

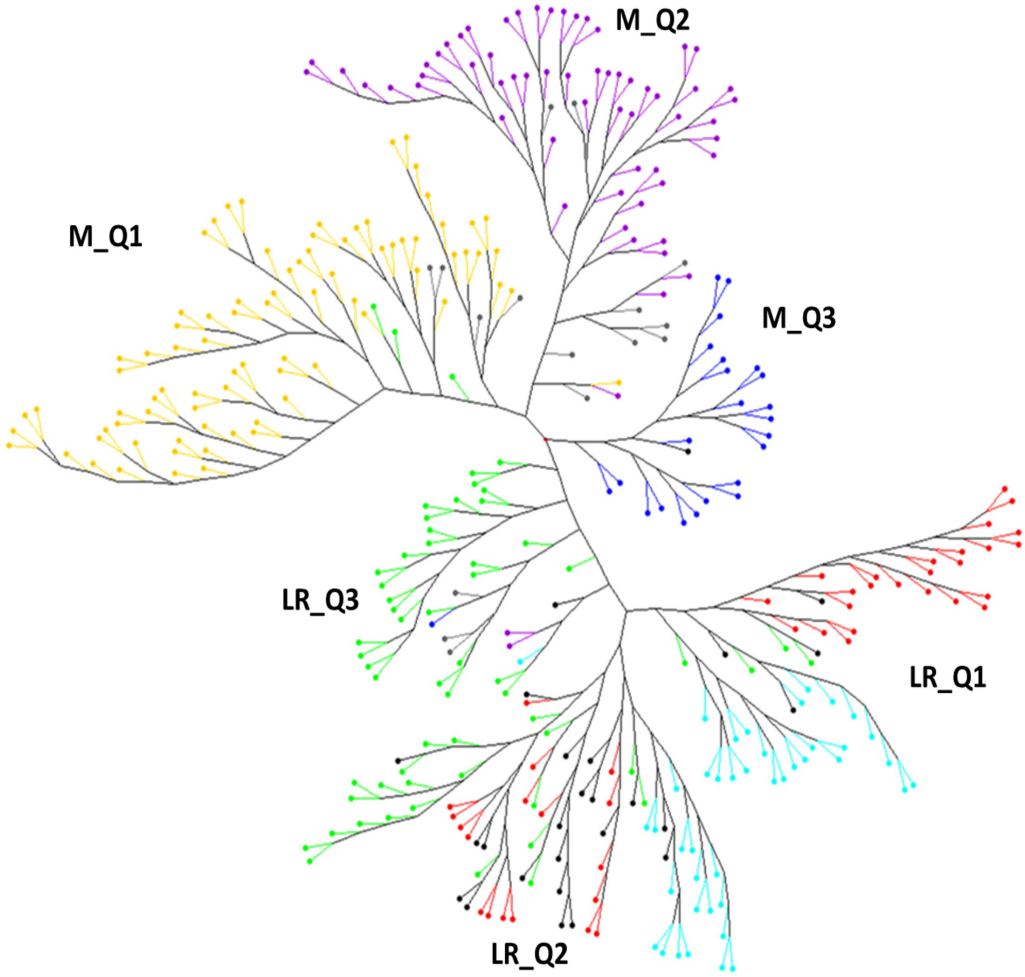

Figure 4. Un-rooted neighbour-joining dendrogram. Colours of branches correspond to the SPs obtained by STRUCTURE analysis. Landraces (LR): red, SP1; green, SP2; blue, SP3. Modern cultivars (M): yellow, SP4; dark blue, SP5; violet, SP6.

France, Italy, Afghanistan, Sudan and Tunisia. This cluster also included the French cultivar 'Boticelli' classified in SP4 but including 37% of the genetic background from SP6. Finally, the third cluster of modern cultivars (M_Q3) contained the group of all elite cultivars from the Balkan Peninsula and eastern Europe (SP5): Serbia (23), Macedonia (2) and Hungary (1). This cluster also included the Syrian landrace 'Salamuni-A', classified as admixed by STRUCTURE.

The clustering of landraces suggested a more complex distribution among the SPs according to the higher frequency of admixture revealed by STRUCTURE. The

first cluster (LR_Q1) showed three branches, the first one with landraces mainly from the western Mediterranean (SP1), with Morocco as the best-represented country of the branch (45% of the accessions). The second country represented in this branch was Algeria (14%), and the remaining accessions from Cyprus, Egypt, Greece, Iraq, Portugal, Spain, Syria, Tunisia and Turkey were represented in a lower frequency (3%-6%). The second and third branches included mainly eastern Mediterranean landraces, with 81% of the accessions coming from Egypt, Iraq, Jordan, Syria and Turkey. The remaining accessions were from western (Algeria, 11% and Tunisia, 4%) and northern (France, 4%) Mediterranean countries.

The second cluster (LR_Q2) was represented by landraces from the three SPs in two main branches. The first branch included a division between genotypes from the eastern (SP3) and western (SP1) Mediterranean SPs. The eastern group was represented mainly by Turkish landraces (46%). An additional 27% was represented by landraces from Cyprus, Iraq, Jordan and Lebanon, and finally the remaining cultivars were previously grouped by structure analysis in the northern Mediterranean SP (SP2). The second group of the branch was composed mainly of landraces from the eastern Mediterranean countries Cyprus (10%), Israel (50%), Jordan (20%) and Syria (10%), although the Bayesian clustering determined that their structure was more similar to that of landraces from the western Mediterranean SP. Only one landrace from Morocco was included in this group. The second cluster included landraces from the western and northern Mediterranean SPs. The first branch included most of the accessions from western Mediterranean countries (Morocco, Portugal, Spain and Tunisia) and one accession from Libya. In this branch, 50% of the accessions showed high levels of admixture. The second branch of the cluster was divided into two groups, the first including mostly landraces classified as western Mediterranean by STRUCTURE. The group included cultivars from Algeria, Greece, Italy, Morocco and Turkey. The second group was mainly composed of northern Mediterranean landraces (71%) from Albania, France, Italy, Macedonia and Serbia. Although included by structure analysis in this SP, Portugal, Spain and Tunisia from the western Mediterranean and Lebanon from the eastern Mediterranean were also included in this group.

Finally, the remaining landraces were included in four branches (LR_Q3) of the main clusters examined above, with 76% corresponding to northern Mediterranean countries (Albania, Bulgaria, Bosnia and Herzegovina, France, Greece, Croatia, Italy, Romania and Serbia). Within these branches, five modern cultivars were also included: the Spanish cultivars 'Montcada' and 'Montserrat', the Turkish cultivars 'Ata 81' and 'Cumhuriyet 75', and the Serbian cultivar 'KG 100'.

5. Discussion

5.1. Genetic diversity

Genetic diversity is essential for plant breeding because it provides new knowledge for improving cultivars. In wheat, the genetic diversity was narrowed down during the second half of the 20th century as a consequence of the introduction of high-yielding improved semi-dwarf cultivars. Several studies, reviewed in Lopes et al. [27], have considered landraces to be a source of lost variability that can provide favourable genes to improve modern cultivars. However, in a recent study of the genetic structure of durum wheat Mediterranean landraces, Soriano et al. [6] reported that the great genetic distance estimated between modern cultivars and landrace populations denotes a low use of durum landraces by durum wheat breeding programmes. The knowledge of the genetic diversity within landrace populations will be of special interest for designing new crosses with commercial varieties in order to widening the variability of the new genotypes. As it is reported in Lopes et al. [27] the monitoring of the genetic diversity of landraces is an approach to increase the frequency of rare alleles in breeding programs to find new allelic variation of genes of interest.

Among the 11196 polymorphic markers in the MED6WHEAT panel with a known genetic position according to the map of Wang et al. [13], the B genome had the highest number of SNPs and the D genome the lowest, as reported by Alipour et al. [7] and Eltaher et al. [11]. The collection of Mediterranean landraces showed the lowest PIC and gene diversity values for the D genome, according to the findings of Lopes et al. [28]. Alipour et al. [7] and Eltaher et al. [11] related the low polymorphism in the D genome to the recent evolutionary history of this genome in comparison with the A and B genomes [29]. However, when the collection of modern cultivars was analysed, we found no differences between PIC values for the three genomes, while the D genome showed a slightly high gene diversity. Trethowan and Mujeeb-Kazi [30] and Jia et al. [31] concluded that higher diversity in the D genome may provide new elite and desirable alleles controlling important traits for dealing with climate change.

The average PIC value for both the landraces and the modern cultivars was 0.30. This value is in agreement with previous studies using bi-allelic markers such as SNP or DArT in either common or durum wheat. In common wheat, Lopes et al. [28] found a PIC value of 0.24 for the WAMI population genotyped with the 9K SNP array. Novoselović et al. [32] characterized a Croatian collection with 1229 DArT markers with an average PIC value among the populations of 0.30. Alipour et al. [7] genotyped a diversity panel of 369 Iranian landraces using 16506 GBS-based SNPs, reporting an average PIC of 0.172. El-Esawi et al. [33], using

Austrian and Belgian wheats, found PIC values of 0.33 and 0.29, respectively, with 1052 DArT markers. Finally, Eltaher et al. [11], in an F3:6 Nebraska winter wheat population genotyped with 25566 SNPs generated by GBS, found a PIC value of 0.25. In durum wheat, similar PIC values have been reported (Baloch et al. [8], 0.26 and 0.30 using DArTseq and SNPs, respectively; Kabbaj et al. [10], 0.32). The SNP markers in our panel were moderately informative according to the classification of average PIC values into the three categories proposed by Botstein et al. [14]: highly informative (PIC>0.5), moderately informative (0.25<PIC<0.5) and slightly informative (PIC<0.25). Previous studies using SSR markers revealed higher levels of polymorphism. In a previous study of our group, Soriano et al. [6], using a panel of 192 durum wheat genotypes (mainly Mediterranean landraces) genotyped with 44 SSR markers, found an expected heterozygosity of 0.71. Similar results with SSRs have been reported in durum wheat [34, 35] and bread wheat [36–38]. The lower PIC value obtained with SNPs or DArTs than with SSR markers may be explained by their bi-allelic nature, which means that the maximum attainable PIC is 0.5 when the two alleles have the same frequency [11, 39].

5.2. Linkage disequilibrium and population structure

Linkage disequilibrium is defined as the non-random association of alleles at different loci and decays rapidly with genetic distance. Thus, determining the LD decay over physical and genetic distance within a panel of genotypes is an important step for determining the resolution and marker density required for association studies. Moreover, LD is influenced by population structure due to stratification and unequal distribution of alleles within groups of genotypes, which can result in false associations [20].

In the current study, the mean r^2 calculated for intra-chromosomal loci was 0.12 and 0.18 with 32% and 46% of the locus pairs in LD for landraces and modern cultivars, respectively. It is well known that crops gradually lose their genetic variability through domestication and breeding, resulting in more uniform cultivars, reducing the recombination rate and affecting LD [40, 41]. The D genome showed the highest r2 for both landraces and modern cultivars. Similar results were reported by Chao et al. [42] and Lopes et al. [28], who explained that higher LD in the D genome was linked to recent introgressions and population bottlenecks in the origin of hexaploid wheat.

The first attempt to dissect the genetic structure of the MED6WHEAT panel showed a clear separation based on historical breeding periods, i.e. landraces vs modern cultivars. Thus, for subsequent analysis of population structure, independent analyses were carried out for the two groups of germplasm. According to a Bayesian-based analysis, without a priori assignment of accessions to populations, the landraces

showed a geographic structure according to the east- ern and northern zones of the Mediterranean Basin, whereas accessions classified as western Mediterranean showed a high level of admixture. This classification denoted a migration from the centre of wheat domestication in the Fertile Crescent to the west of the Mediterranean Basin, as reported by Moragues et al. [43] and Soriano et al. [6] in durum wheat. The higher admixture found in western Mediterranean landraces may be due to the incorporation and fixation of favourable alleles from eastern and northern genetic pools during the migration process. By contrast, for modern cultivars the separation was mainly based on the pedigree of the accessions: CIMMYT/ICARDA, cultivars obtained mainly by French breeding programmes, and accessions from the Balkan Peninsula. These groups may have originated through the sharing of germplasm from different breeding programmes with similar growing conditions, particularly from the shuttle breeding carried out by international centres. For the landrace collection, only 45% of the accessions showed a strong q-value (>0.7), suggesting high levels of admixture among SPs, whereas for the modern cultivars 79% of the accessions showed a strong q-value. Oliveira et al. [44] suggested that admixture is the result of the incorporation of alleles from more than one gene pool because of the spread of wheat from different ancestral populations. Moragues et al. [43] proposed as a possible cause of admixture the exchange of germplasm between different Mediterranean regions during the expansion of the Arabian Empire. The low level of admixture between modern cultivars could be due to the development by breeding programmes of cultivars with specific adaptation to the local environments and the use of different genetic resources.

Based on the defined SPs, the results of PCoA and neighbour-joining were in agreement with those reported by STRUCTURE, showing first a robust separation between the landraces and modern cultivars, and within these main clusters a separation into three genetic SPs.

The origin of the axes in the PCoA showed a mixture between landraces and modern cultivars, as also reported by Oliveira et al. [45]. Landraces from the Balkan Peninsula co-localized with modern cultivars from the same region, and two landraces from Italy and three from France were located close to modern cultivars from those countries. A possible cause of this mixture could be the presence of these landraces in the pedigree of the modern cultivars, as reported for the French landrace 'Mounton a Epi Rouge', which according to Bonjean [46] played an important role in the pedigrees of improved French cultivars between 1965 and 1975.

According to the admixture revealed by STRUCTURE analysis, the modern cultivars showed well-defined clusters with differentiation among SPs in the neighbour-joining tree, whereas for the landraces the branches included accessions

from different SPs. Only four genotypes from a given SP were misclassified within the clusters of modern cultivars. Within M_Q1, formed mainly by French and Italian cultivars, two French and one Italian landraces were also included. Cluster M_Q2, which grouped most cultivars carrying CIMMYT and ICARDA genetic background, also included the French cultivar 'Boticelli', which, although belonging to SP4, has an important genetic background (37%) from SP6. Cluster M_Q3 included a landrace from Syria, classified as admixed, with 30% of its genetic background from the northern Europe SP. The presence of landraces within modern cultivars was reported in a global durum wheat panel by Kabbaj et al. [10], who concluded that the simplest explanation was that they were not true landraces, but old tall cultivars wrongly labelled during the collecting mission by the gene banks. This seems a plausible explanation, considering that the first breeding attempts made by pioneer breeders or entrepreneurial Mediterranean agriculturalists consisted in identifying and isolating the best lines already existing within original wheat landraces [47]. Alternatively, the grouping within elite cultivars was probably due to the fact that they were used in breeding programmes to enlarge the genetic diversity, as reported for grain legumes by Sharma et al. [48].

When landraces were analysed by hierarchical clustering, a higher level of admixed genotypes was found on the basis of the STRUCTURE classification. Although the main groups within the clusters were formed mostly by members of specific SPs, a discrepancy between the classifying methods was observed among groups of the same cluster. As reported for clusters LR_Q1 and LR_Q2, some of the landraces misclassified by STRUCTURE according to their country of origin or with a high level of admixture were grouped by neighbour-joining into clusters containing accessions from the same geographical region. In LR_Q3, five modern cultivars were also grouped with landraces, probably due to the presence in their pedigree of genetic background from landraces or closely related accessions. However, the closed pedigree of most commercial cultivars did not allow us to clarify this.

These results highlight the importance of using different approaches to determine the genetic structure of a germplasm collection. Although the different methods are coincident for the genotypes with strong genetic membership to a given group, they are useful to complement the information provided when accessions show large admixture or gene flow among different geographical regions, as in the case of Mediterranean landraces.

5.3. Gene flow

Genetic differentiation and gene flow provide information about population differentiation. Gene flow homogenizes populations by genetically decreasing variance among populations and increasing variance within populations. In our study,

the analysis of genetic differentiation and gene flow indicated that the majority of the genetic variation was explained by differences among cultivars within genetic SPs. Differentiation of modern cultivar SPs was higher than that of landrace SPs, indicating a lower level of gene exchange among cultivars from different origins. These results are in agreement with population structure and neighbour-joining clustering, in which a higher level of admixture was found for landraces from different geographical regions. Accordingly, it has also been suggested that the low genetic differentiation among SPs is due to seed exchange by farmers, mainly influenced by geographic distances [49, 50]. When two SPs were compared, in general gene flow between landrace SPs was also higher than between modern cultivars. Within landraces, the highest gene flow was found between the northern Mediterranean SP and the western and eastern Mediterranean SPs. The value was lower for the exchange between the western and eastern Mediterranean SPs, supporting the hypothesis of geographic distance. When the modern cultivars were analysed, gene flow showed a higher value between SP4 (France and Italy accessions) and SP5 (Balkan accessions). Cultivars with a CIMMYT/ICARDA origin (SP6) had lower values of gene exchange with the other SPs, probably because of the delivery of improved inbred lines to be released by local programmes through the nurseries that these international centres distribute globally. In the case of modern cultivars, the SPs reflected similarities between the genetic pools managed by the breeding programmes conducted in each specific country. The highest gene flow value was reported between SPs from different periods, i.e. between landraces from the northern Mediterranean (SP2) and modern cultivars released by French and Italian breeding programmes (SP4), suggesting the presence of the genetic background of landraces or old cultivars in the improved modern varieties.

6. Concluding remarks

The current study aimed to explore the presence of genetic and geographic structures in a collection of bread wheat landraces and modern cultivars representing the variability existing for the species in the Mediterranean Basin. The results demonstrated the usefulness of the methodologies employed for achieving this goal. The structure for landraces showed a geographical pattern with different levels of admixture, mainly justified by physical distances between the territories where they were collected, whereas the structure for modern cultivars reflected differences and similarities between the genetic pools managed by the breeding programmes operating in the region.

The results reported in the current study may be of special interest for driving the development of new cultivars with desirable traits for the climatic conditions of the Mediterranean Basin, and for identifying useful molecular markers through genome-wide association studies to assist breeding programmes.

Supporting information

S1 File. List of accessions (DOCX); S2 File. Summary statistics for HT, PIC and LD for each one of the chromosomes (XLSX); S3 File. Linkage disequilibrium plots (TIF).

Acknowledgments

This study was funded by project AGL2015-65351-R of the Spanish Ministry of Economy and Competitiveness. RR is a recipient of a PhD grant from the Spanish Ministry of Economy and Competitiveness. JMS was hired by the INIA-CCAA programme funded by INIA and the Generalitat de Catalunya. The authors acknowledge the contribution of the CERCA Program (Generalitat de Catalunya). Thanks are given to Marta Lopes (previously at CIMMYT, Ankara, Turkey), Miguel Sánchez (ICARDA, Rabat, Morocco) and Dejan Dodig (Maize Research Institute, Zemun Polje, Belgrade, Serbia) for providing part of the modern germplasm used in the study, and the different gene banks for providing landrace populations.

Author Contributions

Conceptualization: Fanny Alvaro, Jose Miguel Soriano.

Data curation: Rubén Rufo, Jose Miguel Soriano.

Formal analysis: Rubén Rufo, Jose Miguel Soriano.

Funding acquisition: Conxita Royo, Jose Miguel Soriano.

Investigation: Rubén Rufo, Jose Miguel Soriano.

Resources: Fanny Alvaro.

Supervision: Jose Miguel Soriano.

Validation: Jose Miguel Soriano.

Writing – original draft: Rubén Rufo, Jose Miguel Soriano.

Writing – review & editing: Rubén Rufo, Fanny Alvaro, Conxita Royo, Jose Miguel Soriano.

7. References

1. Leegood RC, Evans JR, Furbank RT. Food security requires genetic advances to increase farm yields. Nature. 2010; 464: 831. https://doi.org/10.1038/464831d PMID: 20376125

2. Royo C, Maccaferri M, A´ lvaro F, Moragues M, Sanguineti MC, Tuberosa R, et al. Understanding the relationships between genetic and phenotypic structures

of a collection of elite durum wheat accessions. Field Crops Res. 2010; 119: 91–105. https://doi.org/10.1016/j.fcr.2010.06.020

3. Feldman M. Origin of cultivated wheat. In: Bonjean AP, Angus WJ (eds) The world wheat book: a history of wheat breeding. Lavoisier Publishing, Paris. 2001; pp 3–56.

4. Nazco R, Villegas D, Ammar K, Peña RJ, Moragues M, Royo C. Can Mediterranean durum wheat landraces contribute to improved grain quality attributes in modern cultivars? Euphytica. 2012; 185: 1–17. https://doi.org/10.1007/s10681-011-0588-6

5. Moragues M, Zarco-Hernández J, Moralejo MA, Royo C. Genetic diversity of glutenin protein subunits composition in durum wheat landraces [Triticum turgidum ssp. turgidum convar. durum (Desf.) MacKey] from the Mediterranean Basin. Gen. Res. and Crop Evol. 2006; 53: 993–1002. https://doi.org/10.1007/s10722-004-7367-3

6. Soriano JM, Villegas D, Aranzana MJ, García del Moral LF, Royo C. Genetic structure of modern durum wheat cultivars and Mediterranean landraces matches with their agronomic performance. PLoS ONE. 2016; 11(8): e0160983. https://doi.org/10.1371/journal.pone.0160983 PMID: 27513751

7. Alipour H, Bihamta MR, Mohammadi V, Peyghambari SA, Ba G, Zhang G. Genotyping-by-Sequencing (GBS) revealed molecular genetic diversity of Iranian wheat landraces and cultivars. Front. Plant Sci. 2017; 8: 1293. https://doi.org/10.3389/fpls.2017.01293 PMID: 28912785

8. Baloch FS, Alsaleh A, Shahid MQ, C¸ iftc¸i V, Sáenz de Miera LE, Aasim M, et al. A whole genome DArT-seq and SNP analysis for genetic diversity assessment in durum wheat from Central Fertile Crescent. PLoS ONE. 2017; 12: e0167821. https://doi.org/10.1371/journal.pone.0167821 PMID: 28099442

9. Joukhadar R, Daetwyler HD, Bansal UK, Gendall AR, Hayden MJ. Genetic diversity, population structure and ancestral origin of Australian wheat. Front. Plant Sci. 2017; 8: 2115. https://doi.org/10.3389/ fpls.2017.02115 PMID: 29312381

10. Kabbaj H, Sall AT, Al-Abdallat A, Geleta M, Amri A, Filali-Maltouf A, et al. Genetic diversity within a global panel of durum wheat (Triticum durum) landraces and modern germplasm reveals the history of alleles exchange. Front. Plant Sci. 2017; 8: 1277. https://doi.org/10.3389/fpls.2017.01277 PMID: 28769970

11. Eltaher S, Sallam A, Belamkar V, Emara HA, Nower AA, Salem KFM, et al. Genetic diversity and population structure of F3:6 Nebraska winter wheat genotypes using genotyping-by-sequencing. Front. Genet. 2018; 9: 76. https://doi.org/10.3389/fgene.2018.00076 PMID: 29593779

12. Royo C, Nazco R, Villegas D. The climate of the zone of origin of Mediterranean durum wheat (Triticum durum Desf.) landraces affects their agronomic performance. Genetic Resources and Crop Evolution. 2014; 61: 1345–1358. https://doi.org/10.1007/s10722-014-0116-3

13. Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, et al. Characterization of polyploid wheat genomic diversity using a high-density 90000 single nucleotide polymorphism array. Plant Biotechnology Journal. 2014; 12: 787–796. https://doi.org/10.1111/pbi.12183 PMID: 24646323

14. Botstein D, White RL, Sholnick M, David RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 1980; 32: 314–331. PMID: 6247908

15. Marshall TC, Slate J, Kruuk LEB, Pemberton JM. Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology. 1998; 7: 639–655. https://doi.org/10.1046/j.1365- 294x.1998.00374.x PMID: 9633105

16. Nei M. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA. 1973; 70: 3321–3323. https://doi.org/10.1073/pnas.70.12.3321 PMID: 4519626

17. Yeh FC, Boyle TJB. Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belgian Journal of Botany. 1997; 129: 157.

18. McDonald BA, McDermott JM. Population genetics of plant pathogenic fungi. Bioscience. 1993; 43: 311–319. https://doi.org/10.2307/1312063

19. Bradbury P J, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics. 2007; 23: 2633–2635. https://doi.org/10.1093/bioinformatics/btm308 PMID: 17586829

20. Pritchard J, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000; 155: 945–959. PMID: 10835412

21. Evanno G, Regnaut S, Goude J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005; 14: 2611–2620. https://doi.org/10.1111/j.1365- 294X.2005.02553.x PMID: 15969739

22. Earl DA, von Holdt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUC- TURE output and implementing the Evanno method. Conservation Genet. Resour. 2012; 4: 359. https://doi.org/10.1007/s12686-011-9548-7

23. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics. 2012; 28: 2537–2539. https://doi.org/10.1093/ bioinformatics/bts460 PMID: 22820204

24. Sokal R, Michener CA. statistical method for evaluating systematic relationships. Science bulletin. 1958; 38: 22.

25. Perrier X, Flori A, Bonnot F. Data analysis methods. In: Hamon P, Seguin M, Perrier X, Glaszmann JC (Eds) Genetic diversity of cultivated tropical plants. Enfield Science Publishers, Montpellier. 2003: pp 43–76.

26. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 1987; 4: 406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454 PMID: 3447015

27. Lopes MS, El-Basyoni I, Baenziger S, Singh S, Royo C, Ozbek K, et al. Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. Journal of Experimental Botany. 2015a; 66: 3477–3486. https://doi.org/10.1093/jxb/erv122 PMID: 25821073

28. Lopes MS, Dreisigacker S, Peña RJ, Sukumaran S, Reynolds MP. Genetic characterization of the wheat association mapping initiative (WAMI) panel for dissection of complex traits in spring wheat. Theor. Appl. Genet. 2015b; 128: 453–464. https://doi.org/10.1007/s00122-014-2444-2 PMID: 25540818

29. Berkman PJ, Visendi P, Lee HC, Stiller J, Manoli S, Lorenc MT, et al. Dispersion and domestication shaped the genome of bread wheat. Plant Biotechnol. J. 2013; 11: 564–71. https://doi.org/10.1111/pbi. 12044 PMID: 23346876

30. Trethowan R, Mujeeb-Kazi A. Novel germplasm resources for improving environmental stress tolerance of hexaploid wheat. Crop Sci. 2008; 48: 1255–1265. https://doi.org/10.2135/cropsci2007.08.0477

31. Jia J, Zhao S, Kong X, Li Y, Zhao G, He W, et al. Aegilops tauschii draft genome sequence reveals a gene repertoire for wheat adaptation. Nature. 2013; 496: 91–95. https://doi.org/10.1038/nature12028 PMID: 23535592

32. Novoselović D, Bentley AR, Simek R, Dvojković K, Sorrells ME, Gosman N, et al. Characterizing Croatian wheat germplasm diversity and structure in a European context by DArT markers. Front. Plant Sci. 2016. 7; 184. https://doi.org/10.3389/fpls.2016.00184 PMID: 26941756

33. El-Esawi MA, Witczak J, Abomohra AE, Ali HM, Elshikh MS, Ahmad M. Analysis of the genetic diversity and population structure of Austrian and Belgian wheat germplasm within a regional context based on DArT markers. Genes. 2018; 9: pii: E47. https://doi.org/10.3390/genes9010047 PMID: 29361778

34. Achtar S, Moualla MY, Kalhout A, Roder MS, MirAli N. Assessment of genetic diversity among Syrian durum (Triticum ssp. durum) and bread wheat (Triticum aestivum L.) using SSR markers. Russ. J. Genet. 2010; 46: 1320–1326. https://doi.org/10.1134/S1022795410110074

35. Le Couviour F, Faure S, Poupard D, Flodrops Y, Dubreuil P, Praud S. Analysis of genetic structure in a panel of elite wheat varieties and relevance for association mapping. Theor. Appl. Genet. 2011; 123: 715–727. https://doi.org/10.1007/s00122-011-1621-9 PMID: 21667038

36. Zhang DD, Bai GH, Zhu CS, Yu JM, Carver BF. Genetic diversity, population structure and linkage dis- equilibrium in U.S. elite winter wheat. The Plant Genome. 2010; 3: 117–127. https://doi.org/10.3835/ plantgenome2010.03.0004

37. Hao CY, Wang LF, Ge HM, Dong YC, Zhang XY. Genetic diversity and linkage disequilibrium in Chinese bread wheat (Triticum aestivum L.) revealed by SSR markers. PLoS ONE. 2011; 6: 1–13. https:// doi.org/10.1371/journal.pone.0017279 PMID: 21365016

38. Chen X, Min D, Yasir TA, Hu YG. Genetic diversity, population structure and linkage disequilibrium in elite Chinese winter wheat investigated with SSR markers. PLoS ONE. 2012; 7: e44510. https://doi. org/10.1371/journal.pone.0044510 PMID: 22957076

39. Chesnokov YV, Artemyeva AM. Evaluation of the measure of polymorphism information of genetic diversity. Agricultural Biology. 2015; 5: 571–578. https://doi.org/10.15389/agrobiology.2015.5.571eng

40. Flint-García SA, Thornsberry JM, Buckler ES. Structure of linkage disequilibrium in plants. Annual Review of Plant Biology. 2003; 54: 357–374. https://doi.org/10.1146/annurev.arplant.54.031902. 134907 PMID: 14502995

41. Laidò G, Marone D, Russo MA, Colecchia SA, Mastrangelo AM, De Vita P, et al. Linkage disequilibrium and genome-wide association mapping in tetraploid wheat (Triticum turgidum L.). PLoS ONE. 2014; 9 (4): e95211. https://doi.org/10.1371/journal.pone.0095211 PMID: 24759998

42. Chao S, Dubcovsky J, Dvorak J, Luo MC, Baezinge SP, Matnyazov R et al. Population- and genome- specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (Triticum aestivum L.). BMC Genomics. 2010; 11: 727. https://doi.org/10.1186/1471-2164-11-727 PMID: 21190581

43. Moragues M, Moralejo M, Sorrells ME, Royo C. Dispersal of durum wheat landraces across the Mediterranean basin assessed by AFLPs and microsatellites. Genetic Resources and Crop Evolution. 2007; 54: 1133–1144. https://doi.org/10.1007/s10722-006-9005-8

44. Oliveira HR, Campana MG, Jones H, Hunt HV, Leigh F, Redhouse D, et al. Tetraploid wheat landraces in the Mediterranean basin: taxonomy, evolution and genetic diversity. PLoS ONE. 2012; 7: e37063. https://doi.org/10.1371/journal.pone.0037063 PMID: 22615891

45. Oliveira HR, Hagenblad J, Leino MW, Leigh FJ, Lister DL, Peña-Chocarro L et al. Wheat in the Mediter- ranean revisited–tetraploid wheat landraces assessed with elite bread wheat Single Nucleotide Polymorphism markers. BMC Genetics. 2014; 15: 54. https://doi.org/10.1186/1471-2156-15-54 PMID: 24885044

46. Bonjean AP. French wheat pool. In: Bonjean AP, Angus WJ (eds) The world wheat book: a history of wheat breeding. Lavoisier Publishing, Paris. 2001; pp 127–165.

47. Royo C, Soriano JM, Álvaro F. Wheat: a crop in the bottom of the Mediterranean diet pyramid. In: Fuerst-Bjelis B (Eds) Mediterranean Identities—Environment, Society, Culture. Intechopen, London. 2017: pp 381–399. https://doi.org/10.5772/intechopen.69184

48. Sharma S, Upadhyaya HD, Varshney RK, Gowda CII. Pre-breeding for diversification of primary gene pool and genetic enhancement of grain legumes. Front. Plant Sci. 2013; 4: 309. https://doi.org/10. 3389/fpls.2013.00309 PMID: 23970889

49. Parzies HK, Spoor W, Ennos RA. Inferring seed exchange between farmers from population genetic structure of barley landrace Arabi Aswad from Northern Syria. Genet. Resour. Crop Evol. 2004; 51: 471–478. https://doi.org/10.1023/B:GRES.0000024157.67531.88

50. Ben-Romdhane M, Riah L, Selmi A, Jardak R, Bouajila A, Ghorbel A, et al. Low genetic differentiation and evidence of gene flow among barley landrace populations in Tunisia. Crop Sci. 2017; 57: 1585– 1593. https://doi.org/10.2135/cropsci2016.05.0298

Chapter 2:

Exploring the Genetic Architecture of Root-Related Traits in Mediterranean Bread Wheat Landraces by Genome-Wide Association Analysis

Rubén Rufo[1], Silvio Salvi[2], Conxita Royo[1] and Jose Miguel Soriano[1]

[1]Sustainable Field Crops Programme, IRTA (Institute for Food and Agricultural Research and Technology), 25198 Lleida, Spain.
[2]Department of Agricultural and Food Sciences, University of Bologna, Viale Fanin 44, 40127 Bologna, Italy.

Published in Agronomy (2020) 10, 613

doi: 10.3390/agronomy10050613

Exploring the Genetic Architecture of Root-Related Traits in Mediterranean Bread Wheat Landraces by Genome-Wide Association Analysis

1. Abstract

Background: Roots are essential for drought adaptation because of their involvement in water and nutrient uptake. As the study of the root system architecture (RSA) is costly and time-consuming, it is not generally considered in breeding programs. Thus, the identification of molecular markers linked to RSA traits is of special interest to the breeding community. The reported correlation between the RSA of seedlings and adult plants simplifies its assessment. Methods: In this study, a panel of 170 bread wheat landraces from 24 Mediterranean countries was used to identify molecular markers associated with the seminal RSA and related traits: seminal root angle, total root number, root dry weight, seed weight and shoot length, and grain yield (GY). Results: A genome-wide association study identified 135 marker-trait associations explaining 6% to 15% of the phenotypic variances for root related traits and 112 for GY. Fifteen QTL hotspots were identified as the most important for controlling root trait variation and were shown to include 31 candidate genes related to RSA traits, seed size, root development, and abiotic stress tolerance (mainly drought). Co-location for root related traits and GY was found in 17 genome regions. In addition, only four out of the fifteen QTL hotspots were reported previously. Conclusions: The variability found in the Mediterranean wheat landraces is a valuable source of root traits to introgress into adapted phenotypes through marker-assisted breeding. The study reveals new loci affecting root development in wheat.

Keywords: drought stress; association mapping; root system architecture; QTL hotspot; seminal root

2. Introduction

Wheat is the most widely cultivated crop in the world, covering around 219 million ha (Faostat 2017, http://www.fao.org/faostat/). It is a staple food for humans, as it provides 18% of daily human intake of calories and 20% of protein (http://www.fao.org/faostat/). Global wheat demand is estimated to increase by 60% by the year 2050 [1], so wheat production will need to rise by 1.7% per year until then. Achieving this objective is a great challenge under the current climate change scenario, as the prediction models estimate a precipitation decrease of 25% to 30% and a temperature increase of 4°C to 5°C for the Mediterranean region

[2]. It is well known that wheat production is greatly affected by environmental stresses such as drought and heat [3] that negatively affect yield and grain quality [4]. Drought is considered the greatest environmental constraint to yield and yield stability in rainfed production systems [5]. Environmental effects on yield in the Mediterranean Basin have been estimated at 60% for bread wheat [6] and 98% for durum wheat [7]. The expected effects of climate change and the declining availability of water and chemical fertilizers will require the release of cultivars with an enhanced genetic capacity to maintain acceptable yield levels and yield stability under harmful environmental conditions [8,9]. To cope with the challenges of climate change, breeders are particularly challenged to stretch the adaptability and performance stability of new cultivars, so many improvement programs are focusing on breeding for adaptation [10].

Plants respond and adapt to water deficit using various strategies that have evolved at several levels of function and are components of the conceptual framework developed by Reynolds et al. [11], which defines drought resistance in terms of dehydration escape, tolerance, and avoidance. Traits defining root system architecture (RSA) are critical for wheat adaptation to drought environments and non-optimal nutritional supply conditions [12]. Besides, water-use efficiency (WUE) can be significantly increased by optimizing the anatomy and growth features of roots [13]. Root traits are critical for drought tolerance due to its role in plant performance and the acquisition of nutrients and water from dry soils [14]. The wheat plant includes two types of roots: seminal (embryonal) and nodal (crown or adventitious or adult root system). The seminal roots are the first to penetrate the soil and remain functional during the whole plant cycle [9,15]. A correlation between seminal and adult roots in terms of size, dry-weight, or even specific architectural features have been reported [9,13]. Since the evaluation of RSA features in the field is very difficult, expensive, and time-consuming when a large number of genotypes need to be phenotyped, several studies have been carried out at early growth stages to allow an optimal screening of RSA traits [8,12,16–18]. Maccaferri et al. [9] observed that among RSA traits, those involving the root structure and related to the uptake of nutrients and water are root length, surface area and volume, and the number of roots, while root diameter is significantly associated with drought tolerance. Another RSA trait of interest in wheat is the seminal root angle (SRA), whose features suggest that narrow angles could lead to deeper root growth to obtain water from deeper soil layers and hence maintain higher yields [5,13].

Identifying quantitative trait loci (QTLs) and applying marker-assisted selection is of particular interest for RSA because the trait is important but difficult to phenotype. In the last few years, genome-wide association studies (GWAS) have become very popular because of their use of germplasm collections with wider variability than the classical bi-parental crosses. These collections allow many recombination

events to be detected, making the association between genotype and phenotype more accurate. Collections of landraces are an ideal subject of GWAS [19] since they are genetically diverse repositories of unique traits that have evolved in local environments characterized by a wide range of biotic and abiotic conditions. Several studies have shown that Mediterranean wheat landraces possess a wide genetic background for root architecture, yield formation, stress tolerance, and quality traits [17–22]. In the current study, a GWAS for three RSA traits and two related traits was performed on a panel of 170 bread wheat (*Triticum aestivum* L.) landraces from 24 Mediterranean countries with the following goals: (1) to detect differences in RSA among genetic subpopulations previously distinguished in the panel, (2) to identify correlations among RSA and grain yield under rainfed conditions, and (3) to identify molecular markers and candidate genes linked to root-related traits and candidate gene models for the associations.

3. Materials and Methods

3.1. Plant Material

A germplasm collection of 170 bread wheat (*Triticum aestivum* L.) genotypes from the MED6WHEAT IRTA panel described by Rufo et al. [23] was used in this study. The panel was genotyped and characterized using the Illumina Infinium 15K Wheat SNP Chip at Trait Genetics GmbH (Gatersleben, Germany), and markers were ordered according to the SNP map developed by Wang et al. [24]. The collection was previously structured into three subpopulations (SPs) matching their geographical origin [23]: western (SP1, WM), northern (SP2, NM), and eastern Mediterranean (SP3, EM) (Supplementary Materials, Table S1). Additionally, the cultivars 'Arthur Nick', 'Anza', 'Soissons', and 'Chinese Spring' were included as checks.

3.2. Root Morphology and Statistical Analysis

Root analysis was performed following the protocol described by Canè et al. [8], which was slightly modified in the current study (Figure 1). Ten representative seeds were randomly chosen from each genotype, weighed, sterilized in a 10% sodium hypochlorite solution for 5–10 min, washed thoroughly in distilled water and placed on hydrated filter paper in a 140 mm Petri dish at 28°C for 24 h. Subsequently, five seedlings were selected on the basis of a normal seminal root emergence and were spaced 8 cm from each other on a filter paper sheet placed on a vertical black rectangular polycarbonate plate (42.5 × 38.5 cm). Finally, each plate was covered with another wet sheet of filter paper. Distilled water was used for the plantlets' growth. The plantlets were grown in a growth chamber for 14 days at 22°C under a 16-h light photoperiod. In addition to the ten seed weight (SW), four other traits were scored for each genotype: total root number (TRN), shoot length (SL) from

the seed to the tip of the longest leaf and SRA, obtained using a digital camera following the methodology described in Canè et al. [8]. The images were processed with ImageJ software [25]. The angle between the two external roots of each plantlet was measured at a distance of 3.5 cm from the tip of the seed. Finally, the roots were desiccated at 70°C for 24 h to obtain the root dry weight (RDW).

The experimental design followed a randomized complete block with two replications in time. Means of five observational units for each genotype were used for TRN, RDW, and SL, while only three observational units were used for SRA because the two external ones were considered as border plantlets for root angle.

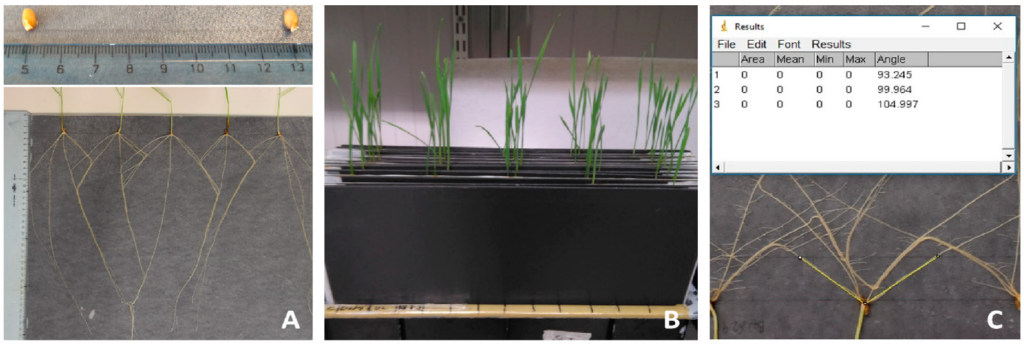

Figure 1. Experimental setup for the analysis of seminal root traits. Seeds were placed 8 cm apart on moist filter paper (A) and kept in a box with distilled water in a growth chamber for 14 days at 22 °C under a 16-h light photoperiod (B). (C) Example of seminal root angle measurement, using ImageJ software.

3.3. Grain Yield

Field experiments were carried out in 2016, 2017, and 2018 harvesting seasons in Gimenells, Lleida, north-east Spain (41°38' N and 0°22' E, 260 m a.s.l) under rainfed conditions. The experiments followed a non-replicated augmented design with two replicated checks (the cultivars 'Anza' and 'Soissons') and plots of 3.6 m^2. The experimental design is shown in Supplementary Materials, Figure S1. Sowing density was adjusted to 250 germinable seeds m^2. Weeds and diseases were controlled following standard practices at the site. The anthesis date was determined in each plot. Grain yield (GY, t ha^{-1}) was determined by mechanically harvesting the plots at maturity and expressed on a 12% moisture level.

3.4. Statistical Analysis

Phenotypic data for GY was fitted to a linear mixed model with the check cultivars as fixed effects and the row number, column number and cultivar as random effects

following the SAS PROC MIXED procedure:

$$y = X\beta + Z\gamma + \varepsilon \quad (1)$$

where β is an unknown vector of fixed-effects parameters with known design matrix X, γ is an unknown vector of random-effects parameters with known design matrix Z, and ε is an unknown random error vector whose elements are no longer required to be independent and homogeneous.

Restricted maximum likelihood was used to estimate the variance components and to produce the best linear unbiased predictors (BLUPs) for the traits of each cultivar and year with the SAS-STAT statistical package (SAS Institute Inc, Cary, NC, USA).

Analyses of variance (ANOVA) were performed for the root traits, considering the genotypes and the replication as random effects in the model. Additionally, for a subset of 55 of the 141 structured landraces, selected as having an SP membership q>0.8 (WM, 17; NM, 15; EM, 23), the sum of squares of the cultivar effect in the ANOVAs was partitioned into differences between SPs and differences within them. ANOVA for grain yield was performed for the complete collection, considering genotype, year, and the combination of genotype and year the sources of variation. Least squares means were calculated and compared using the Tukey HSD test at P<0.05. Pearson correlation coefficients among root traits were computed. Repeatability (H) was calculated on a mean basis across two replications following the formula described by Harper [26] $r=(B - W)/(B +((n -1)W))$, where n is the number of genotypes and B and W the two variances from the ANOVA table: between (B) and within (W). Frequency distributions, ANOVAs, the Tukey test, and the Pearson correlation coefficients were calculated using the JMP v13.1.0 statistical package (SAS Institute Inc, Cary, NC, USA).

3.5. Genome-Wide Association Analysis

A GWAS was performed for the mean of measured root traits and from the BLUPs for GY per year and across years with TASSEL 5.0 software [27]. A mixed linear model (MLM) was conducted using the information of the genetic structure reported in Rufo et al. [23] as the fixed effect and a kinship (K) matrix, calculated using Haploview [28], as the random effect (Q+K model) at the optimum compression level. In addition, the anthesis date was incorporated as a cofactor in the analysis. As reported in other studies [29–32], an adjusted –log10 P>3 was established as a threshold for considering a marker-trait association (MTA) statistically significant. A moderate threshold at –log10 P>2.5 was also established for GY. Confidence intervals (CI) for MTAs were calculated for each chromosome according to the linkage disequilibrium (LD) decay reported by Rufo et al. [23]. In order to simplify

the MTA information, the associations were grouped into QTL hotspots when at least two MTAs belonging to different traits overlapped their CIs. Circular Manhattan plots were performed using the R package "CMplot" (http://www.r-project.org).

