

THESE

En vue de l'obtention du : **DOCTORAT**

Centre de Recherche : Biotechnologies Végétales et Microbiennes, Biodiversité et Environnement
Structure de Recherche : Équipe de Microbiologie et Biologie Moléculaire
Discipline : Biologie
Spécialité : Bio-informatique, Génomique et Cartographie génétique des plantes

Présentée et soutenue le 01 / 10 / 2022 par :

Issame FAROUK

Construction of high-density linkage map, Quantitative Trait Loci (QTL) detection, Single Nucleotide Polymorphism (SNP) annotation, and dissection of grain yield QTL in the durum population Lahn/Cham1 tested in contrasting environments

Devant le Jury :

Zine El Abidine TRIQUI	PES, Université Mohammed V, Faculté des Sciences-Rabat	Président / Rapporteur
Majida HAFIDI	PES, Université Moulay Ismail, Faculté des Sciences- Meknès	Rapportrice / Examinatrice
Leila MEDRAOUI	PH, Université Mohammed V, Faculté des Sciences-Rabat	Rapportrice / Examinatrice
Abdelkarim FILALI-MALTOUF	Expert, PES, Académie Hassan II des sciences et techniques-Rabat	Examineur
Ahmed AMRI	Expert, Docteur, ICARDA-Rabat	Invité
Abderrahmane LYAMANI	Expert, Docteur, UM6P-Benguerir	Invité
Miloudi M NACHIT	Expert, Docteur, ICARDA-Rabat	Co-Directeur de thèse
Fatima GABOUN	Expert, Docteur, Institut National de Recherche Agronomique-Rabat	Co-Directeur de thèse
Bouchra BELKADI	PES, Université Mohammed V, Faculté des Sciences-Rabat	Directeur de thèse

Année Universitaire : 2021/2022

Dédicaces

A mes chers parents: Mon cher et illustre père, ma chère et tendre mère, acceptez l'hommage de ma gratitude, qui grande qu'elle puisse être, ne sera jamais à la hauteur de vos éloquences et vos dévouements.

A mes chères sœurs, en témoignage de leur support moral et leur fraternité sans égale.

A mes petites filles et ma femme pour leur amour infini.

A mes amis: Karim Rabah, Chaimaa Yatrib, Meryem Zaïm, Abdellah Benbia, Houcine pour leur aide sans cesse.

Ainsi Qu'à mes collègues : Hicham Ibn Abdelouahab, Faïssel Ben Tourtou, Ahmed El Yousfy, Abdelmadjid Regti, Soufiane Chibane, Zakaria El Mouden, Abd El Hak Dahou, Amaal Chafyq pour leur soutien permanent.

A tous ceux qui me sont chers.

Je dédie ce travail.

Aknowledgements

The work carried out and presented in this thesis is the result of a triple collaboration between The International Center for Agricultural Research in the Dry Areas (ICARDA), the National Institute of Agronomic Research of Rabat (INRA), and the Mohammed V University of Rabat.

This work constitutes a continuation of what has been achieved in the durum breeding program for Mediterranean dryland started about four decades ago by the International Maize and Wheat Improvement Center (CIMMYT) and ICARDA under the supervision of Dr. **Miloudi M NACHIT**. Thus, I take this opportunity to address my deepest thanks to the team of the molecular biology and biotechnology laboratory under the direction of Dr. **NACHIT** at ICARDA who worked on the development of the Lhan/Cham1 durum wheat population which was the basis of this work.

This study was carried out under the supervision of the professor **Bouchra BELKADI** from the Faculty of Sciences (FSR) of the University Mohammed V-Rabat, Doctor **Miloudi M NACHIT** from ICARDA, and Doctor **Fatima GABOUN** from INRA.

The results obtained within the framework of this thesis have been subjected to several publications in peer-reviewed and indexed journals and have been presented in international events in oral or poster communications forms.

I am indebted to my supervisor Mrs **Bouchra BELKADI**, Professeur de l'enseignement superieur at the Faculty of Sciences of Rabat for accepting the supervision of this thesis, for her continued guidance, advice, and for her endless support. I am grateful to her for her supervision, mentorship, and friendship.