3.6. Gene Annotation

Gene annotation within the CIs of the QTL hotspots was performed using the gene models for high-confidence genes reported for the wheat genome sequence [33], available at https://wheat-urgi. versailles.inra.fr/Seq-Repository/Assemblies. Markers flanking the CIs were used to estimate physical distances from genetic distances.

4. Results

4.1. Phenotypic Data of Root Traits

A summary of the genetic variation of the root traits is shown in Table 1. The genotypes showed a low coefficient of variation (CV) with a narrow range of variation among traits, from 10.4 for SW to 18.8 for RDW and repeatability (H) ranging from 48.5% for RDW to 75.4% for SW.

Table 1. Statistics of the seminal root traits.

Table	TRN (N)	RDW (mg)	SRA (°)	SW (g)	SL (cm)
Min	3.2	43	53.1	0.27	14.9
Max	5.4	20.5	125.9	0.63	30.2
Mean	4.4	11.7	98.6	0.48	22.0
SD	0.5	2.2	13.3	0.05	2.8
CV (%)	10.9	18.8	13.5	10.4	12.7
H (%)	52.0	48.5	70.0	75.4	50.0

SD, standard deviation; CV, coefficient of variation; H, repeatability; TRN, total root number; RDW, root dry weight; SRA, seminal root angle; SW, seed weight; SL, seed length.

The ANOVA (Table 2) for cultivars with a high membership coefficient ($q>0.8$) showed that for all traits the total variability was mainly explained by the genotype effect, in a range from 63.5% for SL to 88.8% for SW. When the sum of squares of the genotype effect was partitioned into differences between and within SPs, the results revealed that the genetic variability was mainly explained by differences within SPs in a range from 47.8% for TRN to 71.8% for SRA (Table 2). Differences between SPs were statistically significant for SRA, TRN, SW, and SL, in a range from 6.0% of the genotype effect for SL to 25.3% for TRN (Table 2). The sum of squares within SPs was partitioned into western (WM), northern (NM), and eastern (EM) effects, being statistically significant for SRA (40.8%), TRN (28.3%), SW (38.3%),

and SL (26.8%) in the western SP, SRA (34.4%) in the northern SP and RDW (53.6%) and SW (55.4%) in the eastern SP.

Table 2. Percentage of the sum of squares of the ANOVA in a set of 55 bread wheat landraces structured into three genetic subpopulations with membership coefficient q>0.8.

Source of variation	*df*	TRN	RDW	SRA	SW	SL
Replicate	1	0.1	0.7	0.0	0.1	0.3
Genotype	54	73.1***	64.8*	82.1***	88.8***	63.5*
Between SPs	2	25.3***	0.6	10.3***	18.9***	6.0*
Within SPs	52	47.8*	64.2*	71.8***	69.9***	57.5*
WM	16	28.3***	31.6	40.8***	38.3***	26.8*
NM	14	16.1	14.8	34.4***	6.3	25.2
EM	22	55.6	53.6*	24.8	55.4***	48.0
Replicate x Genotype	54	26.8	34.5	17.9	11.1	36.2
Total	108					

WM, western Mediterranean; NM, northern Mediterranean; EM, eastern Mediterranean; TRN, total root number; RDW, root dry weight; SRA, seminal root angle; SW, seed weight; SL, seed length. *P<0.05, ***P<0.001.

The ANOVA for grain yield revealed that the genotype effect was the most important in the phenotypic expression of traits, accounting for 59% of the total phenotypic variation, whereas the year effect accounted only for 5%. The interaction accounted for almost 36% of the phenotypic variation although it was not significant (Table 3).

Table 3. Percentage of the sum of squares for grain yield of the ANOVA in the collection of 170 bread wheat landraces.

Source of variation	*df*	Grain yield	*P*
Genotype	169	59.2	< 0.001
Year	2	5.1	< 0.001
Genotype x Year	338	35.7	No significant
Total	509		

The landraces from northern Mediterranean countries showed the highest number of seminal roots with a root angle not statistically different from the western Mediterranean ones. On the other hand, eastern Mediterranean landraces showed the lowest number of roots but the widest angle. These landraces reported the lowest SW and the longest shoots. No differences were reported for RDW among the three SPs (Table 4).

Correlation coefficients between root traits were calculated, showing highly significant correlation coefficients between RDW and SW and RDW and SL (r=0.47 and 0.45 respectively; P<0.0001). Moderate significant correlations were reported

Table 4. Means comparison of seminal root traits measured in a set of 55 Mediterranean wheat landraces structured into three genetic subpopulations [23] with q>0.8. Means within columns with different letters are significantly different at P<0.05 following a Tukey test.

	TRN (N)	RDW (mg)	SRA (°)	SW (g)	SL (cm)
Northern Mediterranean	4.7 a	0.011 a	98.5 b	0.50 a	20.8 b
Western Mediterranean	4.3 b	0.011 a	96.2 b	0.49 a	21.4 ab
Eastern Mediterranean	4.0 c	0.011 a	106.5 a	0.45 b	22.5 a

TRN, total root number; RDW, root dry weight; SRA, seminal root angle; SW, seed weight; SL, seed length.

for TRN with RDW, SW and SRA (r=0.20, 0.28 and 0.28, respectively), and for SW with SL (r=0.27). Finally, a negative correlation coefficient (r=−0.12) was found between SRA and SW (Figure 2). GY showed a moderate significant correlation with TRN and SW (r=0.28 and 0.29, respectively; P<0.0005).

4.2. Marker-Trait Associations

After filtering for duplicated patterns, missing values, and minor frequency alleles,

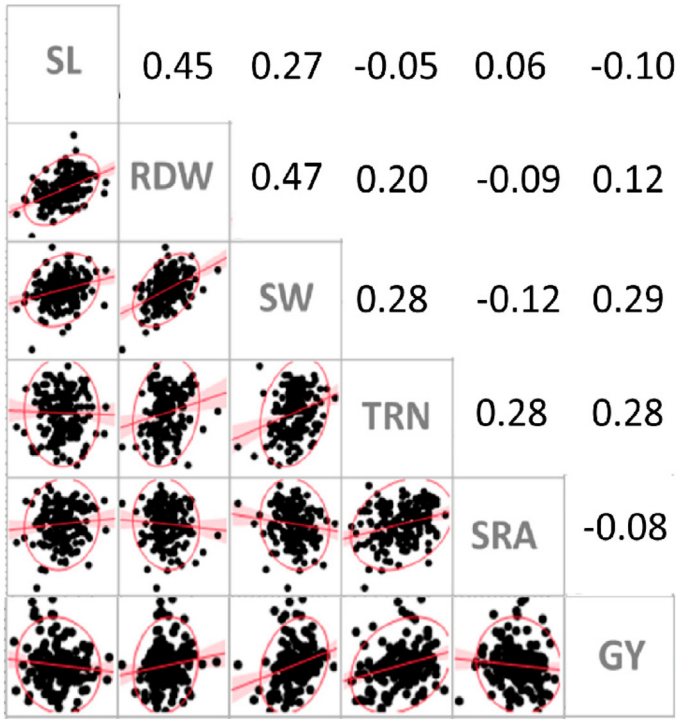

Figure 2. Correlations between seminal root traits and grain yield. On the right side are shown the values of the correlation coefficients (r). SL, seed length; RDW, root dry weight; SW, seed weight; TRN, total root number; SRA, seminal root angle; GY, grain yield.

a total of 10,458 SNPs were used to genotype the panel of 170 wheat landraces [23].

The results of the GWAS for root related traits are reported in Figure 3 and Supplementary Materials, Table S2. Using a common threshold of −log10 P>3, as reported by other authors [29–32], a total of 135 MTAs were identified for the analyzed traits. Of these, 50 MTAs corresponded to SW, 39 to RDW, 18 to SL, 17 to SRA, and 11 to TRN. The A and B genomes harbored 46% and 48% of MTAs, respectively, whereas the D genome harbored only 6% of MTAs. The number of MTAs per chromosome ranged from 1 in chromosomes 4D, 5D, and 6D to 14 in chromosome 1B, with a mean of 7 MTAs per chromosome. Most of the MTAs (88%) showed a phenotypic variance explained (PVE) by each MTA in a range of 5% to 10%, and only 2% showed a PVE higher than 15%. Among traits, the PVE mean was stable in a range of 7% (SL) to 9% (RDW).

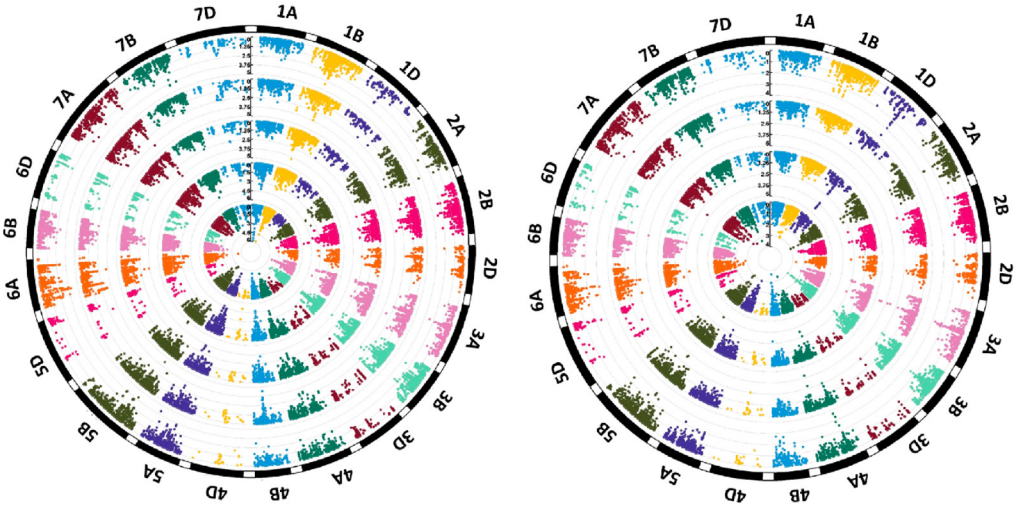

Figure 3. GWAS for root related traits (left circle) and grain yield for 3 years and across years (right circle). From the inside out, root traits correspond to RDW, SW, TRN, SRA, and SL, whereas for GY corresponds to 2016, 2017, 2018 harvesting seasons and the mean across years.

In order to identify and summarize the genomic regions most involved in trait variation, QTL hotspots were defined when two or more MTAs from different traits were grouped together within the same LD block. LD was previously estimated for locus pairs in each chromosome, and its decay was set to 1 to 10 cM depending on the chromosome [23]. Using this approach, 15 QTL hotspots grouping 43 MTAs were identified (Table 5), while 92 MTAs remained as singletons.

The results of the GWAS for GY are reported in Figure 3 and Supplementary Materials, Table S3. A common threshold of −log10 P>3, detected a total of 40

MTAs, thus a moderate threshold at –log10 P>2.5 was applied, increasing the number of significant associations to 112. Of these, 32 MTAs corresponded to the year 2016, 30 to 2017, 18 to 2018, and 32 across years. The A and B genomes harbored 43% and 38% of MTAs, respectively, whereas the D genome harbored only 18% of MTAs. The number of MTAs per chromosome ranged from 1 in chromosomes 3D and 6B to 16 in chromosome 1D. Chromosomes 1A, 4D, 5D, and 7D did not show any association. All of MTAs showed a phenotypic variance explained (PVE) by each MTA in a range from 5% to 11%. Most of the MTAs with a PVE>8% were located on chromosome 1D (76%, 13 out of 17), whereas the percentage increased to 80% among MTAs with a PVE>10% (4 out of 5).

In order to identify and summarize the genomic regions with a pleiotropic effect for root traits and grain yield, QTL hotspots were defined as previously but including the MTAs for GY. Using this approach, 17 QTL hotspots grouping 81 MTAs were identified (Table 6). From them, five were in common with those reported only with root traits (rootQTL1B.3, rootQTL2A.2, rootQTL3B.2, rootQTL6A.1, and rootQTL6A.2). GY shared 8 genomic regions with SW and 9 with RDW, 4 with SL, and 3 with SRA, whereas no regions were in common with TRN. In 59% of these genomic regions, GY co-localize with only one root trait, whereas the other 41% co-localize with two different root traits.

Table 5. Root QTL hotspots. Positions are indicated in cm.

QTL hotspot	MTAs	Trait	Chromosome	Peak	CI left	CI right
root QTL1B.1	2	RDW, SW	1B	70.6	69.6	71.6
root QTL1B.2	2	RDW, TRN	1B	77.5	75.9	79.1
root QTL1B.3	2	RDW, SL	1B	83.0	81.4	84.6
root QTL2A.1	2	RDW, SW	2A	47.8	46.7	48.9
root QTL2A.2	2	RDW, SW	2A	104.1	103.6	104.6
root QTL2A.3	2	SW, SL	2A	177.5	176.9	178.2
root QTL2B.1	2	SW, SL	2B	109.5	109.0	110.0
root QTL3B.1	3	RDW, SW, TRN	3B	62.3	61.4	63.2
root QTL3B.2	2	SRA, SW	3B	80.6	79.6	81.5
root QTL5A.1	2	RDW, SL	5A	56.5	56.0	57.0
root QTL5B.1	2	RDW, SL	5B	95.7	94.5	96.9
root QTL6A.1	2	RDW, SL	6A	45.8	40.0	51.6
root QTL6A.2	2	RDW, TRN	6A	76.7	70.7	82.6
root QTL6A.3	2	RDW, SW	6A	138.4	132.3	144.6
root QTL7A.1	3	RDW, SRA, TRN	7A	216.6	215.3	218.0

TRN, total root number; RDW, root dry weight; SRA, seminal root angle; SW, seed weight; SL, seed length.

In order to identify the most useful markers for selecting for the root traits, extreme phenotypes were identified in the upper and lower 10th percentile of genotypes within the collection for each trait (Figure 4). Among the most significant

Table 6. QTL hotspots including grain yield. Positions are indicated in cm.

QTL hotspot	MTAs	Trait	Chromosome	Peak	CI left	CI right
QTL yield/root_1B.1	3	GY, SRA	1B	8.4	7.4	9.4
QTL yield/root_1B.2	3	GY, SW	1B	43.9	42.9	44.9
QTL yield/root_1B.3	3	GY, SW	1B	63.5	61.5	65.5
QTL yield/root_1B.4	8	GY, RDW, SL	1B	83.3	82.3	84.3
QTL yield/root_2A.1	8	GY, RDW, SW	2A	104.1	103.6	104.6
QTL yield/root_2A2	3	GY, SRA	2A	151.3	150.8	151.9
QTL yield/root_3A.1	8	GY, RDW, SL	3A	84.3	81.9	86.7
QTL yield/root_3B.1	5	GY, SW	3B	72.8	70.8	74.8
QTL yield/root_3B.2	3	GY, SRA, SW	3B	80.5	79.6	81.5
QTL yield/root_4B.1	4	GY, SW	4B	76.6	74.1	79.2
QTL yield/root_5B.1	3	GY, SL	5B	57.8	56.8	58.9
QTL yield/root_5B.2	7	GY, RDW	5B	77.3	75.9	78.8
QTL yield/root_5B.3	2	GY, RDW	5B	176.2	175.2	177.2
QTL yield/root_6A.1	5	GY, RDW, SL	6A	45.6	39.6	51.6
QTL yield/root_6A.2	10	GY, RDW, SW	6A	76.6	70.7	82.6
QTL yield/root_7A.1	3	GY, RDW, SW	7A	135	133.2	136.8
QTL yield/root_7B.1	3	GY, RDW	7B	70.0	67.8	72.3

TRN, total root number; RDW, root dry weight; SRA, seminal root angle; SW, seed weight; SL, seed length; GY, grain yield.

MTAs for each trait, markers with different alleles between extreme genotypes were identified (Table 7, Figure 5). The frequency of the most common allele among genotypes from the upper 10th percentile ranged from 78% for RDW to 88% for SW, while for the lower 10th percentile it ranged from 65% for TRN and SRA to 92% for RDW (Figure 5).

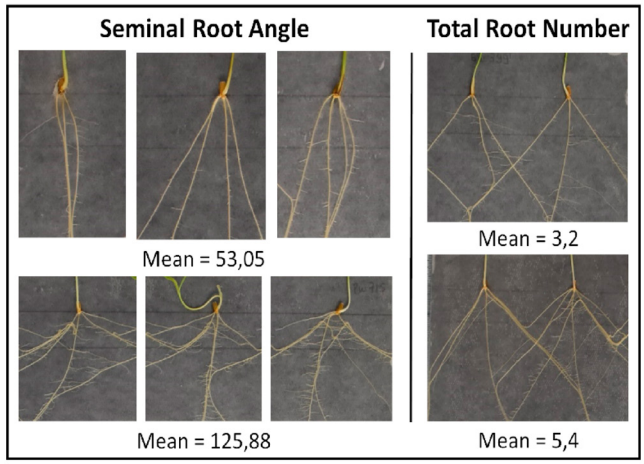

Figure 4. Extreme phenotypes for SRA and TRN. The means correspond for 3 observational units of the genotype for SRA and 5 observational units of the genotype for TRN.

4.3. Gene Annotation

As reported in Supplementary Materials, Table S4, a total of 1489 gene models

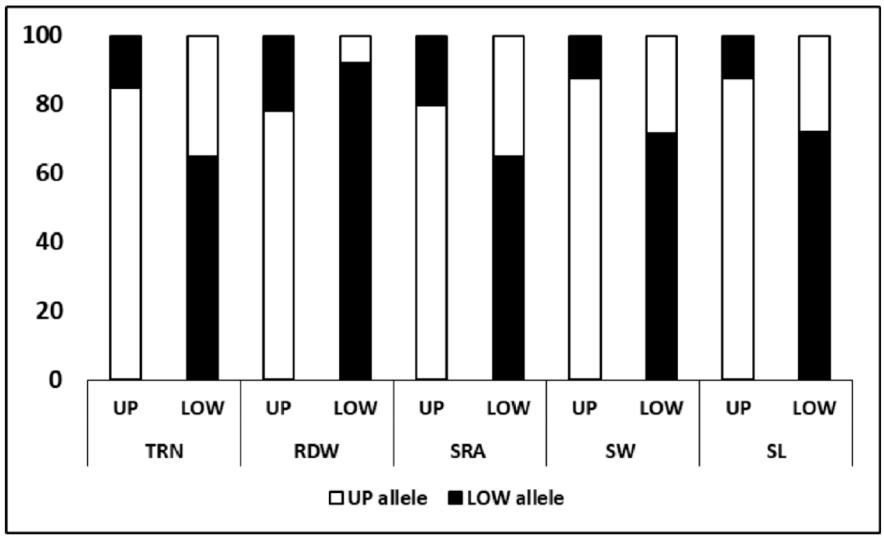

Figure 5. Marker allele frequency means from landraces within the upper and lower 10th percentile for the analyzed traits. All significant markers shown in Table 5 are included. TRN, total root number; RDW, root dry weight; SRA, seminal root angle; SW, seed weight; SL, seed length.

were identified within the 15 QTL hotspots using the high-confidence gene annotation from the wheat genome sequence [33]. Genetic distances were converted into physical distances using the position of common flanking markers on the genetic map [24] and the genome sequence. The number of gene models ranged from 224 in rootQTL_2A.2 to 9 in rootQTL_5B.1. Based on the high number of gene models, a selection was made according to gene families involved in root traits, growth and development, and abiotic stress resistance (Table 8). Thus, 31 gene families with a total of 96 gene models remained for subsequent analysis. Among them, F-box and zinc finger family proteins were identified in 12 of the 15 QTL hotspots, whereas 10 gene families were present in only one QTL hotspot. Among chromosomes with QTL hotspots, chromosome 2A had the highest number of gene models (22), whereas chromosomes 5A and 5B had the lowest number (4).

5. Discussion

Breeding for drought adaptation is one of the main challenges to be addressed in the coming years in order to increase wheat production and ensure sufficient food supply in the current scenario of climate change. Roots are crucial in this adaptation, as they are responsible for water and nutrient uptake. The wide morphological plasticity of the root system to different soil conditions and the role of root traits in drought environments are well known [34,35]. Wheat roots reduce their growth in water-limited conditions but increase the water uptake rate, extracting the water

Table 7. Selected significant markers from the GWAS with different allele composition for the upper (UP) and lower (LOW) 10th percentile of genotypes. Different letters on the UP and LOW 10th phenotype indicate that means are significantly different at P<0.01 following a Tukey test.

Trait	Phenotype			Marker	Chr	Position	R^2	Most Frequent Allele			
	Mean	UP 10th	LOW 10th					UP	Freq	LOW	Freq
TRN (N)	4.4	5.2[a]	3.4[b]	Excalibur_c8613_1266	1B	78.1	0.09	T	0.70	C	0.80
				Tdurum_contig12995_792	5B	98.3	0.12	T	1.00	Y	0.50
RDW (g)	0.012	0.016[a]	0.007[b]	RAC875_c8482_160	1A	95.2	0.14	A	0.60	A	0.90
				GENE-0392_97	1A	95.5	0.11	T	0.60	C	0.70
				RAC875_c39401_502	1A	104.1	0.15	T	1.00	C	1.00
				BS00023071_51	1B	95.5	0.15	G	1.00	A	1.00
				Kukri_c14877_303	6A	79.1	0.14	C	0.70	T	1.00
SRA (º)	98.65	119.98[a]	69.88[b]	Kukri_c34453_332	4B	70.9	0.11	A	0.60	C	0.70
				BS00028183_51	5B	51.0	0.11	T	1.00	C	0.60
SW (g)	0.487	0.59[a]	0.37[b]	RAC875_c37540_583	2B	139.5	0.14	T	1.00	C	0.77
				Ku_c101046_1063	4B	77.7	0.11	Y	0.50	C	0.80
				BS00094770_51	4D	80.4	0.12	C	1.00	A	0.60
				GENE-2220_165	7D	161.1	0.12	T	1.00	T	0.70
SL (cm)	22.04	27.72[a]	16.80[b]	Tdurum_contig8158_269	1B	82.4	0.09	A	0.80	G	0.90
				RFL_Contig5277_888	2A	177.6	0.09	C	1.00	C	0.60
				BobWhite_c6966_236	5A	130.9	0.09	T	0.80	C	0.90
				BS00023161_51	5B	48.3	0.08	T	0.90	Y	0.50
				RAC875_rep_c106439_1159	6A	48.1	0.09	A	0.88	C	0.70

Chr, chromosome. TRN, total root number. RDW, root dry weight. SRA, seminal root angle. SW, seed weight. SL, shoot length.

Table 8. Selected gene model families

Description	N QTL hotspots	Function
F-box family protein	12	Salt and drought stress responses
RING/FYVE/PHD zinc finger protein	12	Salt and drought stress responses
MYB-related transcription factor	8	Salt and drought stress responses
NAC domain-containing proteins	8	Induced by biotic and abiotic stresses
Cytochrome P450 family protein	5	Involved in seed size
BZIP transcription factor	5	Regulated by abiotic stress
Ethylene-responsive transcription factor	4	Induced by biotic and abiotic stresses
Calmodulin	4	Heat shock transduction pathway
Peroxidase	4	Root growth
ABC transporter	4	Control root development
Nucleoside triphosphate hydrolase	3	Associated with drought stress
E3 ubiquitin-protein ligase	3	Associated with drought stress
Glycine-rich protein	2	Enhance drought stress tolerance
Xyloglucan endotransglucosylase/hydrolase	2	Response dehydration, salinity, cold
Aquaporin	2	Drought stress tolerance
Expansin protein	2	Drought tolerance in wheat
Trihelix transcription factor	2	Stomatal development, drought
VQ motif family protein	2	Involved in seed size
Heat shock family protein	2	Induced by abiotic stress
Protein root UVB sensitive 6	2	Early seedling morphogenesis
SAUR-like auxin-responsive family protein	2	Maintain growth during abiotic stress
Bax inhibitor-1 family protein	1	Tolerance to abiotic stresses
Formin-like protein	1	Structure organization in drought stressed plants
Late embryogenesis abundant protein	1	Participate in drought response
Cell wall invertase	1	Downregulated by drought
Senescence regulator	1	Related to drought stress
Plastid-lipid associated protein PAP/fibrillin	1	Induced by drought
Protein STAY-GREEN LIKE, chloroplastic	1	Improves drought resistance
PI-PLC X domain-containing protein	1	Induced by abiotic stresses
Histidine-containing phosphotransfer protein	1	Enhance tolerance to drought stress
Phospholipase D	1	Enhance drought stress tolerance

from deep soil layers [36]. The shape and spatial arrangement of the RSA can provide a growth advantage and increasing yield performance during periods of water scarcity [37]. Thus, it is necessary to increase the knowledge of the genetics of root architecture in order to improve wheat yield stability under stress conditions by introgressing favorable alleles through breeding programs.

The current study evaluated root-related traits in a collection of Mediterranean bread wheat landraces representative of the variability existing for the species in the Mediterranean Basin [23] with the aim of providing QTL information for these traits regarding seminal roots. Seminal roots are important for early vigor and crop establishment in dryland areas because they explore the soil for nutrients and water [38]. Moreover, it has been reported that under drought stress, seminal roots activity is more important than that of nodal roots [39]. Additionally, field phenotyping of hundreds of genotypes is a complex and expensive task. As the root geometry of adult plants is strongly related to the SRA [5], it may be assumed that genotypes that

differ in root architecture at an early developmental stage would also differ in the field at later growth stages, when nutrient and/or water capture become critical for yield performance [8].

The range of variation for the traits analyzed in the present study (from 10.9% for TRN to 18.8% for RDW) is in agreement with those reported for elite durum wheat cultivars by Canè et al. [8], who explained this variability as an adaptive value for the environmental conditions of the region of origin of the cultivars. Moreover, the high repeatability found for the traits supports the approach followed to analyze the seminal roots under controlled conditions.

Landraces from the eastern Mediterranean Basin showed the widest SRA, the lowest SW, the longest SL, and the lowest number of roots. According to previous studies in durum wheat [18,40], landraces from southeastern Mediterranean countries corresponding to the warmest and driest areas of the Mediterranean Basin, reported more grains per unit area and lighter grains than those developed in cooler and wetter zones of the region. Although it has been reported that in water-limited environments a vigorous root system could have benefits at the beginning of the growing season because it offers a more efficient water capture [41], no significant differences were observed for RDW among the SPs in the current study. Moreover, our results for SRA are in agreement with those reported by Roselló et al. [18], who found that durum wheat landraces from the eastern Mediterranean have the widest root angle, which probably allows them to cover a larger soil area and be more efficient in water uptake than landraces that originated in wetter areas.

Although not significant, probably due to the very early stage when the root traits were measured, the correlation between SRA and SW was negative. The same result was also reported by Canè et al. [8], who suggested that it could be due to the influence of the root angle on the distribution of the roots on soil layers and, therefore, the water uptake from deeper layers. On the other hand, the correlation between RDW and SW was positive, in agreement with the findings of Fang et al. [42], thus indicating the effectiveness of greater root mass for obtaining more soil water for plant growth and grain filling in drought. Seedling growth has also been related to SW in wheat [43]. The vertical distribution of the root system can have a strong effect on yield [44], so mass root concentrated in upper layers can be more effective for resource capture, while roots in deeper layers have more access to deep water.

The complexity of the genetic control of root traits was confirmed with 135 marker-trait associations identified in the current study. Their distribution across genomes was similar in the A and B genomes (46% and 48%, respectively), leaving only 6% of MTAs in the D genome. These results agree with the lower genetic diversity and higher LD found in the D genome, as reported previously [23].

According to Chao et al. [45], the different levels of diversity in wheat genomes could be due to different rates of gene flow from the ancestors of wheat, since polyploidy bottleneck resulting from speciation reduced diversity and increased the levels of LD in the D genome in comparison with the A and B genomes.

In order to simplify and to integrate closely linked MTAs in a consensus region, QTL hotspots were identified based on the results of LD decay reported in [23]. LD decay was used to define the CIs for the QTL hotspots. Following this approach, 43 MTAs were grouped in 15 QTL hotspots. The genomic position of QTL hotspots was compared with previous studies reporting meta-QTLs for root traits [46] and MTAs from GWAS studies in order to detect previously identified regions controlling root traits. Among the 15 QTL hotspots, only rootQTL6A.3 was located in the same region of a previously mapped meta-QTL, RootMQTL74 [46]. When compared with MTA-QTLs reported by [18] in durum wheat Mediterranean landraces, the QTL hotspot rootQTL6A.3 corresponded to the MTA-QTLs mtaq-6A.3 and mtaq-6A.6. This hotspot was also in the same region of a major SRA QTL identified by Alahmad et al. [47] and by a QTL controlling root growth angle identified by Maccaferri et al. [9], who also found a QTL for grain weight that is located in a common region with the hotspot rootQTL2A.2, which includes an MTA for SW. rootQTL3B.1 shared a common position with an MTA reported by Ayalew et al. [48] on chromosome 3B under stress conditions. rootQTL7A.1, including an MTA for RDW, was located in a similar position as MLM-RDWB-10 reported by Li et al. [49] and associated with RDW at the booting stage. Finally, no genomic regions were shared with the study carried out by Beyer et al. [50]. Only four of the 15 QTL hotspots identified in this work had been detected previously, suggesting the importance of wheat Mediterranean landraces for the identification of new loci controlling root-related traits.

As reported in previous studies, at early developmental stages [8,18] the co-location of MTAs for grain yield and root related traits within the same QTL hotspot suggests their pleiotropic effect, however, deeper analyses should be necessary to confirm it. In durum wheat elite cultivars, Canè et al. [8] found that 30% of the QTLs affecting root system architecture were included within QTLs for agronomic traits. More recently, Roselló et al. [18] using a collection of Mediterranean durum wheat landraces found that 45% of QTL hotspots for root related traits were mapped in similar regions to yield-related traits reported for the same collection of landraces.

From a breeding standpoint, exploiting genetic diversity from local landraces is a valuable approach for recovering and broadening allelic variation for traits of interest [19]. Therefore, identifying the genotypes showing the extreme phenotypes within the pool of Mediterranean landraces and the associated markers provide the opportunity for introgressing suitable traits in elite cultivars by marker-assisted

breeding using the most recent technologies to speed the process.

The availability of a high-quality reference wheat genome sequence [33] enabled us to quickly identify gene models corresponding to QTLs. Thus, the genetic position of the CIs of the QTL hotspots was projected into physical distances on the reference sequence to search for putative candidate gene models. To narrow the number of candidates, only gene models involved in the development and abiotic stress according to the literature were taken into consideration. Therefore, of 1489 gene models identified within the 15 QTL hotspots, only 31 gene families were selected.

F-box and zinc finger family proteins were the most represented, each one appearing in 12 hotspots. F-box proteins play important roles in plant development and abiotic stress responses via the ubiquitin pathway [51] and the ABA signaling pathway [52]. In wheat, the F-box protein TaFBA1 is involved in plant hormone signaling and response to abiotic stresses and is expressed in all plant organs, including roots [53]. The overexpression of TaFBA1 in transgenic tobacco reported by Li et al. [54] to improve heat tolerance resulted in increased root length in the transgenic plants. Zinc finger proteins are involved in several processes, such as regulation of plant growth and development, and response to abiotic stresses [46]. In Arabidopsis and rice, they play a role in tolerance to drought and salt stresses [55], while in wheat the overexpression of TaZFP34 enhances root-to-shoot ratio during plant adaptation to drying soil [56].

Other kinds of gene models found in a high number of QTL hotspots were MYB transcription factors and NAC domain-containing proteins, each of them presents in 8 hotspots. MYB domain-containing transcription factors are involved in salt and drought stress adaptation in wheat. Some examples in wheat are the genes TaMyb1, TaMYBsdu1, and TaMYB33. The expression of TaMyb1 in roots is strongly related to responses to abiotic stresses [57]. The gene TaMYBsdu1 was found to be upregulated in leaves and roots of wheat under long-term drought stress [58]. Finally, the overexpression of TaMYB33 in Arabidopsis enhances tolerance to drought and salt stresses [59]. NAC domain-containing proteins have been described to play many important roles in abiotic stress adaptation [46]. Xie et al. [60] reported that NAC1 promoted the development of lateral roots. Similarly, He et al. [61] found that the expression of AtNAC2 in response to salt stress led to an increase in the development of lateral roots. Xia et al. [62] demonstrated that the gene TaNAC4 is a transcriptional activator involved in wheat's response to biotic and abiotic stresses.

Proteins belonging to the cytochrome P450 family and bZIP transcription factors were present in five QTL hotspots. The first class of proteins belongs to one of the largest families of plant proteins, with genes affecting important traits

for crop improvement such as TaCYP78A3, which is involved in the control of seed size [63]. bZIP transcription factors are involved in abiotic stress response [64]. In Arabidopsis, it has been observed that the overexpression of TabZIP14-B, involved in salt and freezing tolerance, hindered root growth in transgenic plants in comparison with the control plants [65].

Other proteins involved in root growth and development are the peroxidases and ABC transporters that were identified in four QTL hotspots. Extracellular peroxidases are involved in plant defense reactions against biotic and abiotic stresses through the generation of reactive oxygen species in wounded root cells [66]. In Arabidopsis, the ABC transporter AtPGP4 is expressed mainly during early root development, and its loss of function enhances lateral root initiation and root hair development [67]. Gaedeke et al. [68] reported a new member of the ABC transporter superfamily of Arabidopsis thaliana, AtMRP5. Using reverse genetics, these authors found that the recessive allele mrp5 exhibited decreased root growth and increased lateral root formation. In addition to peroxidases and ABC transporters, other proteins identified in four QTLs were the ethylene-responsive transcription factors (ERFs), found to be involved in the response to abiotic stresses. In wheat, the ERF TaERFL1a is induced in wheat seedlings in response to salt, cold, and water deficiency [69].

Other family proteins involved in drought stress, seed size, or early development were represented in a lower number of QTL hotspots. Among them, aquaporins are known to affect drought tolerance influencing the capacity of roots to take up the soil water [70]. The expansins were suggested to be involved in root development, as the overexpression of the wheat expansin TaEXPB23 improved drought tolerance by stimulating the growth of the root system in tobacco [71].

6. Conclusions

The exploitation of unexplored genetic variation present in local landraces can potentially contribute to breeding programs aimed at enhancing drought tolerance in wheat. Roots are crucial for adaptation to drought stress because they are the plant organ responsible for water and nutrient uptake and interaction with soil microbes. Thus, designing and developing novel root system ideotypes could be one of the targets of wheat breeding for the coming years. The variability found in the Mediterranean wheat landraces together with the newly identified QTL hotspots shows landraces as a valuable source of favorable root traits to introgress into adapted phenotypes through marker-assisted breeding. Among the different marker trait associations, those reported in extreme genotypes could result as a starting point to develop new mapping populations to fine map the corresponding traits.

Supplementary Materials: Supplementary materials can be found at http://www.mdpi.com/2073-4395/10/5/613/s1. Figure S1: Field scheme of the experimental

design for grain yield, Table S1: List of accessions, Table S2: Significant GWAS results for root related traits, Table S3: Significant GWAS results for grain yield, Table S4: Gene models identified within the 15 root QTL hotspots.

Author Contributions: Conceptualization, S.S. and J.M.S.; Data curation, R.R.; Formal analysis, R.R.; Funding acquisition, C.R. and J.M.S.; Investigation, R.R., S.S. and J.M.S.; Methodology, R.R., S.S. and J.M.S.; Project administration, C.R. and J.M.S.; Supervision, S.S. and J.M.S.; Writing—original draft, R.R.; Writing—review & editing, S.S., C.R. and J.M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by projects AGL2015-65351-R and RTA2015-00072-C03-01 of the Spanish Ministry of Economy and Competitiveness. R.R. is a recipient of a PhD grant from the Spanish Ministry of Economy and Competitiveness.

Acknowledgments: The authors acknowledge the contribution of the CERCA programme (Generalitat de Catalunya). Thanks are given to the group of Agricultural Genetics of University of Bologna for technical support.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

EM	Eastern Mediterranean
GWAS	Genome wide association study
GY	Grain yield
MTA	Marker-trait association
NM	Northern Mediterranean
QTL	Quantitative trait locus
RDW	Root dry weight
RSA	Root system architecture
SL	Shoot length
SP	Sub-population
SRA	Seminal root angle
SW	Seed weight
TRN	Total root number
WM	Western Mediterranean

7. References

1. Leegood, R.C.; Evans, J.R.; Furbank, R.T. Food security requires genetic advances to increase farm yields. Nature 2010, 464, 831.

2. Giorgi, F.; Lionello, P. Climate change projections for the Mediterranean region. Glob. Planet. Chang. 2008, 63, 90–104.

3. Uga, Y.; Kitomi, Y.; Ishikawa, S.; Yano, M. Genetic improvement for root growth angle to enhance crop production. Breed. Sci. 2015, 65, 111–119.

4. Kulkarni, M.; Soolanayakanahally, R.; Ogawa, S.; Uga, Y. Drought Response in Wheat: Key Genes and Regulatory Mechanisms Controlling Root System Architecture and Transpiration Efficiency. Front. Chem. 2017, 5, 1–13.

5. Manschadi, A.M.; Hammer, G.L.; Christopher, J.T.; de Voil, P. Genotypic variation in seedling root architectural traits and implications for drought adaptation in wheat (Triticum aestivum L.). Plant Soil 2008, 303, 115–129.

6. Sanchez-Garcia, M.; Álvaro, F.; Martín-Sánchez, J.A.; Sillero, J.C.; Escribano, J.; Royo, C. Breeding effects on the genotype×environment interaction for yield of bread wheat grown in Spain during the 20th century. Field Crop. Res. 2012, 126, 79–86.

7. Royo, C.; Maccaferri, M.; Álvaro, F.; Moragues, M.; Sanguineti, M.C.; Tuberosa, R.; Maalouf, F.; del Moral, L.F.G.; Demontis, A.; Rhouma, S.; et al. Understanding the relationships between genetic and phenotypic structures of a collection of elite durum wheat accessions. Field Crop. Res. 2010, 119, 91–105.

8. Canè, M.A.; Maccaferri, M.; Nazemi, G.; Salvi, S.; Francia, R.; Colalongo, C.; Tuberosa, R. Association mapping for root architectural traits in durum wheat seedlings as related to agronomic performance. Mol. Breed. 2014, 34, 1629–1645.

9. Maccaferri, M.; El-Feki, W.; Nazemi, G.; Salvi, S.; Canè, M.A.; Colalongo, M.C.; Stefanelli, S.; Tuberosa, R. Prioritizing quantitative trait loci for root system architecture in tetraploid wheat. J. Exp. Bot. 2016, 67, 1161–1178.

10. Bhatta, M.; Morgounov, A.; Belamkar, V.; Baenziger, P. Genome-Wide Association Study Reveals Novel Genomic Regions for Grain Yield and Yield-Related Traits in Drought-Stressed Synthetic Hexaploid Wheat. Int. J. Mol. Sci. 2018, 19, 3011.

11. Reynolds, M.P.; Mujeeb-Kazi, A.; Sawkins, M. Prospects for utilising plant-adaptive mechanisms to improve wheat and other crops in drought- and salinity-prone environments. Ann. Appl. Biol. 2005, 146, 239–259.

12. Sanguineti, M.C.; Li, S.; Maccaferri, M.; Corneti, S.; Rotondo, F.; Chiari, T.; Tuberosa, R. Genetic dissection of seminal root architecture in elite durum wheat germplasm. Ann. Appl. Biol. 2007, 151, 291–305.

13. Wasson, A.P.; Richards, R.A.; Chatrath, R.; Misra, S.C.; Prasad, S.V.S.; Rebetzke, G.J.; Kirkegaard, J.A.; Christopher, J.; Watt, M. Traits and selection strategies to improve root systems and water uptake in water-limited wheat crops. J. Exp. Bot. 2012, 63, 3485–3498.

14. Liu, P.; Jin, Y.; Liu, J.; Liu, C.; Yao, H.; Luo, F.; Guo, Z.; Xia, X.; He, Z. Genome-wide association mapping of root system architecture traits in common wheat (Triticum aestivum L.). Euphytica 2019, 215, 1–12.

15. Chochois, V.; Voge, J.P.; Rebetzke, G.J.; Watt, M. Variation in adult plant phenotypes and partitioning among seed and stem-borne roots across Brachypodium distachyon accessions to exploit in breeding cereals for well-watered and drought environments. Plant Physiol. 2015, 168, 953–967.

16. Mace, E.S.; Singh, V.; van Oosterom, E.J.; Hammer, G.L.; Hunt, C.H.; Jordan, D.R. QTL for nodal root angle in sorghum (Sorghum bicolor L. Moench) co-locate with QTL for traits associated with drought adaptation. Theor. Appl. Genet. 2012, 124, 97–109.

17. Ruiz, M.; Giraldo, P.; González, J.M. Phenotypic variation in root architecture traits and their relationship with eco-geographical and agronomic features in a core collection of tetraploid wheat landraces (Triticum turgidum L.). Euphytica 2018, 214, 54.

18. Roselló, M.; Royo, C.; Sanchez-Garcia, M.; Soriano, J.M. Genetic Dissection of the Seminal Root System Architecture in Mediterranean Durum Wheat Landraces by Genome-Wide Association Study. Agronomy 2019, 9, 364.

19. 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. Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. J. Exp. Bot. 2015, 66, 3477–3486.

20. Nazco, R.; Peña, R.J.; Ammar, K.; Villegas, D.; Crossa, J.; Royo, C. Durum wheat (Triticum durum Desf.) Mediterranean landraces as sources of variability for

allelic combinations at Glu-1/Glu-3 loci affecting gluten strength and pasta cooking quality. Genet. Resour. Crop Evol. 2014, 61, 1219–1236.

21. Moragues, M.; Del Moral, L.F.G.; Moralejo, M.; Royo, C. Yield formation strategies of durum wheat landraces with distinct pattern of dispersal within the Mediterranean basin I: Yield components. Field Crop. Res. 2006, 95, 194–205.

22. Roselló, M.; Villegas, D.; Álvaro, F.; Soriano, J.M.; Lopes, M.S.; Nazco, R.; Royo, C. Unravelling the relationship between adaptation pattern and yield formation strategies in Mediterranean durum wheat landraces. Eur. J. Agron. 2019, 107, 43–52.

23. Rufo, R.; Alvaro, F.; Royo, C.; Soriano, J.M. From landraces to improved cultivars: Assessment of genetic diversity and population structure of Mediterranean wheat using SNP markers. PLoS ONE 2019, 14, e0219867.

24. Wang, S.; Wong, D.; Forrest, K.; Allen, A.; Chao, S.; Huang, B.E.; Maccaferri, M.; Salvi, S.; Milner, S.G.; Cattivelli, L.; et al. Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. Plant Biotechnol. J. 2014, 12, 787–796.

25. Abràmoff, M.D.; Magalhães, P.J.; Ram, S.J. Image Processing with ImageJ, 2nd ed.; Packt Publishing: Birmingham, UK, 2004; Volume 11, ISBN 9781785889837.

26. Harper, D.G.C. Some comments on the repeatability of measurements. Ringing Migr. 1994, 15, 84–90.

27. Bradbury, P.J.; Zhang, Z.; Kroon, D.E.; Casstevens, T.M.; Ramdoss, Y.; Buckler, E.S. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics 2007, 23, 2633–2635.

28. Barrett, J.C.; Fry, B.; Maller, J.; Daly, M.J. Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 2005, 21, 263–265.

29. Wang, S.-X.; Zhu, Y.-L.; Zhang, D.-X.; Shao, H.; Liu, P.; Hu, J.-B.; Zhang, H.; Zhang, H.-P.; Chang, C.; Lu, J.; et al. Genome-wide association study for grain yield and related traits in elite wheat varieties and advanced lines using SNP markers. PLoS ONE 2017, 12, e0188662.

30. Mangini, G.; Gadaleta, A.; Colasuonno, P.; Marcotuli, I.; Signorile, A.M.; Simeone, R.; De Vita, P.; Mastrangelo, A.M.; Laidò, G.; Pecchioni, N.; et al. Genetic dissection of the relationships between grain yield components by genome-wide association mapping in a collection of tetraploid wheats. PLoS ONE 2018, 13, e0190162.

31. Condorelli, G.E.; Maccaferri, M.; Newcomb, M.; Andrade-Sanchez, P.; White, J.W.; French, A.N.; Sciara, G.; Ward, R.; Tuberosa, R. Comparative Aerial and Ground Based High Throughput Phenotyping for the Genetic Dissection of NDVI as a Proxy for Drought Adaptive Traits in Durum Wheat. Front. Plant Sci. 2018, 9.

32. Sukumaran, S.; Reynolds, M.P.; Sansaloni, C. Genome-Wide Association Analyses Identify QTL Hotspots for Yield and Component Traits in Durum Wheat Grown under Yield Potential, Drought, and Heat Stress Environments. Front. Plant Sci. 2018, 9, 81.

33. The International Wheat Genome Sequencing Consortium (IWGSC); Appels, R.; Eversole, K.; Stein, N.; Feuillet, C.; Keller, B.; Rogers, J.; Pozniak, C.J.; Choulet, F.; Distelfeld, A.; et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science 2018, 361, eaar7191.

34. Christopher, J.; Christopher, M.; Jennings, R.; Jones, S.; Fletcher, S.; Borrell, A.; Manschadi, A.M.; Jordan, D.; Mace, E.; Hammer, G. QTL for root angle and number in a population developed from bread wheats (Triticum aestivum) with contrasting adaptation to water-limited environments. Theor. Appl. Genet. 2013, 126, 1563–1574.

35. Paez-Garcia, A.; Motes, C.; Scheible, W.-R.; Chen, R.; Blancaflor, E.; Monteros, M. Root Traits and Phenotyping Strategies for Plant Improvement. Plants 2015, 4, 334–355.

36. Asseng, S.; Ritchie, J.T.; Smucker, A.J.M.; Robertson, M.J. Root growth and water uptake during water deficit and recovering in wheat. Plant Soil 1998, 201, 265–273.

37. Rogers, E.D.; Benfey, P.N. Regulation of plant root system architecture: Implications for crop advancement. Curr. Opin. Biotechnol. 2015, 32, 93–98.

38. Reynolds, M.; Tuberosa, R. Translational research impacting on crop productivity in drought-prone environments. Curr. Opin. Plant Biol. 2008, 11, 171–179.

39. Manschadi, A.M.; Christopher, J.T.; Hammer, G.L.; Devoil, P. Experimental and modelling studies of drought-adaptive root architectural traits in wheat (Triticum aestivum L.). Plant Biosyst. Int. J. Deal. Asp. Plant Biol. 2010, 144, 458–462.

40. Royo, C.; Nazco, R.; Villegas, D. The climate of the zone of origin of Mediterranean durum wheat (Triticum durum Desf.) landraces affects their agronomic performance. Genet. Resour. Crop Evol. 2014, 61, 1345–1358.

41. Liao, M.; Palta, J.A.; Fillery, I.R.P. Root characteristics of vigorous wheat improve early nitrogen uptake. Aust. J. Agric. Res. 2006, 57, 1097.

42. Fang, Y.; Du, Y.; Wang, J.; Wu, A.; Qiao, S.; Xu, B.; Zhang, S.; Siddique, K.H.M.; Chen, Y. Moderate Drought Stress Affected Root Growth and Grain Yield in Old, Modern and Newly Released Cultivars of Winter Wheat. Front. Plant Sci. 2017, 8, 672.

43. Aparicio, N.; Villegas, D.; Araus, J.L.; Blanco, R.; Royo, C. Seedling development and biomass as affected by seed size and morphology in durum wheat. J. Agric. Sci. 2002, 139, 143–150.

44. King, J.; Gay, A.; Sylvester-Bradley, R.; Bingham, I.; Foulkes, J.; Gregory, P.; Robinson, D. Modelling Cereal Root Systems for Water and Nitrogen Capture: Towards an Economic Optimum. Ann. Bot. 2003, 91, 383–390.

45. Chao, S.; Dubcovsky, J.; Dvorak, J.; Luo, M.-C.; Baenziger, S.P.; Matnyazov, R.; Clark, D.R.; Talbert, L.E.; Anderson, J.A.; Dreisigacker, S.; et al. Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (Triticum aestivum L.). BMC Genomics 2010, 11, 727.

46. Soriano, J.M.; Alvaro, F. Discovering consensus genomic regions in wheat for root-related traits by QTL meta-analysis. Sci. Rep. 2019, 9, 10537.

47. Alahmad, S.; El Hassouni, K.; Bassi, F.M.; Dinglasan, E.; Youssef, C.; Quarry, G.; Aksoy, A.; Mazzucotelli, E.; Juhász, A.; Able, J.A.; et al. A major root architecture QTL responding to water limitation in durum wheat. Front. Plant Sci. 2019, 10, 436.

48. Ayalew, H.; Liu, H.; Börner, A.; Kobiljski, B.; Liu, C.; Yan, G. Genome-Wide Association Mapping of Major Root Length QTLs Under PEG Induced Water Stress in Wheat. Front. Plant Sci. 2018, 9, 1759.

49. Li, L.; Peng, Z.; Mao, X.; Wang, J.; Chang, X.; Reynolds, M.; Jing, R. Genome-wide association study reveals genomic regions controlling root and shoot traits at late growth stages in wheat. Ann. Bot. 2019, 124, 993–1006.

50. Beyer, S.; Daba, S.; Tyagi, P.; Bockelman, H.; Brown-Guedira, G.; Mohammadi, M. Loci and candidate genes controlling root traits in wheat seedlings—A wheat root GWAS. Funct. Integr. Genomics 2019, 19, 91–107.

51. Jia, Y.; Gu, H.; Wang, X.; Chen, Q.; Shi, S.; Zhang, J.; Ma, L.; Zhang, H.; Ma, H. Molecular cloning and characterization of an F-box family gene CarF-box1 from chickpea (Cicer arietinum L.). Mol. Biol. Rep. 2012, 39, 2337–2345.

52. Koops, P.; Pelser, S.; Ignatz, M.; Klose, C.; Marrocco-Selden, K.; Kretsch, T. EDL3 is an F-box protein involved in the regulation of abscisic acid signalling in Arabidopsis thaliana. J. Exp. Bot. 2011, 62, 5547–5560.

53. Zhou, S.; Sun, X.; Yin, S.; Kong, X.; Zhou, S.; Xu, Y.; Luo, Y.; Wang, W. The role of the F-box gene TaFBA1 from wheat (Triticum aestivum L.) in drought tolerance. Plant Physiol. Biochem. 2014, 84, 213–223.

54. Li, Q.; Wang, W.; Wang, W.; Zhang, G.; Liu, Y.; Wang, Y.; Wang, W. Wheat F-Box Protein Gene TaFBA1 Is Involved in Plant Tolerance to Heat Stress. Front. Plant Sci. 2018, 9.

55. Xu, D.-Q.; Huang, J.; Guo, S.-Q.; Yang, X.; Bao, Y.-M.; Tang, H.-J.; Zhang, H.-S. Overexpression of a TFIIIA-type zinc finger protein gene ZFP252 enhances drought and salt tolerance in rice (Oryza sativa L.). FEBS Lett. 2008, 582, 1037–1043.

56. Chang, H.; Chen, D.; Kam, J.; Richardson, T.; Drenth, J.; Guo, X.; McIntyre, C.L.; Chai, S.; Rae, A.L.; Xue, G.-P. Abiotic stress upregulated TaZFP34 represses the expression of type-B response regulator and SHY2 genes and enhances root to shoot ratio in wheat. Plant Sci. 2016, 252, 88–102.

57. Lee, T.G.; Jang, C.S.; Kim, J.Y.; Kim, D.S.; Park, J.H.; Kim, D.Y.; Seo, Y.W. A Myb transcription factor (TaMyb1) from wheat roots is expressed during hypoxia: Roles in response to the oxygen concentration in root environment and abiotic stresses. Physiol. Plant. 2007, 129, 375–385.

58. Rahaie, M.; Xue, G.-P.; Naghavi, M.R.; Alizadeh, H.; Schenk, P.M. A MYB gene from wheat (Triticum aestivum L.) is up-regulated during salt and drought stresses and differentially regulated between salt-tolerant and sensitive genotypes. Plant Cell Rep. 2010, 29, 835–844.

59. Qin, Y.; Wang, M.; Tian, Y.; He, W.; Han, L.; Xia, G. Over-expression of TaMYB33 encoding a novel wheat MYB transcription factor increases salt and drought tolerance in Arabidopsis. Mol. Biol. Rep. 2012, 39, 7183–7192.

60. Xie, Q.; Frugis, G.; Colgan, D.; Chua, N.H. Arabidopsis NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. Genes Dev. 2000, 14, 3024–3036.

61. He, X.-J.; Mu, R.-L.; Cao, W.-H.; Zhang, Z.-G.; Zhang, J.-S.; Chen, S.-Y. AtNAC2, a transcription factor downstream of ethylene and auxin signaling pathways, is involved in salt stress response and lateral root development. Plant J. 2005, 44, 903–916.

62. Xia, N.; Zhang, G.; Liu, X.-Y.; Deng, L.; Cai, G.-L.; Zhang, Y.; Wang, X.-J.; Zhao, J.; Huang, L.-L.; Kang, Z.-S. Characterization of a novel wheat NAC transcription factor gene involved in defense response against stripe rust pathogen infection and abiotic stresses. Mol. Biol. Rep. 2010, 37, 3703–3712.

63. Ma, M.; Wang, Q.; Li, Z.; Cheng, H.; Li, Z.; Liu, X.; Song, W.; Appels, R.; Zhao, H. Expression of TaCYP78A3, a gene encoding cytochrome P450 CYP78A3 protein in wheat (Triticum aestivum L.), affects seed size. Plant J. 2015, 83, 312–325.

64. Sornaraj, P.; Luang, S.; Lopato, S.; Hrmova, M. Basic leucine zipper (bZIP) transcription factors involved in abiotic stresses: A molecular model of a wheat bZIP factor and implications of its structure in function. Biochim. Biophys. Acta Gen. Subj. 2016, 1860, 46–56.

65. Zhang, L.; Zhang, L.; Xia, C.; Gao, L.; Hao, C.; Zhao, G.; Jia, J.; Kong, X. A Novel Wheat C-bZIP Gene, TabZIP14-B, Participates in Salt and Freezing Tolerance in Transgenic Plants. Front. Plant Sci. 2017, 8.

66. Minibayeva, F.V.; Gordon, L.K.; Kolesnikov, O.P.; Chasov, A.V. Role of extracellular peroxidase in the superoxide production by wheat root cells. Protoplasma 2001, 217, 125–128.

67. Santelia, D.; Vincenzetti, V.; Azzarello, E.; Bovet, L.; Fukao, Y.; Düchtig, P.; Mancuso, S.; Martinoia, E.; Geisler, M. MDR-like ABC transporter AtPGP4 is involved in auxin-mediated lateral root and root hair development. FEBS Lett. 2005, 579, 5399–5406.

68. Gaedeke, N.; Klein, M.; Kolukisaoglu, U.; Forestier, C.; Müller, A.; Ansorge, M.; Becker, D.; Mamnun, Y.; Kuchler, K.; Schulz, B.; et al. The Arabidopsis thaliana ABC transporter AtMRP5 controls root development and stomata movement. EMBO J. 2001, 20, 1875–1887.

69. Gao, T.; Li, G.-Z.; Wang, C.-R.; Dong, J.; Yuan, S.-S.; Wang, Y.-H.; Kang, G.-Z. Function of the ERFL1a Transcription Factor in Wheat Responses to Water Deficiency. Int. J. Mol. Sci. 2018, 19, 1465.

70. Javot, H. The Role of Aquaporins in Root Water Uptake. Ann. Bot. 2002, 90, 301–313.

71. Li, A.X.; Han, Y.Y.; Wang, X.; Chen, Y.H.; Zhao, M.R.; Zhou, S.M.; Wang, W. Root-specific expression of wheat expansin gene TaEXPB23 enhances root growth and water stress tolerance in tobacco. Environ. Exp. Bot. 2015, 110, 73–84.

Chapter 3:

Using Unmanned Aerial Vehicle and Ground-Based RGB Indices to Assess Agronomic Performance of Wheat Landraces and Cultivars in a Mediterranean-Type Environment

Rubén Rufo[1], Jose Miguel Soriano[1], Dolors Villegas[1], Conxita Royo[1] and Joaquim Bellvert[2]

[1]Sustainable Field Crops Programme, IRTA (Institute for Food and Agricultural Research and Technology), 25198 Lleida, Spain.
[2]Efficient Use of Water in Agriculture Programme, IRTA, 25198 Lleida, Spain.

Published in Remote Sensing (2021) 13, 1187

doi: org/10.3390/rs13061187

Using Unmanned Aerial Vehicle and Ground-Based RGB Indices to Assess Agronomic Performance of Wheat Landraces and Cultivars in a Mediterranean-Type Environment

1. Abstract

The adaptability and stability of new bread wheat cultivars that can be successfully grown in rainfed conditions are of paramount importance. Plant improvement can be boosted using effective high-throughput phenotyping tools in dry areas of the Mediterranean basin, where drought and heat stress are expected to increase yield instability. Remote sensing has been of growing interest in breeding programs since it is a cost-effective technology useful for assessing the canopy structure as well as the physiological traits of large genotype collections. The purpose of this study was to evaluate the use of a 4-band multispectral camera on-board an unmanned aerial vehicle (UAV) and ground-based RGB imagery to predict agronomic traits as well as quantify the best estimation of leaf area index (LAI) in rainfed conditions. A collection of 365 bread wheat genotypes, including 181 Mediterranean landraces and 184 modern cultivars, was evaluated during two consecutive growing seasons. Several vegetation indices (VI) derived from multispectral UAV and ground-based RGB images were calculated at different image acquisition dates of the crop cycle. The modified triangular vegetation index (MTVI2) proved to have a good accuracy to estimate LAI (R^2=0.61). Although the stepwise multiple regression analysis showed that grain yield and number of grains per square meter (NGm^2) were the agronomic traits most suitable to be predicted, the R^2 were low due to field trials were conducted under rainfed conditions. Moreover, the prediction of agronomic traits was slightly better with ground-based RGB VI rather than with UAV multispectral VIs. NDVI and GNDVI, from multispectral images, were present in most of the prediction equations. Repeated measurements confirmed that the ability of VIs to predict yield depends on the range of phenotypic data. The current study highlights the potential use of VI and RGB images as an efficient tool for high-throughput phenotyping under rainfed Mediterranean conditions.

Keywords: high-throughput phenotyping; drought stress; UAV imagery; ground-based RGB image; vegetation indices; phenology; grain yield; biomass

2. Introduction

Wheat is the main crop around the world and provides 18% of the global human intake of calories and 20% of protein (http://www.fao.org/faostat/accessed on

14 December 2020). Since global wheat demand is predicted to increase by 60% by the year 2050, there is an urgent need to raise wheat production by 1.7% per year until then [1]. Therefore, the rate of genetic improvement required in the next decades is higher than that achieved so far [2]. Given the limitations imposed by the soil availability for agricultural uses, most increases rely on the release of improved cultivars with enhanced yield potential and stability under variable environmental conditions. Drought stress during the grain filling period, originating from a combination of water deficit and high temperatures, is the main constraint on wheat yield in semi-arid environments, such as the Mediterranean Basin [3], which has been identified as one of the regions most sensitive to the effects of climate change. A reduction of 20% in yearly precipitation and a mean temperature increase of 4°C have been predicted for this area by climate change models (http://www.ipcc.ch/ accessed on 14 December 2020) [4]. For this reason, breeding programs are focusing on the adaptability and stability of new cultivars that can be successfully grown in dry areas [5]. There is a general agreement that phenotyping is currently the bottleneck for further yield increases in breeding programs [6]. The availability of cost-effective technologies able to phenotype large number of plots in a rapid, cost-effective, and high spatial resolution way is essential for genetic progress [7]. In recent years, high-throughput phenotyping (HTP) has been increasingly used in plant breeding to estimate traits such as yield, green biomass, plant height, and leaf area index (LAI) [8–10]. Among the different approaches used for field HTP, remote sensing permits nonintrusive, nondestructive, high-throughput monitoring of agronomic, physiological, and architectural plant traits [11]. In HTP, this approach is mostly through spectral vegetation indices (VI), which are obtained from the formulation of different wavelengths mostly located at the visible, red-edge, and near-infrared [12]. Usually, these indices are calculated from multispectral cameras installed on-board an unmanned aerial vehicle (UAV), with the main advantage being the capacity for screening hundreds of plots in a short period of time [13,14]. Various authors have stressed the suitability of using VI measured early in the season for grain yield forecasting [15], although anthesis and milk grain development have been shown to be more useful for yield appraisal in wheat [16,17]. Some of them have shown a root mean square error (RMSE) ranging from 0.57 to 0.97 t/ha for predicting yield in wheat [18,19]. Other methodologies also use machine-learning regressions, chemometrics, radiative transfer models, photogrammetry, or hybrid approaches to estimate vegetation traits [20–22]. On the other hand, far-infrared (thermal) radiation and LIDAR sensors have been respectively used to estimate plant water status [23] and to characterize the architectural features [24].

Red-green-blue (RGB) imagery, obtained from conventional digital cameras, has also been reported to be a suitable method to calculate vegetation indices for wheat breeding in water-limited environments [25]. Conventional digital cameras are more

affordable, portable, and easy to use, being a cost-effective way to obtain images of a large number of samples with minimum effort [26]. Moreover, their use has also been proposed in breeding programs for assessing plant traits such as green biomass since the calculation of vegetation indices is based on simple methods that can obtain data automatically from a high number of images [25]. Some studies have demonstrated that vegetation indices derived from RGB cameras are also able to give the same or better results as those obtained from multispectral images [9,27]. Kefauver et al. [27] compared UAV and field-based high- throughput phenotyping using RGB cameras for assessing nitrogen use efficiency (NUE) in barley. It was found that the regression models explained 77.8% and 71.6% of the variance in yield from UAV and ground data, respectively, while combining the datasets led to an increase in the explanation of variance to 82.7%. Gracia-Romero et al. [9] compared the performance of RGB images acquired from ground and aerial cameras to estimate yield in maize under different levels of phosphorus fertilization. The authors found that, in general, ground-based RGB indices correlated in a comparable way with grain yield.

Most studies comparing the performance of RGB and multispectral images for the assessment of wheat traits have been conducted on sets of semidwarf cultivars grown in well-irrigated fields, where the expression of the yield potential and the range of phenotypic values are maximized, or under different irrigation treatments [28]. However, information is lacking regarding the suitability of remote sensing images to predict agronomic traits of wheats with contrasting canopy architectures under rainfed conditions. The current study examines the performance of VIs obtained at different dates from a 4-band multispectral camera (Parrot Sequoia) on-board UAV and those obtained from ground-based RGB images to assess agronomic traits of large panels of bread wheat landraces and modern cultivars adapted to Mediterranean conditions.

3. Materials and Methods

3.1. Experimental Field Setup and Agronomic Data Recording

A collection of 365 bread wheat (*Triticum aestivum* L.) genotypes from the MED6WHEAT IRTA-panel [29] was used in this study. The collection consisted of 181 landraces and 184 modern cultivars from 24 and 19 Mediterranean countries, respectively (Table S1). Field experiments were conducted at Gimenells, Lleida (41°38' N and 0°22' E, 260 m a.s.l) under rainfed conditions for two consecutive growing seasons, 2016–2017 and 2017–2018. Experiments followed a nonreplicated augmented design with two replicated checks (cv. 'Anza' and 'Soissons') and plots of 3.6 m^2 (1.2 m wide x 3 m long) with eight rows spaced 0.15 m apart. The seed rate was adjusted to 250 germinable seeds per m2 and the plots were kept free of weeds

and diseases. The sowing dates were 21 November 2016 and 15 November 2017.

Phenology was assessed based on the scale of Zadoks et al. [30]. A growth stage (GS) was considered to have been achieved when at least 50% of the plants reached it. The following six GS were determined at each plot: stem elongation or when the first node was detectable (S, GS31); booting, determined when boots swollen (B, GS45); heading (H, GS55); anthesis (A, GS65); medium milk-grain development (M, GS75); and hard-dough grain development (D, GS87). Meteorological data were recorded from a weather station placed in the experimental field.

The following agronomic traits were measured: yield, biomass, number of spikes per square meter (NSm^2), number of grains per square meter (NGm^2), and thousand kernel weight (TKW). The NSm^2, NGm^2, and TKW were obtained from samples collected at maturity one week before harvest from 1-m-long central row of each plot. After harvesting, plants were stored in a glasshouse in paper sacks at room temperature during five months until processing. Subsequently, samples were processed as dry matter after drying them at 70°C for 24 h to determine the aboveground biomass (t/ha). The plots were mechanically harvested at maturity, and the grain yield (GY, t/ha) is expressed on a 12% moisture basis. The fraction of intercepted photosynthetically active radiation (fiPAR) was measured from 13:00 to 15:00 (local time) at each image acquisition date in 64 different plots of each landrace and modern set of genotypes using a portable ceptometer (AccuPAR model LP-80, decagon devices Inc., Pullman, WA, USA). Measurements were collected in clear sky conditions. Two measurements per plot were recorded by placing the ceptometer in a horizontal position at ground level. A fixed tripod connected to the sensor allowed us to collect the incident radiation above the plants. These measurements were also used to obtain the leaf area index (LAI) using the Norman-Jarvis model [31], and assuming a leaf area distribution parameter for wheat as 0.96.

3.2. Remote Sensing Images Acquisition

During the first growing season, both ground-based RGB and multispectral UAV images were acquired on the following three dates: 28 March (128 days after sowing, DAS); 21 April (151 DAS), and 19 May (179 DAS). Figure 1 shows the color of the different genotypes in the field at the three image acquisition dates. The adverse meteorological conditions during the spring of the second year hindered image capturing at the early growth stages. Therefore, images were collected on April 17 (153 DAS) and May 18 (184 DAS), to match the main growth stages of the crop. Table 1 summarizes the growth stages of the genotypes included in the panel at each image acquisition occasion.

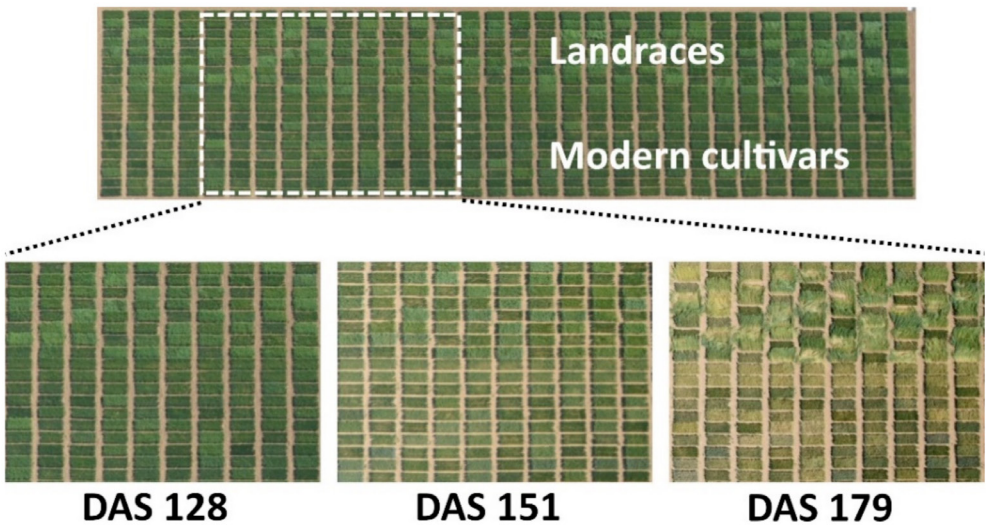

Figure 1. Field view of both collection sets, landraces and modern cultivars, at each image acquisition date of the growing season 2016–2017. DAS, days after sowing.

3.2.1. Ground-Based RGB Vegetation Indices

Ground-based RGB images were collected in clear-sky conditions from 12:00 to 14:00 (local time) over the two years at the same day as UAV multispectral image acquisition. Ground-based RGB images were taken following the methodology reported by Casadesús mand Villegas [26]. A digital camera (Sony Alpha A5000, TYO, JPN) was used, with an objective Sony 16–50 mm at the minimum focal length, 19.8 megapixels of resolution, fixed aperture of F3.5, shutter speed of 1/250, without flash, and the aperture in automatic. When the plants were shorter than 120 cm, pictures were taken by holding the camera at 150 cm, approximately 50 cm from the border of the plot and oriented downwards. Once the average plot height exceeded 120 cm (which was the case with some landraces), it was necessary to use a camera stick at 170–190 cm. Three pictures were obtained per plot without stopping, covering the central rows of each plot in a zenithal plane. All the images were 1152 768 pixels, saved in JPEG format and processed with open-source BreedPix v0.2 software [25]. RGB indices were calculated based on properties of color related to the "greenness" of the canopy. Ten vegetation indices (VIs) were calculated following the protocol described in Casadesús et al. [25] (Table 2). As described in Kefauver et al. [27], hue, intensity, and saturation are the components of the HIS (hue–intensity–saturation) color space. Similar to intensity is the parameter lightness in both CIE-Lab and CIE-Luv color spaces, defined by the Commission Internationale de l'Éclairage (CIE), where a* and u* represent a color in an axis from green to red and b* and v* from yellow to blue according to the human visual system.

Table 1. Number and percentage of genotypes showing each growth stage at each image acquisition occasion.

Landraces 2016–2017				
Date	**Days after Sowing**	**Growth Stage**	**Number of Genotypes**	**(%)**
28 March 2017	128	Stem elongation	181	100
21 April 2017	151	Booting	95	53
		Heading	53	29
		Anthesis	29	16
		Milk development	4	2
19 May 2017	179	Milk development	52	29
		Dough development	129	71
Modern 2016–2017				
28 March 2017	128	Stem elongation	169	92
		Booting	15	8
21 April 2017	151	Booting	8	4
		Heading	72	39
		Anthesis	45	25
		Milk development	59	32
19 May 2017	179	Dough development	184	100
Landraces 2017–2018				
17 April 2018	153	Stem elongation	97	53
		Booting	83	46
		Heading	1	1
18 May 2018	184	Milk development	109	60
		Dough development	72	40
Modern 2017–2018				
17 April 2018	153	Stem elongation	26	14
		Booting	126	69
		Heading	32	17
18 May 2018	184	Milk development	66	36
		Dough development	118	64

Table 2. Red-green-blue (RGB) vegetation indices, based on different color properties, used in the study.

Parameter	Definition	Reference
Intensity	Brightness of the image from black to white	[32]
Hue	Color tint	
Saturation	Amount of tint	
Lightness	Overall albedo from the HIS color space	
a*	Red–green spectrum of chromaticity	
u*		
b*	Yellow–blue color spectrum	
v*		
GA	Green area	[25]
GGA	Greener area	

a* and u* represent a color in an axis from green to red and b* and v* from yellow to blue according to the human visual system.

3.2.2. Multispectral Images Acquired with the UAV

The UAV used for the multispectral image acquisition was the DJI S800 EVO hexacopter (Nanshan, CHN) (Figure 2a). Flight altitude was 40 m above ground level (AGL). The multispectral camera used was a Parrot Sequoia (Parrot, Paris,

France) with a 1.2 mega-pixel sensor yielding a resolution of 1280 x 960 pixels. Horizontal, vertical, and diagonal field of view (HFOV, VFOV, and DFOV) provided by the optical focal length were 61.9°, 48.5°, and 73.7°, respectively. The camera included four individual image sensors with filters centered at the wavelengths and full-width half-max bandwidths (FWHM) of 550±40 (green), 660±40 (red), 735±10 (red edge) and 790±40 nm (near infrared), respectively. The Parrot Sequoia camera includes a separate sunshine sensor that measures solar irradiance in the same spectral bands as the four image sensors. Flight plans were designed for 80% image overlap along flight paths. In addition to the radiometric corrections made by the internal solar irradiance sensor, corrections were conducted through in situ spectral measurements with black-and-white ground calibration targets, bare soil, and wheat plots using the JAZ-3 Ocean Optics STS VIS spectrometer (Ocean Optics, Inc., Dunedin, FL, USA) with a wavelength response from 350 to 800 nm (Figure 2b, Table S2). In 2017, data from white calibration targets was not used due to saturation problems (Table S2). The calibration of the spectrometer measurements was taken using a reference panel (white color Spectralon and dark) laid on the ground as targets before and after the flights. Image orthorectification was completed using ground control points (GCP). The position of the center of each GCP was acquired with a handheld GPS (Global Positioning System) (Geo7x, Trimble GeoExplorer series, Sunnyvale, CA, USA). All images were mosaicked using the Agisoft Photoscan Professional software (Agisoft LLC., St. Petersburg, Russia) and pixel-based georectification was done with the software QGIS version 3.2.0 (USA, http://www.qgis.org). The collected multispectral images were used to calculate several vegetation indices (VI), which were carefully selected based on the relationship to certain specific features of plant physiology [33] (Table 3).

Figure 2. (a) Unmanned aerial vehicle (UAV) Hexacopter DJI S800 EVO used to collect the multispectral images of the experimental plots; (b) reference targets used for the geometric and radiometric calibrations.

Table 3. Vegetation spectral indices evaluated in this study.

Vegetation Index	Equation	Reference
NDVI	$(R_{790} - R_{660})/(R_{790} + R_{660})$	[33]
RDVI	$(R_{790} - R_{660})/\sqrt{R_{790} + R_{660}}$	[34]
MSAVI	$\frac{1}{2}\left[2R_{790} + 1 - \sqrt{(2R_{790} + 1)^2 - 8(R_{790} - R_{660})}\right]$	[35]
MTVI2	$\frac{[1.5(1.2(R_{790} - R_{550}) - 2.5(R_{660} - R_{550}))]}{\sqrt{(2R_{790}+1)^2 - (6R_{790} - 5\sqrt{R_{660}}) - 0.5}}$	[36]
TCARI/OSAVI	$\frac{3[(R_{735} - R_{660}) - 0.2(R_{735} - R_{550})(R_{735}/R_{660})]}{(1+0.16)\frac{(R_{790} - R_{660})}{(R_{790} + R_{660} + 0.16)}}$	[37]
GNDVI	$(R_{790} - R_{550})/(R_{790} + R_{550})$	[38]

3.3. Statistical Analysis

Restricted maximum likelihood (REML) was used to estimate the variance components and produce the best linear unbiased predictors (BLUPs) for agronomical traits, VIs, and RGB indices, following the MIXED procedure of the SAS-STAT statistical package (SAS Institute, Inc., Cary, NC, USA). To assess differences between genotypes, years, and flight occasions, one-way ANOVAs were conducted separately for the 181 landraces and the 184 modern cultivars. LAI measurements were regressed with all the VIs described previously using aggregated data of the two growing seasons for landrace (N=320) and modern (N=320) panels separately and joining both panels (N=640). Stepwise linear regression models were fit to the relationships between genotypic means for agronomic traits as dependent variables and UAV or RGB vegetation indices calculated at each flight occasion as independent ones. Since 12 landrace cultivars were considered outliers for its VI values, stepwise linear regression was conducted on 169 landraces and 184 modern cultivars. To assess the relationship between agronomic traits (yield, biomass, NSm^2, NGm^2, and TKW) and VIs, both the landrace and modern sets were randomly and equally divided into two independent groups: one for training purposes called training dataset and the other as an evaluation group for the prediction accuracy called test dataset. All the statistical analyses and randomly splitting data for predictive modelling were carried out using the JMP v13.1.0 statistical package (SAS Institute, Inc.), considering a significance level of $p<0.05$.

4. Results

4.1. Environmental Conditions

The experimental site is representative of the Mediterranean climate, characterized by an uneven distribution of rainfall during the season, low temperatures in winter

that rise sharply in spring, and high temperatures continuing until the end of the crop cycle (Figure 3). The first growing season had less rainfall (105 mm) than the second one (269 mm) during the growth cycle from sowing (December) to maturity (June). Moreover, water scarcity was significantly higher in the 2016–2017 growing season than in the 2017–2018 growing season, mostly during the grain-filling periods, which received 5 mm and 147 mm of rainfall, respectively.

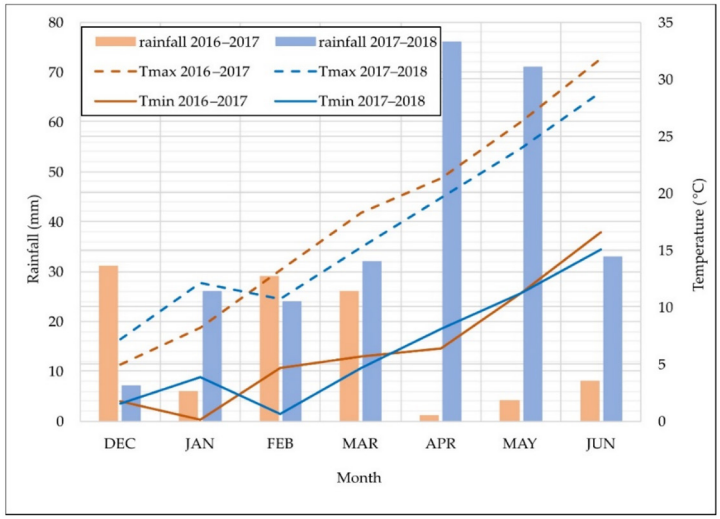

Figure 3. Monthly rainfall (mm), and minimum (Tmin) and maximum (Tmax) temperatures during the growth cycle of each growing season.

4.2. Agronomic Performance

The number of genotypes, minimum, maximum, mean, and standard deviation (SD) values for each dataset, trait, and growing season are shown in Table 4. The analysis of variance (ANOVA) for the agronomic traits was performed separately for landraces and modern genotypes (Table 5). Given that the year effect was significant for all traits in the two types of germplasm (except for biomass in the landrace set), the results are presented independently for each growing season. The percentage of variability for all traits explained by genotype was much higher than that explained by the year or by the year x genotype interaction. The contribution to total variation by the year effect was lower than that of the year x genotype interaction for all traits, except for NGm^2 in both landrace and modern genotypes and for TKW in modern genotypes. F-values showed that all agronomic traits, except biomass, differed significantly between landraces and modern genotypes. All the evaluated traits, except thousand kernel weight (TKW), were higher in 2018 than in 2017 in the whole collection. Grain yield, NGm^2, TKW, and biomass were also higher in modern cultivars than in landraces in both years. The evaluated traits had a higher coefficient of variability (CV) in both years.

Table 4. Main descriptive statistics for yield (t/ha), biomass at ripening (t/ha), number of spikes per square meter (NSm^2), number of grains per square meter (NGm^2), and thousand kernel weight (TKW, g) for the sample datasets used in the models. N, number of genotypes; Min, minimum values; Max, maximum values; SD, standard deviation.

Set	Agronomic Traits	Training					Test				
		N	Min	Max	Mean	SD	N	Min	Max	Mean	SD
Landrace 2016–2017	Yield (t/ha)	84	3.0	8.3	5.0	0.9	85	3.2	8.5	5.2	0.9
	Biomass (t/ha)		8.6	24.5	15.5	3.5		6.5	24.5	15.9	3.4
	NSm^2		386	761	544	73		381	686	542	63
	NGm^2		7438	20,154	12,764	2481		8180	23,003	13,082	2742
	TKW (g)		27.0	51.6	38.9	5.2		23.3	52.7	39.8	5.1
Landrace 2017–2018	Yield (t/ha)	84	3.6	7.2	5.5	0.7	85	4.1	9.0	5.8	0.9
	Biomass (t/ha)		7.1	29.9	16.1	4.7		6.7	33.9	16.9	5.2
	NSm^2		372	824	580	94		345	889	583	95
	NGm^2		13,035	22,227	16,357	1645		12,917	24,836	17,130	2650
	TKW (g)		19.9	49.3	33.9	6.3		17.9	49.0	34.6	7.5
Modern 2016–2017	Yield (t/ha)	92	7.1	11.8	9.5	0.9	92	6.5	11.7	9.4	1.1
	Biomass (t/ha)		8.5	22.9	16.4	2.9		10.2	22.9	16.3	3.0
	NSm^2		253	820	486	117		280	813	471	108
	NGm^2		14,276	31,452	22,630	3628		12,520	33,852	22,170	4251
	TKW (g)		31.3	58.8	42.8	5.2		32.6	58.1	43.3	5.1
Modern 2017–2018	Yield (t/ha)	92	6.9	12.0	10.0	1.0	92	7.3	12.4	10.0	1.0
	Biomass (t(ha)		10.4	39.0	19.2	4.6		6.2	29.4	19.6	3.9
	NSm^2		200	973	583	149		220	920	585	142
	NGm^2		17,002	34,191	26,848	3706		17,752	41,629	26,718	4082
	TKW (g)		29.7	51.1	37.7	4.2		24.4	51.3	38.1	4.6

4.3. LAI Prediction through Vegetation Indices

Estimates of LAI were carried out with aggregated data of the two growing seasons for landraces and modern sets separately. Although LAI measurements were regressed with all the VIs reported in Tables 2 and 3, only the NDVI, GNDVI, modified triangular vegetation index (MTVI2), GA, GGA, Hue, a*, and u* showed significant relationships ($p<0.001$) (Table 6). Despite the lower R^2 values for landraces, LAI predictions for both panels showed similar slopes for the relation between observed LAI and estimated LAI. Thus, LAI was assessed for the whole collection, joining data from landraces and modern genotypes of the two growing seasons. The highest R^2 for LAI estimates using UAV multispectral images was obtained with the MTVI2 ($R^2=0.61$), which showed a RMSE of 1.17. On the other hand, Hue was the ground-based RGB index with the highest R^2 ($R^2=0.45$) and a RMSE of 1.40.

Then, the LAI of all plots was estimated through MTVI2, considering the growth stage of each genotype at each flight occasion. LAI varied significantly between the set of genotypes and years ($p<0.001$) at each flight occasion and growth stage. Figure 4 shows that LAI was higher in 2018 than in 2017 for both landraces and modern cultivars. In the first growing season (2016–2017), landraces had LAI values significantly higher than those of modern cultivars at 128 DAS and 151–153 DAS,

Table 5. Analysis of variance performed separately for 181 landraces and 184 modern cultivars and values for grain yield, biomass, number of spikes per square meter (NSm^2), number of grains per square meter (NGm^2), and thousand kernel weight (TKW) for each growing season. SS, sum of squares. CV, coefficient of variability. **p<0.01. ***p<0.001.

		Landraces					Modern				
		Yield (t/ha)	Biomass (t/ha)	NSm^2	NGm^2	TKW (g)	Yield (t/ha)	Biomass (t/ha)	NSm^2	NGm^2	TKW (g)
SS Year (%)		8.4	0.8	5.3	38.3	15.0	6.2	14.2	14.2	23.9	22.5
SS Genotype (%)		63.7	52.6	55.9	40.1	64.8	64.7	42.9	50.6	62.5	64.2
SS Year × Genotype (%)		27.9	46.6	38.8	21.5	20.1	29.1	42.9	35.2	13.6	13.3
F year		50.3 ***	2.8	23.1 ***	296.7 ***	125.6 ***	38.8 ***	61.0 ***	74.0 ***	322.1 ***	309.2 ***
F genotype		2.3 ***	1.1	1.4 **	1.8 ***	3.2 ***	2.2 ***	1.0	1.4 **	4.6 ***	4.8 ***
CV (%)	2016–2017	17.9	22.2	12.5	20.2	13.1	10.7	18.0	23.5	17.6	11.9
	2017–2018	14.7	30.1	16.2	13.4	20.1	9.9	22.1	24.9	14.5	11.6
Mean	2016–2017	5.1	15.7	543	12,923	39.4	9.5	16.4	479	22,400	43.0
	2017–2018	5.6	16.5	582	16,746	34.3	10.0	19.4	584	26,783	37.9
Minimum	2016–2017	3.0	6.5	381	7438	23.3	6.5	8.5	253	12,520	31.3
	2017–2018	3.6	6.7	345	12,917	17.9	6.9	6.2	200	17,002	24.4
Maximum	2016–2017	8.5	24.5	761	23,003	52.7	11.8	22.9	820	33,852	58.8
	2017–2018	9.0	33.9	889	24,835	49.3	12.4	39.0	973	41,629	51.3

Table 6. Statistically significant ($p<0.001$) relationships between leaf area index (LAI) measured with the ceptometer and vegetation indices (VIs) obtained from UAV multispectral and RGB images. Calculations have been made with aggregated data of the two growing seasons and image acquisition occasions joining germplasm collections and for landraces and modern sets separately. ns, no significant. RMSE, root mean square error.

Method	VI	Equation	R^2	RMSE	Equation	R^2	RMSE	Equation	R^2	RMSE
		Landraces + *Modern* (N = 640)			*Landraces* (N = 320)			*Modern* (N = 320)		
UAV Multispectral	NDVI	y = 11.63x − 5.55	0.38	1.48	y = 10.74x − 4.49	0.16	1.45	y = 11.06x − 5.22	0.43	1.47
	GNDVI	y = 8.89x − 2.54	0.18	1.70	ns			y = 9.42x − 3.35	0.26	1.67
	MTVI2	y = 7.45x − 1.01	0.61	1.17	y = 7.11x − 0.72	0.39	1.24	y = 7.58x − 1.10	0.66	1.12
Ground-based RGB	GA	y = 7.18x − 1.36	0.41	1.43	y = 8.04x − 2.00	0.20	1.41	y = 6.71x − 1.09	0.45	1.43
	GGA	y = 4.52x − 1.86	0.39	1.45	y = 4.47x + 2.07	0.29	1.33	y = 4.27x + 1.89	0.38	1.52
	Hue	y = 0.09x − 2.91	0.45	1.40	y = 0.10x − 3.59	0.33	1.29	y = 0.08x − 2.44	0.45	1.44
	a*	y = 0.18x − 2.04	0.22	1.66	ns			y = −0.18x + 1.96	0.21	1.73
	u*	y = 0.19x − 3.05	0.3	1.57	y = -0.16x + 3.54	0.15	1.46	y = −0.18x + 2.99	0.3	1.63

but similar at 178–184 DAS (Figure 4a). Maximum LAI values for landraces and modern cultivars in 2017 were obtained at the booting and stem elongation stages, respectively (Figure 4b). The LAI of landraces in 2017 was significantly higher than modern cultivars until anthesis, when it decreased significantly until the values were lower than those estimated in the modern panel. Therefore, the LAI of modern cultivars started declining later than in landraces. In 2018, the LAI of landraces and modern cultivars had a similar pattern throughout the growing season without significant differences between them, except at the hard dough-grain stage, where the LAI of landraces was slightly lower than that of modern cultivars (Figure 4b).

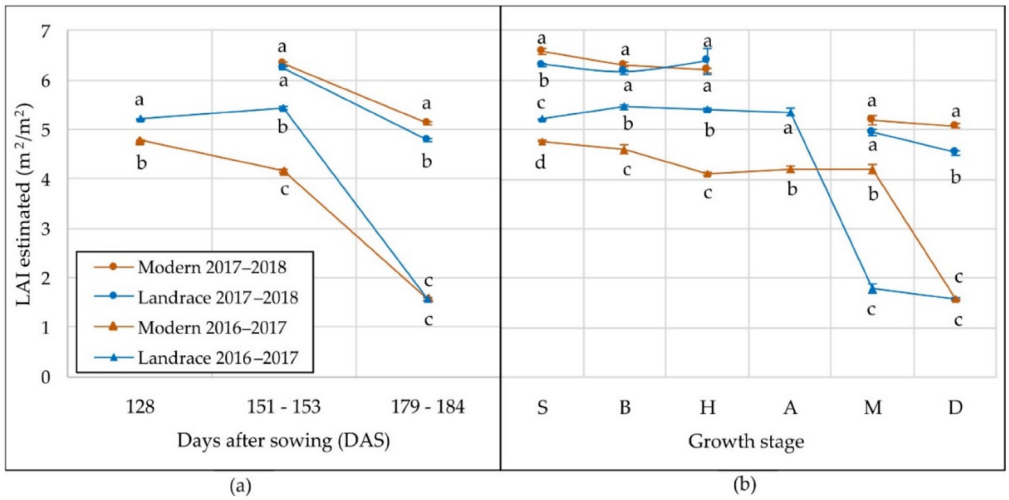

Figure 4. Mean values in 2017 and 2018 of leaf area index estimated through MTVI2 for landraces and modern cultivars at: (a) each date of image acquisition expressed in days after sowing (DAS), and (b) each growth stage. S, stem elongation; B, booting; H, heading; A, anthesis; M, milk-grain development; D, hard dough-grain development. Different letters at each date or growth stage indicate significant differences at $p \leq 0.01$ using Tukey's honest significant difference test.

4.4. Performance of Stepwise Regression Models

Table 7 shows the main statistics of the models built to estimate the different agronomic traits with UAV multispectral and RGB VIs for each year and germplasm set. Scatter plots for the relation between estimated and observed agronomic traits on the test dataset based on Table 7 equations are shown in Figures S1 and S2. The results indicate that the training models developed from multispectral images were significant for all traits, germplasm sets, and years, with the exception of NSm^2 for the landraces set in 2018. Grain yield and NGm^2 were the traits showing the highest R^2 in both germplasm collections. Most of the equations developed with multispectral VI had in common the NDVI and GNDVI indices, although in some cases MTVI and MSAVI also appeared. The models constructed with RGB-VI were also statistically significant in all cases except for biomass and NSm^2 for the

landraces set in 2018. Yield was also one of the most predictive traits.