I have to express my sincerest and deep gratitude and thanks also to Mr **Miloudi M NACHIT**, Docteur expert at ICARDA for his continuous involvement, supervision, encouragements, patience, guidance, advice, and strong support. He was more than supervisor; he was a source of inspiration.

I gratefully recognize the help of Mrs **Fatima GABOUN**, Docteur expert at INRA for her advice, guidance, for his strong support in several critical moments.

I had the pleasure of working under the supervision of a very special team. I would like to thank this team again for their strong support, friendship, high human qualities, and

relationship, and for giving me the opportunity to conduct this work. Words will not be enough to express my gratitude to them for their scientific and moral support.

Also, sincere appreciation and thanks to Mr **Zine El Abidine TRIQUI**, Professeur de l'enseignement superieur at the Faculty of Sciences of Rabat for having agreed to chair the jury.

I have to thank Mrs **Majida HAFIDI**, Professeur de l'enseignement superieur at the Faculty of Sciences of Meknès for having agreed to evaluate this work and come from Meknès to discuss it.

Special thanks are addressed to Mrs **Laila MEDRAOUI**, Professeur Habilité at the Faculty of Sciences of Rabat for having agreed to evaluate and discuss this work.

Sincere appreciations and thanks are extended to Mr **Abdelkarim FILALI-MALTOUF**, Professeur de l'enseignement superieur at the Hassan II Academy of Science and Technology of Rabat for having agreed to evaluate and discuss this work.

Sincere appreciations and thanks are also extended to Mr **Ahmed AMRI**, Docteur expert at ICARDA for having reviewed and evaluated my thesis.

I would like to express special words of thanks to Mr **Abderrahmane LYAMANI**, Docteur expert at Mohamed 6 Polytechnic University of Benguérir for having reviewed and evaluated this work.

It is a great opportunity and privilege for me to present my thesis work to a great jury. Special thanks to every member of this jury for their constructive critical comments.

Abstract

Durum wheat, *Triticum turgidum* L. var *durum*, constitutes with other cereal crops the basis of human nutrition. In the durum breeding programs, grain yield is required in developing high productive and stable durum cultivars. The grain yield is quantitatively inherited and influenced by the environment. The genetic linkage maps are crucial for studying and understanding the grain yield traits. The primary objective of this study was to construct a high-density genetic linkage map using a durum wheat population developed from the cross between two parents (Lahn and Cham1) that are widely grown in the Mediterranean basin. The Lahn/Cham1 population consisted of 112 Recombinant inbred lines (RIL) and tested in 11 contrasting environments (favorable, cold, dry, and hot). Then, the study aims to detect the QTLs and annotate the SNPs linked to grain yield, and to identify the composition in candidate genes of the four QTLs harbored by the chromosome 2A. The map was constructed with 247 SSRs and enriched with 1425 SNPs, covering about 6122 cM. The QTLs mapping in Multi-Environment Trials (MET) allowed the detection of one hundred and twenty-six QTLs along all durum wheat chromosomes. The analysis of 583 SNPs sequences selected from the detected QTL regions showed that a set of 122 SNP sequences were annotated of which 53% of the candidate genes found were involved in stresses tolerance, 29.5% in plant development and growth, and 3.3% in cell transport. Moreover, 1.6% of candidate genes were retrotransposon and transposon, 2.4% had unknown function, and 9.8% were related with other cellular processes. The results also showed that 66.7% of the candidate genes harbored on 4B chromosome were involved in stresses tolerance and 33.3% in plant development and growth. As for the QTLs mapping for specific environments, nine QTLs were detected, four of which are harbored by the chromosome 2A. The DNA sequences of these four QTLs were used for homology search and 546 candidate genes were identified of which some are present in several QTLs (F-box gene family, hydroxyproline-rich glycoprotein-like, retrotransposons / transposons and others candidate genes). Overall, this study provided information on employing SNPs markers to detect candidate genes linked with grain yield trait in durum wheat in contrasting environments (dry, cold, hot). Both chromosomes 2A and 4B contributed significantly to grain yield. The SNPs annotated and the candidate genes identified in this work may be used to develop molecular markers useful to track the genes of interest in Marker-Assisted Selection (MAS) in durum breeding programs.