Test models obtained with the corresponding dataset also showed the highest R^2 for yield and NGm^2, using either multispectral or RGB VIs. However, R^2 tended to be slightly lower for the latter. For multispectral VI, the maximum R^2 obtained to predict yield in landraces and modern cultivars was 0.36 and 0.43, respectively, which corresponded to an RMSE of 0.28 t/ha and 0.39 t/ha. In addition, the maximum R^2 for NGm^2 predictions through multispectral VIs was 0.19 and 0.38 for landraces and modern genotypes, respec- tively, corresponding with RMSE values of 768 and 1835 grains/m^2 (Table 7). Considering training and test model values together, the highest R^2 for yield was obtained in modern genotypes, being higher for the growing season 2016-2017 (R^2=0.43 and R^2=0.37 through UAV and RGB imagery, respectively) than in the next growing season (R^2=0.29 and R^2=0.45 through UAV and RGB imagery, respectively).

Table 8 shows the training and test statistics for the five agronomic traits obtained with aggregated datasets of the two growing seasons for landraces and modern cultivars. Scatter plots for the relation between estimated and observed agronomic traits on the test dataset based on Table 8 equations are shown in Figures S3 and S4. In general, the models fitted better for modern cultivars. The test models for most agronomic traits were not significant in the set of landraces. In general, both test and training models improved when the data from two growing seasons were analyzed together. Grain yield and NGm^2 were again the traits that showed the highest R^2, using either UAV multispectral and RGB VIs (Table 8). For these two traits, despite the R^2 of training models being higher in modern cultivars, the RMSE tended to be lower in landraces. For the models built with multispectral VI, the RMSE in yield predictions ranged from 0.26 to 0.32 t/ha and from 0.34 to 0.38 t/ha, for landraces and modern cultivars, respectively. The models built with ground-based RGB VIs had RMSE values ranging from 0.28–0.50 t/ha and 0.39–0.54 t/ha for landraces and modern cultivars, respectively. The highest R^2 for yield training models of landraces were obtained with ground-based RGB VI, testing data with the dataset corresponding to 2018 (R^2=0.30). In contrast, training model of yield in modern genotypes had the highest R^2 using UAV multispectral VI, testing data in the dataset of 2017 (R^2=0.51).

5. Discussion

The current study evaluates the suitability of using a 4-band multispectral camera (Parrot Sequoia) on-board UAV and ground-based RGB images to predict yield in wheat under a rainfed Mediterranean-type environment. Despite remote sensing methods being nondestructive and cost-efficient approaches based on the information provided by visible and near-infrared (VIS-NIR) radiation reflection

Table 7. Training and test statistics of the models for the estimations of agronomic traits through UAV multispectral and RGB VIs for each germplasm set and growing season. *p<0.05. **p<0.01. N, number of genotypes; R^2, determination coefficient; RMSE, root mean standard error; Yield (t/ha); NSm^2, number of spikes per square meter; NGm^2, number of grains per square meter; TKW, thousand kernel weight (g); I, intensity; L, lightness; S, saturation. Number after each VI means the flight occasion: 1, 128 DAS; 2, 151-153 DAS; 3, 179-184 DAS.

		UAV Multispectral						Ground-Based RGB					
Set	Traits	Training			Test			Training			Test		
		N	Equation	R^2	N	R^2	RMSE	N	Equation	R^2	N	R^2	RMSE
Landraces 2016–2017	Yield	84	−26.69 + 31.89GNDVI_1 + 5.98NDVI_3	0.18 **	85	0.18 **	0.45	84	40.45 + 2.26GA_3 + 12.77S_3 − 0.91L_1 − 0.15b*_2	0.45 **	85	0.28 **	0.66
	Biomass		−130.27 + 29.05MSAVI_2 + 135.54GNDVI_2	0.18 **		ns	1.41		−17.65 + 40.58GGA_1	0.11 **		ns	1.04
	NSm^2		−2167.28 + 2958.35GNDVI_2	0.15 *		ns	27.22		−731.73 + 8062.69I_3 − 58.49L_3 − 64.60a*_2	0.24 **		0.10 **	37.05
	NGm^2		−73937 + 9059.90NDVI_3 + 92920GNDVI_1	0.27 **		0.08 **	1204		34,836 − 544.90L_2	0.12 **		ns	867
	TKW		172.87 - 173.71GNDVI_2 + 40.10GNDVI_3	0.15**		ns	1.86		−115.24 + 614.25I_2 - 106.53I_3	0.17**		ns	2.06
Landraces 2017–2018	Yield	84	0.006 + 7.01GNDVI_3	0.18 **	85	0.36 **	0.28	84	−4.53 + 0.08Hue_2 - 0.18a*_3	0.33 **	85	0.25 **	0.34
	Biomass		−26.24 + 46.70MSAVI_2	0.10 **		ns	1.34		ns	ns		ns	ns
	NSm^2		ns	ns		ns	ns		ns	ns		ns	ns
	NGm^2		−2,167,724 + 2,631,346GNDVI_2 + 12,018GNDVI_3	0.24 **		0.19 **	768		19,077 − 25,561I_2 − 453.53a*_3	0.10 **		0.16 **	590
	TKW		−11.97 + 57.72RDVI_2	0.15 **		ns	2.22		52.23 − 209.47S_2 − 1.72a*_2	0.29 **		0.11 **	2.92
Modern 2016–2017	Yield	92	−9.23 + 8.09NDVI_3 + 19.68MTVI_2	0.28 **	92	0.43 **	0.39	92	4.93 − 0.15a*_3 − 0.27u*_1	0.34 **	92	0.37 **	0.49
	Biomass		−63.23 + 21.86MSAVI_3 + 102.20MTVI_2	0.28 **		0.11 **	1.48		0.28 + 0.20Hue_2	0.24 **		0.22 **	1.14
	NSm^2		−13857 + 11,036TCARI/OSAVI_2 + 15,315GNDVI_2	0.22 **		0.16 **	56.49		916.14 - 33.34L_1 - 44.65a*_1	0.21 **		0.18 **	55.83
	NGm^2		−221,638 + 272,852GNDVI_2	0.33 **		0.35 **	1863		3,4087 + 5840.71GGA_3 − 9,9931I_1 − 1085.73a*_2	0.45 **		0.45 **	1780
	TKW		292.96 − 279.44GNDVI_2	0.17 **		0.11 **	2.23		−11.24 + 26.68GA_2 + 163.80I_1 + 2.27a*_2 + 0.59v*_3	0.36 **		0.11 **	5.17
Modern 2017–2018	Yield	92	−13 + 25.25NDVI_3	0.29 **	92	0.24 **	0.38	92	−1.04 − 16.13GA_2 + 9.05GA_3 + 0.19Hue_2	0.45 **	92	0.22 **	0.54
	biomass		−143.78 + 177.75MSAVI_2	0.07 **		ns	1.28		−21.58 − 55.66I_3 + 0.59Hue_2	0.12 **		ns	1.84
	NSm^2		10,710 − 14,953MTVI2_3 + 2604GNDVI_2	0.22 **		0.08 **	72.94		−594.57 − 84.23u*_2	0.06 *		0.06 *	37.29
	NGm^2		−52,520 + 34,921MSAVI_3 + 59,546GNDVI_2	0.49 **		0.38 **	1835		34,497 − 447.30L_3 − 2198.99u*_3	0.32 **		0.21 **	1825
	TKW		82.24 − 52.04GNDVI_2	0.15 **		0.10 **	1.55		57.85 − 0.87b*_2	0.14 **		0.04 *	1.34

Table 8. Training and Test statistics of the models for the estimations of agronomic traits through UAV multispectral and RGB VIs aggregating the data of the two growing seasons for landraces and modern cultivars. *p<0.05. **p<0.01. N, number of genotypes; R^2, determination coefficient; RMSE, root mean standard error; Yield (t/ha); NSm^2, number of spikes per square meter; NGm^2, number of grains per square meter; TKW, thousand kernel weight (g); I, intensity; L, lightness; S, saturation. Number after each VI means the flight occasion: 1, 128 DAS; 2, 151-153 DAS; 3, 179-184 DAS.

UAV Multispectral													
Set	Traits	Training			Test 2016–2017			Test 2017–2018			Test 2016–2017+2017–2018		
		N	Equation	R^2	N	R^2	RMSE	N	R^2	RMSE	N	R^2	RMSE
Landraces 2016–2017+2017–2018	Yield	168	0.30 + 11.26NDVI_3 − 5.11MTVI2_3	0.25 **	85	0.17 **	0.27	85	0.27 **	0.26	170	0.28 **	0.32
	Biomass		−50.03 + 35.23MTVI2_2 + 38.30GNDVI_2	0.11 **		ns	-		ns	-		ns	-
	NSm^2		ns	ns		ns	-		ns	-		ns	-
	NGm^2		−713.21 + 21637GNDVI_3	0.44 **		ns	545		0.17 **	980.73		0.42 **	1376
	TKW		−80.01 + 29.67RDVI_2 + 109.65GNDVI_2	0.16 **		ns	0.99		ns	-		0.14 **	2.20
Modern 2016–2017+2017–2018	Yield	184	−5.19 + 9.47GNDVI_2 + 13.63NDVI_3 - 6.46MTVI2_3	0.30 **	92	0.51 **	0.34	92	0.33 **	0.38	184	0.46 **	0.38
	Biomass		−63.23 + 102.20MTVI2_2 + 21.86MSAVI_3	0.28 **		0.11 **	1.48		0.01	2.38		0.20 **	17.73
	NSm^2		−1906.39 + 1047.64NDVI_2 + 1265.34GNDVI_2 + 477.01GNDVI_3	0.27 **		0.24 **	46.45		0.10 **	55.94		0.29 **	60.88
	NGm^2		−69278 + 20,340NDVI_2 + 22,033NDVI_3 + 66,300GNDVI_2	0.54 **		0.49 **	1419		0.32 **	2058.47		0.53 **	2091
	TKW		92.05 − 37.35GNDVI_2 − 26.21GNDVI_3	0.31**		0.19 **	1.13		0.10 **	1.67		0.33 **	2.42
Ground-based RGB													
Landraces 2016–2017+2017–2018	Yield	168	3.91 − 0.07L_2 − 0.10a*_2 − 0.14b*_2 + 0.05Hue_3 + 9.19S_3	0.36 **	85	0.21 **	0.50	85	0.30 **	0.28	170	0.28 **	0.42
	biomass		ns	ns		ns	-		ns	-		ns	-
	NSm^2		412.15 − 13.54u*_2 − 84.37GGA_3	0.09 **		ns	-		ns	-		0.09 **	24.31
	NGm^2		25,069 + 6675.08GGA_2 − 34,739I_2 - 292.79L_2 + 11,005GA_3 + 426.02u*_3	0.50 **		ns	-		0.14 **	804.53		0.39 **	1604
	TKW		32.37 − 27.23GGA_2 - 3.93b*_2 + 4.06v*_2 + 0.13Hue_3	0.28 **		ns	1.94		0.04	2.50		0.15 **	2.98
Modern 2016–2017+2017–2018	Yield	184	10.64 − 4.95GA_3 − 7.68I_3 + 0.07Hue_3 − 0.22u*_3	0.28 **	92	0.27 **	0.51	92	0.24 **	0.39	184	0.30 **	0.54
	Biomass		−0.28 + 0.20Hue_2	0.24 **		0.09 **	1.90		ns	-		0.20**	2.09
	NSm^2		238.83 − 18.41a*_2	0.19 **		0.20 **	30.24		0.05*	30.87		0.26 **	50.66
	NGm^2		21,688 - 37,911I_3 + 143.07Hue_3 − 373.30a*_2	0.45 **		0.45 **	1314		0.19 **	1257.45		0.45 **	2221
	TKW		24.78 + 39.88GA_2 + 0.63a*_2 + 1.04u*_2	0.36**		0.28 **	1.45		0.11 **	1.31		0.36 **	2.37

[39], care should be taken to standardize measurements across different plant architectures and sun elevation [6]. The light intensity, temperature, cloud cover, wind speed, and timing of measurements can also affect the accuracy of the estimation of traits evaluated in the field [40]. Digital photography is also a promising approach due to the use of conventional cameras as a low-cost sensor to get the image and open-source software to process the data from it [25].

The large year effect for the assessed traits found in the current study may be attributed to the contrasting water availability in the two years of the experimental fields, which doubled in 2018 compared to the preceding year. The largest differences were observed in April and beyond, coinciding with the grain-filling period, which likely was the main cause of the lower yield, spike number, and grain number recorded in 2017 compared with 2018. It is well known that water scarcity after anthesis has significant effects on yield and yield components [12,33,41]. The heaviest kernels observed in 2017 were most probably a consequence of the compensation between yield components, since a lower NGm^2 was observed in 2017. It has been shown that the value of each component strongly depends on the values of the components defined previously, and NGm^2 is defined before TKW [42]. The number of grains and their weight are established sequentially during plant development, with the potential number of grains being determined before anthesis, and the grain weight being fixed after it [42,43]. This is in accordance with the heaviest grains being obtained in the current study in 2017, the year with the lowest grain number. The high yields achieved in the two years are in agreement with those reported in previous studies at the same experimental site [44], where the high yields could be attributed to the high soil fertility (about 3% organic matter) and the superficial subsoil water layer at this site [45]. The CVs obtained in the current study for the analyzed traits are within the normal ranges reported for water-limited environments [10]. Moreover, the largest variability of agronomic traits found in landraces when compared with modern varieties is in agreement with the results of previous studies conducted in durum wheat [41,46].

The remotely sensed estimates of LAI in both landraces and modern cultivars were higher in 2018 (Figure 3), as well as grain yield, which may be mostly explained by the higher rainfall received during the grain-filling period in that year. As reported by Villegas et al. [47], drought severely affects the total above-ground biomass due to a decrease in the rate of growth. Although water stress affects the growth of wheat, the effects are less harmful at the early stages of the crop cycle than during the grain-filling period [48]. In 2018, the LAI of the landrace was slightly lower than that of the modern cultivar only at the end of the growing season (the dough development stage), suggesting that, under well-watered conditions, the vegetative growth capacity of the latter is higher or senescence of the former starts earlier. Villegas et al. [42] and Royo et al. [49] reported similar conclusions. In 2017, however,

the LAI of landraces was significantly higher than that of modern genotypes until anthesis, which could be due to the higher resistance to water stress of landraces [42,50], their superior water use efficiency before flowering [51], and their large root system [52], which is able to exploit deeper soil layers. Figure 3 shows that, despite the LAI of landraces in 2017 being higher than that of modern genotypes until anthesis, it started declining earlier than the latter. This anticipated decrease of LAI in the landrace genotypes could be partially explained by a higher water demand of landraces, a consequence of their larger canopy, which could not be fulfilled at the end of the growing season, leading to the anticipated senescence. Moreover, it could be partially attributed to the greater potential of modern cultivars, compared with the landraces, to use water during grain filling to achieve yield increases [51]. It is also important to mention that differences in remotely sensed estimates of LAI between phenological stages could also be influenced by differences in the chlorophyll content. It is well known that chlorophyll content decreases during senescence and as a consequence, also those VIs that uses bands mostly placed in the NIR and green regions [53]. Therefore, it may happen that plants with the same LAI at different growing stages had different value of a VI due to differences in the chlorophyll content. Despite of this, Din et al. [54] reported that the MTVI2 was one of the most consistent VIs to change through phenological stages. However, it is possible that the estimates of the low LAI values at the end of the growing season could also be affected by a low chlorophyll content due to senescence, as previously mentioned.

A number of studies have estimated agronomic traits such as grain yield or biomass through UAV multispectral and RGB imagery in wheat and other cereals, but the majority of them have been conducted in irrigated environments [9,16,55] or under a wide range of phenotypic variability resulting from varying growing conditions [9,56,57]. A proper assessment of agronomic traits through remote sensing is expected when phenotypic variability is present. This usually occurs in experiments conducted under irrigated conditions, where genotypes are allowed to express their full potential, thus, maximizing differences between them [17], or when a wide range of phenotypic values results from treatments varying the agronomic management [9,56,57]. However, studies conducted in wheat under rainfed conditions are scarce and the precision of the assessments obtained on them is lower. A study by Kyratzis et al. [12], conducted on durum wheat, obtained R^2 values of 0.43 for the relationships between NDVI and yield at different growth stages, which are comparable to the values reported here.

In this study, MTVI2 was the best VI to estimate LAI through multispectral imagery (R^2=0.61). On the other hand, estimates of LAI through RGB VIs showed slightly lower R^2, with Hue being the best predictor (R^2=0.45). It is widely known that some vegetation indices, such as NDVI, show saturation when LAI

reports high values [12,36,56]. Furthermore, estimating green LAI through the NDVI has several limitations since it is affected, for instance, by factors such as soil background, canopy shadows, atmospheric conditions, and variations in leaf chlorophyll concentration [58]. Haboudane et al. [36] stated that improved VIs such as MTVI were more sensitive to chlorophyll variations and, therefore, responded better to LAI changes. In addition, it has been reported that MTVI2 is better than other VIs mitigating this saturation effect in wheat with LAI values ranging from 2 to 8 [54,56,59]. Despite LAI was obtained through destructive measurements, results from our study had similar LAI values and the regression with the MTVI2 showed a RMSE of 1.17. This RMSE agrees with values obtained by Xing et al. [59], who reported RMSE values ranging from 1.1 to 1.6 when using different VIs calculated with a spectrometer and Sentinel 2 imagery. In particular, the RMSE of MTVI2 obtained by these authors was 1.26 and 1.16 using the spectrometer and Sentinel 2 imagery, respectively, which agreed with the RMSE obtained in our study. Hassan et al. [60] also exhibited a strong relationship between VIs and LAI measured with an AccuPAR LP-80 ceptometer with values ranging from 2 to 5.5.

The current study demonstrated that predictions of yield could be properly obtained using both multispectral and RGB VI, with the R^2 of the latter tending to be higher. Although models differed depending on the type of germplasm and the trait to be assessed, NDVI and GNDVI were the VIs mostly entered in all of the prediction equations obtained through UAV multispectral imagery, thus, confirming the feasibility of using such structural VIs to assess different agronomic traits in wheat [14,17,39,61]. On the other hand, as mentioned above, ground-based RGB imagery showed better estimations than UAV multispectral imagery for the prediction set. RGB indices such as GA, GGA, a*, and u* have been proven to be more suitable for predicting higher yield due to their capacity to calculate a combination of physiological components related to biomass [25,26]. Kefauver et al. [27] and Gracia-Romero et al. [9] reported the feasibility of using RGB VI to estimate different agronomic traits. In this study, a positive and negative contribution of GA and a*, respectively, at the last image acquisition date (DAS 179-184) were present in most of the algorithms for predicting yield. This confirms that the indices that performed better in assessing differences in yield were the ones related to canopy greenness and, thus, to vegetation cover [62]. GA quantifies the green pixels of the total pixels in the image, and, thus, is reliable to use for estimating the fraction of vegetation cover [63]. As most of the carbohydrates for grain filling are formed after heading, a larger leaf area or vegetation cover is positively correlated with grain yield, determining the future number of grains and their weight [14,25]. Accordingly, a* and u* measurements are also related to 'greenness', where the values go from high negative (green) to low negative or even positive values (lack of green). Furthermore, Rezzouk et al. [64] observed that ground-based RGB

imagery presented a higher resolution than aerial images, since they found that the number of pixels per plot decreased drastically when acquiring images aerially. In our case, this was probably not the case since the pixel resolution of RGB and UAV multispectral imagery were <1 cm and 5 cm, respectively. In addition, the use of relatively low-cost RGB sensors could be a feasible alternative to multispectral cameras from UAV measurements for plant phenotyping [57].

The lower R^2 observed between VI and yield in landraces than in modern cultivars when the data of each year were analyzed separately could be partially due to the different size and structure of the canopy of both types of germplasm, as landraces were much taller and had a different canopy architecture, which probably saturated the VI at high LAI values. However, in all cases the R^2 values were 0.22. GNDVI and NDVI were the VIs entered into the equations to estimate yield, showing in all cases positive relationships with it. This is in agreement with previous studies showing positive correlations between yield and VI in different environments [65,66], as negative relationships are more frequent under severe water stress conditions [67,68]. Yield predictions in modern genotypes through UAV multispectral VIs varied between the training and test datasets, mostly for the growing season 2016–2017. The R^2 of the later was slightly higher, suggesting that the model is able to improve yield predictions on dry years. This could be explained because during years with water scarcity, the variability between genotypes in traits related to leaf biochemical properties or canopy structural attributes, which can explain a part of the yield, could be higher. Biomass and the number of spikes per unit area could not be assessed in landraces in a reliable way as, although some models were statistically significant, they accounted for a small fraction of the observed variability. However, in modern cultivars predictions of biomass were year-dependent as models accounted from 11% to 28% of the observed variation in 2017 but 12% in 2018. This could be due to the saturation of VI when LAI>5, which was the case in the two germplasm sets in 2018 and in the landraces in 2017, as shown in Figure 3b. Despite this, the significant predictions of biomass were always obtained through the MSAVI index, which seeks to address some of the limitations of NDVI when applied to areas with a high degree of exposed soil surface. It was not surprising that the number of spikes per unit area could not be properly estimated through VI, as the reflectance of the spikes probably caused some distortion in measurements made in the visible and near-infrared ranges, as demonstrated in previous studies [69]. Estimations of NSm^2 with RGB indices were not properly assessed. The number of grains per unit area was better estimated in modern cultivars than in landraces, with both VI and RGB indices. This was not surprising given the strong relationship between the number of grains per unit area and yield in semidwarf cultivars [70,71]. Again, the relationships between VI and RGB indices with grain weight were more consistent in modern cultivars than in landraces. The negative correlations between grain

weight and GNDVI revealed by some prediction models suggest that the plants with higher biomass produced lighter grains, likely as a consequence of competition between plants for allocating photosynthates in vegetative and productive structures during grain filling.

Repeated measurements of the whole collection acquired on different dates throughout the growing season are necessary to improve the prediction of agronomic traits [61]. According to this, in our study, predictions of agronomic traits improved when information from different flights was analyzed together (Table 8). The highest R^2 for grain yield predictions (R^2=0.51) was obtained for modern genotypes in 2017 using combined data from flights acquired on DAS 151 (heading, anthesis, and milk development) and 179 (dough development). Despite R^2 being slightly lower in 2018, in all cases the RMSE varied between 0.26 and 0.38 t/ha, which demonstrates the suitability of the models developed. It has been proven in several studies [9,33,72] that higher variability within a population can increase the determination coefficient and, therefore, the predictive ability of the model.

5.1. Conclusions

The efficiency of breeding programs and the agronomic research will increase considerably depending on the reliability of models for HTP. This study demonstrated the potential of a 4-band multispectral camera (Parrot Sequoia) and RGB images for assessing agronomic traits—particularly yield and grain number per unit area—in bread wheat grown in a Mediterranean-type environment. However, the suitability of the models proved to be specific, as their consistency depended on the canopy structure, leaf dimensions and orientation, and the environmental conditions during vegetative growth, which poses a difficulty for their general use in a random crop season. Thus, uniformity in the crop cycle among cultivars seems to be essential to improve prediction models minimizing environmental effects. The results of the current study demonstrate that the predictive value of the models developed for semidwarf varieties increased when the data of more than one crop season were aggregated to build them. For future studies, the assessment of biophysical parameters earlier during the growing season will improve the accuracy of LAI estimates, particularly when values are low, but not because of a reduction in the chlorophyll content caused by the senescence. This leads to the conclusion that more research is needed to generate series of data from multiple years and growing stages in order to improve the reliability of the predictions obtained with the models developed from the UAV 4-band multispectral (Parrot Sequoia) and RGB cameras. In addition, the use of machine learning techniques should be addressed.

Supplementary Materials: The following are available online at https://www.mdpi.com/2072-4 292/13/6/1187/s1, Table S1: List of accessions. Figures

S1–S4: ANOVA and Scatterplots for the agronomic traits. Table S2: Spectrometer calibration.

Author Contributions: Conceptualization, J.M.S., C.R. and J.B.; methodology, R.R., J.M.S., D.V. and J.B.; software, R.R. and J.B.; formal analysis, R.R., J.M.S., D.V. and J.B.; investigation, R.R. and J.B.; writing—original draft preparation, R.R.; writing—review and editing, R.R., J.M.S., D.V., C.R. and J.B.; supervision, J.M.S. and J.B.; project administration, J.M.S. and C.R.; funding acquisition, J.M.S., C.R. and J.B. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by projects PID2019-109089RB-C31 and RTI2018-099949-R-C21 (Ministerio de Ciencia e Innovación, Spain). R.R. is a recipient of a PhD grant from the Spanish Ministry of Economy and Competitiveness.

Data Availability Statement: The data presented in this study are available within the manuscript and in the Supplementary Materials.

Acknowledgments: The authors acknowledge the contribution of the CERCA Program (Generalitat de Catalunya).

Conflicts of Interest: The authors declare no conflict of interest.

6. References

1. Leegood, R.C.; Evans, J.R.; Furbank, R.T. Food security requires genetic advances to increase farm yields. Nature 2010, 464, 831.
2. Fischer, R.A.T.; Edmeades, G.O. Breeding and Cereal Yield Progress. Crop Sci. 2010, 50, S-85–S-98.
3. Araus, J.L.; Slafer, G.A.; Reynolds, M.P.; Royo, C. Plant breeding and drought in C3 cereals: What should we breed for? Ann. Bot. 2002, 89, 925–940.
4. Bates, B.C.; Kundzewicz, Z.W.; Wu, S.; Palutikof, J.P. Climate Change and Water. In The American Midland Naturalist; Tech- nical Paper of the Intergovernmental Panel on Climate Change; Intergovernmental Panel on Climate Change Secretariat: Geneva, Switzerland, 2008; ISBN 9789291691234.
5. Bhatta, M.; Morgounov, A.; Belamkar, V.; Baenziger, P. Genome-Wide Association Study Reveals Novel Genomic Regions for Grain Yield and Yield-Related Traits in Drought-Stressed Synthetic Hexaploid Wheat. Int. J. Mol. Sci. 2018, 19, 3011.
6. Araus, J.L.; Cairns, J.E. Field high-throughput phenotyping: The new crop breeding frontier. Trends Plant Sci. 2014, 19, 52–61.

7. Reynolds, M.; Tuberosa, R. Translational research impacting on crop productivity in drought-prone environments. Curr. Opin. Plant Biol. 2008, 11, 171–179.

8. Xie, C.; Yang, C. A review on plant high-throughput phenotyping traits using UAV-based sensors. Comput. Electron. Agric. 2020, 178, 105731.

9. Gracia-Romero, A.; Kefauver, S.C.; Vergara-Díaz, O.; Zaman-Allah, M.A.; Prasanna, B.M.; Cairns, J.E.; Araus, J.L. Comparative Performance of Ground vs. Aerially Assessed RGB and Multispectral Indices for Early-Growth Evaluation of Maize Performance under Phosphorus Fertilization. Front. Plant Sci. 2017, 8, 2004.

10. Herrmann, I.; Bdolach, E.; Montekyo, Y.; Rachmilevitch, S.; Townsend, P.A.; Karnieli, A. Assessment of maize yield and phenology by drone-mounted superspectral camera. Precis. Agric. 2020, 21, 51–76.

11. White, J.; Andrade-Sanchez, P.; Gore, M. Field-based phenomics for plant genetics research. Field Crop. Res. 2012, 133, 101–113.

12. Kyratzis, A.C.; Skarlatos, D.P.; Menexes, G.C.; Vamvakousis, V.F.; Katsiotis, A. Assessment of vegetation indices derived by UAV imagery for durum wheat phenotyping under a water limited and heat stressed Mediterranean environment. Front. Plant Sci. 2017, 8, 1114.

13. Araus, J.L.; Kefauver, S.C.; Zaman-Allah, M.; Olsen, M.S.; Cairns, J.E. Translating High-Throughput Phenotyping into Genetic Gain. Trends Plant Sci. 2018, 23, 451–466.

14. Gracia-Romero, A.; Kefauver, S.C.; Fernandez-Gallego, J.A.; Vergara-Díaz, O.; Nieto-Taladriz, M.T.; Araus, J.L. UAV and Ground Image-Based Phenotyping: A Proof of Concept with Durum Wheat. Remote Sens. 2019, 11, 1244.

15. Aparicio, N.; Villegas, D.; Casadesus, J.; Araus, J.L.; Royo, C. Spectral vegetation indices as nondestructive tools for determining durum wheat yield. Agron. J. 2000, 92, 83–91.

16. Aparicio, N.; Villegas, D.; Araus, J.L.; Casadesús, J.; Royo, C. Relationship between growth traits and spectral vegetation indices in durum wheat. Crop Sci. 2002, 42, 1547–1555.

17. Royo, C.; Aparicio, N.; Villegas, D.; Casadesus, J.; Monneveux, P.; Araus, J.L. Usefulness of spectral reflectance indices as durum wheat yield predictors under contrasting Mediterranean conditions. Int. J. Remote Sens. 2003, 24, 4403–4419.

18. Zhao, Y.; Potgieter, A.B.; Zhang, M.; Wu, B.; Hammer, G.L. Predicting Wheat Yield at the Field Scale by Combining High- Resolution Sentinel-2 Satellite Imagery and Crop Modelling. Remote Sens. 2020, 12, 1024.

19. Zhou, X.; Kono, Y.; Win, A.; Matsui, T.; Tanaka, T.S.T. Predicting within-field variability in grain yield and protein content of winter wheat using UAV-based multispectral imagery and machine learning approaches. Plant Prod. Sci. 2020, 1–15.

20. Berger, K.; Rivera Caicedo, J.P.; Martino, L.; Wocher, M.; Hank, T.; Verrelst, J. A survey of active learning for quantifying vegetation traits from terrestrial earth observation data. Remote Sens. 2021, 13, 287.

21. Bellvert, J.; Nieto, H.; Pelechá, A.; Jofre-Cˇ ekalovic´, C.; Zazurca, L.; Miarnau, X. Remote Sensing Energy Balance Model for the Assessment of Crop Evapotranspiration and Water Status in an Almond Rootstock Collection. Front. Plant Sci. 2021, 12, 288.

22. Mahajan, G.R.; Das, B.; Murgaokar, D.; Herrmann, I.; Berger, K.; Sahoo, R.N.; Patel, K.; Desai, A.; Morajkar, S.; Kulkarni, R.M. Monitoring the Foliar Nutrients Status of Mango Using Spectroscopy-Based Spectral Indices and PLSR-Combined Machine Learning Models. Remote Sens. 2021, 13, 641.

23. Elsayed, S.; Elhoweity, M.; Ibrahim, H.H.; Dewir, Y.H.; Migdadi, H.M.; Schmidhalter, U. Thermal imaging and passive reflectance sensing to estimate the water status and grain yield of wheat under different irrigation regimes. Agric. Water Manag. 2017, 189, 98–110.

24. Deery, D.; Jimenez-Berni, J.; Jones, H.; Sirault, X.; Furbank, R. Proximal Remote Sensing Buggies and Potential Applications for Field-Based Phenotyping. Agronomy 2014, 4, 349–379.

25. Casadesús, J.; Kaya, Y.; Bort, J.; Nachit, M.M.; Araus, J.L.; Amor, S.; Ferrazzano, G.; Maalouf, F.; Maccaferri, M.; Martos, V.; et al. Using vegetation indices derived from conventional digital cameras as selection criteria for wheat breeding in water-limited environments. Ann. Appl. Biol. 2007, 150, 227–236.

26. Casadesús, J.; Villegas, D. Conventional digital cameras as a tool for assessing leaf area index and biomass for cereal breeding. J. Integr. Plant Biol. 2014, 56, 7–14.

27. Kefauver, S.C.; Vicente, R.; Vergara-Díaz, O.; Fernandez-Gallego, J.A.; Kerfal, S.; Lopez, A.; Melichar, J.P.E.; Serret Molins, M.D.; Araus, J.L. Comparative UAV

and field phenotyping to assess yield and nitrogen use efficiency in hybrid and conventional barley. Front. Plant Sci. 2017, 8, 1733.

28. Gomez-Candon, D.; Bellvert Rios, J.; Royo, C. Performance of the two-sourceenergy balance (TSEB) model as a tool for monitoring the response of durum wheat to drought by high-throughput field phenotyping. Front. Plant Sci. 2021. (under revision).

29. Rufo, R.; Alvaro, F.; Royo, C.; Soriano, J.M. From landraces to improved cultivars: Assessment of genetic diversity and population structure of Mediterranean wheat using SNP markers. PLoS ONE 2019, 14, e0219867.

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

31. Norman, J.M.; Jarvis, P.G. Photosynthesis in Sitka Spruce (Picea sitchensis (Bong.) Carr.). III. Measurements of Canopy Structure and Interception of Radiation. J. Appl. Ecol. 1974, 11, 375.

32. Trussell, H.J.; Saber, E.; Vrhel, M.; Joel, H. Color Image Processing: Basics and Special Issue Overview. IEEE Signal Process. Mag. 2005, 22, 14–22.

33. Gonzalez-Dugo, V.; Hernandez, P.; Solis, I.; Zarco-Tejada, P.J. Using high-resolution hyperspectral and thermal airborne imagery to assess physiological condition in the context of wheat phenotyping. Remote Sens. 2015, 7, 13586–13605.

34. Roujean, J.L.; Breon, F.M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. Remote Sens. Environ. 1995, 51, 375–384.

35. Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. A modified soil adjusted vegetation index. Remote Sens. Environ. 1994, 48, 119–126.

36. Haboudane, D.; Miller, J.R.; Pattey, E.; Zarco-tejada, P.J.; Strachan, I.B. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. Remote Sens. 2004, 90, 337–352.

37. Haboudane, D.; Miller, J.R.; Tremblay, N.; Zarco-Tejada, P.J.; Dextraze, L. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. Remote Sens. Environ. 2002, 81, 416–426.

38. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sens. Environ. 1996, 58, 289–298.

39. Lobos, G.A.; Matus, I.; Rodriguez, A.; Romero-Bravo, S.; Araus, J.L.; del Pozo, A. Wheat genotypic variability in grain yield and carbon isotope discrimination under Mediterranean conditions assessed by spectral reflectance. J. Integr. Plant Biol. 2014, 56, 470–479.

40. Costa, J.M.; Grant, O.M.; Chaves, M.M. Thermography to explore plant-environment interactions. J. Exp. Bot. 2013, 64, 3937–3949.

41. Royo, C.; Nazco, R.; Villegas, D. The climate of the zone of origin of Mediterranean durum wheat (Triticum durum Desf.) landraces affects their agronomic performance. Genet. Resour. Crop Evol. 2014, 61, 1345–1358.

42. García Del Moral, L.F.; Rharrabti, Y.; Elhani, S.; Martos, V.; Royo, C. Yield formation in Mediterranean durum wheats under two contrasting water regimes based on path-coefficient analysis. Euphytica 2005, 146, 203–212.

43. Brocklehurst, P.A. Factors controlling grain weight in wheat. Nature 1977, 266, 348–349.

44. Roselló, M.; Villegas, D.; Álvaro, F.; Soriano, J.M.; Lopes, M.S.; Nazco, R.; Royo, C. Unravelling the relationship between adaptation pattern and yield formation strategies in Mediterranean durum wheat landraces. Eur. J. Agron. 2019, 107, 43–52.

45. Moragues, M.; Del Moral, L.F.G.; Moralejo, M.; Royo, C. Yield formation strategies of durum wheat landraces with distinct pattern of dispersal within the Mediterranean basin I: Yield components. Field Crop. Res. 2006, 95, 194–205.

46. Royo, C.; Dreisigacker, S.; Ammar, K.; Villegas, D. Agronomic performance of durum wheat landraces and modern cultivars and its association with genotypic variation in vernalization response (Vrn-1) and photoperiod sensitivity (Ppd-1) genes. Eur. J. Agron. 2020, 120, 126129.

47. Villegas, D.; Aparicio, N.; Blanco, R.; Royo, C. Biomass accumulation and main stem elongation of durum wheat grown under Mediterranean conditions. Ann. Bot. 2001, 88, 617–627.

48. Moragues, M.; García Del Moral, L.F.; Moralejo, M.; Royo, C. Yield formation strategies of durum wheat landraces with distinct pattern of dispersal within the

Mediterranean basin: II. Biomass production and allocation. Field Crop. Res. 2006, 95, 182–193.

49. Royo, C.; Aparicio, N.; Blanco, R.; Villegas, D. Leaf and green area development of durum wheat genotypes grown under Mediterranean conditions. Eur. J. Agron. 2004, 20, 419–430.

50. Soriano, J.M.; Villegas, D.; Sorrells, M.E.; Royo, C. Durum wheat landraces from east and west regions of the mediterranean basin are genetically distinct for yield components and phenology. Front. Plant Sci. 2018, 9, 1–9.

51. Subira, J.; Álvaro, F.; García del Moral, L.F.; Royo, C. Breeding effects on the cultivar environment interaction of durum wheat yield. Eur. J. Agron. 2015, 68, 78–88.

52. Subira, J.; Ammar, K.; Álvaro, F.; García del Moral, L.F.; Dreisigacker, S.; Royo, C. Changes in durum wheat root and aerial biomass caused by the introduction of the Rht-B1b dwarfing allele and their effects on yield formation. Plant Soil 2016, 403, 291–304.

53. Adamsen, F.J.; Pinter, P.J.; Barnes, E.M.; LaMorte, R.L.; Wall, G.W.; Leavitt, S.W.; Kimball, B.A. Measuring Wheat Senescence with a Digital Camera. Crop Sci. 1999, 39, 719–724.

54. Din, M.; Zheng, W.; Rashid, M.; Wang, S.; Shi, Z. Evaluating hyperspectral vegetation indices for leaf area index estimation of Oryza sativa L. at diverse phenological stages. Front. Plant Sci. 2017, 8, 820.

55. Condorelli, G.E.; Maccaferri, M.; Newcomb, M.; Andrade-Sanchez, P.; White, J.W.; French, A.N.; Sciara, G.; Ward, R.; Tuberosa, R. Comparative Aerial and Ground Based High Throughput Phenotyping for the Genetic Dissection of NDVI as a Proxy for Drought Adaptive Traits in Durum Wheat. Front. Plant Sci. 2018, 9, 893.

56. Yao, X.; Wang, N.; Liu, Y.; Cheng, T.; Tian, Y.; Chen, Q.; Zhu, Y. Estimation of wheat LAI at middle to high levels using unmanned aerial vehicle narrowband multispectral imagery. Remote Sens. 2017, 9, 1304.

57. Vergara-Díaz, O.; Zaman-Allah, M.A.; Masuka, B.; Hornero, A.; Zarco-Tejada, P.; Prasanna, B.M.; Cairns, J.E.; Araus, J.L. A Novel Remote Sensing Approach for Prediction of Maize Yield under Different Conditions of Nitrogen Fertilization. Front. Plant Sci. 2016, 7, 666.

58. Zarco-Tejada, P.J.; Ustin, S.L.; Whiting, M.L. Temporal and spatial relationships between within-field yield variability in cotton and high-spatial hyperspectral remote sensing imagery. Agron. J. 2005, 97, 641–653.

59. Xing, N.; Huang, W.; Xie, Q.; Shi, Y.; Ye, H.; Dong, Y.; Wu, M.; Sun, G.; Jiao, Q. A transformed triangular vegetation index for estimating winter wheat leaf area index. Remote Sens. 2020, 12, 16.

60. Hassan, M.; Yang, M.; Rasheed, A.; Jin, X.; Xia, X.; Xiao, Y.; He, Z. Time-Series Multispectral Indices from Unmanned Aerial Vehicle Imagery Reveal Senescence Rate in Bread Wheat. Remote Sens. 2018, 10, 809.

61. Gizaw, S.A.; Garland-Campbell, K.; Carter, A.H. Evaluation of agronomic traits and spectral reflectance in Pacific Northwest winter wheat under rain-fed and irrigated conditions. Field Crop. Res. 2016, 196, 168–179.

62. Gracia-Romero, A.; Vergara-Díaz, O.; Thierfelder, C.; Cairns, J.; Kefauver, S.; Araus, J. Phenotyping Conservation Agriculture. Management Effects on Ground and Aerial Remote Sensing Assessments of Maize Hybrids Performance in Zimbabwe. Remote Sens. 2018, 10, 349.

63. Lukina, E.V.; Stone, M.L.; Raun, W.R. Estimating vegetation coverage in wheat using digital images. J. Plant Nutr. 1999, 22, 341–350.

64. Rezzouk, F.Z.; Gracia-Romero, A.; Kefauver, S.C.; Gutiérrez, N.A.; Aranjuelo, I.; Serret, M.D.; Araus, J.L. Remote sensing techniques and stable isotopes as phenotyping tools to assess wheat yield performance: Effects of growing temperature and vernalization. Plant Sci. 2020, 295, 110281.

65. Bort, J.; Casadesus, J.; Nachit, M.M.; Araus, J.L. Factors affecting the grain yield predicting attributes of spectral reflectance indices in durum wheat: Growing conditions, genotype variability and date of measurement. Int. J. Remote Sens. 2005, 26, 2337–2358.

66. Gutierrez, M.; Reynolds, M.P.; Raun, W.R.; Stone, M.L.; Klatt, A.R. Spectral Water Indices for Assessing Yield in Elite Bread Wheat Genotypes under Well-Irrigated, Water-Stressed, and High-Temperature Conditions. Crop Sci. 2010, 50, 197–214.

67. Lopes, M.S.; Saglam, D.; Ozdogan, M.; Reynolds, M. Traits associated with winter wheat grain yield in Central and West Asia. J. Integr. Plant Biol. 2014, 56, 673–683.

68. Rutkoski, J.; Poland, J.; Mondal, S.; Autrique, E.; Pérez, L.G.; Crossa, J.; Reynolds, M.; Singh, R. Canopy temperature and vegetation indices from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat. G3 Genes Genomes Genet. 2016, 6, 2799–2808.

69. Shibayama, M.; Wiegand, C.L.; Richardson, A.J. Diurnal patterns of bidirectional vegetation indices for wheat canopies. Int. J. Remote Sens. 1986, 7, 233–246.

70. Perry, M.W.; D'Antuono, M.F. Yield improvement and associated characteristics of some Australian spring wheat cultivars introduced between 1860 and 1982. Aust. J. Agric. Res. 1989, 40, 457–472.

71. Donmez, E.; Sears, R.G.; Shroyer, J.P.; Paulsen, G.M. Genetic Gain in Yield Attributes of Winter Wheat in the Great Plains. Crop Sci. 2001, 41, 1412–1419.

72. Ferrio, J.P.; Bertran, E.; Nachit, M.; Royo, C.; Araus, J.L. Near infrared reflectance spectroscopy as a potential surrogate method for the analysis of δ13C in mature kernels of durum wheat. Aust. J. Agric. Res. 2001, 52, 809–816.

Chapter 4:

Identification of quantitative trait loci hotspots affecting agronomic traits and high-throughput vegetation indices in rainfed wheat

Rubén Rufo[1], Andrea López[1], Marta da Silva[1], Joaquim Bellvert[2] and Jose Miguel Soriano[1]

[1]Sustainable Field Crops Programme, IRTA (Institute for Food and Agricultural Research and Technology), 25198 Lleida, Spain.
[2]Efficient Use of Water in Agriculture Programme, IRTA, 25198 Lleida, Spain.

Published in Frontiers in Plant Science (2021) 12:735192
doi: 10.3389/fpls.2021.735192

Identification of Quantitative Trait Loci Hotspots Affecting Agronomic Traits and High-Throughput Vegetation Indices in Rainfed Wheat

1. Abstract

Understanding the genetic basis of agronomic traits is essential for wheat breeding programmes to develop new cultivars with enhanced grain yield under climate change conditions. The use of high-throughput phenotyping (HTP) technologies for the assessment of agronomic performance through drought-adaptive traits opens new possibilities in plant breeding. HTP together with a genome-wide association study (GWAS) mapping approach can become a useful method to dissect the genetic control of complex traits in wheat to enhance grain yield under drought stress. This study aimed to identify molecular markers associated with agronomic and remotely sensed vegetation index (VI)-related traits under rainfed conditions in bread wheat and to use an *in silico* candidate gene (CG) approach to search for upregulated CGs under abiotic stress. The plant material consisted of 170 landraces and 184 modern cultivars from the Mediterranean basin. The collection was phenotyped for agronomic and VI traits derived from multispectral images over 3 and 2 years, respectively. GWAS identified 2579 marker–trait associations (MTAs). The QTL overview index statistic detected 11 QTL hotspots involving more than one trait in at least 2 years. A candidate gene analysis detected 12 CGs upregulated under abiotic stress in six QTL hotspots and 46 downregulated CGs in 10 QTL hotspots. The current study highlights the utility of VI to identify chromosome regions that contribute to yield and drought tolerance under rainfed Mediterranean conditions.

Keywords: Wheat, yield components, vegetation indices, marker trait association, candidate genes

2. Introduction

Wheat (*Triticum aestivum* L.) is the most common cultivated crop worldwide. It is grown on 216 million hectares of land, and its global production of 765 million tons of grain provides 19% of the calories and 21% of the protein in the human diet (Faostat 2019, http://www.fao.org/faostat). To cover the expected food demand of a world population that will increase up to 60% by 2050, wheat production needs to increase by 1.7% per year (Leegood et al., 2010). Achieving this objective will not be easy considering the expected negative effects of climate change on wheat yield, particularly in areas such as the Mediterranean basin, where a rise in temperatures by 3–5°C and a decrease in annual rainfall by 25–30% have been predicted (Giorgi and

Lionello 2008). An increasing frequency and severity of terminal drought stress will reduce grain weight, grain quality, and wheat yield (Araus et al., 2002; Slafer et al., 2005; Kulkarni et al., 2017). Therefore, there is a need to improve the identification of genotypes able to maintain acceptable levels of yield and yield stability in semiarid environments, which have been identified as the most sensitive to the effects of climate change (Rufo et al., 2021). The release of improved cultivars with enhanced drought adaptation will be critical for breeding programmes focusing on wheat adaptability and stability under rainfed conditions (Graziani et al., 2014; Bhatta et al., 2018).

The recent progress in high-throughput phenotyping (HTP) based on the use of multispectral images acquired from unmanned aerial vehicles (UAVs) has increasingly improved the assessment of agronomic traits (Gracia-Romero et al., 2017; Xie and Yang 2020; Gomez-Candon et al., 2021; Rufo et al., 2021) on large germplasm collections in a rapid, cost-effective, and high spatial resolution way (Duan et al., 2017), as it allows the estimation of various plant traits using nonintrusive and nondestructive technology (White et al., 2012; Rufo et al., 2021). Remote sensing has attracted growing interest in breeding programmes since it can deliver detailed information about biophysical crop traits in many situations to cope with the current phenotyping bottleneck (Araus and Cairns 2014; Juliana et al., 2019; Bellvert et al., 2021). Some studies have demonstrated the use of vegetation indices (VI) to indirectly detect wheat plants under water stress due to a decrease in vegetative growth (Condorelli et al., 2018). Others have demonstrated the use of energy balance models to estimate the actual water status (Gomez-Candon et al., 2021). When VIs are derived from multispectral cameras, they are obtained from the combination of wavelengths located at the visible, red-edge and near-infrared (NIR) regions of the light spectrum (Kyratzis et al., 2017). These wavelengths allow discerning differences in vegetative greenness, rate of senescence, photosynthetic efficiency and stay green duration (Stenberg et al., 2004; Babar et al., 2006; Lopes and Reynolds 2012). It has been stated that anthesis and milk grain are the most suitable growth stages for the assessment of agronomic traits on a plot-by-plot basis (Aparicio et al., 2002; Royo et al., 2003). The use of HTP as a suitable and accurate predictor of agronomic traits such as phenology, grain filling duration, biomass and yield will provide unique opportunities to increase the power of QTL discovery by increasing the number of genotypes included in the analysis (Juliana et al., 2019). This method will increase the frequency of rare alleles of potential interest to improve wheat adaptation to various environmental conditions.

The dissection of the genetic and molecular basis of complex traits such as yield and drought stress tolerance through complementary approaches such as QTL mapping and genome-wide association studies (GWAS) or association mapping (AM) is essential in breeding programmes. GWAS is based on linkage disequilibrium

(LD) (Flint-Garcia et al., 2003), and it is a powerful approach that provides high mapping resolution due to the higher recombination events analysed in comparison with biparental mapping (Soriano et al., 2017; Qaseem et al., 2019). AM has been used to identify genomic regions related to drought and heat tolerance in durum and bread wheat (Maccaferri et al., 2016; Valluru et al., 2017). Several studies have been conducted to investigate the genetic basis of grain yield and yield-related traits in bread wheat under water stress conditions using association mapping (Edae et al., 2014; Gizaw et al., 2018; Qaseem et al., 2019; Mérida-García et al., 2020). The release of genome sequences for emmer wheat (Avni et al., 2017), bread wheat (IWGSC 2018) and durum wheat (Maccaferri et al., 2019) and the availability of open databases of RNA-seq experiments (Ramírez-González et al., 2018) have made it possible to use a candidate gene (CG) approach to find targets within QTL intervals without performing new functional studies.

The aim of the current study was to identify molecular markers linked to important agronomic traits, VIs and plant features related to drought resistance assessed by HTP, to define the most important QTL hotspots for such traits and to perform *in silico* detection of the underlying CG in those genomic regions.

3. Materials and methods

3.1. Plant material and field trials

A germplasm collection of 354 bread wheat (*Triticum aestivum* L.) genotypes from the MED6WHEAT IRTA panel described in Rufo et al., (2019) was used in this study, of which 170 corresponded to landraces and 184 to modern varieties collected and adapted to 24 and 19 Mediterranean countries, respectively (Supplementary Table S1). The panel is structured into 6 genetic subpopulations (SP) and 38 genotypes that remained admixed (Rufo et al., 2019). SP1: west Mediterranean landraces (43 accessions); SP2: north Mediterranean landraces (59 accessions); SP3: east Mediterranean landraces (42 accessions); SP4: France-Italy modern germplasm (82 accessions); SP5: Balkan modern varieties (24 accessions); and SP6: CIMMYT-ICARDA derived varieties (62 accessions).

The field trials were conducted at Gimenells, Lleida (41°38' N and 0°22' E, 260 m.a.s.l), northeastern Spain, under rainfed conditions for three consecutive seasons: 2016, 2017 and 2018. Average minimum and maximum monthly temperatures and rainfall were calculated from daily data recorded for a weather station close to the experimental fields. Climatic data (rainfall and temperature) for the last 15 years corresponding to the weather station in Gimenells, Lleida (Spain) were downloaded from https://ruralcat.gencat.cat/web/guest/agrometeo.estacions. Experiments followed a non-replicated augmented design with two replicated checks (cv. 'Anza'

and 'Soissons') at a ratio of 1:4 between checks and tested genotypes and in 3.6 m^2 plots with eight rows spaced 0.15 m apart. The sowing density was adjusted to 250 germinable seeds m^{-2}, and the sowing dates were 02 December 2015, 21 November 2016, and 15 November 2017, whereas harvesting dates 07 July 2016, 05 July 2017, and 05 July 2018. Weeds and diseases were controlled following standard practices at the site.

3.2. Agronomic data

The following traits were measured across the three years (2016, 2017 and 2018) according to the protocol described in Rufo et al., (2021): grain yield (GY, t ha^{-1}), number of spikes per m^2 (NSm^2), number of grains per m^2 (NGm^2), thousand kernel weight (TKW, g), aboveground biomass at physiological maturity (biomass, t DM ha^{-1}), harvest index (HI), plant height (PH, cm) and early vigour (estimated as green area, GA). The HI was calculated as the ratio between grain and plant weights in a 1-m long row sample. PH was measured at maturity for three main stems per plot and was measured from the soil to the top of the spike, excluding the awns. Early vigour was calculated by integrating the green area (GA) values obtained by ground-based RGB images taken every fourteen days as described in (Casadesús and Villegas 2014) from emergence until the detection of the first node. Finally, days from sowing to anthesis (DSA, GS65) and grain filling duration (GFD, GS87) were measured on each plot based on the growth stage (GS) scale of Zadoks et al., (1974). Growth stages were achieved when at least 50% of the plants in each plot reached them.

3.3. Image acquisition

Image acquisition was conducted with a multispectral camera (Parrot Sequoia) (Parrot, Paris, France) installed onboard an UAV (DJI S800 EVO hexacopter, Nanshan, CHN). Images were acquired during two years, on 21 April and 19 May 2017 and 17 April and 18 May 2018. Flights were always conducted at approximately 12:00 solar time under sunny conditions and with a wind speed below 12 m/s. The UAV flew at a height of 40 m agl (above ground level) and with a flight plan of 80/60 frontal and side overlap. The multispectral camera has four spectral bands located at wavelengths of 550 ± 40 nm (green), 660 ± 40 nm (red), 735 ± 10 nm (red edge), and 790 ± 40 (near infrared). The camera yields a resolution of 1280 x 960 pixels. All images were radiometrically corrected through an external incident light sensor that measured the irradiance levels of light at the same bands as the sensor, as well as with in situ spectral measurements in ground calibration targets (black, white, soil and grass). Spectral measurements were conducted with a Jaz spectrometer (Ocean Optics, Inc., Dunedin, FL, USA). Jaz has a wavelength response from 200 to 1100 nm and an optical resolution of 0.3 to 10.0 nm. The calibration of the spectrometer

measurements was taken using a reference panel (white colour SpectralonTM). Geometrical correction was conducted by using ground control points (GCPs) and measuring the position in each with a handheld global positioning system (GPS) (Geo7x, Trimble GeoExplorer series, Sunnyvale, CA, USA). All images were mosaicked using Agisoft Photoscan Professional version 1.6.2 (Agisoft LLC., St. Petersburg, Russia) software and geometrically and radiometrically corrected with QGIS 3.2.0 (USA, http://www.qgis.org). Then, six spectral vegetation indices (VIs) were carefully selected based on their significance in relation to certain plant physiology features in wheat (Table 1). In addition, the leaf area index (LAI) was measured using a portable ceptometer (AccuPAR model LP-80, decagon devices Inc., Pullman, WA, USA) from 13:00 to 15:00 (local time) on each image acquisition date in 64 different plots of each set of landrace and modern set. Then, the LAI was estimated for each plot in the whole collection through the MTVI2, following the methodology described by Rufo et al., (2021) and Gomez-Candón et al., (2021). All VIs were assessed in 2017 and 2018 through UAV multispectral images at two growth stages: (1) when all the plots reached anthesis (A) (VI_A) and (2) post anthesis (PA) at the milk and dough developmental stages (VI_PA).

Table 1. Spectral vegetation indices assessed in this study.

Vegetation index	Band centre (nm)	Spectral band	References
Structural indices			
Normalized Difference Vegetation Index (NDVI)	790, 660	NIR, Red	Rouse et al. 1974
Renormalized Difference Vegetation Index (RDVI)	790, 660	NIR, Red	Roujean and Breon 1995
Improved Soil Adjusted Vegetation Index (MSAVI)	790, 660	NIR, Red	Qi et al. 1994
Modified Triangular Vegetation Index (MTVI2)	790, 660, 550	NIR, Red, Green	Haboudane et al. 2004
Green Normalized Difference Vegetation Index (GNDVI)	790, 550	NIR, Green	Gitelson et al. 1996
Chlorophyll indices			
Transformed CARI/Optimized Soil adjusted Vegetation Index (TCARI/OSAVI)	790, 735, 660, 550	NIR, Red Edge, Red, Green	Haboudane et al. 2002

3.4. Genotyping

The panel was genotyped with 13177 SNP markers using the Illumina Infinium 15K Wheat SNP Array at Trait Genetics GmbH (Gatersleben, Germany), and 11196 markers were ordered according to the SNP map developed by Wang et al., (2014). To reduce the risk of errors in further analyses, markers and accessions were analysed for the presence of duplicated patterns and missing values. After excluding markers with more than 25% missing values and with a minor allele frequency (MAF) lower than 5%, a total of 10090 SNPs were used for mapping purposes.

3.5. Statistical analyses

Phenotypic data were fitted to a linear mixed model considering the check cultivars as the fixed effect, and the row and column number and accessions as

random in the model for each environment following the MIXED procedure of the SAS-STAT statistical package (SAS Institute Inc., Cary, NC, USA)

$$y=x\beta+z\gamma+\varepsilon$$

where β is an unknown vector of fixed-effects parameters with known design matrix X, γ is an unknown vector of random-effects parameters with known design matrix Z, and ε is an unknown random error vector whose elements are no longer required to be independent and homogeneous. Restricted maximum likelihood (REML) was used to estimate the variance components and to produce the best linear unbiased predictors (BLUPs) for agronomic traits and VIs (Supplementary Table S2).

To assess differences between years and genetic subpopulations, one-way ANOVAs were conducted for the whole collection. The broad sensed heritability (H^2) was estimated following Knapp et al. (1985):

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_E^2 + \sigma_{GE}^2}$$

where σ^2_G is the genotypic variance, σ^2_E is the variance due to the environmental (year) effect, and σ^2_{GE} is the is the variance for the interaction of genotype with environment.

Least squares means were calculated and compared using the Tukey HSD test. Pearson correlation coefficients were calculated among the evaluated traits. Mean phenotypic values across the three years were used to perform a hierarchical cluster analysis by the Ward method (Ward 1963). Analyses of variance and mean differences were carried out using the JMP v14.2.0 statistical package (SAS Institute, Inc., Cary, NC, USA), considering a significance level of alpha = 0.05.

3.6. Marker trait associations

A GWAS with 10090 SNP markers was conducted on the whole germplasm collection using Tassel 5.0 software (Bradbury et al., 2007) for all agronomic and VI traits per year and across the three growing seasons. A mixed linear model (MLM) was fitted using a principal component analysis (PCA) matrix with 6 principal components as the fixed effect and a kinship (k) matrix as the random effect (PCA + K model) at the optimum compression level based on the groups defined by the kinship matrix. Compression levels range from "no compression" (compression = 1) when each genotype belongs to its own group, to "maximum compression" (compression = *n*) when all genotypes belong to the same group. In addition, the anthesis date was incorporated as a cofactor in the analysis, as reported in previous studies (Crowell et al., 2016; Condorelli et al., 2018; Soriano et al., 2021). Manhattan plots were generated using the R script CMplot (https://github.com/YinLiLin/

CMplot). A false discovery rate (FDR) threshold (Benjamini and Hochberg 1995) was established at $-\log_{10} p > 4.8$ ($p < 0.05$), using 3696 markers according to the results of the LD decay (Rufo et al. 2019). Besides, a frequently used threshold was established at $-\log_{10} P > 3$, as previously reported in the literature (Wang et al., 2014, 2020; Mangini et al., 2018; Sukumaran et al., 2018; Condorelli et al., 2018). Confidence intervals (CIs) for MTAs were estimated for each chromosome according to the LD decay reported by Rufo et al., (2019) using the formula reported in Chardon et al., (2004):

$$S_i^2 = \left(\frac{CI}{3.92}\right)^2$$

where CI corresponded with the LD decay for each chromosome. To simplify the MTA information, the associations were grouped into QTL hotspots. To define a hotspot, the density of MTAs along the chromosome was calculated as the QTL overview index (Chardon et al., 2004) for each cM of the genetic map reported by Wang et al., (2014):

$$U = \frac{\frac{nbQTL}{nbE}}{Total\ length\ of\ map}$$

where nbQTL is the number of QTLs and nbE the total number of experiments.

3.7. Gene annotation and in silico gene expression analysis

Gene annotation for the target region of QTL hotspots was performed using the gene models for high-confidence genes reported for the wheat genome sequence (IWGSC 2018), available at https://wheat-urgi.versailles.inra.fr/Seq-Repository/Annotations. Physical distances were estimated using the genetic distances from the markers flanking the CIs of each QTL hotspot.

In silico expression analysis and the identification of upregulated gene models were carried out using the RNA-seq data available at http://www.wheat-expression.com/ (Ramírez-González et al., 2018) for the following studies: 1) drought and heat stress time-course in seedlings, 2) spikes with water stress, and 3) seedlings with PEG to simulate drought.

Gene Ontology (GO) data were retrieved from the high-confidence gene annotation at https://wheat-urgi.versailles.inra.fr/Seq-Repository/Annotations.

4. Results

4.1. Environmental conditions

The experimental site has a typical Mediterranean climate characterized by an irregular pattern of yearly rainfall distribution, low temperatures in winter that

rise sharply in spring and high temperatures continuing until the end of the crop cycle. Figure 1 represents a graphical summary of the rainfall and maximum and minimum temperatures during the crop cycle across the three years of field trials. All the weather variables were representative of long-term data from the region for each growing season, although 2017 was considered exceptionally dry due to the low rainfall received. The last year (2018) was characterized as the wettest from December (sowing) to June (physiological maturity) with 269 mm of rainfall, whereas the first and second growing seasons with 207 mm and 105 mm of rainfall were rather dry, respectively, suffering severe water scarcity during the grain filling period with only 5 mm of precipitation.

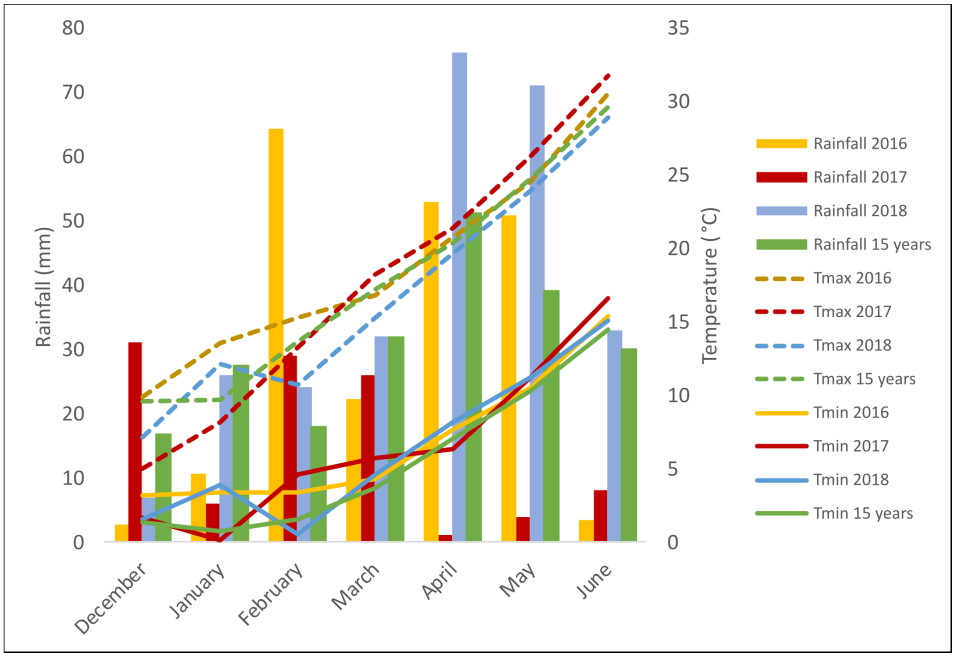

Figure 1. Monthly rainfall (mm) and minimum (Tmin) and maximum (Tmax) temperatures during the growth cycle of each growing season.

4.2. Phenotypic analyses

A summary of the genetic variation is shown in Table 2 and supplementary table S3 for agronomic and VIs-related traits. Agronomic traits showed coefficients of variation (CV) ranging from 36.6% for grain yield to 6.3% for days to anthesis. VIs showed higher CVs during post anthesis with values at anthesis ranging from 17.4% for LAI to 2.1% for NDVI, and post anthesis ranging from 54.0% for LAI to 15.3% for GNDVI. Agronomic traits showed higher values for heritability than VI for most of the traits. For the agronomic traits heritability ranged from 0.9 for yield to 0.1 for GA, whereas for VIs heritability ranged between 0.41 to 0.05 for TCARI/OSAVI and RDVI respectively at anthesis, and between 0.46 to 0.03 for TCARI/

OSAVI and MTVI2 respectively during post anthesis.

Table 2. Summary statistics of the agronomic traits, leaf area index (LAI), and VIs.

	Yield	HI	Biomass	NSm^2	NGm^2	TKW	PH	GA	GS65	GFD
Min	0.5	0.1	4.2	200	1122	12.7	70.2	16.2	132	18
Max	14.4	0.9	33.9	973.3	51548	58.8	163.7	56.6	180	71
Mean	7.7	0.4	17.2	541.9	20199	38.7	105.2	35.8	158.2	34
SD	2.8	0.1	3.9	115.4	7087	6.6	18	9	10	5.3
CV (%)	36.6	22.2	22.7	21.3	35.1	17.1	17.1	25	6.3	15.7
h^2	0.9	0.8	0.4	0.5	0.8	0.6	0.8	0.1	0.3	0.5
	LAI	**NDVI**	**RDVI**	**MSAVI**	**MTVI2**	**TCARI/OSAVI**	**GNDVI**			
Anthesis										
Min	0.1	0.8	0.55	0.61	0.67	-7.26	0.77			
Max	7.67	1	0.87	0.95	1	0.22	0.94			
Mean	5.53	0.94	0.69	0.8	0.88	-0.23	0.87			
SD	0.96	0.02	0.09	0.11	0.12	0.02	0.03			
CV (%)	17.4	2.1	13	13.7	13.6	8.7	3.4			
h^2	0.18	0.4	0.05	0.05	0.13	0.41	0.13			
Postanthesis										
Min	0.72	0.48	0.29	0.28	0.24	-3.89	0.5			
Max	6.62	0.95	0.76	0.86	0.86	0.44	0.92			
Mean	3.26	0.77	0.54	0.61	0.6	-0.03	0.72			
SD	1.76	0.13	0.13	0.18	0.21	0.01	0.11			
CV (%)	54	16.9	24.1	29.5	35	40	15.3			
h^2	0.04	0.08	0.06	0.04	0.03	0.46	0.09			

SD, standard deviation; CV, coefficient of variability; h^2, heritability; GS65, number of days from sowing to anthesis (A); GFD, grain filling duration; HI, harvest index; NSm^2, number of spikes per square meter; NGm^2, number of grains per square meter; TKW, thousand kernel weight; PH, plant height; GA, green area.

The results of the analyses of variance (ANOVAs) for the agronomic traits measured during the three growing seasons are shown in Table 3. The percentage of variability explained by year was the highest for GA (81.6%) and GS65 (67.9%), while the sum of squares of SP was the highest for yield (76.5%), NGm^2 (65.0%), HI (62.6%), PH (61.4%) and GS87 (59.0%). Finally, the highest percentage explained by the interaction between year and SP was found for biomass, NSm^2 and TKW, reaching 86.2%, 84.9% and 71.0%, respectively. Significant differences were found between SPs for all traits. The year and the year x SP interaction was also significant for all traits, except for HI.

Table 4 shows the results of the ANOVA for the VIs and LAI estimated through the MTVI2 at the anthesis and PA stages during 2017 and 2018. Differences between subpopulations and between years, as well as the year x SP interaction, were statistically significant for all traits in both years. The sum of squares for year accounted for 1.3% (NDVI) to 92.9% (RDVI) of the variation at anthesis, whereas at PA, the percentages ranged from 10.0% (TCARI/OSAVI) to 92.3% (LAI). The percentages of the total variation explained by SP ranged from 2.3% (RDVI) to 11.0% (TCARI/OSAVI) at anthesis, while they ranged from 1.0% (MTVI2) to 11.2% (TCARI/OSAVI) PA. Year was the most important for explaining the

Table 3. Analysis of variance for grain yield, harvest index (HI), biomass, number of spikes per square metre (NSm^2), number of grains per square metre (NGm^2), thousand kernel weight (TKW), plant height (PH), green area (GA), number of days from sowing to anthesis (A) (GS65), and grain filling duration (GFD, GS87) for the three years of field trials.

SS (%)	Yield	HI	Biomass	NSm^2	NGm^2	TKW	PH	GA	GS65	GFD
Year	1.3**	0.1	5.2**	8.0**	6.5**	10.5**	2.6**	81.6**	67.9**	8.7**
SP	76.5**	62.6**	8.6**	7.1**	65.0**	18.5**	61.4**	2.2**	30.4**	59.0**
Year x SP	22.2**	37.3	86.2**	84.9*	28.5**	71.0*	36.0**	16.2**	1.7**	32.3**

SS, sum of squares; SP, subpopulation. *p < 0.01; **p < 0.001.

Table 4. Analyses of variance for the LAI estimated through MTVI2 and all the VIs calculated at the anthesis (A) and postanthesis (PA) stages in 2017 and 2018.

SS (%)	LAI	NDVI	RDVI	MSAVI	MTVI2	TCARI/OSAVI	GNDVI
Anthesis							
Year	62.6**	1.3*	92.9**	88.6**	71.1**	15.6**	67.7**
SP	8.9**	8.8**	2.3**	2.7**	8.9**	11.0**	8.0**
Year x SP	28.5**	89.9**	4.8**	8.7**	20**	73.4**	24.3**
Post anthesis							
Year	92.3**	85.3**	88.2**	90.8**	91.8**	10.0**	83.4**
SP	1.1**	2.9**	2.2**	1.4**	1.0**	11.2**	4.8**
Year x SP	6.6**	11.8**	9.6**	7.8**	7.2**	78.8**	11.8**

SS, sum of squares; SP, subpopulation. *p < 0.01; **p < 0.001.

variations in LAI, RDVI, MSAVI, MTVI2 and GNDVI in the two growth stages. SP explained the least percent of variation at both growth stages for all traits. The year x SP interaction accounted for 4.8% (RDVI) to 89.9% (NDVI) of the model variance at anthesis, with the highest values for NDVI and TCARI/OSAVI. The variance explained by the year x SP interaction at PA ranged from 6.6% (LAI) to 78.8% (TCARI/OSAVI).

The mean values of phenotypic traits for each year and SP are shown in Table 5. Yearly means showed that the highest yield was in 2016, a year in which the yield components NSm^2, NGm^2 and TKW reached intermediate values between those obtained in the two subsequent years. The shortest duration of the preanthesis period and the longest GFD were also observed in 2016. On the other hand, the lowest yield, NSm^2 and NGm^2, the heaviest grains and the shortest GFD were observed in 2017. GA reached the highest value in 2016, which was characterized as the wettest year during the period from January-March, i.e., the stem elongation stage, when the trait was measured. In contrast, 2017 was the driest year in the same period, which showed the lowest value for GA. In 2017, PH showed maximum

Table 5. Mean values of grain yield, harvest index (HI), biomass, number of spikes per square metre (NSm^2), number of grains per square metre (NGm^2), thousand kernel weight (TKW), plant height (PH), green area (GA), number of days from sowing to anthesis (A) (GS65), and grain filling duration (GFD, GS87) in a set of 170 landraces and 184 modern cultivars of bread wheat for each growing season and genetic subpopulation.

	Yield (t/ha)	HI	Biomass (t/ha)	NSm^2	NGm^2	TKW (g)	PH (cm)	GA	GS65	GFD
2016	8.0^{a}	0.35^{a}	17.5^{a}	532^{b}	20836^{a}	38.6^{b}	105.8^{b}	46.7^{a}	150^{c}	36^{a}
2017	7.4^{c}	0.36^{a}	16.1^{b}	509^{b}	17877^{b}	41.3^{a}	108.4^{a}	26.6^{c}	155^{b}	32^{c}
2018	7.9^{b}	0.36^{a}	18.1^{a}	583^{a}	21966^{a}	36.2^{c}	101.6^{c}	35.8^{b}	169^{a}	34^{b}
SP1	5.2^{e}	0.29^{b}	16.6^{c}	556^{ab}	14065^{de}	38.2^{bc}	120.4^{a}	37.5^{a}	159^{b}	33^{c}
SP2	5.9^{d}	0.30^{b}	16.7^{c}	534^{b}	15354^{d}	39.2^{b}	121.9^{a}	35.9^{bc}	163^{a}	31^{c}
SP3	4.4^{f}	0.29^{b}	14.9^{d}	569^{ab}	13613^{e}	32.6^{d}	114.9^{b}	33.5^{d}	159^{b}	33^{c}
SP4	10.8^{a}	0.44^{a}	18.5^{a}	568^{a}	28045^{a}	39.3^{b}	87.1^{d}	35.7^{c}	158^{b}	35^{b}
SP5	9.2^{b}	0.43^{a}	18.0^{abc}	475^{c}	21872^{b}	43.2^{a}	92.1^{cd}	37.9^{a}	160^{b}	35^{b}
SP6	9.7^{b}	0.42^{a}	18.0^{ab}	493^{c}	23877^{b}	42.0^{a}	95.7^{c}	38.1^{a}	152^{c}	37^{a}
AD	6.7^{c}	0.31^{b}	17.1^{bc}	550^{ab}	18414^{c}	36.9^{c}	111.4^{b}	36.8^{ab}	157^{b}	35^{b}

Data for each subpopulation represent the mean values across the 3 years. Different letters at each growing season or subpopulation indicate significant differences at $p < 0.01$ using Tukey's honest significant difference test.

SP1, West Mediterranean landraces; SP2, North Mediterranean landraces; SP3, East Mediterranean landraces; SP4, Modern cultivars from France and Italy; SP5, Modern cultivars from Balkans; SP6, Modern cultivars from CIMMYT and ICARDA; AD, admixed genotypes.

values but biomass showed the lowest values at maturity. Finally, in 2018, biomass, the number of spikes and grains per unit area showed high values, and the cycle until anthesis was the longest.

Significant differences in agronomic traits between subpopulations highlighted the division of the whole set into landraces and modern cultivars (Table 5). Modern SPs (SP4, SP5 and SP6) showed higher values of grain yield and yield components, HI, and biomass than landrace SPs. The highest value for grain yield was observed for SP4, in agreement with its higher number of spikes and grains per unit area. SP4 showed the lowest grain weight among modern SPs but was not significantly different from the heaviest grains observed in landraces (SP1 and SP2). As expected, landraces were taller than modern cultivars. SP3 showed the lowest value for GA. For phenology, SP2 took the longest time to reach the anthesis stage, whereas SP6 took the shortest time. In contrast, the GFD was the shortest for SP2 and the longest for SP6. Modern SPs showed a longer GFD than landraces.

The mean values of the VIs and LAI (estimated by MTVI2) at anthesis and at PA for 2017 and 2018 and the different SPs are shown in Table 6. All traits had higher values at anthesis, except for TCARI/OSAVI. For all traits, differences between years were statistically significant at the two stages. The LAI, RDVI and MSAVI showed the highest mean values in 2018. The mean values for TCARI/OSAVI were the

Table 6. Mean values of the LAI estimated by MTVI2 and all the VIs at anthesis (A) and postanthesis (PA) in 2017 and 2018 as well as for each genetic subpopulationand the group of admixed genotypes from a set of 170 landraces and 184 modern bread wheat cultivars.

	LAI	NDVI	RDVI	MSAVI	MTVI2	TCARI/OSAVI	GNDVI
Anthesis							
2017	4.77^{b}	0.95^{a}	0.59^{b}	0.70^{b}	0.78^{a}	0.03^{a}	0.9^{a}
2018	6.29^{a}	0.94^{a}	0.78^{a}	0.91^{a}	0.98^{b}	-0.51^{b}	0.84^{b}
SP1	5.88^{a}	0.96^{a}	0.71^{a}	0.83^{a}	0.92^{a}	-0.08^{a}	0.87^{b}
SP2	5.86^{a}	0.96^{a}	0.71^{a}	0.83^{ab}	0.92^{ab}	-0.08^{a}	0.87^{b}
SP3	5.75^{a}	0.95^{ab}	0.69^{b}	0.81^{c}	0.91^{ab}	-0.02^{a}	0.87^{b}
SP4	5.27^{b}	0.94^{bc}	0.67^{c}	0.78^{d}	0.85^{c}	-0.63^{b}	0.89^{a}
SP5	5.24^{b}	0.93^{c}	0.67^{c}	0.78^{d}	0.84^{c}	-0.21^{a}	0.87^{bc}
SP6	5.20^{b}	0.94^{c}	0.67^{c}	0.78^{d}	0.84^{c}	-0.21^{a}	0.86^{c}
AD	5.70^{a}	0.95^{a}	0.69^{b}	0.81^{bc}	0.90^{b}	-0.09^{a}	0.87^{b}
Post anthesis							
2017	1.56^{b}	0.65^{b}	0.42^{b}	0.43^{b}	0.39^{b}	0.07^{a}	0.62^{b}
2018	4.95^{a}	0.89^{a}	0.67^{a}	0.79^{a}	0.8^{a}	-0.13^{b}	0.82^{a}
SP1	3.21^{bc}	0.76^{b}	0.56^{ab}	0.62^{ab}	0.61^{ab}	0.06^{ab}	0.70^{cd}
SP2	3.32^{b}	0.80^{a}	0.57^{a}	0.64^{a}	0.64^{a}	-0.08^{c}	0.73^{b}
SP3	2.91^{d}	0.74^{cd}	0.54^{c}	0.60^{b}	0.59^{bc}	0.14^{a}	0.68^{d}
SP4	3.50^{a}	0.79^{a}	0.55^{bc}	0.62^{b}	0.60^{b}	-0.20^{d}	0.75^{a}
SP5	3.36^{ab}	0.77^{bc}	0.53^{c}	0.60^{b}	0.59^{bc}	-0.01^{abc}	0.73^{b}
SP6	3.13^{c}	0.74^{d}	0.51^{d}	0.57^{c}	0.57^{c}	0.03^{abc}	0.70^{c}
AD	3.20^{bc}	0.76^{bc}	0.54^{c}	0.60^{b}	0.59^{b}	-0.01^{bc}	0.71^{c}

Years followed by different letters indicate significant differences at $p < 0.01$ using Tukey's honest significant difference test.

SP1, West Mediterranean landraces; SP2, North Mediterranean landraces; SP3, East Mediterranean landraces; SP4, Modern cultivars from France and Italy; SP5, Modern cultivars from Balkans; SP6, Modern cultivars from CIMMYT and ICARDA; AD, admixed genotypes.

highest in 2017. The year 2017 showed the highest values of MTVI2 and GNDVI at anthesis, but these VIs and NDVI were minimal at PA the same year. Due to saturation of the reflectance, NDVI became insensitive at high LAI values (LAI > 3) in both years at anthesis. LAI, NDVI, RDVI, MSAVI and MTVI2 significantly differed between landrace and modern cultivar SPs at anthesis, with higher values being recorded in the landraces. However, no pattern was found for VI traits among SPs PA. SP2 and SP4 had higher mean values for all traits PA, with the exception of TCARI/OSAVI.

Correlation coefficients between agronomic traits, VIs and LAI were calculated (Figure 2), showing highly significant coefficients among agronomic traits as yield with NGm^2 ($r = 0.90$) and PH ($r = -0.71$). Interestingly, when analysed VIs related traits against agronomic traits, highly significant coefficients ($r > 0.61$) were found between GS65 and RDVI, MTVI2, GNDVI, LAI, NDVI_PA and MSAVI_PA, and

between GA and MSAVI, GNDVI, LAI, MTVI2 and RDVI_A. In addition, NGm^2 and PH showed a moderate significant correlation with GNDVI_PA and NDVI_A, respectively (r = 0.46).

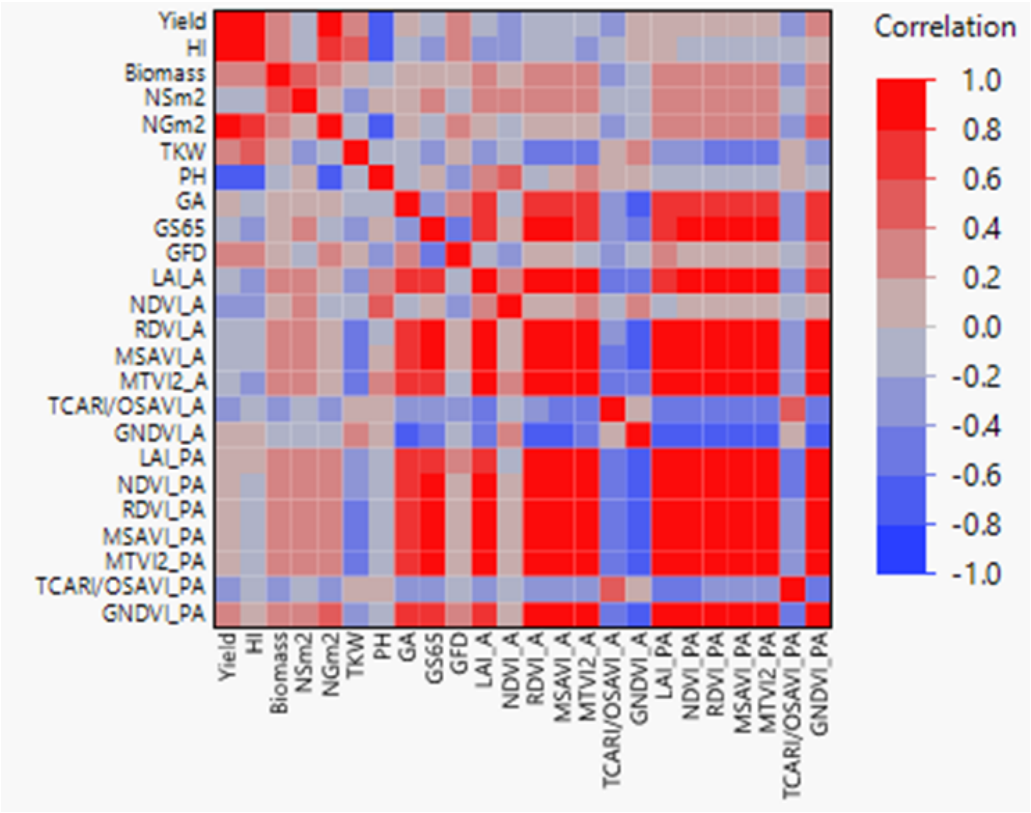

Figure 2. Pearson correlations between agronomic traits, vegetation indices (VIs) and LAI. GS65, number of days from sowing to A; GFS, grain filling duration; HI, harvest index; NSm2, number of spikes per square meter; NGm2, number of grains per square meter; TKW, thousand kernel weight; LAI, leaf area index; PH, plant height; GA, green area; A, anthesis; PA, postanthesis. Significant correlatoins at $P < 0.0001$ were established for $r > 0.45$ and $r < -0.45$.

To quantify the relation between trait variation and population structure, multiple linear regressions were carried out between population structure (q_i) coefficients (Rufo et al. 2019) (Table 7) and phenotypic performance for landrace and modern sets separately and both sets combined. The landrace R^2 values ranged from 0.10 for MSAVI_A to 0.39 for GA, while the modern R^2 values ranged from 0.10 for MTVI2_A to 0.64 for GNDVI_A. When the regressions were conducted on the combined data set, the R^2 values ranged from 0.11 for biomass to 0.60 for NGm^2. The traits yield and GNDVI_PA showed high R^2 values (> 0.35) for each set separately and for the combined set. The highest R^2 values were found in modern

Table 7. Relationship between trait variation and population structure (q-values) for landrace and modern sets separately and the combined set.

Trait	R^2 LR vs. MOD N = 354	R^2 LR N = 170	R^2 MOD N = 184
GS65	-	-	0.47
GFD	-	-	0.16
Yield	0.59	0.37	0.43
HI	0.47	-	-
Biomass	0.11	-	-
NSm^2	-	0.12	0.18
NGm^2	0.58	-	0.42
TKW	-	0.34	0.11
LAI_A	0.20	-	-
NDVI_A	-	-	-
RDVI_A	0.20	-	-
MSAVI_A	0.22	0.10	-
MTVI2_A	0.29	-	0.10
TCARI/OSAVI_A	0.19	-	0.13
GNDVI_A	0.39	-	0.64
LAI_PA	0.20	-	0.39
NDVI_PA	0.14	0.23	0.50
RDVI_PA	-	0.11	0.40
MSAVI_PA	-	0.13	0.44
MTVI2_PA	-	0.11	0.39
TCARI/OSAVI_PA	0.20	0.17	0.17
GNDVI_PA	0.51	0.39	0.60
PH	0.48	-	0.33
GA	-	0.39	-

GS65, number of days from sowing to anthesis; GFD, grain filling duration; HI, harvest index; NSm^2, number of spikes per square meter; NGm^2, number of grains per square meter; TKW, thousand kernel weight; PH, plant height; GA, green area; A, anthesis; PA, postanthesis, LR, landrace, MOD, modern; N, number of genotypes.

set regressions for GNDVI_A, GNDVI_PA, NDVI_PA and GS65. Among the components of yield, TKW showed the highest R^2 values in landrace set regressions, while in modern set regressions, NGm^2 showed the highest R^2 values.

The bidimensional clustering shown in Figure 3 represents the relationships

among accessions and their mean phenotypic performances (3 years for agronomic traits and 2 years for VIs). The horizontal cluster grouped accessions according to their phenotypic similarity based on the traits in the vertical cluster. Horizontal clustering separated two main clusters: cluster A was composed only of landraces, and cluster B included modern cultivars and two landraces: cv 'TRI 11548' from Iraq and cv '1170' from Turkey. Cluster A was characterized by lower yield and yield components, except NSm^2, lower biomass, a shorter GFD but longer GS65, and taller plants than cluster B, but cluster A had higher values for VIs at anthesis, except for GNDVI_A.

Each of these two clusters was separated into two subclusters, A1 and A2 for landraces and B1 and B2 for modern cultivars. Subcluster A1 was represented mainly by south Mediterranean landraces (77%), including those from east and west regions, whereas subcluster A2 contained most of the north Mediterranean landraces (62%). East Mediterranean landraces were in a single cluster within A1, whereas west Mediterranean landraces were distributed in other clusters within A1. Differences among subclusters A1 and A2 were due to higher NSm2 and TCARIOSAVI_PA in A1 and longer cycles until anthesis in A2, along with higher values for GA and VIs at PA. Regarding modern cultivars, subcluster B1 was composed mainly of genotypes carrying the CIMMYT/ICARDA genetic background (SP6) (62%) and included the two landraces allocated to cluster B mentioned above. Moreover, subcluster B2 included 91% of the cultivars from SP4 (French and Italian modern cultivars). Most of the modern Balkan cultivars (SP5) were grouped in subcluster B1. Subcluster B1 was characterized by higher TCARIOSAVI_PA and GA, whereas subcluster B2 was characterized by higher NSm^2, GNDVI_A and the rest of the VIs assessed in PA.

4.3. Marker-trait associations

A summary of the results of the GWAS for all traits per year and for the mean values across years is reported in Figure 4. Due to the low number of marker trait associations (MTA) showing significance above a FDR threshold at of $-\log_{10} P > 4.8$, a common threshold of $-\log_{10} P > 3$, as reported in the literature (Wang et al., 2017, 2020; Mangini et al., 2018; Sukumaran et al., 2018; Condorelli et al., 2018), reported a total of 2579 MTAs (Supplementary Table S4). Manhattan plots for each of the traits and year are represented in supplementary figure S1. The year 2017 presented the highest number of MTAs, 74% of the total number of MTAs, whereas 2018 and the mean across years presented the lowest number of MTAs, 3% and 4%, respectively (Figure 4A). During 2016, only MTAs related to agronomic traits were reported, accounting for 19% of the total MTAs across years, as no multispectral images were captured during that year.