Keywords: QTL, SNP, grain yield, candidate genes, linkage map, stress tolerance

Résumé

Le blé dur, *Triticum turgidum* L. var *durum*, constitue avec les autres cultures céréalières la base de l'alimentation humaine. Dans les programmes de sélection de blé dur, le rendement en grain est recherché pour développer des variétés de blé dur hautement productives avec un rendement stable. Le rendement en grains est hérité quantitativement et influencé par l'environnement. Les cartes génétiques sont cruciales pour étudier et comprendre le caractère du rendement en grains. L'objectif principal de cette étude est de construire une carte génétique à haute densité à partir d'une population de blé dur issue du croisement entre deux parents (Lahn et Cham1) largement cultivés dans le bassin méditerranéen. La population Lahn/Cham1 est composée de 112 lignées consanguines recombinantes (RIL) et testée dans 11 environnements différents (favorable, froid, sec et chaud). Ensuite, l'étude a visé la détection des QTL et l'annotation des SNP liés au rendement en grain, et à l'identification de la composition des quatre QTL localisées sur le chromosome 2A en gènes candidats. La carte génétique a été construite avec 247 SSR et enrichie de 1425 SNP, couvrant environ 6122 cM. La cartographie des QTL en Multi-Environnements Trials (MET) a permis la détection de cent vingt-six QTL le long de tous les chromosomes du blé dur. L'analyse de 583 séquences SNPs sélectionnées depuis les régions QTL détectées a montré qu'un ensemble de 122 séquences SNP ont été annotées dont 53% des gènes candidats trouvés étaient impliqués dans la tolérance aux stress, 29,5% dans le développement et la croissance des plantes, et 3,3% dans le transport cellulaire. De plus, 1,6 % des gènes candidats trouvés étaient des rétrotransposons et des transposons, 2,4 % avaient une fonction inconnue et 9,8 % étaient liés à d'autres processus cellulaires. Les résultats ont également montré que 66,7 % des gènes candidats localisées sur le chromosome 4B étaient impliqués dans la tolérance aux stress et 33,3 % dans le développement et la croissance des plantes. Quant à la cartographie des QTL pour des environnements spécifiques, 9 QTL ont été détectés, dont quatre sont localisées sur le chromosome 2A. Les séquences d'ADN de ces quatre QTL ont été utilisées pour la recherche d'homologie, 546 gènes candidats ont été identifiés dont certains sont présents dans plusieurs QTL (F-box gene family, hydroxyproline-rich glycoprotein-like, des rétrotransposons/transposons et d'autres gènes candidats). Cette étude a fourni des informations sur l'utilisation de marqueurs SNP pour identifier des gènes candidats liés au caractère de rendement en grains chez le blé dur dans des environnements différents (sec, froid, chaud). Les chromosomes 2A et 4B contribuent significativement au rendement en grain. Les SNP annotés et les gènes candidats identifiés dans ce travail peuvent être utilisés pour développer des marqueurs moléculaires utiles pour suivre les gènes d'intérêt dans la sélection assistée par marqueurs (SAM) dans les programmes d'amélioration du blé dur.

Mots-clés : QTL, SNP, rendement en grains, gènes candidats, carte génétique, tolérance aux stress

Résumé étendu

Le blé dur, *Triticum turgidum* L. var *durum*, constitue avec les autres cultures céréalières la base de l'alimentation humaine. Dans les programmes de sélection de blé dur, le rendement en grain est recherché pour développer des variétés de blé dur hautement productives avec un rendement stable. C'est un caractère complexe, quantitatif, gouverné par des Locus de Caractères Quantitatifs (LCQ) "ou QTL pour quantitative trait loci" et influencé par les interactions de ces QTLs entre eux et par les interactions entre ces derniers et l'environnement. La construction des cartes génétiques est une étape cruciale pour étudier et comprendre les caractères complexes y compris le rendement en grains. Dans ce sens, et pour bien comprendre mécanismes du fonctionnement du rendement en grains, cette étude avait comme objectif principal la construction d'une carte génétique suffisamment dense à partir d'une population de blé dur issue du croisement entre deux parents (Lahn et Cham1) largement cultivés dans le bassin méditerranéen. Cette population est composée de 112 lignées consanguines recombinantes (RIL) et testée dans 11 environnements différentes (favorable, froid, sec et chaud). Ensuite, l'étude a visé l'identification des QTL sur le génome du blé dur et l'annotation des SNP liés au rendement en grain, et enfin l'identification de la composition en gènes candidats des QTL localisées sur le chromosome 2A.

La population Lahn/Cham1 a été génotypée avec 251 marqueurs moléculaires (essentiellement des SSR). Ces marqueurs ont servi de piliers pour construire la carte génétique étant donné que plusieurs cartes construites par d'autres auteurs contiennent ces marqueurs. Cette population a été génotypée également avec 9400 Single Nucleotide Polymorphism (SNP), ces SNPs ont servi pour la saturation de la carte génétique. Nous avons utilisé le logiciel de cartographie génétique (QTL IciMapping software version 4.1.0.0). Le même logiciel a servi pour l'identification des QTLs sur le génome du blé dur en utilisant deux approches de détection: la première en Multi-Environnements Trials (MET) qui consiste en une analyse sur tout le génome en prenant en charge tous les environnements simultanément, la deuxième en Biparental analysis, qui consiste en une analyse sur tout le génome mais d'une manière spécifique pour chaque environnement séparément. Après l'identification des QTLs liés au rendement en grains, pour la première approche (MET) et en utilisant des outils et des approches Bioinformatiques, les SNPs contenus dans les

régions chromosomiques des QTLs ont été sélectionnés et les gènes candidats auxquels ces SNPs sont liés ont été recherchés, la fonction de ces gènes a été également déterminée. Concernant la deuxième méthode, la composition en gènes candidats des QTLs a été identifiés ainsi que la fonction assurée par ces gènes a été déterminée en utilisant des outils et des approches Bioinformatiques.

La carte génétique a été construite avec 247 SSR et enrichie de 1425 SNP, couvrant environ 6122 cM. 117 marqueurs SSR (47.36%) et 598 SNP (41.96%) sont localisés sur les chromosomes du génome A ayant ainsi une taille totale de 2902.61 cM avec une densité moyenne de 4.06 cM entre deux marqueurs adjacents, le chromosome 2A contient le nombre le plus élevé en marqueurs (152 dont 133 SNP), il couvre 508.84 cM avec 3.31 cM entre deux marqueurs adjacents le chromosome 5A contient le nombre le plus bas en marqueurs (37 dont 24 SNP) avec une taille de 151.31 cM de taille avec 4.09 cM entre deux marqueurs adjacents. Tandis que 130 marqueurs SSR (52.64%) et 827 SNP (58.04%) sont localisés sur les chromosomes du génome B avec 3219.61% cM comme taille totale pour ce génome avec une moyenne de 3.36 cM entre deux marqueurs adjacents. Le chromosome 2B contient le nombre le plus élevé en SNP (205) tandis que le chromosome 7B contient le nombre le plus bas en SNP (23).

La cartographie des QTL en (MET) a permis la détection de cent vingt-six QTL le long de tous les chromosomes du blé dur. L'analyse de 583 séquences SNPs sélectionnées depuis les régions QTL détectées a montré qu'un ensemble de 122 séquences SNP ont été annotées dont 53% des gènes candidats trouvés étaient impliqués dans la tolérance aux stress, 29.5% dans le développement et la croissance des plantes, et 3.3% dans le transport cellulaire. De plus, 1.6 % des gènes candidats trouvés étaient des rétrotransposons et des transposons, 2.4 % avaient une fonction inconnue et 9.8 % étaient liés à d'autres processus cellulaires. Les résultats ont également montré que 66.7 % des gènes candidats localisées sur le chromosome 4B étaient impliqués dans la tolérance aux stress et 33.3 % dans le développement et la croissance des plantes. Quant à la cartographie des QTL pour des environnements spécifiques (BIP analysis), 9 QTL ont été détectés, dont quatre sont localisées sur le chromosome 2A. Les séquences d'ADN de ces quatre QTL ont été utilisées pour la recherche d'homologie, 546 gènes candidats ont été identifiés dont certains sont présents dans plusieurs QTL (F-box gene family, hydroxyproline-rich glycoprotein-like,

des rétrotransposons/transposons et d'autres gènes candidats). Cette étude a fourni des informations sur l'utilisation de marqueurs SNP pour identifier des gènes candidats liés au caractère de rendement en grains chez le blé dur dans des environnements différentes (sec, froid, chaud). Les chromosomes 2A et 4B contribuent significativement au rendement en grain. Les SNP annotés et les gènes candidats identifiés dans ce travail peuvent être utilisés pour développer des marqueurs moléculaires utiles pour suivre les gènes d'intérêt dans la sélection assistée par marqueurs (SAM) dans les programmes d'amélioration du blé dur.