The number of MTAs per chromosome for all years and for the mean values

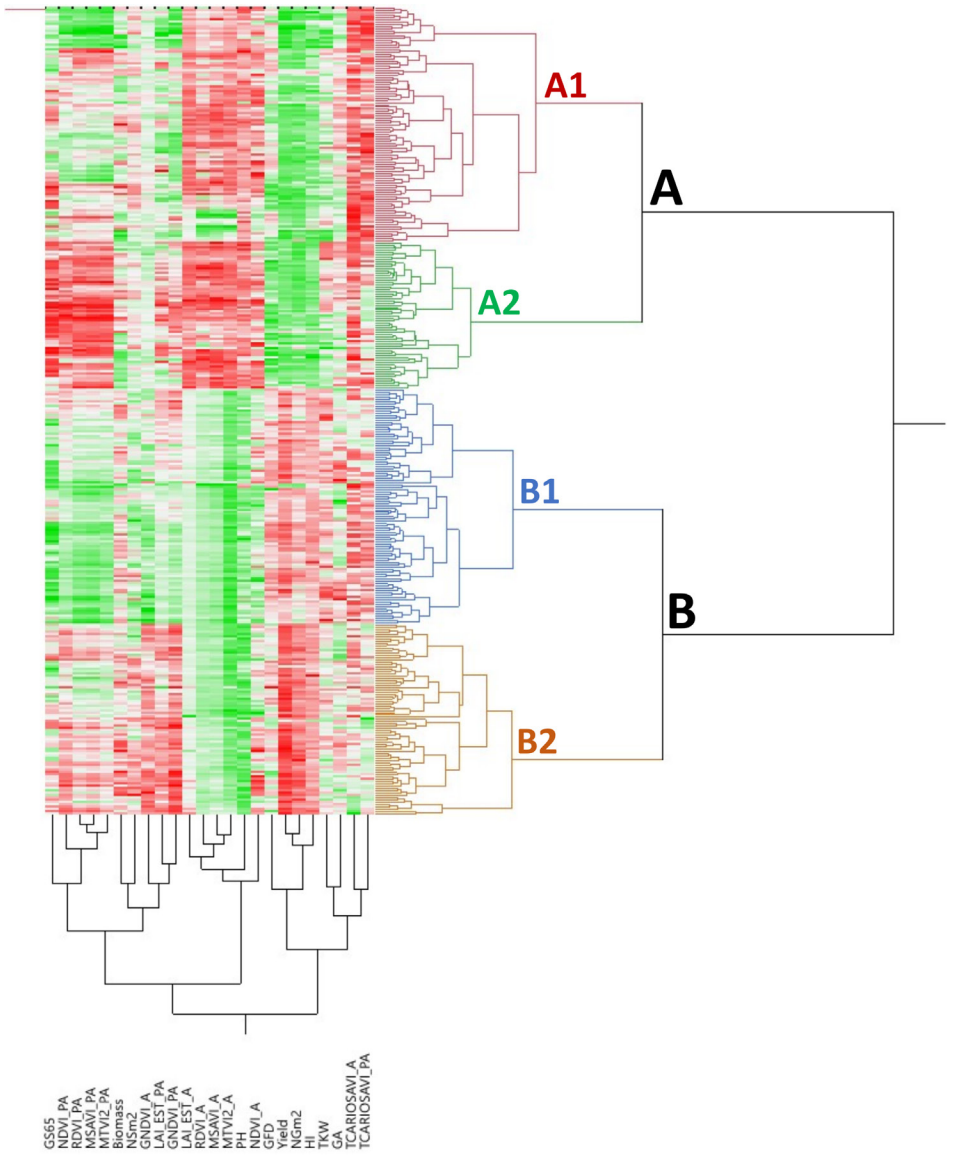

Figure 3. Bidimensional clustering showing the phenotypic relationships between the 354 bread wheat genotypes based on the analysed traits indicated in the vertical cluster at bottom. Red and green colours in the columns indicate high and low values, respectively. Dark, higher values; light, lower values; white, intermediate values.

across years ranged from 9 on chromosome 4D to 354 on chromosome 1B (Figure 4B). Genome B accounted for 48% of the total MTAs, followed by genomes A and D with 41% and 11%, respectively. The percentage of MTAs with a phenotypic variance explained (PVE) lower than 0.10 was 97.5%, which agreed with the highly quantitative nature of the analysed traits (Figure 4C).

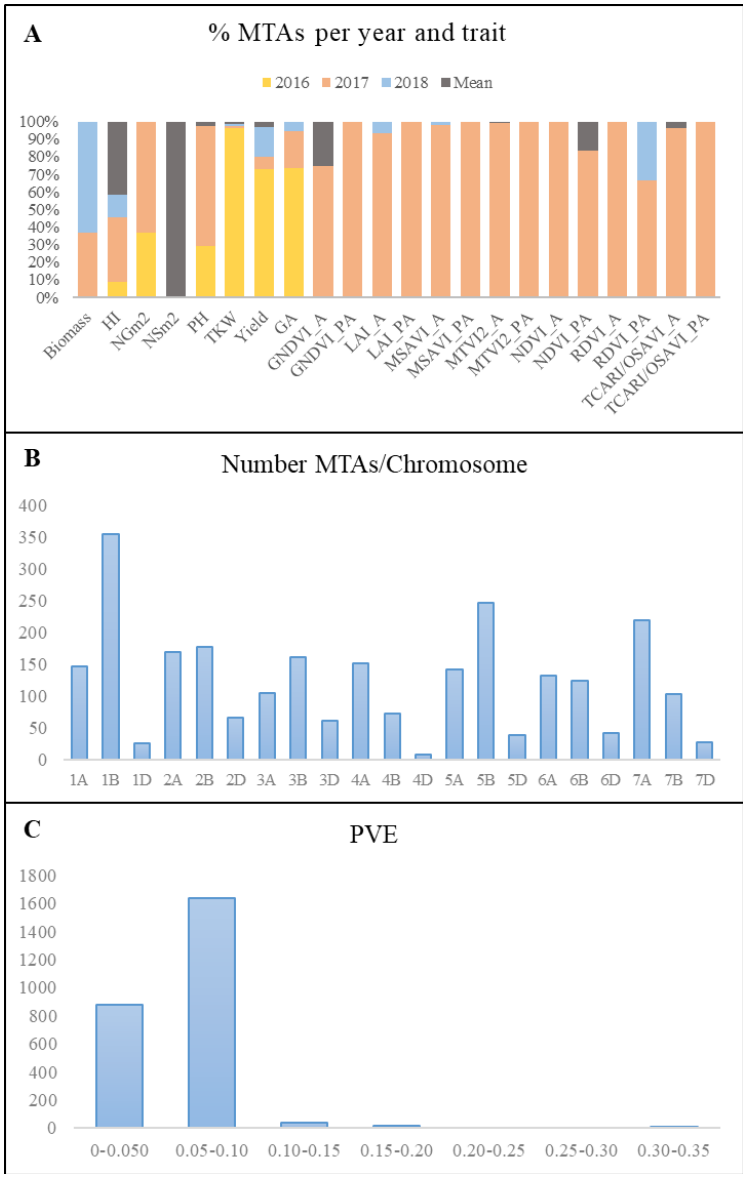

Figure 4. Summary of MTAs. (A) Percentage of MTAs per year and trait. (B) Number of MTAs per chromosome. (C) Phenotypic variance explained (PVE). MTAs, marker-trait associations; HI, harvest index; LAI, leaf area index estimated by MTVI2; NSm^2, number of spikes per square metre; NGm^2, number of grains per square metre; TKW, thousand kernel weight (g); PH, plant height; GA, green area from emergence until the first node; A, anthesis stage; PA, post anthesis.

A total of 815 MTAs were identified for seven agronomic traits (Supplementary Table S5). Yield showed the highest number of MTAs (368), most of them (268) from 2016, whereas only one association was found for NSm^2 with the mean across years. MTAs for TKW were found mainly during 2016 (96%), and those for PH

were found mainly during 2017 (68%).

A total of 1764 MTAs over $-\log_{10} P > 3$ were identified for 15 VI traits (Supplementary Table S5). Among them, 1718 were detected at or before anthesis (GA), and only 46 MTAs were identified PA. Ninety-six percent of the MTAs were identified during 2017, which was the year characterized by the lowest rainfall. TCARIOSAVI_A was the trait with the highest number of MTAs (1243), followed by MTVI2_A with 350.

To identify the genomic regions most involved in trait variation, QTL hotspots were identified using the QTL overview index defined by Chardon et al., (2004) for each cM of the genetic map reported by Wang et al., (2014). Confidence intervals were calculated using the LD decay for each chromosome reported by Rufo et al., (2019).

A total of 209 peaks were identified using the mean of the overview index across the 21 chromosomes (0.7) as the threshold, whereas using a high threshold (3.5), a total of 41 peaks were detected (Figure 5). These 41 peaks were reduced to 28 QTL hotspots (Supplementary Table S6), 12 in genomes A and B and 4 in genome D. To simplify the search for candidate genes, QTL hotspots were excluded when 1) the centromere was included within the hotspot or the CI was higher than 35 Mb and 2) MTAs corresponded only to one year of field experiments. Eleven QTL hotspots grouping 295 MTAs remained for subsequent analysis (Table 8). As shown in Figure 5, hotspots defined by the QTL overview index correspond to genome regions with a higher number of MTAs.

4.4. In silico analysis of candidate genes

A search for CGs to study the relative gene expression levels under abiotic stress conditions and different tissues and developmental stages was performed within the QTL hotspot regions reported in Table 8 using the positions of flanking markers in the 'Chinese Spring' reference genome (IWGSC 2018) at https://wheat-urgi.versailles.inra.fr/Tools/JBrowse. A total of 1342 gene models were detected, and to classify this information, Gene Ontology (GO) for 1025 of the gene models (76%) was downloaded from https://wheat-urgi.versailles.inra.fr/Seq-Repository/Annotations (Figure 6, Supplementary Table S7). Seven hundred ninety-one CG were classified according to molecular function (MF), 183 according to biological process (BP) and 51 according to cellular component (CC). The most represented CGs according to molecular function were 'protein binding' (31%), 'protein kinase activity' (13%) and 'nucleic acid binding' (11%). According to BP, 30% of the CGs were involved in 'defence response' and 19% in 'transport'. Finally, according to cellular component, 27% of the product of CGs were in the nucleus, 22% in the membrane and 14% in the cytoplasm and cell wall.

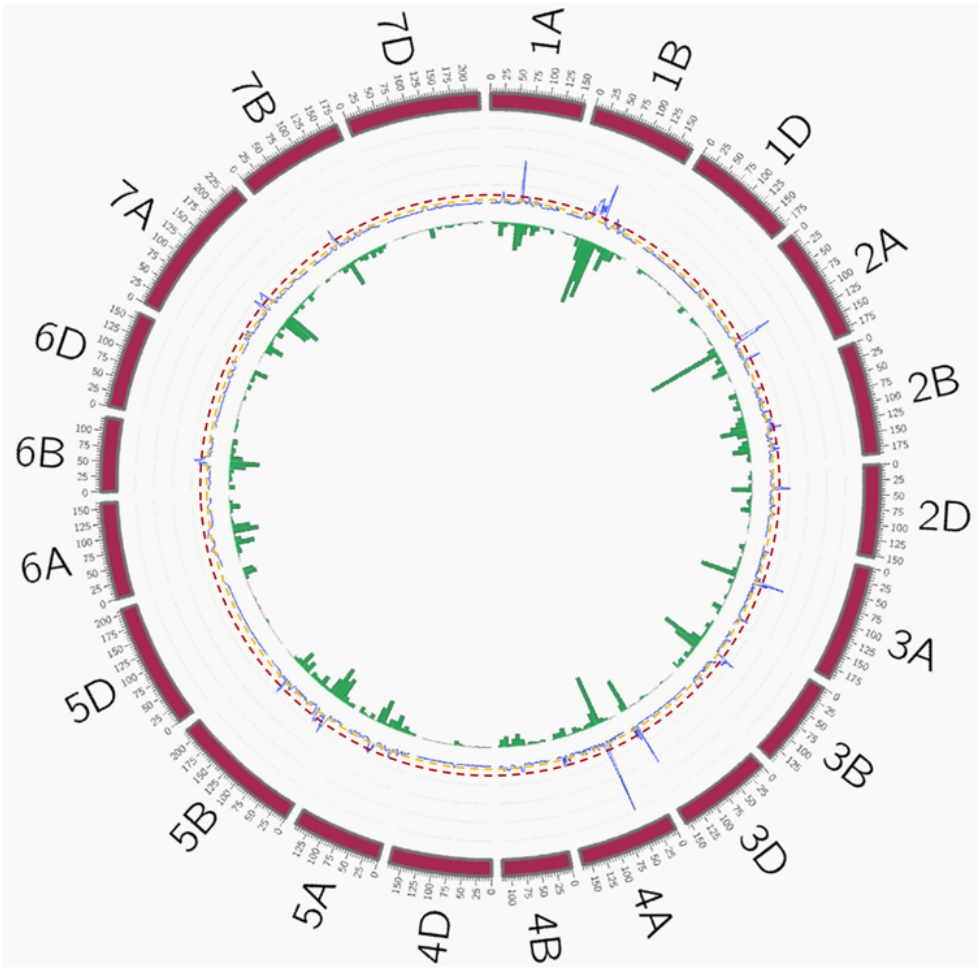

Figure 5. QTL overview index. The index values are represented along chromosomes as a blue line. Yellow and red dashed lines represent the thresholds for average (0.7) and higher values (3.5), respectively. Green bars below the QTL overview index represent the number of significant MTAs per 10 cM ($-\log_{10}P > 3$). QTl, quantitative trait loci; MTAs, marker-trait associations.

Subsequently, a search for differentially expressed genes (DEGs) under three abiotic stress conditions as reported in http://www.wheat-expression.com was carried out. These conditions included 1) drought and heat stress time-course in seedlings, 2) spikes with water stress, and 3) seedlings treated with PEG to simulate drought, and DEGs were analysed in four tissues (roots, shoots/leaves, spikes and grains) during different developmental phases (seedling, vegetative and reproductive). A total of 12 CGs that were upregulated under abiotic stress were found in 6 QTL hotspots and 46 downregulated in 10 QTL hotspots (Figure 7).

Among the different upregulated DEGs, a defensin in hotspot QTL1A.1 showed the highest expression under abiotic stress conditions and was expressed in most of

Table 8. QTL hotspots identified for agronomic and remotely sensed VI-related traits.

QTL Hotspot	Position (cM)	MTAs	Max -log*P*	Env	Number of Traits	Left marker	Position (bp)	Right marker	Position (bp)	CI (Mb)
QTL_1A.1	24-29-31	18	18.44	2	2	BS00056550_51	7294564	Kukri_c29655_239	9579957	2.3
QTL_1B.2	107-114-124	52	6.79	4	6	BS00072791_51	629262942	RFL_Contig2971_282	652455350	23.2
QTL_2A.2	149-151-153	25	5.52	2	6	IAAV880	755788335	Tdurum_contig50839_593	758514679	2.7
QTL_2D.1	39-41-45	29	6.2	2	6	Kukri_c16477_181	61979485	BS00067584_51	79416689	17.4
QTL_3A.2	176-177-178	9	5.91	2	5	Excalibur_c77321_69	737299578	Tdurum_contig31235_99	739397160	2.1
QTL_3D.1	142-143-144	32	6.98	2	5	Kukri_rep_c87658_1436	613696030	wsnp_Ex_c13629_21411429	609166802	4.5
QTL_4A.2	150-151-153	8	5.26	2	5	Excalibur_c74390_108	733915685	RAC875_c11702_1015	739524596	5.6
QTL_5B.2	54-61-64	51	6.22	3	6	BobWHIte_c47103_84	457342562	BS00039492_51	487602616	30.3
QTL_5B.3	68-69-69	10	5.37	3	3	GENE-3574_643	519148851	TA002629-0202	526576640	7.4
QTL_5B.4	159-161-163	21	18.17	3	8	wsnp_Ku_c3151_5892200	680852782	RAC875_c278_1801	683143220	2.3
QTL_7A.1	110-114-119	40	6.48	2	4	BS00024786_51	79542753	Kukri_rep_c101532_1046	84767559	5.2

QTL Hotspot	Traits
QTL_1A.1	GNDVI_A, TCARI/OSAVI_A
QTL_1B.2	Yield, HI, MTVI2_A, PH, TCARI/OSAVI_A, TKW
QTL_2A.2	Yield, HI, LAI_A, MTVI2_A, TCARI/OSAVI_A, TKW
QTL_2D.1	Yield, MSAVI_A, MTVI2_A, PH, RDVI_A, TCARI/OSAVI_A
QTL_3A.2	Yield, LAI_PA, MTVI2_A, PH, TCARI/OSAVI_A
QTL_3D.1	Yield, LAI_PA, MTVI2_A, PH, TCARI/OSAVI_A
QTL_4A.2	Yield, MSAVI_A, PH,RDVI_A, TCARI/OSAVI_A
QTL_5B.2	Yield, HI, MSAVI_A,RDVI_A, TCARI/OSAVI_A, TKW
QTL_5B.3	Yield, HI, TCARI/OSAVI_A
QTL_5B.4	Yield, HI, LAI_A, MSAVI_A, MTVI2_A, PH, RDVI_A, TCARI/OSAVI_A
QTL_7A.1	Yield, MTVI2_A, PH, TCARI/OSAVI_A

Positions are indicated in centimorgans (cMs) and base pairs (bp).

MTAs, marker-trait associations; Env, number of environments; CI, confidence interval; HI, harvest index; TKW, thousand kernel weight; LAI, leaf area index; PH, plant height; A, anthesis; PA, postanthesis.

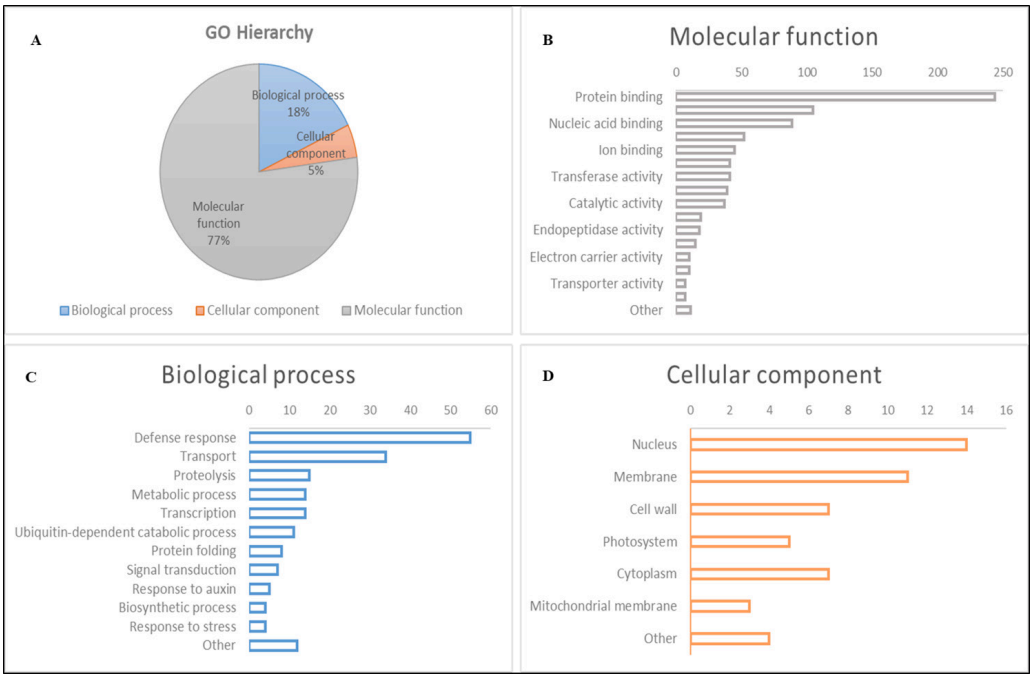

Figure 6. Gene Ontology (GO) classification of gene models within QTL hotspots. (A) GO hierarchy. (B) Molecular function. (C) Biological process. (D) Cellular component.

the tissues and all the developmental phases; it also showed the highest expression levels for each of the phases. All DEGs reported expression in the spikes with a range from 0.02 tpm for cytochrome b in QTL1B.2 to 3.48 tpm for defensin in QTL1A.1. Only four DEGs were expressed in the roots and five in the leaves/shoots and grain. Only zinc finger protein-like 1 in QTL2A.2 was expressed in all four plant tissues, showing the highest expression in roots. Regarding the developmental phase, no expression was reported in any stage for two DEGs, cytochrome b in QTL1B.2 and enoyl-[acyl-carrier-protein] reductase in QTL3D.1. The reproductive phase had the highest number of DEGs (9 out of 10 showing expression), whereas 6 were expressed in the seedlings and only 4 were expressed during the vegetative phase.

Among the downregulated DEGs, the hotspot QTL1B.2 showed the highest number of downregulated DEGS (16), whereas QTL2A.2 did not show any of them. Three DEGs showed expression in all tissues, whereas any of them were downregulated in all of tissues. Six DEGs were expressed only in roots under non stressed conditions, 3 in leaves/shoots, 7 in spikes and 9 in grains. Whereas DEGs non expressed under abiotic stress in only one tissue corresponded to 1 in roots and 6 in grains. According to the developmental phase, 12 DEGs were expressed in all of them, whereas any DEG was downregulated in all of them. One DEG was expressed only in the seedling stage at normal conditions, whereas 2 were expressed

Figure 7. Upregulated and downregulated CGs under abiotic stress conditions in four tissues and three developmental phases. Values are based on $\log_2$ tpm. CGs, candidate genes; tpm, transcripts per million.

QTL hotspot	Gene model	Description	Position (bp)	No stress	Abiotic stress	Roots	Leaves/shoots	Spike	Grain	Seedling	Vegetative	Reproductive
1A.1	TraesCS1A01G013600	Defensin	7642717..7643050		4.78	1.41	1.77	3.48		2.40	2.02	2.38
1B.2	TraesCS1B01G405800	NADH-quinone oxidoreductase subunit H1	633730327..633730890		1.95		0.47	0.73	0.18	0.69		0.57
	TraesCS1B01G413000	Cytochrome b	638717625..638718569		0.92			0.02	0.11			
	TraesCS1B01G413100	Histone deacetylase 2	638729835..638730206	0.14	1.77		0.76	0.47		0.44	0.97	0.13
	TraesCS1B01G413200	Cytochrome c oxidase subunit 3	638748168..638748964	0.16	2.96			1.72	1.68			1.30
2A.2	TraesCS2A01G550300	Zinc finger protein-like 1	757806098..757809625		1.19	2.60	0.07	2.01	1.30	0.60	1.39	1.50
	TraesCS2A01G552400	MYB transcription factor	758509544..758511071		2.14		0.28	1.49		0.96		0.58
3D.1	TraesCS3D01G539400	RNA-binding protein	611408366..611409476	0.05	1.10	0.34		1.11				0.39
	TraesCS3D01G541300	Enoyl-[acyl-carrier-protein] reductase	612604634..612605155		1.00			0.95				
5B.2	TraesCS5B01G286000	Squamosa promoter binding-like protein	471401342..471406922	0.07	2.01			2.15				1.12
	TraesCS5B01G289100	Double stranded RNA binding protein 3	474522122..474523435	0.06	1.21			1.35	0.32			0.63
7A.1	TraesCS7A01G128300	Histone-lysine N-methyltransferase	82194366..82207882	0.01	0.91	1.20		0.66		0.27	0.37	
1A.1	TraesCS1A01G018000	Receptor-like kinase	9359231..936372	0.80			1.66					1.04
	TraesCS1A01G018500	Receptor-like kinase	9549121..9551796	2.11			2.97	0.71		1.14	1.14	2.29
1B.2	TraesCS1B01G398600	Dirigent protein	629702739..629704071	1.56		0.56	2.43			4.02		
	TraesCS1B01G402900	O-acyltransferase WSD1	632654884..632660030	1.46				3.29				1.95
	TraesCS1B01G403100	Potassium voltage-gated channel	632751553..632752549	1.23		0.91	1.26	1.61		1.32	1.15	1.06
	TraesCS1B01G409200	Cysteine protease	636737982..636739203	3.24				5.07				3.72
	TraesCS1B01G413600	DUF1639 family protein	638899900..638914175	0.81					2.89			1.29
	TraesCS1B01G420300	transcription factor-like protein	643278646..643279326	2.32					4.36			2.77
	TraesCS1B01G421400	3-hydroxy-3-methylglutaryl coenzyme A reductase	645280128..645280388	2.30		5.28					3.52	1.37
	TraesCS1B01G421500	Clavata3/ESR (CLE) gene family member	645293982..645294707	1.80		4.79					3.30	0.25
	TraesCS1B01G421600	Clavata3/ESR (CLE) gene family member	645303491..645304149	1.55		4.53					3.19	
	TraesCS1B01G421700	CLAVATA3/ESR (CLE)-related protein 1	645307915..645308829	1.77		4.75					3.42	
	TraesCS1B01G422800	L-lactate dehydrogenase 2	645566647..645567091	0.93		2.23	0.97	0.42		1.14	1.34	0.39
	TraesCS1B01G423200	Cytochrome P450 family protein, expressed	645886704..645888602	1.27				0.71	3.03			1.74
	TraesCS1B01G424900	Fasciclin-like arabinogalactan protein	649368588..649369391	1.66		3.52	0.92	1.04		2.56	2.29	0.39
	TraesCS1B01G425800	Beta purothionin	652047013..652048013	3.14					5.22			3.62
	TraesCS1B01G426000	Beta purothionin	652297912..652298728	5.93					8.01			6.41
	TraesCS1B01G426100	Beta purothionin	652301394..652302238	9.01					11.08			9.48
2D.1	TraesCS2D01G112300	Universal stress protein	62289147..62289911	2.16		3.67	1.79	0.98	1.43	0.65	3.51	0.82
	TraesCS2D01G115500	Uveal autoantigen with coiled-coil domains	64063087..64064626	0.96		2.31	0.72	0.41	0.17	1.07	1.33	0.59
	TraesCS2D01G116900	Zinc finger CCCH domain-containing protein	65717112..65718011	1.89					3.97			2.37
	TraesCS2D01G119700	Calcium-dependent lipid-binding	68921493..68924295	1.72		2.11	1.06	2.68		0.92	1.37	1.76
	TraesCS2D01G125900	Short-chain dehydrogenase/reductase	73312058..73313069	0.97			1.18	1.30		0.54	0.66	0.93
	TraesCS2D01G126700	Tyrosine-specific transport protein, putative	73583386..73588101	1.07			1.84	0.37		1.17	1.02	0.86
	TraesCS2D01G127000	NAD(P)H dehydrogenase	73819387..73820088	1.45		3.99	0.66			1.78	2.87	
	TraesCS2D01G129100	Lectin receptor kinase	74998913..7500118	1.74			2.44	1.07		0.77		2.01
	TraesCS2D01G131800	Auxin-induced in root cultures protein 12	76802063..76802689	1.46		4.23				1.68	2.87	
	TraesCS2D01G134200	Protein FLOWERING LOCUS T	79066232..79068117	2.55			3.14	2.43		1.61	0.34	2.83
3A.2	TraesCS3A01G521900	Receptor-like protein kinase	738236580..738240286	1.16			2.11					1.48
3D.1	TraesCS3D01G534600	Ankyrin repeat family protein	609289397..609291597	1.35			2.31					1.74
	TraesCS3D01G537900	Ankyrin repeat family protein	610929362..610931074	1.49		2.64	1.42	1.18		0.48	2.08	1.13
	TraesCS3D01G538100	P-glycoprotein 6	611007093..611007449	1.13		0.00	0.91	1.34	1.43	1.91	0.21	0.94
	TraesCS3D01G541400	Regulator of chromosome condensation	612763465..612764794	1.53					3.59			2.01
4A.2	TraesCS4A01G485000	Acid invertase 1	738960964..738964520	3.58			3.82	4.11				4.03
	TraesCS4A01G485700	Beta-fructofuranosidase 1	739151693..739154579	1.13					3.21			1.61
5B.2	TraesCS5B01G282100	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenas	466973656..466975035	5.16					7.24			5.64
	TraesCS5B01G288600	Bidirectional sugar transporter SWEET	473871423..473872731	2.36				3.95	1.92			2.86
	TraesCS5B01G291700	Pentatricopeptide repeat	477081079..477082648	1.09		1.81	0.57	1.25	0.60		0.85	1.16
	TraesCS5B01G292600	Sugar transporter, putative	478125645..478129800	1.49				3.28				1.94
	TraesCS5B01G294700	Beta-glucosidase	478916711..478921170	1.53				3.35				2.01
5B.3	TraesCS5B01G336800	ALWAYS EARLY 2	520083186..520083754	0.92				2.77				1.42
	TraesCS5B01G337200	Plant thionin family protein	520552153..520552374	1.99				3.82				2.47
5B.4	TraesCS5B01G518700	Cyclic AMP-dependent transcription factor ATF-5	681699307..681699895	1.62				3.45				2.10
	TraesCS5B01G520200	Ammonium transporter	682536466..682538840	1.42		1.18	2.14			0.91	0.75	1.46
	TraesCS5B01G520600	Serine/threonine-protein kinase	682817804..682820398	1.64		2.34	1.94	1.08		0.99	2.39	1.11
7A.1	TraesCS7A01G131000	CASP-like protein	84690365..84690931	1.17		3.74				1.50	2.48	

only in the vegetative and 22 in the reproductive stages.

5. Discussion

The current study was conducted under typical Mediterranean environmental conditions, with a pattern of increasing temperatures during the spring and an irregular distribution of rainfall across years. A GWAS panel of 354 bread wheat genotypes, including Mediterranean landraces and modern cultivars, was grown for three years under these conditions in northeastern Spain. Given that the decrease in the genetic diversity of wheat occurred during the second half of the 20th century, associated with the introduction of high-yielding semidwarf cultivars (Autrique et al., 1996), landraces are considered a natural reservoir of genetic variation within the species and an invaluable source of new alleles to widen the genetic variability in breeding populations, particularly for traits regulating adaptation to suboptimal environments (Lopes et al., 2015). Recent studies have demonstrated the scarce use of wheat landraces in breeding programmes in the past, as suggested by the high genetic diversity and defined population structure among landrace and modern cultivar subpopulations (Soriano et al., 2016; Rufo et al., 2019).

5.1. Phenotypic performance

The high heritability reported for the agronomic traits, reaching 0.9 for yield and 0.8 for HI, NGm^2 and PH indicated that genetic differentiation among landraces and modern cultivars played a predominant role in determining the variation for these traits.

The ANOVAs showed a large effect of SP on the phenotypic expression of the agronomic traits, whereas year showed the largest effect for most of the VIs, followed by the year x SP interaction, with the SP effect being the lowest. The variability in agronomic traits was mostly caused by the different agronomic performances of wheat landraces and modern cultivars, as reported in previous studies (Soriano et al., 2016; Royo et al., 2020). On the other hand, the high year effect on VIs was likely due to the contrasting water availabilities during the two years in which images were acquired by UAV in the experimental fields. This was not an unexpected result given that the decrease in the rate of growth of wheat caused by drought stress results in a severe reduction in total aboveground biomass (Royo et al., 2004; Villegas et al., 2014).

Yearly variation in weather conditions, particularly water input, resulted in a yield range from 7.4 t/ha in 2017, the driest year, to 7.9 and 8.0 t/ha during the years with higher rainfall. Even with the low water input, the average experimental yields were higher than expected in a severe drought environment. The numbers of spikes and grains per unit area were the highest in 2018, the wettest year, but were the lowest

in 2017. Grain weight showed the opposite pattern, suggesting that under drier and hotter conditions, cultivars filled their grains at a higher rate (1.29 g/day in 2017 and 1.06 g/day in 2016 and 2018), thus showing a shorter GFD in 2017. The high yields recorded, considering the rainfed conditions of the field trials, could be attributed to the high soil fertility (approximately 3% of organic matter) and the superficial subsoil water layer at this site (Royo et al., 2021).

From a genetic viewpoint, a clear separation was observed between landraces and modern cultivars for most of the agronomic traits, which can be attributed to the improvement achieved by breeding. As expected, yield was negatively correlated to PH as reported previously by Royo et al., (2020). Among landraces, those from northern Mediterranean countries characterized by high rainfall and lower temperatures (Royo et al., 2014) showed higher yields due to an increase in the number of grains per unit area and grain weight. These genotypes showed longer cycles until anthesis and a shorter grain filling duration, although this last trait was not statistically significant. Landraces from eastern Mediterranean countries showed lower yields, a lower number of grains and lighter grains but an increase in the number of spikes per unit area compared with landraces from northern Mediterranean countries. Similar results for east Mediterranean landraces were previously reported in durum wheat by Soriano et al., (2018) and Roselló et al., (2019), suggesting an adaptation of landraces from this area to warmer environments, which has been associated with the allelic constitution of vernalization and photoperiod genes (Royo et al., 2020). The results of the current study are in agreement with previous research reporting a tendency for wheat to increase the number of ear-bearing tillers as an adaptation strategy under heat stress (Hütsch et al., 2019) and to increase the number of spikes per unit area in genotypes adapted to dry and warm areas compared with genotypes adapted to wetter and colder areas (Royo et al., 2014, 2020). Among modern cultivars, significant differences were mainly found between SP4 (cultivars from France and Italy) and the other two SPs (Balkans and CIMMYT-ICARDA-derived germplasm). These results suggest that breeding in France and Italy was in the direction of increasing yield through increasing the number of spikes and grains per unit area, whereas the other SPs showed higher TKW. In addition, the regression results of the modern set suggested a high impact of genetic population structure on the number of grains per unit area. Cultivars derived from CIMMYT and ICARDA germplasms reached anthesis earlier, up to eight days earlier compared with Balkan cultivars and six days earlier compared to French and Italian cultivars, which was in line with the high R^2 values obtained in the relation between the modern set structure and GS65. This earliness can help these cultivars from warmer regions avoid heat stress at the end of flowering.

All traits related to HTP showed significant differences between years before and after anthesis, showing higher values for most of the VIs in 2018 than in the

previous year. These higher values agree with the rainfall recorded for both years, which was significantly lower in 2017 than in 2018, mostly during the grain filling period. Furthermore, the difference in the mean values between growth stages was much higher in 2017. This result could be explained by the water scarcity particularly affecting the PA stage, which results in an important loss of chlorophyll content during the grain filling period; therefore, VIs using bands mostly placed in the near-infrared (NIR) and green regions showed lower values (Adamsen et al., 1999). It was supported by the high and positive correlations values between GA and GNDVI and LAI at post anthesis stage, which was the case in 2018. Even though water stress affects the growth of wheat, the effects are higher during the grain filling period (Moragues et al., 2006). Thus, the LAI and GNDVI values decreased at the end of the growing cycle due to a low chlorophyll content associated with senescence during the grain filling period (Rufo et al., 2021). In addition, Gitelson et al., (2002) reported that the sensitivity of the green band was higher than that of the red band when the vegetation fraction was more than 60%, so vegetation indices using green wavelengths perform better at high LAI values, which in wheat under Mediterranean conditions are the highest at booting (Aparicio et al., 2000; Royo et al., 2004; Kyratzis et al., 2017; Rufo et al., 2021). This agreed with the high and positive correlations values between GS65 and GNDVI and LAI, indicating that more days until anthesis provides a high green LAI at post anthesis stage in wet years as 2018. TCARI/OSAVI were higher PA for both years. This agreed with the results of Zarco-Tejada et al., (2005), who reported that in advanced growth stages, chlorophyll indices such as TCARI performed better due to being less sensitive to the loss of turgor and leaf drop. In fact, these authors also stated that the different patterns of the indices across growth stages suggested that chlorophyll-related indices are more suitable closer to harvest, while structural indices related to canopy light scattering and growth are better for early stages.

Differences in the mean values of SPs were found in the two growth stages, with the highest values mainly found at anthesis based on the differences among years, thus highlighting the effect of PA senescence on the chlorophyll content. Landraces and modern cultivars showed significant differences in the LAI and structural VIs at anthesis, and these values were higher in the landraces. As reported in previous studies in durum wheat (García Del Moral et al., 2005; Soriano et al., 2018), landraces are characterized by their tolerance to water scarcity and their superior water use efficiency before anthesis compared to modern cultivars (Subira et al., 2015). Subpopulations showing the highest mean values for the LAI and VIs at PA were those including landraces from the north of the Mediterranean basin (SP2) and modern cultivars from France and Italy (SP4). Landraces from SP2 are better adapted to colder and wetter environments than landraces originating in the southern part of the Mediterranean basin. This adaptation pattern has been associated with

the greatest early soil coverage and more aboveground biomass along the whole cycle length (Royo et al., 2014, 2021). For this reason, the canopy remains green much longer in landraces from northern Mediterranean countries than in those from southern Mediterranean countries (Royo et al., 2014). The same pattern was found in modern cultivars, with GNDVI values remaining higher than those of landraces after anthesis and being significantly different among modern subpopulations. These results agreed with those from the relationship between structure and GNDVI_PA, where the modern set showed the highest R^2 values according to the differences found in GNDVI mean values among modern subpopulations. These results and the capacity to discern between landrace and modern SPs regarding the VI values at anthesis proved the accuracy of HTP in characterizing populations. Several studies have stated the potential of remote sensing for assessing agronomic traits by screening hundreds of plots in a short period of time, minimizing replications (Araus et al., 2018; Juliana et al., 2019; Gracia-Romero et al., 2019). Furthermore, various authors have stressed the suitability of using VIs measured early in the season for grain yield forecasting (Aparicio et al., 2000).

Bidimensional clustering was helpful to jointly visualize the results obtained by Tukey's tests. Moreover, clustering of agronomic and HTP data revealed similarity with the separation obtained by Rufo et al., (2019) using SNP markers and SPs defined based on the structured collection. In both cases, a clear differentiation among landraces and modern cultivars was observed, which resulted in separation into two main clusters. Within the landrace cluster, (A) separation was observed between landraces from northern and southern Mediterranean countries, thus including landraces from SP2 in one cluster and those from SP1 and SP3 in the other cluster, with different groupings among them. Modern cultivars of SP6 (CIMMYT-ICARDA) clustered separately from the French and Italian cultivars (SP4), whereas modern cultivars from the Balkans grouped mostly with SP6. Although these two SPs were separated genetically, no significant differences were found for the agronomic traits, except for phenology, and regarding the VIs, no differences were found at anthesis. Two landraces (TRI 11548 and 1170) were included within modern cultivars from CIMMYT-ICARDA and the Balkans. These two landraces were characterized by a longer GFD, higher HI and lower number of spikes per unit area than the average for landraces. Landrace TRI 11548 from Iraq also showed higher yield and grain weight than other landraces, so it probably resulted from a selection made in an early landrace population.

5.2. Marker trait associations

Dissecting the genetic basis of complex traits in plant breeding is essential to tackle molecular-based approaches for crop improvement. Several efforts have been previously made to identify QTLs and MTAs associated with traits of interest to

carry out marker-assisted selection (MAS) approaches and the introgression of alleles of commercial interest in adapted phenotypes.

The highest number of MTAs related to agronomic traits was found in 2016, while 96% of MTAs related to VIs, GA and the LAI were identified in 2017. It has been reported that under contrasting conditions, the G×E interaction could affect the identification of stable associations among different environments (Mwadzingeni et al., 2017), which could explain the difference in the number of significant associations among the three years of field trials. The highest number of associations for yield and TKW in 2016 could be due to the moderate amount of water input (rainfall) during the spring, together with the longest grain filling duration, as reported in previous studies where grain weight predominantly enhanced yield in wet environments (García Del Moral et al., 2003; Moragues et al., 2006; Royo et al., 2006). Moreover, Royo et al., (2000) found that genotypes with longer GFDs could have greater opportunities to increase grain weight in favourable growing seasons than in warmer and drier seasons. The elevated number of VI-related MTAs found in the driest year (2017) could be explained by the higher variability in traits related to leaf biochemical properties or canopy structural attributes within the set of genotypes grown in environments with water scarcity (Rufo et al., 2021). The highest number of MTAs was identified for PH in 2017, when the CV was higher for this trait. Qaseem et al., 2019 suggested that taller genotypes under drought stress could increase yield accumulation and convert more assimilates into grain. Of the 1764 MTAs detected for VIs, 1718 were found at anthesis, with 1243 for TCARI/OSAVI. This result could be explained by the significant differences found between landraces and modern SPs at anthesis for traits related to HTP. The highest variability was found when comparing SP4 with the rest of the SPs for TCARI/OSAVI, which could explain the elevated number of MTAs for this trait. The distribution of the MTAs across genomes agreed with the results of Rufo et al., (2020), with a similar number in the A and B genomes (41 and 48%, respectively) and the remaining 11% in the D genome. These results are consistent with those of previous studies (Chao et al., 2010; Wang et al., 2014; Gao et al., 2015), which attributed these values to the lower genetic diversity and higher LD found in the D genome of bread wheat compared with genomes A and B (Rufo et al., 2019)

5.3. QTL hotspots

To reduce the complexity of the high number of identified MTAs, QTL hotspots were defined using the QTL overview index proposed by Chardon et al., (2004). Although this statistic was initially used for classical biparental QTL analysis, we adapted it to GWAS using the confidence intervals of the MTAs as the distance of LD decay for each of the chromosomes. As reported in Figure 5, QTL hotspots defined by the high-value threshold of the overview index corresponded to genome

regions with a higher MTA density, thus supporting the suitability of this approach in GWAS. To identify genome regions previously mapped in locations similar to our QTL hotspots and to detect new loci controlling agronomic traits and VIs, a comparison with previous GWAS studies and/or meta-QTL analysis reporting yield and VI-related traits was conducted. Seven of the eleven QTL hotspots have been described previously in the literature. When compared with the meta-QTL analysis reported by Liu et al., (2020) in bread wheat, the QTL hotspots QTL1B.2 and QTL2D.1 were located in similar positions as MQTL1B.7 and MQTL1B.8 and MQTL2D.3 and MQTL2D.4, respectively, controlling grain yield, grain number and TKW under drought and heat stress. QTL1B.2 was also in the homologous region of QTL IWB50693 in durum wheat controlling spike length (Anuarbek et al., 2020), QSN.caas-1BL controlling NSm^2 identified by Gao et al., (2015) in bread wheat and IWB3330 controlling the normalized chlorophyll pigment ratio index (NCPI) identified by Gizaw et al., (2018) in bread wheat. Gao et al., (2015) also found three QTLs for TKW, chlorophyll content and NDVI located in a common region with the hotspot QTL5B.2. This QTL hotspot was also detected in a similar region as QTL yield/root_5B.1 controlling grain yield and shoot length identified by Rufo et al., (2020). QTL1B.2 and QTL5B.2 were found to have homology with several studies. This was an expected result, since they were the longest hotspots including the highest number of MTAs. The genomic regions for QTL hotspots QTL5B.4 and QTL7A.1 were also found in common with three QTLs identified by Anuarbek et al., (2020) controlling the number of fertile spikes and TKW in durum wheat under rainfed conditions. Hotspots QTL1A.1 and QTL2A.2 shared a common position with mtaq-1A.2 reported by Roselló et al., (2019a) in durum wheat and QTL yield/root_2A2 identified by Rufo et al., (2020) in bread wheat, respectively, controlling root-related traits and grain yield in bread wheat.

The detection of these regions in common with other studies opens the opportunity to pyramid different QTLs with pleiotropic effects in future breeding approaches. Moreover, the use of the reference genome sequence makes it possible to rapidly identify common molecular markers to be used in MAS.

5.4. Candidate genes

Gene annotation from the 'Chinese spring' reference genome sequence (IWGSC 2018) allowed us to identify 1342 gene models within the eleven QTL hotspots. Candidate gene mining was performed by searching for DEGs upregulated and downregulated under drought conditions in different tissues and developmental stages through in silico analysis at http://www.wheat-expression.com.

Four candidate genes that were upregulated under drought stress have been previously reported in the literature to be involved in stress resistance. Among

them, in QTL hotspot 1A.1, a defensin protein (TraesCS1A01G013600) was found to show the highest expression under drought stress. According to Kumar et al., (2019), although defensins are mainly involved in antifungal responses, the defensin gene Ca-AFP from chickpea in transgenic Arabidopsis plants was overexpressed under drought stress and induced a higher germination rate, root length and plant biomass. Two gene models enhancing drought and heat stress tolerance were found in QTL hotspot 2A.2: the gene model TraesCS2A01G550300 encoding a zinc finger protein, as reported by (Yoon et al., 2014) in poplar, and the gene model TraesCS2A01G552400 encoding a MYB transcription factor, which was described by Zhao et al., (2018). TaMYB31 from wheat is transcriptionally induced by drought stress in transgenic Arabidopsis plants (Zhao et al., (2018). Finally, in QTL hotspot 5B.2, a squamosa-binding protein was identified (TraesCS5B01G286000). These protein families have been found to be involved in several biological processes. Cao et al., (2019), in expression studies of the Squamosa binding protein from wheat TaSPL16, found that this gene was highly expressed in young panicles but expressed at low levels in seeds, in agreement with the expression profile of TraesCS5B01G286000 found in our study. The ectopic expression of TaSPL16 in Arabidopsis produced a delay in the emergence of vegetative leaves and early flowering and affected yield-related traits. Other gene models upregulated under drought stress, as reported in the RNA-seq analysis from Ramírez-González et al., (2018), such as NADH-quinone oxidoreductase, cytochrome b, histone deacetylase 2, RNA-binding protein, enoyl-[acyl-carrier-protein] reductase, double stranded RNA binding protein 3 and histone-lysine N-methyltransferase, have not been related to drought stress tolerance in the literature, and further experiments are required to assess their expression under drought stress conditions. On the other side, 46 gene models were shown to be downregulated under drought stress. However, the decrease in the expression level seems more associated to the breakdown of the physiological functions due to drought than a causal effect in response to the stress.

6. Conclusions

The use of local landraces in breeding programs is considered a valuable approach to broadening the genetic variability of crops lost during the breeding process and improving traits of commercial importance. The results reported in the present study evidenced the selection for grain yield, HI, NGm^2, and GNDVI_PA during the breeding process. Whereas differentiation among landraces were found for agronomic and VIs (grain yield, TKW, GNDVI_PA, and GA), in modern cultivars SPs differentiation weremainly due to GS65, GNDVI, and NDVI_PA. The use of a statistical approach as the QTL overview index for the definition of QTL hotspots resulted successful for the identification of consensus genome regions including most of the stable marker trait associations across years. The results of this study

will be useful for our wheat breeding program by the selection of the appropriate genotypes carrying favorable alleles for the differential traits that will be useful for designing new crosses.