Mots-clés : QTL, SNP, rendement en grains, gènes candidats, carte génétique, tolérance aux stress

Acronyms

AFLP	: Amplified Fragment Length Polymorphism
ANOVA	: Analysis Of Variance
BIP	: Biparental analysis
BLAST	: Basic Local Alignment Search Tool
CIM	: Composite interval mapping
DArT	: Diversity Arrays Technology
DDJB	: DNA Databank of Japan
EBI	: European Bioinformatics Institute
EMBL	: European Molecular Biology Laboratory
GO	: Gene Ontology
ICIM	: Inclusive Composite Interval Mapping
IM	: Interval Mapping
INSDC	: International Nucleotide Sequence Database Consortium
LOD	: Likelihood
MAS	: Marker-Assisted Selection
MET	: Multi-Environmental traits
MIM	: Multiple Interval Mapping
MQM	: Multiple QTL Mapping
NCBI	: National Center for Biotechnology Information
NLM	: National Library of Medicine
ORF	: Open Reading Frame
PDP	: Protein Data Bank
QTL	: Quantitative Trait Loci
RAP-DB	: Rice Annotation Project database
RFLP	: Restriction Fragments Length Polymorphism
RIL	: Recombinant Inbred Line
SNP	: Single Nucleotide Polymorphisms
SSP	: Seed Storage Proteins
SSR	: Simple Sequence Repeat
URGI	: Unit� de Recherches en G�nomique Info

Figures list

Figure 1	: Geographical location of the fertile crescent	4
Figure 2	: Divergence species of the <i>Poaceae</i> family and polyploidization events	7
Figure 3	: Evolution and genetic relationship between durum and bread wheat and related diploid grasses, with examples of spikes and seeds	8
Figure 4	: National production of cereals (Maize, durum wheat, barley, bread wheat) over the two last decades	10
Figure 5	: Geographical distribution of the cereal sector	11
Figure 6	: Durum Lahn/Cham1 genetic linkage map with the grain yield QTLs detected using MET (mentioned in black color) and BIP (mentioned in other colors) approaches	29
Figure 7	: Genetic linkage map of 2A chromosome QTLs mapping for grain yield of tolerance to drought, heat, and cold	54
Figure 8	: Number of candidate genes identified in the EP, INC, RF, and RQ environments for Rice and Brachypodium with the QTL sizes	56
Figure 9	: Genetic linkage map of 4B chromosome with QTLs and SNPs annotation	58
Article 2		
Figure 1	: Candidate genes identification in specific and stressed environments	60
Figure 2	: Candidate genes identification in multi-environmental analysis	62
Figure 3	: Annotation and similarity study on Venn diagram of SNP	62
Figure 4	: Functional classification of candidate genes co-located with QTLs linked to grain yield	63

Tables list

Article 1

Table 1	: Testing sites, environments (crop rotation) with growing seasons, sowing dates, annual rainfall, added irrigations, and mean temperatures at tillering and flowering stages for the field trials of Lahn/Cham1 population	42
Table 2	: Statistical parameters of grain yield (kg/ha) for RIL mean, maximum, and minimum yielders (kg/ha), and parents across testing environment	43
Table 3	: Assignment and distribution of molecular markers, %, cM size, and coverage across the 14 A and B genomes for SSR and SNP enriched maps in the durum Lahn/Cham1 population	44
Table 4	: Significant QTL regions in contrasting environments for Lahn/Cham1 population with their chromosomal location, position, flanking markers, LOD scores, additive effect, and phenotypic variations.	45
Table 5	: Significant QTL regions for Multienvironment trials with their chromosomes, position, flanking markers, LOD scores, additive effect, and phenotypic variations	46
Figure 6	: Comparison of detected grain yield QTLs harbored by A and B chromosomes in different wheat mapping populations	50