Using *in silico* approaches allowed gene mining in QTL hotspots, thus facilitating CG identification.

Data Availability Statement: The original contributions presented in the study are included in the article/SupplementaryMaterial, further inquiries can be directed to the corresponding author/s.

Author Contributions: JS: conceptualization and supervision. RR, AL, ML, JB, and JS: methodology. RR, JB, and JS: formal analysis. RR, AL, and JS: data curation. RR: writing—original draft preparation. RR, ML, JB, and JS: writing—review and editing. ML and JS: project administration and funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the projects AGL2015-65351-R and PID2019-109089RB-C31 from the Spanish Ministry of Science and Innovation.

Acknowledgments: Authors acknowledge Conxita Royo for his helpful revision of the manuscript and the contribution of the CERCA Program (Generalitat de Catalunya).

Supplementary Material: The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021. 735192/full#supplementary-material

Supplementary Figure 1: Manhattan plots for the analyzed traits during the 3 years and the mean across years.

Supplementary Table 1: Genotypes used in this study.

Supplementary Table 2: Phenotypic values of the analyzed traits for the 3 years of trials.

Supplementary Table 3: Boxplots and summary statistics for the analyzed traits. Coefficients of variation for agronomic traits among the 3 years are included.

Supplementary Table 4: Total number of marker–trait associations (MTAs) per year and trait with –log10 P > 3.

Supplementary Table 5: Marker trait associations (MTAs).

Supplementary Table 6: QTL hotspots.

Supplementary Table 7: Gene models within QTL hotspots.

7. References

Adamsen FJ, Pinter PJ, Barnes EM, et al (1999) Measuring Wheat Senescence with a Digital Camera. Crop Sci 39:719–724. https://doi.org/10.2135/cropsci1999.0011183X003900030019x

Anuarbek S, Id SA, Pecchioni N, et al (2020) Quantitative trait loci for agronomic traits in tetraploid wheat for enhancing grain yield in Kazakhstan environments. https://doi.org/10.1371/journal.pone.0234863

Aparicio N, Villegas D, Araus JL, et al (2002) Relationship between growth traits and spectral vegetation indices in durum wheat. Crop Sci 42:1547–1555. https://doi.org/10.2135/cropsci2002.1547

Aparicio N, Villegas D, Casadesus J, et al (2000) Spectral vegetation indices as nondestructive tools for determining durum wheat yield. Agron J 92:83–91. https://doi.org/10.2134/agronj2000.92183x

Araus JL, Cairns JE (2014) Field high-throughput phenotyping: The new crop breeding frontier. Trends Plant Sci. 19:52–61

Araus JL, Kefauver SC, Zaman-Allah M, et al (2018) Translating High-Throughput Phenotyping into Genetic Gain. Trends Plant Sci. 23:451–466

Araus JL, Slafer GA, Reynolds MP, Royo C (2002) Plant breeding and drought in C3 cereals: What should we breed for? Ann Bot 89:925–940. https://doi.org/10.1093/aob/mcf049

Autrique E, Nachit MM, Monneveux P, et al (1996) Genetic Diversity in Durum Wheat Based on RFLPs, Morphophysiological Traits, and Coefficient of Parentage. Crop Sci 36:735–742. https://doi.org/10.2135/cropsci1996.0011183X003600030036x

Avni R, Nave M, Barad O, et al (2017) Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. Science (80-) 357:93–97. https://doi.org/10.1126/science.aan0032

Babar MA, Reynolds MP, van Ginkel M, et al (2006) Spectral Reflectance to Estimate Genetic Variation for In-Season Biomass, Leaf Chlorophyll, and Canopy Temperature in Wheat. Crop Sci 46:1046–1057. https://doi.org/10.2135/cropsci2005.0211

Bellvert J, Nieto H, Pelechá A, et al (2021) Remote Sensing Energy Balance Model for the Assessment of Crop Evapotranspiration and Water Status in an Almond Rootstock Collection. Front Plant Sci 12:288. https://doi.org/10.3389/fpls.2021.608967

Bhatta M, Morgounov A, Belamkar V, Baenziger P (2018) Genome-Wide Association Study Reveals Novel Genomic Regions for Grain Yield and Yield-Related Traits in Drought-Stressed Synthetic Hexaploid Wheat. Int J Mol Sci 19:3011. https://doi.org/10.3390/ijms19103011

Bradbury PJ, Zhang Z, Kroon DE, et al (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Budak H, Kantar M, Yucebilgili Kurtoglu K (2013) Drought tolerance in modern and wild wheat. Sci. World J. 2013

Cao R, Guo L, Ma M, et al (2019) Identification and functional characterization of Squamosa promoter binding protein-like gene Taspl16 in wheat (Triticum aestivum l.). Front Plant Sci 10:212. https://doi.org/10.3389/fpls.2019.00212

Casadesús J, Villegas D (2014) Conventional digital cameras as a tool for assessing leaf area index and biomass for cereal breeding. J Integr Plant Biol 56:7–14. https://doi.org/10.1111/jipb.12117

Chao S, Dubcovsky J, Dvorak J, et al (2010) Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (Triticum aestivum L.). BMC Genomics 11:727. https://doi.org/10.1186/1471-2164-11-727

Chardon F, Virlon B, Moreau L, et al (2004) Genetic architecture of flowering time in maize as inferred from quantitative trait loci meta-analysis and synteny conservation with the rice genome. Genetics 168:2169–2185. https://doi.org/10.1534/genetics.104.032375

Condorelli GE, Maccaferri M, Newcomb M, et al (2018) Comparative Aerial and Ground Based High Throughput Phenotyping for the Genetic Dissection of NDVI as a Proxy for Drought Adaptive Traits in Durum Wheat. Front Plant Sci 9:893. https://doi.org/10.3389/fpls.2018.00893

Crowell S, Korniliev P, Falcão A, et al (2016) Genome-wide association and high-resolution phenotyping link Oryza sativa panicle traits to numerous trait-specific QTL clusters. Nat Commun 7:1–14. https://doi.org/10.1038/ncomms10527

Duan T, Chapman SC, Guo Y, Zheng B (2017) Dynamic monitoring of NDVI in wheat agronomy and breeding trials using an unmanned aerial vehicle. F Crop Res 210:71–80. https://doi.org/10.1016/j.fcr.2017.05.025

Edae EA, Byrne PF, Haley SD, et al (2014) Genome-wide association mapping of yield and yield components of spring wheat under contrasting moisture regimes. Theor Appl Genet 127:791–807. https://doi.org/10.1007/s00122-013-2257-8

Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of Linkage Disequilibrium in Plants. Annu Rev Plant Biol 54:357–374. https://doi.org/10.1146/annurev.arplant.54.031902.134907

Gao F, Wen W, Liu J, et al (2015) Genome-Wide Linkage Mapping of QTL for Yield Components, Plant Height and Yield-Related Physiological Traits in the Chinese Wheat Cross Zhou 8425B/Chinese Spring. Front Plant Sci 6:1099. https://doi.org/10.3389/fpls.2015.01099

García Del Moral LF, García Del Moral MB, Molina-Cano JL, Slafer GA (2003) Yield stability and development in two- and six-rowed winter barleys under Mediterranean conditions. F Crop Res 81:109–119. https://doi.org/10.1016/S0378-4290(02)00215-0

García Del Moral LF, Rharrabti Y, Elhani S, et al (2005) Yield formation in Mediterranean durum wheats under two contrasting water regimes based on path-coefficient analysis. Euphytica 146:203–212. https://doi.org/10.1007/s10681-005-9006-2

Giorgi F, Lionello P (2008) Climate change projections for the Mediterranean region. Glob Planet Change 63:90–104. https://doi.org/10.1016/j.gloplacha.2007.09.005

Gitelson AA, Kaufman YJ, Merzlyak MN (1996) Use of a green channel in remote sensing of global vegetation from EOS- MODIS. Remote Sens Environ 58:289–298. https://doi.org/10.1016/S0034-4257(96)00072-7

Gitelson AA, Kaufman YJ, Stark R, Rundquist D (2002) Novel algorithms for remote estimation of vegetation fraction. Remote Sens Environ 80:76–87. https://doi.org/10.1016/S0034-4257(01)00289-9

Gizaw SA, Godoy JG V., Pumphrey MO, Carter AH (2018) Spectral Reflectance for Indirect Selection and Genome-Wide Association Analyses of Grain Yield and Drought Tolerance in North American Spring Wheat. Crop Sci 58:2289–2301. https://doi.org/10.2135/cropsci2017.11.0690

Gomez-Candon D, Bellvert Rios J, Royo C (2021) Performance of the two-sourceenergy balance (TSEB) model as a tool for monitoring the response of durum wheat to drought by high-throughput field phenotyping. Front Plant Sci

Gracia-Romero A, Kefauver SC, Fernandez-Gallego JA, et al (2019) UAV and Ground Image-Based Phenotyping: A Proof of Concept with Durum Wheat. Remote Sens 11:1244. https://doi.org/10.3390/rs11101244

Gracia-Romero A, Kefauver SC, Vergara-Díaz O, et al (2017) Comparative Performance of Ground vs. Aerially Assessed RGB and Multispectral Indices for Early-Growth Evaluation of Maize Performance under Phosphorus Fertilization. Front Plant Sci 8:2004. https://doi.org/10.3389/fpls.2017.02004

Graziani M, Maccaferri M, Royo C, et al (2014) QTL dissection of yield components and morpho-physiological traits in a durum wheat elite population tested in contrasting thermo-pluviometric conditions. Crop Pasture Sci 65:80. https://doi.org/10.1071/CP13349

Haboudane D, Miller JR, Pattey E, et al (2004) Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies : Modeling and validation in the context of precision agriculture. Remote Sens 90:337–352. https://doi.org/10.1016/j.rse.2003.12.013

Haboudane D, Miller JR, Tremblay N, et al (2002) Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. Remote Sens Environ 81:416–426. https://doi.org/10.1016/S0034-4257(02)00018-4

Hütsch BW, Jahn D, Schubert S (2019) Grain yield of wheat (Triticum aestivum L.) under long-term heat stress is sink-limited with stronger inhibition of kernel setting than grain filling. J Agron Crop Sci 205:22–32. https://doi.org/10.1111/jac.12298

Juliana P, Montesinos-López OA, Crossa J, et al (2019) Integrating genomic-enabled prediction and high-throughput phenotyping in breeding for climate-resilient bread wheat. Theor Appl Genet 132:177–194. https://doi.org/10.1007/s00122-018-3206-3

Kulkarni M, Soolanayakanahally R, Ogawa S, Uga Y (2017) Drought Response in Wheat : Key Genes and Regulatory Mechanisms Controlling Root System Architecture and Transpiration Efficiency. 5:1–13. https://doi.org/10.3389/fchem.2017.00106

Kumar M, Yusuf MA, Yadav P, et al (2019) Overexpression of chickpea defensin gene confers tolerance to water-deficit stress in Arabidopsis thaliana. Front Plant Sci 10:290. https://doi.org/10.3389/fpls.2019.00290

Kyratzis AC, Skarlatos DP, Menexes GC, et al (2017) Assessment of vegetation indices derived by UAV imagery for durum wheat phenotyping under a water limited and heat stressed Mediterranean environment. Front Plant Sci 8:. https://doi.org/10.3389/fpls.2017.01114

Leegood RC, Evans JR, Furbank RT (2010) Food security requires genetic advances to increase farm yields. Nature 464:831

Liu H, Mullan D, Zhang C, et al (2020) Major genomic regions responsible for wheat yield and its components as revealed by meta-QTL and genotype–phenotype association analyses. Planta 252:65. https://doi.org/10.1007/s00425-020-03466-3

Lopes MS, El-Basyoni I, Baenziger PS, et al (2015) Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. J Exp Bot 66:3477–3486. https://doi.org/10.1093/jxb/erv122

Lopes MS, Reynolds MP (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. https://doi.org/10.1093/jxb/ers071

Maccaferri M, El-Feki W, Nazemi G, et al (2016) Prioritizing quantitative trait loci for root system architecture in tetraploid wheat. J Exp Bot 67:1161–1178. https://doi.org/10.1093/jxb/erw039

Maccaferri M, Harris NS, Twardziok SO, et al (2019) Durum wheat genome highlights past domestication signatures and future improvement targets. Nat Genet 51:885–895. https://doi.org/10.1038/s41588-019-0381-3

Mangini G, Gadaleta A, Colasuonno P, et al (2018) Genetic dissection of the relationships between grain yield components by genome-wide association mapping in a collection of tetraploid wheats. PLoS One 13:e0190162. https://doi.org/10.1371/journal.pone.0190162

Mérida-García R, Bentley AR, Gálvez S, et al (2020) Mapping agronomic and quality traits in elite durum wheat lines under differing water regimes. Agronomy 10:144. https://doi.org/10.3390/agronomy10010144

Moragues M, García Del Moral LF, Moralejo M, Royo C (2006) Yield formation strategies of durum wheat landraces with distinct pattern of dispersal within the Mediterranean basin: II. Biomass production and allocation. F Crop Res 95:182–193. https://doi.org/10.1016/j.fcr.2005.02.008

Mwadzingeni L, Shimelis H, Rees DJG, Tsilo TJ (2017) Genome-wide association analysis of agronomic traits in wheat under drought-stressed and non-stressed conditions. PLoS One 12:e0171692. https://doi.org/10.1371/journal.pone.0171692

Qaseem MF, Qureshi R, Shaheen H, Shafqat N (2019) Genome-wide association analyses for yield and yield-related traits in bread wheat (Triticum aestivum L.) under pre-anthesis combined heat and drought stress in field conditions. PLoS One 14:e0213407. https://doi.org/10.1371/journal.pone.0213407

Qi J, Chehbouni A, Huete AR, et al (1994) A modified soil adjusted vegetation index. Remote Sens Environ 48:119–126. https://doi.org/10.1016/0034-4257(94)90134-1

Ramírez-González RH, Borrill P, Lang D, et al (2018) The transcriptional landscape of polyploid wheat. Science (80-) 361:. https://doi.org/10.1126/science.aar6089

Roselló M, Royo C, Sanchez-Garcia M, Soriano JM (2019a) Genetic Dissection of the Seminal Root System Architecture in Mediterranean Durum Wheat Landraces by Genome-Wide Association Study. Agronomy 9:364. https://doi.org/10.3390/agronomy9070364

Roselló M, Villegas D, Álvaro F, et al (2019b) Unravelling the relationship between adaptation pattern and yield formation strategies in Mediterranean durum wheat landraces. Eur J Agron 107:43–52. https://doi.org/10.1016/j.eja.2019.04.003

Roujean JL, Breon FM (1995) Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. Remote Sens Environ 51:375–384. https://doi.org/10.1016/0034-4257(94)00114-3

Rouse JW, Haas RH, Schell JA, et al (1974) Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation. NASA/GSFC, Greenbelt, MD

Royo C, Abaza M, Blanco R, Garcia del Moral LF (2000) Triticale grain growth and morphometry as affected by drought stress, late sowing and simulated drought stress. Aust J Plant Physiol 27:1051–1059. https://doi.org/10.1071/pp99113

Royo C, Ammar K, Villegas D, Soriano JM (2021) Agronomic, physiological and genetic changes associated with evolution, migration and modern breeding in durum wheat. Front Plant Sci 1–25. https://doi.org/10.3389/fpls.2021.674470

Royo C, Aparicio N, Blanco R, Villegas D (2004) Leaf and green area development of durum wheat genotypes grown under Mediterranean conditions. Eur J Agron 20:419–430. https://doi.org/10.1016/S1161-0301(03)00058-3

Royo C, Aparicio N, Villegas D, et al (2003) Usefulness of spectral reflectance indices as durum wheat yield predictors under contrasting Mediterranean conditions. Int J Remote Sens 24:4403–4419. https://doi.org/10.1080/0143116031000150059

Royo C, Dreisigacker S, Ammar K, Villegas D (2020) Agronomic performance of durum wheat landraces and modern cultivars and its association with genotypic variation in vernalization response (Vrn-1) and photoperiod sensitivity (Ppd-1) genes. Eur J Agron 120:126129. https://doi.org/10.1016/j.eja.2020.126129

Royo C, Fanny AE, Lvaro Ae A´, et al (2006) Genetic changes in durum wheat yield components and associated traits in Italian and Spanish varieties during the 20th century. Euphytica 259–270. https://doi.org/10.1007/s10681-006-9327-9

Royo C, Nazco R, Villegas D (2014) The climate of the zone of origin of Mediterranean durum wheat (Triticum durum Desf.) landraces affects their agronomic performance. Genet Resour Crop Evol 61:1345–1358. https://doi.org/10.1007/s10722-014-0116-3

Rufo R, Alvaro F, Royo C, Soriano JM (2019) From landraces to improved cultivars: Assessment of genetic diversity and population structure of Mediterranean wheat using SNP markers. PLoS One 14:. https://doi.org/10.1371/journal.pone.0219867

Rufo R, Salvi S, Royo C, Soriano J (2020) Exploring the Genetic Architecture of Root-Related Traits in Mediterranean Bread Wheat Landraces by Genome-Wide Association Analysis. Agronomy 10:613. https://doi.org/10.3390/agronomy10050613

Rufo R, Soriano JM, Villegas D, et al (2021) Using Unmanned Aerial Vehicle and Ground-Based RGB Indices to Assess Agronomic Performance of Wheat Landraces and Cultivars in a Mediterranean-Type Environment. Remote Sens 13:1–18

Slafer GA, Araus JL, Royo C, García Del Moral LF (2005) Promising eco-physiological traits for genetic improvement of cereal yields in Mediterranean environments. Ann Appl Biol 146:61–70. https://doi.org/10.1111/j.1744-7348.2005.04048.x

Soriano JM, Malosetti M, Roselló M, et al (2017) Dissecting the old Mediterranean durum wheat genetic architecture for phenology, biomass and yield formation

by association mapping and QTL meta-analysis. PLoS One 12:1–19. https://doi.org/10.1371/journal.pone.0178290

Soriano JM, Sansaloni C, Ammar K, Royo C (2021) Labelling Selective Sweeps Used in Durum Wheat Breeding from a Diverse and Structured Panel of Landraces and Cultivars. Biology (Basel) 10:1–20

Soriano JM, Villegas D, Aranzana MJ, et al (2016) Genetic structure of modern durum wheat cultivars and mediterranean landraces matches with their agronomic performance. PLoS One 11:e0160983. https://doi.org/10.1371/journal.pone.0160983

Soriano JM, Villegas D, Sorrells ME, Royo C (2018) Durum wheat landraces from east and west regions of the mediterranean basin are genetically distinct for yield components and phenology. Front Plant Sci 9:1–9. https://doi.org/10.3389/fpls.2018.00080

Stenberg P, Rautiainen M, Manninen T, et al (2004) Reduced simple ratio better than NDVI for estimating LAI in Finnish pine and spruce stands. Silva Fenn 38:3–14. https://doi.org/10.14214/sf.431

Subira J, Álvaro F, García del Moral LF, Royo C (2015) Breeding effects on the cultivar×environment interaction of durum wheat yield. Eur J Agron 68:78–88. https://doi.org/10.1016/j.eja.2015.04.009

Sukumaran S, Reynolds MP, Sansaloni C (2018) Genome-Wide Association Analyses Identify QTL Hotspots for Yield and Component Traits in Durum Wheat Grown under Yield Potential, Drought, and Heat Stress Environments. Front Plant Sci 9:81. https://doi.org/10.3389/fpls.2018.00081

Valluru R, Reynolds MP, Davies WJ, Sukumaran S (2017) Phenotypic and genome-wide association analysis of spike ethylene in diverse wheat genotypes under heat stress. New Phytol 214:271–283. https://doi.org/10.1111/nph.14367

Wang J, Ding B, Guo Y, et al (2014a) Overexpression of a wheat phospholipase D gene, TaPLDα, enhances tolerance to drought and osmotic stress in Arabidopsis thaliana. Planta 240:103–115. https://doi.org/10.1007/s00425-014-2066-6

Wang S-X, Zhu Y-L, Zhang D-X, et al (2017) Genome-wide association study for grain yield and related traits in elite wheat varieties and advanced lines using SNP markers. PLoS One 12:e0188662. https://doi.org/10.1371/journal.pone.0188662

Wang S, Wong D, Forrest K, et al (2014b) Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. Plant Biotechnol J 12:787–796. https://doi.org/10.1111/pbi.12183

Wang X, Guan P, Xin M, et al (2020) Genome-wide association study identifies QTL for thousand grain weight in winter wheat under normal- and late-sown stressed environments. Theor Appl Genet 134:143–157. https://doi.org/10.1007/s00122-020-03687-w

Ward JH (1963) Hierarchical Grouping to Optimize an Objective Function. J Am Stat Assoc 58:236–244. https://doi.org/10.1080/01621459.1963.10500845

White J, Andrade-Sanchez P, Gore M (2012) Field-based phenomics for plant genetics research. F Crop Res 133:101–113

Xie C, Yang C (2020) A review on plant high-throughput phenotyping traits using UAV-based sensors. Comput. Electron. Agric. 178:105731

Yoon S-K, Park E-J, Choi Y-I, et al (2014) Response to drought and salt stress in leaves of poplar (Populus alba × Populus glandulosa): Expression profiling by oligonucleotide microarray analysis. Plant Physiol Biochem 84:158–168. https://doi.org/10.1016/J.PLAPHY.2014.09.008

Zadoks JC, Chang TT, Konzak CF (1974) A decimal code for the growth stages of cereals. Weed Res 14:415–421. https://doi.org/10.1111/j.1365-3180.1974.tb01084.x

Zarco-Tejada PJ, Ustin SL, Whiting ML (2005) Temporal and spatial relationships between within-field yield variability in cotton and high-spatial hyperspectral remote sensing imagery. Agron J 97:641–653. https://doi.org/10.2134/agronj2003.0257

Zhao Y, Cheng X, Liu X, et al (2018) The Wheat MYB Transcription Factor TaMYB31 Is Involved in Drought Stress Responses in Arabidopsis. Front Plant Sci 9:1426. https://doi.org/10.3389/fpls.2018.01426

Discussion

Discussion

1. Introduction

Plants were among the first species selected for the studies that led to the birth of genetics and during the last century plant breeding provided the basis for the present levels of food production. However, agriculture will have to meet important demands in the near future. According to a new United Nations report (https://www.un.org/en/global-issues/population), the world's population is expected to increase by 2 billion persons in the next 30 years, from 7.7 billion currently to 9.7 billion in 2050, and could arrive at nearly 11 billion around 2100. With these numbers, the production of sufficient, safe, and healthy food for an increasing human population is a huge challenge. To cover the expected food demand of world population, wheat production needs to raise by 1.7% per year until 2050 (Leegood et al. 2010). But this production also must meet the need for a reduced impact of agriculture in a changing environment, the cultivated land will be dramatically reduced due to the current agricultural management causing a strong soil degradation and an over-exploitation of natural resources. Thus, plant scientists in the current century have the objective to increase the crop productivity per unit area, enhancing sustainability and preserving biodiversity (Colasuonno et al. 2021). Moreover, the impact of agricultural practices on climate change has redesigned the breeding paradigm: the new improved varieties must be able to produce with the minimum environmental effects in an adverse scenario of climate change, which it fits with the concept of 'sustainable agricultural ecosystems' (Royo et al. 2017). In other words, the genetic adaptation of crops to environmental conditions and not the opposite.

On a climate change scenario, to incorporate resilience in the new varieties is one of the main issues for the breeding programs. To tackle this objective, exploring the use of underexploited germplasm, as landraces or wild relatives, is of great interest to provide new favourable alleles. Enhancing grain yield under rainfed conditions is being achieved thanks to drought-tolerant and yield-stable germplasm, as well as through the exploitation of the genetic diversity present in wild wheats and traditional varieties (Crespo-Herrera et al. 2018). In this regard, the Mediterranean wheat landraces are considered a powerful pool of genetic resources to improve the resistance to biotic and abiotic stresses since are considered to hold the largest genetic variability within the species (Royo et al. 2017). Incorporating novel alleles associated with tolerance to abiotic stress and resistance to the major diseases of wheat from this primary gene pool into elite varieties will be essential to prevent the stagnation of wheat production (De Vita and Taranto 2019). Therefore, identifying the genotypes showing the extreme phenotypes within the pool of Mediterranean landraces and the associated markers provide the opportunity for introgressing

suitable traits in elite cultivars by marker-assisted breeding using the most recent technologies to speed the process.

Wheat crop models have been used over decades to predict growth and yield as influenced by the growing environment and agronomic practices, and to identify desirable traits leading to the specification of crop ideotypes (Gouache et al. 2015; Ramirez-Villegas et al. 2015; Tao et al. 2017). In recent years, modelling has become important for supporting plant breeding, in particular in designing ideotypes, for target environments and future climatic conditions (Semenov et al. 2014). One of the most significant effects of climate change and global warming trend is phenology fitting, which supposes the advance of the phenological phases of crops (Rezaei et al. 2018). Therefore, by modulating the sowing dates and choosing the most suitable genotypes, it is possible to adapt the crop to develop each growth stage under the best environmental conditions, especially adjusting flowering time to escape from terminal drought.

A better understanding of the physiology and genetic basis of drought-adaptive traits will be necessary to improve drought resistance by genomics approaches. Use of new breeding technologies to re-design high quality crops for a precise control of the water regime with the adoption of high-throughput platforms will streamline the collection of good phenotypic data while increasing the cost-effectiveness of phenotyping (Tuberosa 2012). Several technologies are already having an impact in plant breeding: Molecular markers are being used routinely for breeders; sequences of the main cultivated plants are becoming available, and resequencing of varieties allows massive genotyping and the discovery and use of complex genetic characters; methods for phenotyping are also being developed based on image analysis. They may become useful to follow the state of crops in the field helping farmers to take decisions. Cost-effective phenotyping will become increasingly strategic for further dissecting drought-adaptive traits of interest for farming under drought-prone conditions (Tuberosa 2012). In the last few years, the advances in next-generation sequencing (NGS) technologies has reduced the costs of DNA sequencing and the huge availability of sequence data on web databases allow for drastic reductions in the time required to identify candidate genes models involved in abiotic stress resistance. Furthermore, the recent progress in genome editing allows the efficient and precise modification of genes in almost all plant species.

Currently, drought-prone zones account for 15 million ha, considered as suboptimal conditions for wheat production (Crespo-Herrera et al. 2018). Then, multiple site trials and the evaluation of elite germplasm in international research centres have been essential to target these regions (Singh et al. 2007; Braun et al. 2010). Varieties derived from the Green Revolution are available without personal or corporate intellectual property rights (IPRs), thanks to the exchange of germplasm

between institutions (Royo et al. 2009). Most of the research developed in the last decades has been conducted by the public sector, such as the wheat breeding program at CIMMYT, which periodically evaluates the production of large collection of germplasm and the breeding progress in specific environments (Crespo-Herrera et al. 2018). Besides, superior lines with yield stability across multiple locations must be tested under the prevailing conditions of each site following local agronomic practices. In this context, the future of wheat breeding programmes will rely on international cooperation between multidisciplinary teams with the aim to share the knowledge acquired in multiple disciplines to meet the challenges of researchers in the next few decades.

In this context, the research conducted by this PhD Thesis was addressed to provide scientific knowledge and useful tools that can help in the development of the next generation of superior bread wheat varieties with resilience to the increased drought expected in the next decades to occur during grain filling in a Mediterranean-type environment.

To achieve that goal, a bread wheat germplasm collection, the MED6WHEAT IRTA-panel, composed by 170 landraces from 24 Mediterranean countries and 184 modern varieties cultivated in 19 countries in the region was used. The whole or part of the panel has been used to address all the objectives of this PhD Thesis. This panel was selected from a larger collection of landraces and modern cultivars in order to adjust phenology to an optimal range for performing field trials, and it is the first time that has been used for research purposes. In the case of objective 3, only the landraces pool was used to identify and map QTLs controlling traits related to seminal root architecture. The field trials were conducted under rainfed conditions for three consecutive growing seasons (2016, 2017 and 2018). Genotyping was done with 13177 SNPs from the Illumina Infinium 15K Wheat SNP Array at Trait Genetics GmbH (Gatersleben, Germany).

The **first chapter** explored the existence of genetic and/or geographic structures and genetic diversity among wheat landraces from the Mediterranean Basin and modern cultivars representative of the region, as well as the gene flow between the subpopulations (SP) identified.

Chapter 2 analysed the seminal root system architecture (RSA) of landraces. Differences among genetic subpopulations were detected and correlation analysis among RSA and grain yield under rainfed conditions was performed. Finally, QTL hotspots for RSA traits were identified and a search for candidate genes within them was carried out.

Chapter 3 examined the performance of vegetation indices (VIs) obtained at different dates from two different approaches, a 4-band multispectral camera (Parrot

Sequoia) on-board an UAV and ground-based RGB images, to assess agronomic traits of large panels of bread wheat landraces and modern cultivars adapted to Mediterranean conditions.

Finally, the aim of **Chapter 4** was to identify molecular markers linked to important agronomic traits, VIs and plant features related to drought resistance assessed by HTP, the definition of the most important QTL hotspots for such traits and the *in silico* analysis of the underlying candidate genes expressed under different stress conditions and tissues.

2. Agronomic performance of bread wheat cultivars in a Mediterranean-type environment

Mediterranean climate is characterized by an irregular pattern of yearly rainfall distribution, low temperatures in winter that rise sharply in spring and high temperatures continuing until the end of the summer. The Mediterranean Sea is a marginal and semi-enclosed sea located on the western side of a large continental area and surrounded by Europe to the North, Africa to the South and Asia to the East. The Mediterranean region is a transitional zone between dry and wet climates, and in these semiarid areas the direct evaporation from the soil plays an important role on the surface energy balance, with evapotranspiration strongly dependent on available soil moisture (Koster et al. 2004; Seneviratne et al. 2010; Taylor 2015). The climatic trend in the Mediterranean region, analysed during historical periods (Lionello and Scarascia 2018) and by future climate scenarios (Tramblay et al. 2020) is towards a decrease in precipitation amounts and occurrence, associated with an increasing frequency of drought episodes. Soil moisture is considered critical since it strongly influences agricultural droughts and flood generation processes (Mimeau et al. 2021). Although the sensitivity of soil moisture to changes in precipitation and temperature is similar at the different sites, these changes are modulated by the climate characteristics of the different seasons, with a higher sensitivity of soil moisture to precipitation intermittence in dryer and warmer regions (Mimeau et al. 2021).

Royo et al. (2014) pointed out the existence of four climatic zones involving the main wheat-growing areas within the Mediterranean Basin, ranging from warm and dry conditions in the south to cool and wet in the north. During the dispersal of wheat along the Mediterranean Basin there was a progressive adaptation of traditional varieties or local landraces specifically to these climatic zones from east to west (Royo et al. 2014; Soriano et al. 2016).

2.1. Adaptive traits of the MED6WHEAT IRTA-panel subpopulations

The genotypes included in the MED6WHEAT IRTA-panel were representative of the variability existing in the species in the Mediterranean Basin. Care was taken to have enough number of genotypes representatives of different drought resistance capacities, having a balanced number of them with common constitutive traits between the four climatic zones identified in the Mediterranean Basin for wheat cultivation (Royo et al. 2014). The study conducted in the **first Chapter** of this PhD Thesis showed a clear separation based on historical breeding periods, identifying two groups of germplasm: landraces and modern cultivars. Significant differences for agronomic traits between SP highlighted this division, with the modern cultivars showing higher values of yield and yield components, HI, and biomass (**Chapter 4**). It can be attributable to the improvement achieved by breeding activities. Furthermore, differences between landraces and modern cultivars were pointed out in the study of the estimation of agronomic traits through multispectral and RGB imagery (**Chapter 3**), where predictions showed clearly better values in the modern cultivars. It could be partially due to the different size and structure of the canopy of both types of germplasm, as landraces were much taller and had a different canopy architecture, saturating some of the VIs at high LAI values. The level of admixture within the landrace set was much higher than between the modern cultivars. The incorporation of alleles from more than one gene pool because of the spread of wheat from different ancestral populations into landraces explains their high level of admixture (Oliveira et al. 2012), while the low level between modern cultivars could be due to the development by breeding programmes of cultivars with specific adaptation to the local environments and the use of different genetic resources among breeding pools. Population structure for landraces showed a clearly geographic pattern and SPs were classified as western (SP1), northern (SP2) and eastern Mediterranean (SP3) according to the geographical region of the countries mostly represented in the SP. This classification denoted a migration from the centre of wheat domestication in the Fertile Crescent to the west of the Mediterranean Basin, as reported by Moragues et al. (2006) and Soriano et al. (2016) in durum wheat. Modern cultivars were structured according to the breeding programmes developing them: France/Italy (SP4), Balkan/eastern European countries (SP5) and CIMMYT/ICARDA (SP6). Analysis of genetic differentiation and gene flow indicated that modern cultivars showed higher genetic differentiation, this is lower gene flow among cultivars from different breeding origins. This result agrees with the higher admixture in landraces, which indicates higher genetic exchange among geographic regions.

The results of the study conducted in **Chapter 4** added consistent information

regarding the variability for agronomic traits caused by the different agronomic performance of wheat landraces and modern cultivars across the Mediterranean Basin, as described by multiple linear regression.

Landraces from Eastern Mediterranean countries (SP3), which are closer to the origin of wheat in the Fertile Crescent, demonstrated an adaptation to warmer and drier environmental conditions showing lower yields, lower number of grains and lighter grains, but higher number of spikes per unit area according to results previously reported by Soriano et al. (2018) in durum wheat. A superior number of tillers and spikes is likely associated to the water used by the crop before flowering in dry and warm regions as an adaptation mechanism to heat stress (Hütsch et al. 2019). These landraces also had the lowest early vigour and biomass among landraces SPs, which agreed with Royo et al. (2014), who found that genotypes from south-eastern Mediterranean countries had less biomass, increasing when moving from there to west and northern Mediterranean regions. On the contrary, north Mediterranean cultivars (SP2) showed higher yields achieved by an increase in the number of grains per unit area and grain weight and showed longer cycle till anthesis and lower grain filling duration. In optimal environments, the grain weight contribution to yield is enhanced, while in warmer environments the number of spikes becomes more relevant (García Del Moral et al. 2003; Moragues et al. 2006; Hütsch et al. 2019). Moreover, they showed the highest LAI and VIs values after anthesis, which it is attributed to a canopy that remains green much longer in landraces from northern Mediterranean countries than in those from southern Mediterranean countries, as expected from the wetter and colder climate of the Northern countries (Royo et al. 2014). This climate has been associated with a greatest early soil coverage and more aboveground biomass along the whole cycle length (Royo et al. 2014). Landraces from western Mediterranean countries, where temperatures are lower than in the east and more water is available during crop cycle, particularly after anthesis showed close values of agronomic traits to those from the east, as GFD, HI or the number of the days from sowing to anthesis. It has been proved that the number of days from sowing to the main growth stages consistently increases from eastern and western Mediterranean genotypes to those from the north (Royo et al. 2014), which results an useful strategy to escape from terminal drought stress (Annicchiarico et al. 2009).

Among modern cultivars, the main separation was found between SP4 (French/Italian breeding programmes) and SP6 (CIMMYT/ICARDA germplasm), whereas cultivars from Balkans (SP5) grouped mostly with SP6. Genotypes with a CIMMYT/ICARDA origin reached anthesis earlier and had the longest GFD than the other two SPs (**Chapter 4**), which can help these cultivars from warmer regions to skip heat stress at the end of flowering and to escape from terminal drought (Annicchiarico et al. 2009). Cultivars from SP6 clustered separated from the French and Italian

ones in the bidimensional clustering, probably due to the significative differences found for phenology between them. Besides, SP4 genotypes showed the highest values of yield, number of spikes and number of grains per unit area, suggesting that breeding in France and Italy was in the direction of increasing yield through the increase of the number of spikes and grains per unit area, whereas the other SPs showed higher grain weight.

2.2. Root system architecture and its relation to drought stress

The wide morphological plasticity of the root system to different soil conditions are well known (Christopher et al. 2013; Paez-Garcia et al. 2015). Traits defining root system architecture (RSA) are critical for wheat adaptation to drought environments and non-optimal nutritional supply conditions as reported in Sanguineti et al. (2007). The seminal roots are the first to penetrate the soil and remain functional during the whole plant cycle (Chochois et al. 2015; Maccaferri et al. 2016), being important for early vigour and crop establishment in dryland areas (Reynolds and Tuberosa 2008). Optimization of root features and anatomy will lead to increase water-use efficiency (WUE) (Wasson et al. 2012). **Chapter 2** of the current Thesis studied three RSA traits and two related traits (shoot length and seed weight) in the set of 170 bread wheat landraces detecting differences among the genetic subpopulations previously defined.

The comparison of the mean trait values between the three SPs indicated that Eastern Mediterranean landraces (SP3) showed the lowest number of roots, the longest shoot length (SL), the lowest seed weight (SW) and the widest seminal root angle (SRA). A widest root angle allows them to cover a larger soil area and be more efficient in water uptake than landraces that originated in wetter areas, which agreed with results for RSA in durum wheat obtained by Roselló et al. (2019). On the contrary, Northern Mediterranean cultivars (SP2) from countries characterized by higher rainfall and lower temperatures showed a narrower SRA and a higher number of roots when comparing with Eastern Mediterranean cultivars (SP3). Landraces in SP2 with a narrow SRA showed the highest values of grain weight among landrace SPs, as well as number of grains per unit area (**Chapter 4**), indicating that the higher yields in well-watered environmental conditions are achieved by an increase in the number of grains per unit area and grain weight. It has been proved a positive correlation between the number of seminal roots and grain yield (Canè et al. 2014), which was ascribed to the fact that a higher number of seminal roots provide greater early vigour enhancing water uptake (Blum 1996; Richards 2006, 2008; Reynolds and Tuberosa 2008).

3. Genome based approaches for wheat breeding

The present work focuses on the assessment of agronomic, vegetation indices (VIs)-related and root-related traits in Mediterranean-Type Environment and the identification of genomic regions regulating them using a genome-wide association approach (GWAS). The use of GWAS has become in the last years a common approach to identify molecular markers linked to drought and heat stress resistance in durum and bread wheat (Maccaferri et al. 2016; Valluru et al. 2017), as well as to narrow down the genome regions for candidate gene (CGs) identification due to the release of genome sequences of different wheat species, emmer wheat (Avni et al. 2017), bread wheat (IWGSC, 2018) and durum wheat (Maccaferri et al. 2019). In this thesis only the bread wheat genome sequence has been used to identify genomic regions associated with the analysed traits in **Chapters 2 and 4**. The comparison among the different traits opens the possibility of detecting QTLs with pleiotropic effects in future studies and will be of special interest for the identification of alleles to be introgressed in the breeding programmes. Among the QTL hotspots showed in Figure 1, 48% of VIs hotspots were located in common region with QTL

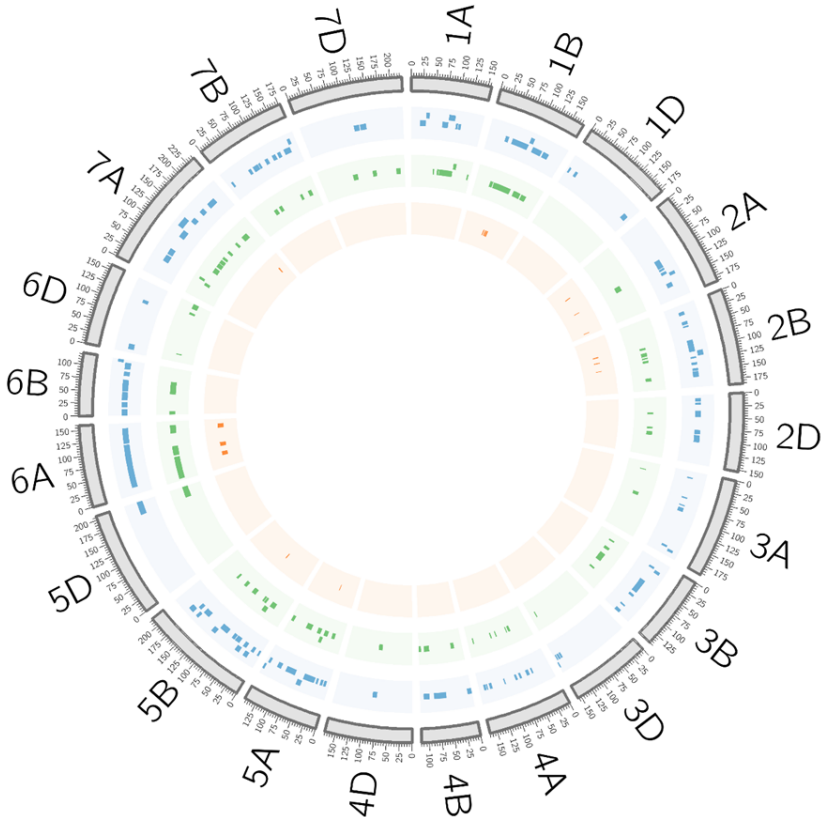

Figure 1. QTL hotspots for VIs traits (blue), agronomic traits (green) and RSA traits (orange).

hotspots for other traits (agronomic and root-related), 75% of the hotspots for agronomic traits shared genomic position with VIs and RSA hotspots, whereas 80% of the latest did it with VIs and agronomic QTL hotspots. This result indicates the possibility of pleiotropic effects of the genes controlling traits related to drought stress and agronomic performance.

The set of 170 bread wheat landraces was used in **Chapter 2** to identify molecular markers associated with the RSA and related traits, as well as grain yield. The identification of genotypes showing extreme phenotypes within the pool of Mediterranean landraces and associated markers provides the opportunity for introgressing suitable traits in elite cultivars. In this Chapter, a total of 135 MTAs were identified using a common threshold of -log10 P>3. To simplify and integrate closely linked MTAs in a consensus region, 15 QTL hotspots were detected based on the results of LD decay reported in **Chapter 1**, which were used to define the confidence intervals (CI) for the QTL hotspots. From them, only 4 shared genome regions with other studies reporting associations for root-related traits, thus suggesting the importance of wheat Mediterranean landraces for the identification of new loci controlling these traits. Comparing the results with the meta-QTL analysis performed by Soriano and Alvaro (2019) only rootQTL6A.3 was in the same region of a previously mapped meta-QTL, the RootMQTL74.

In **Chapter 4**, a GWAS was conducted on the whole germplasm collection of the MED6WHEAT IRTA-panel identifying molecular markers associated with agronomic and remotely sensed vegetation indices (VIs)–related traits under rainfed conditions. A total number of 2579 MTAs were identified using a common threshold of -log10 P>3, and as previously described in **Chapter 2**, 11 QTL hotspots grouping 295 MTAs were detected. In this case a different approach was used. The QTL overview index, defined by Chardon et al. (2004) to discover meta-QTLs, was used. CIs were calculated according to the LD decay, but hotspots were defined by the density of MTAs for each cM of the genetic map. Seven of the eleven QTL hotspots had been described previously in the literature when compared with other GWAS analyses and QTL meta-analysis. In addition, Gao et al. (2015) found three QTLs for TKW, chlorophyll content and NDVI in a common region with the hotspot QTL5B.2, which controls agronomic traits (grain yield, HI and grain weight) and VIs-related traits (MSAVI, RDVI and TCARI/OSAVI). The complexity and highly quantitative nature of grain yield is confirmed based on its presence in most of the QTL hotspots (10 out of 11). All the VIs involved in the QTL hotspots were found before anthesis. It was in agreement with the results found for VI assessment of agronomic traits in **Chapter 3**, which concluded that vegetation indices perform better at high LAI values, which under Mediterranean conditions are maximum at booting stage (Aparicio et al. 2000; Royo et al. 2004).

4. In silico identification of candidate genes involved in stress resistance

The release of genome sequence for bread wheat (IWGSC 2018) has made possible the identification of gene models or CG within QTL intervals putatively involved in drought and agronomic performance without carrying new functional studies. Thus, the physical position of the markers flanking the QTL hotspots CIs was searched on the reference sequence and the annotated gene models were identified.

In **Chapter 2**, 1489 gene models were identified within the 15 QTL hotspots, but only 31 of them involved in plant development and abiotic stress were selected. A deeper analysis was carried out in **Chapter 4** through the web-based tool developed by Ramírez-González et al. (2018), http://www.wheat-expression.com/. Search for differentially expressed genes (DEG) upregulated under abiotic stress conditions was carried out detecting 12 CGs in 6 QTL hotspots. Common gene models in **Chapter 2 and 4** were found. Among them, zinc finger protein and MYB transcription factors, which have been found in a high number of QTL hotspots. Zinc finger proteins are involved in several processes, such as regulation of plant growth and development, and response to abiotic stresses (Chang et al. 2016). Moreover, zinc finger protein-like 1 in QTL2A.2 (**Chapter 4**), although it was expressed in most of the plant tissues, it showed the highest expression in roots. MYB transcription factors are involved in salt and drought stress adaptation in wheat and their expression in roots is strongly associated to responses to abiotic stresses (Lee et al. 2007). In our study, this type of genes has been found mainly expressed in the spike. The CG analysis conducted in **Chapter 4** also detected a defensin protein which showed the highest expression among the rest of CGs under drought stress. According to Kumar et al. (2019) these proteins induced higher germination rate, root length and plant biomass in transgenic Arabidopsis plants under drought stress. Although these findings supported the involvement of CGs in drought stress response, new functional studies in the germplasm used in this PhD thesis should be performed for gene validation.

5. References

Annicchiarico P, Royo C, Bellah F, Moragues M (2009) Relationships among adaptation patterns, morphophysiological traits and molecular markers in durum wheat. Plant Breed 128:164–171. https://doi.org/10.1111/j.1439-0523.2008.01557.x

Aparicio N, Villegas D, Casadesus J, et al (2000) Spectral vegetation indices as nondestructive tools for determining durum wheat yield. Agron J 92:83–91. https://doi.org/10.2134/agronj2000.92183x

Avni R, Nave M, Barad O, et al (2017) Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. Science (80-) 357:93–97. https://doi.org/10.1126/science.aan0032

Blum A (1996) Crop responses to drought and the interpretation of adaptation. In: Drought Tolerance in Higher Plants: Genetical, Physiological and Molecular Biological Analysis. Springer Netherlands, pp 57–70

Braun HJ, Atlin G, Payne T (2010) Multi-location Testing as a Tool to Identify Plan Response to Global Climate Change. CABI Climate ch:115–138. https://doi.org/10.1079/9781845936334.0115

Canè MA, Maccaferri M, Nazemi G, et al (2014) Association mapping for root architectural traits in durum wheat seedlings as related to agronomic performance. Mol Breed 34:1629–1645. https://doi.org/10.1007/s11032-014-0177-1

Chang H, Chen D, Kam J, et al (2016) Abiotic stress upregulated TaZFP34 represses the expression of type-B response regulator and SHY2 genes and enhances root to shoot ratio in wheat. Plant Sci 252:88–102. https://doi.org/10.1016/j.plantsci.2016.07.011

Chardon F, Virlon B, Moreau L, et al (2004) Genetic architecture of flowering time in maize as inferred from quantitative trait loci meta-analysis and synteny conservation with the rice genome. Genetics 168:2169–2185. https://doi.org/10.1534/genetics.104.032375

Chochois V, Voge JP, Rebetzke GJ, Watt M (2015) Variation in adult plant phenotypes and partitioning among seed and stem-borne roots across Brachypodium distachyon accessions to exploit in breeding cereals for well-watered and drought environments. Plant Physiol 168:953–967. https://doi.org/10.1104/pp.15.00095

Christopher J, Christopher M, Jennings R, et al (2013) QTL for root angle and number in a population developed from bread wheats (Triticum aestivum) with contrasting adaptation to water-limited environments. Theor Appl Genet 126:1563–1574. https://doi.org/10.1007/s00122-013-2074-0

Colasuonno P, Marcotuli I, Gadaleta A, Soriano JM (2021) From genetic maps to qtl cloning: An overview for durum wheat. Plants 10:1–25. https://doi.org/10.3390/plants10020315

Crespo-Herrera LA, Crossa J, Huerta-Espino J, et al (2018) Genetic Gains for Grain Yield in CIMMYT's Semi-Arid Wheat Yield Trials Grown in Suboptimal Environments. Crop Sci 58:1890–1898. https://doi.org/10.2135/cropsci2018.01.0017

De Vita P, Taranto F (2019) Durum wheat (Triticum turgidum ssp. durum) breeding to meet the challenge of climate change. In: Advances in Plant Breeding Strategies: Cereals. Springer International Publishing, pp 471–524

Gao F, Wen W, Liu J, et al (2015) Genome-Wide Linkage Mapping of QTL for Yield Components, Plant Height and Yield-Related Physiological Traits in the Chinese Wheat Cross Zhou 8425B/Chinese Spring. Front Plant Sci 6:1099. https://doi.org/10.3389/fpls.2015.01099

García Del Moral LF, García Del Moral MB, Molina-Cano JL, Slafer GA (2003) Yield stability and development in two- and six-rowed winter barleys under Mediterranean conditions. F Crop Res 81:109–119. https://doi.org/10.1016/S0378-4290(02)00215-0

Gouache D, Bogard M, Thepot S, et al (2015) From Ideotypes to Genotypes: Approaches to Adapt Wheat Phenology to Climate Change. Procedia Environ Sci 29:34–35. https://doi.org/10.1016/j.proenv.2015.07.143

Hütsch BW, Jahn D, Schubert S (2019) Grain yield of wheat (Triticum aestivum L.) under long-term heat stress is sink-limited with stronger inhibition of kernel setting than grain filling. J Agron Crop Sci 205:22–32. https://doi.org/10.1111/jac.12298

Koster RD, Dirmeyer PA, Guo Z, et al (2004) Regions of strong coupling between soil moisture and precipitation. Science (80-) 305:1138–1140. https://doi.org/10.1126/science.1100217

Kumar M, Yusuf MA, Yadav P, et al (2019) Overexpression of chickpea defensin gene confers tolerance to water-deficit stress in Arabidopsis thaliana. Front Plant Sci 10:290. https://doi.org/10.3389/fpls.2019.00290

Lee TG, Jang CS, Kim JY, et al (2007) A Myb transcription factor (TaMyb1) from wheat roots is expressed during hypoxia: Roles in response to the oxygen concentration in root environment and abiotic stresses. Physiol Plant 129:375–385. https://doi.org/10.1111/j.1399-3054.2006.00828.x

Leegood RC, Evans JR, Furbank RT (2010) Food security requires genetic advances to increase farm yields. Nature 464:831

Lionello P, Scarascia L (2018) The relation between climate change in the Mediterranean region and global warming. Reg Environ Chang 18:1481–1493. https://doi.org/10.1007/s10113-018-1290-1

Maccaferri M, El-Feki W, Nazemi G, et al (2016) Prioritizing quantitative trait loci for root system architecture in tetraploid wheat. J Exp Bot 67:1161–1178. https://doi.org/10.1093/jxb/erw039

Maccaferri M, Harris NS, Twardziok SO, et al (2019) Durum wheat genome highlights past domestication signatures and future improvement targets. Nat Genet 51:885–895. https://doi.org/10.1038/s41588-019-0381-3

Mimeau L, Tramblay Y, Brocca L, et al (2021) Modeling the response of soil moisture to climate variability in the Mediterranean region. Hydrol Earth Syst Sci 25:653–669. https://doi.org/10.5194/hess-25-653-2021

Moragues M, García Del Moral LF, Moralejo M, Royo C (2006) Yield formation strategies of durum wheat landraces with distinct pattern of dispersal within the Mediterranean basin: II. Biomass production and allocation. F Crop Res 95:182–193. https://doi.org/10.1016/j.fcr.2005.02.008

Oliveira HR, Campana MG, Jones H, et al (2012) Tetraploid wheat landraces in the Mediterranean basin: Taxonomy, evolution and genetic diversity. PLoS One 7:37063. https://doi.org/10.1371/journal.pone.0037063

Paez-Garcia A, Motes C, Scheible W-R, et al (2015) Root Traits and Phenotyping Strategies for Plant Improvement. Plants 4:334–355. https://doi.org/10.3390/plants4020334

Ramírez-González RH, Borrill P, Lang D, et al (2018) The transcriptional landscape of polyploid wheat. Science (80-) 361:. https://doi.org/10.1126/science.aar6089

Ramirez-Villegas J, Watson J, Challinor AJ (2015) Identifying traits for genotypic adaptation using crop models. J. Exp. Bot. 66:3451–3462

Reynolds M, Tuberosa R (2008) Translational research impacting on crop productivity in drought-prone environments. Curr Opin Plant Biol 11:171–179. https://doi.org/10.1016/J.PBI.2008.02.005

Rezaei EE, Siebert S, Hüging H, Ewert F (2018) Climate change effect on wheat phenology depends on cultivar change. Sci Rep 8:1–10. https://doi.org/10.1038/s41598-018-23101-2

Richards RA (2008) Genetic opportunities to improve cereal root systems for dryland agriculture. Plant Prod. Sci. 11:12–16

Richards RA (2006) Physiological traits used in the breeding of new cultivars for water-scarce environments. In: Agricultural Water Management. Elsevier, pp 197–211

Roselló M, Royo C, Sanchez-Garcia M, Soriano JM (2019) Genetic Dissection of the Seminal Root System Architecture in Mediterranean Durum Wheat Landraces by Genome-Wide Association Study. Agronomy 9:364. https://doi.org/10.3390/agronomy9070364

Royo C, Aparicio N, Blanco R, Villegas D (2004) Leaf and green area development of durum wheat genotypes grown under Mediterranean conditions. Eur J Agron 20:419–430. https://doi.org/10.1016/S1161-0301(03)00058-3

Royo C, Elias EM, Manthey FA (2009) Durum Wheat Breeding. In: Cereals. Springer US, pp 199–226

Royo C, Nazco R, Villegas D (2014) The climate of the zone of origin of Mediterranean durum wheat (Triticum durum Desf.) landraces affects their agronomic performance. Genet Resour Crop Evol 61:1345–1358. https://doi.org/10.1007/s10722-014-0116-3

Royo C, Soriano JM, Alvaro F (2017) Wheat: A Crop in the Bottom of the Mediterranean Diet Pyramid. Intech 16:381–399

Sanguineti MC, Li S, Maccaferri M, et al (2007) Genetic dissection of seminal root architecture in elite durum wheat germplasm. Ann Appl Biol 151:291–305. https://doi.org/10.1111/j.1744-7348.2007.00198.x

Semenov MA, Stratonovitch P, Alghabari F, Gooding MJ (2014) Adapting wheat in Europe for climate change. J. Cereal Sci. 59:245–256

Seneviratne SI, Corti T, Davin EL, et al (2010) Investigating soil moisture-climate interactions in a changing climate: A review. Earth-Science Rev. 99:125–161

Singh RP, Huerta-Espino J, Sharma R, et al (2007) High yielding spring bread wheat germplasm for global irrigated and rainfed production systems. In: Euphytica. Springer, pp 351–363

Soriano JM, Alvaro F (2019) Discovering consensus genomic regions in wheat for root-related traits by QTL meta-analysis. Sci Rep 9:10537. https://doi.org/10.1038/s41598-019-47038-2

Soriano JM, Villegas D, Aranzana MJ, et al (2016) Genetic structure of modern durum wheat cultivars and mediterranean landraces matches with their agronomic performance. PLoS One 11:e0160983. https://doi.org/10.1371/journal.pone.0160983

Soriano JM, Villegas D, Sorrells ME, Royo C (2018) Durum wheat landraces from east and west regions of the mediterranean basin are genetically distinct for yield

components and phenology. Front Plant Sci 9:1–9. https://doi.org/10.3389/fpls.2018.00080

Tao F, Rötter RP, Palosuo T, et al (2017) Designing future barley ideotypes using a crop model ensemble. Eur J Agron 82:144–162. https://doi.org/10.1016/j.eja.2016.10.012

Taylor CM (2015) Detecting soil moisture impacts on convective initiation in Europe. Geophys Res Lett 42:4631–4638. https://doi.org/10.1002/2015GL064030

Tramblay Y, Llasat MC, Randin C, Coppola E (2020) Climate change impacts on water resources in the Mediterranean. Reg. Environ. Chang. 20:1–3

Tuberosa R (2012) Phenotyping for drought tolerance of crops in the genomics era. Front Physiol 3 SEP:1–26. https://doi.org/10.3389/fphys.2012.00347

Valluru R, Reynolds MP, Davies WJ, Sukumaran S (2017) Phenotypic and genome-wide association analysis of spike ethylene in diverse wheat genotypes under heat stress. New Phytol 214:271–283. https://doi.org/10.1111/nph.14367

Wasson AP, Richards RA, Chatrath R, et al (2012) Traits and selection strategies to improve root systems and water uptake in water-limited wheat crops. J. Exp. Bot. 63:3485–3498

Conclusions

Conclusions

1. The structure for landraces showed a geographical pattern with different levels of admixture, mainly justified by physical distances between the regions where they were collected, whereas the structure for modern cultivars pointed out differences and similarities between the genetic pools handled by the breeding programs from each region.

2. Differentiation of modern cultivar SPs was higher than landrace SPs, indicating a lower level of gene exchange among different breeding programs. The highest gene exchange was reported between northern Mediterranean landraces and modern cultivars released by French and Italian breeding programs, indicating the presence of the genetic background of landraces in the improved modern varieties.

3. The structure and gene flow of landraces suggested a migration from the centre of wheat domestication in the Fertile Crescent to the west of the Mediterranean Basin. The level of admixture was higher in western Mediterranean landraces maybe due to the incorporation and fixation of favourable alleles from eastern and northern genetic pools during the migration course.

4. Eastern Mediterranean landraces showed the widest seminal root angle, the lowest seed weight, the longest shoot length, and the lowest number of roots.

5. GWAS performed in the landraces set identified 135 marker-trait associations for root related traits, which were grouped in 15 QTL *hotspots*. From them, 11 corresponded to genome regions not previously reported, and 5 shared common positions with *hotspots* for grain yield, thus suggesting a pleiotropic effect among root related traits and grain yield.

6. The modified triangular vegetation index (MTVI2) proved to have a good accuracy to estimate LAI since it was more sensitive to chlorophyll variations and was better than other VIs mitigating the saturation effect in wheat.

7. The assessment of biophysical parameters earlier during the growing season could improve the accuracy of LAI estimates through remote sensing imagery, particularly when values are low.

8. Yield predictions can be obtained through VIs calculated from both UAV multispectral and ground-based RGB images, being the R^2 values of the latter higher.

9. NDVI and GNDVI were the VIs mostly present in the prediction equations of the whole collection obtained through UAV multispectral imagery. The

RGB indices GA, GGA, a*, and u* have been proven to be more suitable for predicting yield due to their capacity to calculate a combination of physiological components related to biomass.

10. Multiple linear regression between trait variation and population structure among landraces and modern cultivars pointed out the selection for grain yield, HI, NGm^2 and GNDVI_PA during the breeding process. Differentiation among landraces was mainly for grain yield, TKW, GNDVI_PA and GA, and among modern cultivars GS65, GNDVI and NDVI_PA.

11. GWAS for agronomic performance and VIs identified 2579 MTAs. The use of the QTL overview index for the definition of QTL *hotspots* was successful for the identification of 11 consensus genomic regions including most of the stable MTAs.

12. Gene annotation of reference genome sequence allowed the identification of 1342 gene models within QTL *hotspots*. Candidate gene mining using *in silico* transcriptomic data allowed the identification of *differentially expressed genes* (DEGs) upregulated and downregulated under drought conditions in different tissues and developmental stages.

Conclusions

1. L'estructura de les varietats tradicionals va mostrar un patró geogràfic amb diferents nivells de barreja, justificat principalment per distàncies físiques entre les regions on van ser recol·lectades, mentre que l'estructura de les varietats modernes va destacar les diferències i semblances entre l'acerb genètic dels programes de millora genètica de cada regió.

2. La diferenciació de les subpoblacions de les varietats modernes va ser més elevada que les de les varietats tradicionals, indicant un menor nivell d'intercanvi genètic entre els diferents programes de millora. El major intercanvi genètic es va donar entre les varietats tradicionals Nord-Mediterrànies i les varietats modernes dels programes de millora francesos i italians, indicant l'ús de les varietats tradicionals en el desenvolupament de les noves varietats.

3. L'estructura i intercanvi genètic de les varietats tradicionals van suggerir una migració del centre de la domesticació del blat en el Creixent Fèrtil cap a l'oest de la conca Mediterrània. El nivell de barreja va ser major en les varietats tradicionals de l'oest del Mediterrani, probablement per la incorporació i fixació d'al·lels favorables provinents de grups genètics de l'est i del nord durant el transcurs de la migració.

4. Les varietats tradicionals Est-Mediterrànies van mostrar l'angle seminal radicular més ample, el menor pes de llavor, la major longitud de tija i el menor nombre d'arrels.

5. L'anàlisi GWAS realitzat a les varietats tradicionals va identificar 135 associacions pels caràcters radiculars, que van ser agrupades en 15 QTL *hotspots*. Entre ells, 11 van correspondre a regions del genoma que no s'havien descrit anteriorment i 5 van compartir posicions comunes amb *hotspots* pel rendiment, indicant un efecte pleiotròpic entre ambdós tipus de caràcters.

6. L'índex de vegetació triangular modificat (MTVI2) va demostrar una bona precisió per estimar el LAI ja que va ser més sensible a les variacions del contingut de clorofil·la i va ser millor que altres índexs de vegetació mitigant l'efecte de saturació en blat.

7. La mesura dels paràmetres biofísics en etapes primerenques durant el cicle de cultiu pot millorar la precisió de les estimacions del LAI a través de la teledetecció, sobretot quan els valors són baixos.

8. Les prediccions del rendiment poden ser obtingudes mitjançant índexs de vegetació calculats a través d'imatges multiespectrals obtingudes amb càmeres

disposades en drons i d'imatges RGB a través de càmeres convencionals a terra, obtenint millors resultats de valors R^2 amb les darreres.

9. Els índexs NDVI i GNDVI van ser els més presents en les equacions de predicció de tota la col·lecció obtingudes a través de les imatges multiespectrals del dron. Els índexs GA, GGA, a* i u* de les imatges RGB van demostrar ser més eficaços per predir rendiment degut a la seva capacitat per calcular una combinació de components fisiològics relacionats amb la biomassa.

10. L'anàlisi de regressió múltiple entre la variabilitat d'un caràcter i l'estructura de la població entre les varietats tradicionals i entre les varietats modernes va mostrar que la millora s'ha dirigit cap a l'increment del rendiment, HI, NGm^2 i GNDVI_PA. La diferenciació entre les varietats tradicionals es va donar, principalment, en rendiment, TKW, GNDVI_PA i GA, mentre que entre les varietats modernes va ser en GS65, GNDVI i NDVI_PA.

11. L'anàlisi GWAS per les variables agronòmiques i pels índexs de vegetació va identificar 2579 associacions. L'ús de l'índex 'QTL overview' per la definició dels QTL *hotspots* va ser eficaç per la identificació d'11 regions genòmiques consens que van incloure la majoria de les associacions estables.

12. L'anotació de la seqüència del genoma de referència va permetre la identificació de 1342 models gènics entre els diferents QTL *hotspots*. La tipificació de gens candidats utilitzant dades transcriptòmiques *in silico* va permetre identificar *differentially expressed genes* (DEGs) sobreexpressats i infraexpressats en condicions de sequera en diferents teixits i estadis de desenvolupament.

Conclusiones

1. La estructura de las variedades tradicionales mostró un patrón geográfico con diferentes niveles de mezcla, justificado principalmente por distancias físicas entre las regiones donde fueron recolectadas, mientras que la estructura de las variedades modernas destacó las diferencias y similitudes entre el acervo genético de los programas de mejora genética de cada región.

2. La diferenciación de las subpoblaciones de las variedades modernas fue más elevada que la de las variedades tradicionales, indicando un menor nivel de intercambio genético entre los diferentes programas de mejora. El mayor flujo génico se dio entre las variedades tradicionales del Norte del Mediterráneo y las variedades modernas de los programas de mejora franceses e italianos, señalando el uso de las variedades tradicionales en el desarrollo de nuevas variedades.

3. La estructura y el intercambio genético de las variedades tradicionales denotaron una migración desde el centro de domesticación del trigo en el Creciente Fértil hacia el oeste de la cuenca Mediterránea. El nivel de mezcla fue mayor en las variedades tradicionales del oeste del Mediterráneo, probablemente por la incorporación y fijación de alelos favorables provenientes de los grupos genéticos del este y del norte durante el proceso de migración.

4. Las variedades tradicionales del este del Mediterráneo mostraron el ángulo seminal radicular más ancho, el menor peso de semilla, la mayor longitud de tallo y el menor número de raíces.

5. El análisis GWAS realizado en las variedades tradicionales identificó 153 asociaciones para los caracteres radiculares, que fueron agrupadas en 15 QTL *hotspots*. Entre ellos, 11 correspondieron a regiones del genoma que no se habían descrito anteriormente y 5 compartieron posiciones comunes con *hotspots* para rendimiento, indicando un efecto pleiotrópico entre ambos tipos de caracteres.

6. El índice de vegetación triangular modificado (MTVI2) demostró una buena precisión para estimar el LAI ya que fue más sensible a las variaciones del contenido de clorofila y fue mejor que otros índices de vegetación mitigando el efecto de saturación en trigo.

7. La medición de los parámetros biofísicos en etapas tempranas durante el ciclo de cultivo puede mejorar la precisión de las estimaciones del LAI a través de la teledetección, sobre todo cuando los valores son bajos.

8. Las predicciones del rendimiento pueden ser obtenidas mediante índices de vegetación calculados a través de imágenes multiespectrales obtenidas con

cámaras en drones e imágenes RGB a través de cámaras convencionales en tierra, obteniendo mejores resultados de valores R^2 con las últimas.

9. Los índices NDVI y GNDVI fueron los más presentes en las ecuaciones de predicción de toda la colección obtenidas a través de las imágenes multiespectrales del dron. Los índices GA, GGA, a* y u* de las imágenes RGB demostraron ser más eficaces para predecir rendimiento debido a su capacidad para calcular una combinación de componentes fisiológicos relacionados con la biomasa.

10. El análisis de regresión múltiple entre la variabilidad de un carácter y la estructura de la población entre las variedades tradicionales y entre las variedades modernas mostraron que la mejora se ha dirigido hacia un incremento en el rendimiento, HI, NGm^2 y GNDVI_PA. La diferenciación entre las variedades tradicionales se dio, principalmente, en rendimiento, TKW, GNDVI_PA y GA, mientras que en las variedades modernas fue en GS65, GNDVI y NDVI_PA.

11. El GWAS para las variables agronómicas y para los índices de vegetación identificó 2579 asociaciones. El uso del índice 'QTL overview' para la definición de los QTL *hotspots* fue eficaz para la identificación de 11 regiones genómicas consenso que incluyó la mayoría de las asociaciones estables.

12. La anotación de la secuencia del genoma de referencia permitió la identificación de 1342 modelos génicos entre los diferentes QTL *hotspots*. La tipificación de genes candidato utilizando datos transcriptómicos *in silico* permitió identificar *differentially expressed genes* (DEGs) sobreexpresados e infraexpresados en condiciones de sequía en diferentes tejidos y estadios de desarrollo.

Annexes

Table 1. List of Landrace accesions used in the PhD Thesis. SP, subpopulation.

Accession	Country	SP	GenBank[1]	Accession Nº
TRI 1667	Albania	SP2	IPK	TRI 1667
TRI 2100	Albania	SP2	IPK	TRI 2100
TRI 1671	Albania	SP2	IPK	TRI 1671
TRI 1313	Bulgaria	SP2	IPK	TRI 1313
TRI 7819	Bulgaria	SP2	IPK	TRI 7819
TRI 7821	Bulgaria	SP2	IPK	TRI 7821
408-IV/61	Bosnia & Herzegovina	SP2	NSGC	NSGC345409
Moriborska	Bosnia & Herzegovina	SP2	VIR	22915
Ranka	Bosnia & Herzegovina	SP2	VIR	38937
TRI 10515	Cyprus	Admixed	IPK	TRI 10515
TRI 10526	Cyprus	SP3	IPK	TRI 10526
TRI 10561	Cyprus	Admixed	IPK	TRI 10561
TRI 10590	Cyprus	SP1	IPK	TRI 10590
TRI 10531	Cyprus	SP3	IPK	TRI 10531
Ali Ben Makhloul	Argelia	SP3	NSGC	NSGC48592
Khalof	Argelia	SP3	NSGC	NSGC67492
Mahon De-mias	Argelia	SP1	NSGC	NSGC263419
MG 17956	Argelia	SP3	NSGC	NSGC470815
MG 17999	Argelia	SP2	NSGC	NSGC470851
MG 18006	Argelia	SP2	NSGC	NSGC470857
MG 18013	Argelia	SP1	NSGC	NSGC470864
MG 18036	Argelia	SP1	NSGC	NSGC470884
MG 18049	Argelia	SP1	NSGC	NSGC470895
Bahatane	Argelia	SP3	CGN	CGN06035
Krelof - A	Argelia	SP2	CGN	CGN11452
Ble' du dahra-baal	Argelia	SP1	VIR	16152

Accession	Country	SP	GenBank[1]	Accession N°
Mahon 7295	Argelia	SP1	ISC	T8801887
Sachah	Egypt	SP2	CGN	CGN06369
Gibson	Egypt	SP3	VIR	21198
Hauch	Egypt	SP3	VIR	20204
Hindiffino crible	Egypt	Admixed	VIR	21979
Hindiffino non crible	Egypt	SP3	VIR	21982
Mokhtar	Egypt	SP1	VIR	46124
Hindi 62	Egypt	SP3	ISC	T8801285
Bladette de Besplas	France	SP2	NSGC	NSGC191706
Ble Blanc de la Reole	France	SP2	NSGC	NSGC48199
Mars Rouge Sans Barbe	France	SP2	NSGC	NSGC192398
Touzelle Belle Abec	France	SP1	NSGC	NSGC191709
Touzelle Blanche Barbu	France	SP2	NSGC	NSGC185381
Touzelle Originario	France	Admixed	NSGC	NSGC191708
Touzelle Rouge de Provence	France	Admixed	NSGC	NSGC184599
Bladette de puylaurens	France	SP2	CGN	CGN05384
Mouton a epi rouge	France	SP2	CGN	CGN05580
Saisette	France	Admixed	CGN	CGN05661
TRI 14046	France	SP2	IPK	TRI 14046
TRI 17938	France	SP2	IPK	TRI 17938
TRI 1425	Greece	SP2	IPK	TRI 1425
TRI 1686	Greece	SP1	IPK	TRI 1686

Accession	Country	SP	GenBank[1]	Accession N°
TRI 17989	Greece	Admixed	IPK	TRI 17989
TRI 2060	Greece	SP1	IPK	TRI 2060
TRI 2071	Greece	Admixed	IPK	TRI 2071
TRI 2129	Greece	SP2	IPK	TRI 2129
986	Croatia	SP2	NSGC	NSGC264963
Croatia 3	Croatia	SP2	NSGC	NSGC11225
Croatia 6	Croatia	SP2	NSGC	NSGC11228
Umarah	Iraq	SP1	CGN	CGN06462
TRI 15277	Iraq	SP3	IPK	TRI 15277
TRI 15292	Iraq	SP3	IPK	TRI 15292
TRI 16079	Iraq	SP3	IPK	TRI 16079
TRI 16084	Iraq	SP2	IPK	TRI 16084
TRI 11548	Iraq	SP3	IPK	TRI 11548
TRI 16080	Iraq	SP3	IPK	TRI 16080
TRI 11528	Iraq	SP3	IPK	TRI 11528
TRI 16063	Iraq	SP3	IPK	TRI 16063
TRI 8358	Iraq	SP3	IPK	TRI 8358
CGN06182	Israel	SP1	CGN	CGN06182
CGN06204	Israel	SP1	CGN	CGN06204
CGN04191	Israel	Admixed	CGN	CGN04191
Palestinskaya	Israel	Admixed	VIR	15818
17310	Israel	SP1	VIR	17310
Cappellina	Italy	SP2	IPK	TRI 13007
TRI 15321	Italy	SP1	IPK	TRI 15321
TRI 14055	Italy	SP2	IPK	TRI 14055
TRI 15226	Italy	Admixed	IPK	TRI 15226
TRI 16900	Italy	SP2	IPK	TRI 16900
TRI 16895	Italy	SP3	IPK	TRI 16895
TRI 14842	Italy	SP1	IPK	TRI 14842
Solina	Italy	SP2	IPK	TRI 13424
TRI 15219	Italy	Admixed	IPK	TRI 15219
TRI 14173	Italy	SP2	IPK	TRI 14173

Accession	Country	SP	GenBank[1]	Accession N°
TRI 16516	Italy	SP2	IPK	TRI 16516
Cappelli	Italy	SP1	IPK	TRI 16574
25	Jordan	SP3	NSGC	NSGC420932
Dorziyeh Karak - B	Jordan	SP3	NSGC	NSGC283147
SY 271	Jordan	Admixed	NSGC	NSGC487289
17411	Jordan	SP1	VIR	17411
Beyrouth 11	Lebanon	SP3	NSGC	NSGC278531
Beyrouth 3	Lebanon	SP2	NSGC	NSGC278533
Salamouni	Lebanon	SP3	NSGC	NSGC182673
TRI 17974	Libya	Admixed	IPK	TRI 17974
TRI 14643	Libya	SP3	IPK	TRI 14643
TRI 14668	Libya	SP3	IPK	TRI 14668
Canivano	Morocco	SP1	CGN	CGN06294
Fez 2	Morocco	SP1	CGN	CGN06049
Recio	Morocco	SP1	CGN	CGN06292
CGN04157 - A	Morocco	Admixed	CGN	CGN04157
CGN04158	Morocco	SP1	CGN	CGN04158
CGN06246	Morocco	SP1	CGN	CGN06246
CGN06247	Morocco	SP1	CGN	CGN06247
CGN06248	Morocco	SP1	CGN	CGN06248
CGN06252	Morocco	SP1	CGN	CGN06252
CGN06255	Morocco	SP1	CGN	CGN06255
CGN06260	Morocco	Admixed	CGN	CGN06260
CGN06264	Morocco	SP1	CGN	CGN06264
CGN06266	Morocco	Admixed	CGN	CGN06266
CGN06269	Morocco	SP1	CGN	CGN06269
CGN06271	Morocco	Admixed	CGN	CGN06271
CGN06284	Morocco	Admixed	CGN	CGN06284
CGN06297	Morocco	SP1	CGN	CGN06297
TRI 18287	Morocco	SP1	IPK	TRI 18287
TRI 18291	Morocco	SP1	IPK	TRI 18291

Accession	Country	SP	GenBank[1]	Accession Nº
TRI 18308	Morocco	SP1	IPK	TRI 18308
309-VII/33	Makedonia	SP2	NSGC	NSGC345309
340-VII/45	Makedonia	SP2	NSGC	NSGC345340
VII/1-B	Makedonia	SP2	NSGC	NSGC362589
Stara bela	Makedonia	SP2	CGN	CGN04172
Magueija	Portugal	SP2	CRF	BGE012669
Santareno	Portugal	SP2	CRF	BGE011900
Temporao de coruche	Portugal	SP2	CRF	BGE012846
Tremes branco	Portugal	SP1	CRF	BGE012703
Bistra	Romania	SP2	Suceava	SVGB5538
Pades	Romania	SP2	Suceava	SVGB7916
Solonetu nou	Romania	SP2	Suceava	SVGB14976
SVGB10195	Romania	SP2	Suceava	SVGB10195
Raton de Belalcazar	Spain	SP1	CRF	BGE011825
Cabezorro	Spain	SP1	CRF	BGE011882
Negrete de Cañaveras	Spain	SP2	CRF	BGE012132
Pelon blanco	Spain	Admixed	CRF	BGE012196
Chamorro de Villadiego	Spain	SP2	CRF	BGE012205
Blat petit de Olot	Spain	SP1	CRF	BGE012870
Candeal	Spain	SP2	CRF	BGE012392
Isla de Fuerteventura	Spain	SP2	CRF	BGE013760
Hembrilla de Jerga	Spain	SP2	CRF	BGE018232
Extremo Sur Argelino	Spain	SP2	CGN	CGN05749
Xeixa Tarragona	Spain	SP1	CRF	BGE018242

Accession	Country	SP	GenBank[1]	Accession N°
41-II/4-B	Serbia	SP2	NSGC	NSGC345043
Crvenica	Serbia	SP2	NSGC	NSGC184168
Piskulja	Serbia	SP2	NSGC	NSGC184188
Legan bez osja	Serbia	SP2	VIR	38803
401	Syria	SP1	NSGC	NSGC94569
Aleppo 21	Syria	Admixed	NSGC	NSGC278540
Aleppo 28	Syria	SP3	NSGC	NSGC278545
Aleppo 32	Syria	SP3	NSGC	NSGC278547
Aleppo 33	Syria	SP3	NSGC	NSGC278548
Damaskus 12	Syria	SP3	NSGC	NSGC278537
Damaskus 8	Syria	SP1	NSGC	NSGC278536
K1140	Syria	Admixed	NSGC	NSGC253959
Kaundouhari	Syria	SP3	NSGC	NSGC182711
TRI 8375	Syria	SP1	IPK	TRI 8375
Salamuni - A	Syria	Admixed	VIR	17172
Allorca	Tunisia	SP2	CGN	CGN05358
Sbei noir	Tunisia	SP1	CGN	CGN06378
TRI 17006	Tunisia	SP3	IPK	TRI 17006
TRI 17002	Tunisia	Admixed	IPK	TRI 17002
Florence 193	Tunisia	SP2	ISC	T8800968
763	Turkey	Admixed	NSGC	NSGC119302
1170	Turkey	Admixed	NSGC	NSGC119309
811 (B)	Turkey	SP3	NSGC	NSGC119305
1552	Turkey	SP2	NSGC	NSGC119325
2103	Turkey	SP3	NSGC	NSGC119348
2933	Turkey	SP3	NSGC	NSGC119366
2936	Turkey	SP3	NSGC	NSGC119369
Edirne	Turkey	SP3	NSGC	NSGC111244
Gemir - B	Turkey	SP3	NSGC	NSGC166257
Kirmizi kiluk - A	Turkey	SP3	NSGC	NSGC165149
Ormece	Turkey	SP3	NSGC	NSGC166545

Accession	Country	SP	GenBank[1]	Accession Nº
Saribasak	Turkey	SP3	NSGC	NSGC165146
T-317	Turkey	SP2	NSGC	NSGC109368
Yazlik	Turkey	SP3	NSGC	NSGC165115
Yumusak	Turkey	SP3	NSGC	NSGC165160

[1]CGN: Wageningen, The Netherlands; CRF, Madrid, Spain; IPK: Gatersleben, Germany; ISC: San Angelo Lodigiano, Italy; NSGC: Aberdeen, ID, USA; Suceava GenBank: Suceava, Romania; VIR, St. Petersburg, Russia.

Table 2. List of Modern accesions used in the PhD Thesis. SP, subpopulation.

Cultivar	Country	SP	Pedigree[1]
Takhar 96	Afghanistan	SP6	veery-7/opata-m-85
Ain abid	Algeria	Admixed	
Adelaide	Canada	Admixed	
Misir-2	Egypt	SP6	super-kauz/baviacora-92
Misir-1	Egypt	SP6	oasis-86/super-kauz//4*bacanora-88/3/2*pastor
Gemmeiza-10	Egypt	SP6	"maya-74/olesen//1160-47/3/bluebird/g11/4/chat/5/crow"
Sakha-69	Egypt	SP6	inia-f-66/rl-4220//siete-cerros-t-66/yaqui-50
Sids-12	Egypt	SP6	"buckbuck//siete-cerros-66/alondra/5/maya-74/olesen//1160.147/3/bluebird/gallo/4/chat/6/maya-74/vulture//cmh-74-a-63014/super-x"
Gemmeiza-11	Egypt	SP6	"bobwhite/kvs//siete-cerros-66/seri-82/3/giza-168/sakha-61"
Sahel-1	Egypt	SP6	ns-732/pima/veery
Sids 1	Egypt	SP6	hd-2172/pavon//1158-57/maya-74
Adagio	France	SP4	
Candelo	France	SP4	
Aviso	France	SP4	moisson/topaze
Belsito	France	SP4	
Innov	France	SP4	ordeal/sideral
Andalou	France	SP4	
Fiorenzo	France	SP4	rabd-88-13/virlor
Sensas	France	SP4	s-0179/s-32203
Adhoc	France	SP4	
Charles peguy	France	SP4	thatcher/vilmorin-27//ariana

Cultivar	Country	SP	Pedigree[1]
Trocadero	France	SP4	baroudeur/bercy
Astral	France	SP4	fortunato/yga/3/florence/aurore//g-4
Aerobic	France	SP4	
Bramante	France	SP4	victo/soissons
Soissons	France	SP4	iena(jena)/(hybride-naturel) hn-35
Isengrain	France	SP4	apollo,deu/soissons
Soberbio	France	SP4	
Cipres	France	SP4	
Bologna	France	SP4	h-89092/h-89136//soissons
Avelino	France	SP4	
Viriato	France	SP4	
Nogal	France	Admixed	norrona/gasser
Premio	France	SP4	
Diamento	France	SP4	
Rgt Somontano	France	SP4	
Lazaro	France	SP4	
Altamira	France	SP4	96248/isengrain
Andino	France	SP4	
Arezzo	France	SP4	
Solehio	France	SP4	isengrain/ornicar
Mecano	France	SP4	
Guadalete	France	SP6	
Bonpain	France	SP6	prinqual/cornette
Equilibre	France	SP4	
Soledad	France	SP4	
Tremie	France	SP4	s-32/moulin
Bastide	France	SP4	fertil/arche
Aubusson	France	SP4	tremie/91-b-294
Garcia	France	SP4	
Kumberri	France	SP4	
Akim	France	SP4	

Cultivar	Country	SP	Pedigree[1]
Camargo	France	SP4	
Exotic	France	SP4	etecho/vivant
CCB Ingenio	France	SP4	
Bueno	France	SP4	
Sublim	France	SP4	
Alhambra	France	SP4	
Bandera	France	SP4	
Inoui	France	SP4	charly/victo
SY Moisson	France	SP4	cappelle-desprez//hybride-80-3/ etoile-de-choisy
Raffy	France	SP4	
Aguila	France	SP4	
Rodrigo	France	SP4	aztec/legion
Sollario	France	SP4	
Alpino	France	SP4	
Galpino	France	SP4	
Carles	France	SP4	
Sorrial	France	SP4	
Royssac	France	SP4	
Rimbaud	France	SP4	
Rvalo	France	SP4	
Sobbel	France	SP4	
Sofru	France	SP4	
Apache	France	SP4	axial/nrpb-84-4233
Illico	France	SP4	ormil/apache
Galopin	France	SP4	
Sobred	France	SP4	
Sokal	France	SP4	
Cezanne	France	SP4	thesee/87-b-29
Craklin	France	SP4	87-b-15/d-136
Paledor	France	SP4	
Botticelli	France	SP4	perico/95-b-343

Cultivar	Country	SP	Pedigree[1]
Eureka	France	SP4	"mironovskaya-808/maris-hunts-man/3/vpm-1/moisson(r-1-5-2)//courtot"
MV Emese	Hungary	SP5	mv-ma/mv-12//f-2098-w-2-21
Masaccio	Italy	SP4	oratorio/genio
Zanzibar	Italy	SP4	frelon/61601//capnor/parador
Palesio	Italy	SP4	pandas/recital
Toskani	Italy	SP4	
Trofeo	Italy	Admixed	bolero/mieti
Anapo	Italy	SP6	eg-52/bel-118
Andana	Italy	SP4	unknown/eridano
Anforeta	Italy	SP6	eg-83/bel-118
Arabia	Italy	SP4	guadalupe/tibet
Carisma	Italy	Admixed	d-29/f-65
Agape	Italy	SP4	serio/tremie
Antille	Italy	SP4	
Tiepolo	Italy	SP4	oracle/calodine
Abate	Ìtaly	SP4	eg-52/eridano
Arz	Lebanon	SP6	"mayo-54-e/lerma-rojo-64//ta-cuari/3/lerma-rojo-64//tezanos-pintos-precoz/ya-qui-54"
Olga	Makedonia	SP5	
Balkania	Makedonia	SP5	
SIete cerros	Mexico	SP6	
Marchouch 8	Morocco	SP6	
Aguilal	Morocco	SP6	sais*2/ks-85241-14
Achtar	Morocco	SP6	hork/yamhill//kalyansona/blue-bird
Nesma	Morocco	SP6	
Arrehane	Morocco	SP6	
KG100	Serbia	SP5	
PKB Arena	Serbia	SP5	

Cultivar	Country	SP	Pedigree[1]
Ana Morava	Serbia	SP5	morava/una
PKB Lepoklasa	Serbia	SP5	
PKB Ratarica	Serbia	SP5	
Zvezdana	Serbia	SP5	ns-63-27//stamena/ns-rana-5
PKB Vizeljka	Serbia	SP5	
Simonida	Serbia	SP5	ns-63-25//rodna/ns-3288
BG Merkur	Serbia	SP5	
BG Carica	Serbia	SP5	
PLB Talas	Serbia	SP5	
BG Vitka	Serbia	SP5	
Vizija	Serbia	SP5	kozara/skopjanka
PKB Mlinarka	Serbia	SP5	
Zlatna	Serbia	SP5	jasenica/rodna
Aleksandra	Serbia	SP5	
Planeta	Serbia	SP5	fillo-9/yazi-6//rabudo-1/shag-14
Pobeda	Serbia	SP5	sremica/balkan
Aurelia	Serbia	SP5	
Zemunska rosa	Serbia	SP5	skopljanka/proteinka
Renesansa	Serbia	SP5	yugoslavia/ns-55-25
Kruna	Serbia	Admixed	
NS 40S	Serbia	SP4	
Chambo	Spain	SP4	
08THES2162	Spain	SP6	
Vejer	Spain	SP6	
Antequera	Spain	SP6	
Conil	Spain	SP6	croc-1/(205)tr.ta//borl-aug-m-95/3/2*milan
Marchena	Spain	SP6	croc-1/(205)tr.ta//borl-aug-m-95/3/2*milan
Tejada	Spain	SP6	chilero/parula//baviacora-92/3/milan/kauz
Babui	Spain	SP6	
Catedral	Spain	SP6	

Cultivar	Country	SP	Pedigree[1]
Eneas	Spain	SP6	
Califa sur	Spain	SP6	
Cartaya	Spain	SP6	kavkaz/buho//kalyansona/blue-bird
Escacena	Spain	SP6	seri-82/rayon-89
Jerezano	Spain	SP6	thornbird//maya-74/nacoza-ri-76/3/rabe/4/milan
Galeon	Spain	SP6	
Kilopondio	Spain	SP6	
Trebujena	Spain	SP6	shearwater/yavaros
Victorino	Spain	SP6	
Alcala	Spain	SP6	
Rinconada	Spain	SP6	
Yecora	Spain	SP6	ciano-67//sonora-64/klein-rendi-dor/3/ii-8156
Cielo	Spain	SP6	
Gazul	Spain	SP6	
Galera	Spain	SP6	
Mapeña	Spain	SP6	tr-353/betres//alcotan/3/rincona-da/4/3*betres
Marca	Spain	SP6	
Anza	Spain	SP6	lerma-rojo-64//norin-10/bre-vor/3/3*andes-enano
Odiel	Spain	SP6	br-5237/cavalier
Trimax	Spain	SP6	
Algido	Spain	Admixed	
Artur Nick	Spain	SP6	
Mulhacen	Spain	SP6	
Platero	Spain	SP6	
Adalid	Spain	Admixed	
Dollar	Spain	SP6	
Montcada	Spain	Admixed	
Montserrat	Spain	Admixed	damiano/montjuich

Cultivar	Country	SP	Pedigree[1]
Idalgo	Spain	SP4	
Santoyo	Spain/France	SP4	
Debeira	Sudan	SP6	"hd-2160/5/tobari-66/ciano-67//bluebird/3/nainari-60*2//tom-thumb/sonora-64/4/hd-1954"
Valbona	Switzerland	Admixed	
Cham-8	Syria	SP6	jupateco-f-73/bluejay//ures-81
Cham-6	Syria	SP6	w-3918-a/jupateco-73
Babaga-3	Syria	SP6	
Cham-4	Syria	SP6	flicker/hork
Hamam-4	Syria	SP6	
Attila	Tunisia	SP6	
Karatopak	Turkey	SP6	tesia-79/veery//seri-82
Ata 81	Turkey	SP6	kavkaz/ciguena
Cumhuriyet 75	Turkey	SP6	"sonora-64*2//tezanos-pintos-precoz/yaqui-54/3/andes-64-a/4/2*frocor//yaqui/kentana"
Efe	Turkey	SP6	
Mane Nick	Turkey	SP6	
Gönen	Turkey	Admixed	8156-reselection/mara//bluebird

[1]Pedigree from http://genbank.vurv.cz/wheat/pedigree/pedigree.asp