Article 2

Table 1	: QTL detected on 2A chromosome with the environmental conditions, QTL peak positions in cM, QTL regions interval size in centimorgan (cM) and in Mega bases (Mb) and with QTLs region flanking markers	61
Table 2	: The environmental conditions, the number of candidate genes identified for the four QTL linked to grain yield of the four stressed environments (EP, INC, RF and RQ) with the QTL interval size, and the number of candidate genes identified for Rice and <i>Brachypodium</i>	61
Table 3	: SNPs annotated and not annotated in A and B genomes	62

List of Contents

Aknowledgements	i
Abstract	iii
Résumé	iv
Résumé étendu	v
Acronyms	viii
Figures list	ix
Tables list	x
Introduction	1
Chapter 1: Literature review	4
1. History and evolution of wheat in the grass family	4
1.1. Wheat domestication history	4
1.1.1. The agrotechnical revolution	5
1.1.2. The domestication revolution	5
1.1.3. The agricultural expansion	5
1.2. Taxonomy	6
1.3. Wheat evolution	6
2. Economic importance	9
3. Genetic and QTL mapping	11
3.1. Genetic linkage map	11
3.2. QTLs mapping	12
3.2.1. Statistical tools and QTL detection	12
3.2.1.1 Single Marker Analysis	13
3.2.1.2 Interval mapping	13
3.2.2. Improvements of QTLs detection methods and position accuracy	14
3.2.2.1. Composite Interval Mapping (CIM)	14
3.2.2.2. Multiple Interval Mapping (MIM)	14
3.2.2.3. Inclusive Composite Interval Mapping (ICIM)	14
4. Bioinformatics	15
4.1. Outlines	15
4.2. Databases	16
4.2.1. Primary sequence databases	16
4.2.2. Secondary (value – added) databases	18
4.2.2.1. NCBI RefSeq database	18
4.2.2.2. Suiss-Prot	18
4.2.2.3. Rice Annotation Project database (RAP-DB)	19

4.2.2.4. Unité de Recherches en Génomique Info (URGI)	19
4.2.2.5. Protein Data Bank (PDB).....	19
4.2.2.6. Gene Ontology (GO)	20
4.3. Data research	20
4.3.1. FASTA.....	21
4.3.2. BLAST.....	21
5. History of wheat genetic linkage map and the importance of its construction in durum breeding programs	22
5.1. Wheat genetic linkage map history.....	22
5.2. The importance of high-density genetic linkage in durum breeding programs.....	24
Chapter II: Detection of grain yield QTLs in the durum population Lahn/Cham1 tested in contrasting environments.....	25
1. Introduction	25
2. Materials and methods	25
2.1. Experimental design.....	25
2.2. Population genotyping	26
2.3. Linkage map construction and QTLs detection.....	26
3. Results.....	26
3.1. RIL performance.....	26
3.2. Linkage map construction	27
3.3.1. In specific environment analysis	28
3.3.2. In Multi-Environmental QTL mapping.....	28
Chapter III: Dissection of Quantitative Trait Loci (QTL), annotation of Single Nucleotide Polymorphism (SNP), and Identification of Candidate Genes for Grain Yield in <i>Triticum turgidum</i> L. var <i>durum</i>.....	53
1. Introduction	53
2. Materials and methods	53
2.1. QTLs dissection and SNPs annotation.....	53
2.1.1. Dissection of QTLs identified on 2A chromosome and candidate genes identification.....	53
2.1.2. SNPs annotation.....	55
3. Results.....	56
3.1. QTLs dissection and candidate genes identification.....	56
3.2. SNPs annotation.....	57
Discussion	66
Conclusion	73
Perspectives.....	75
References.....	77

Introduction

Introduction

Durum and bread wheat (*Triticum durum* and *Triticum aestivum*) constitute the major cereals in global human consumption, especially in the Mediterranean basin. In West Asia and North Africa (WANA region), the durum wheat production represents 50% of the world production and the area covered is more than 85% of the total world area (Nachit et al., 1995). The annual durum wheat world production reported a decrease of 8% in the 2017/2018 season with an estimation of 36.6 million tons compared to the previous year. This is mainly linked to a reduction of durum production in the North American continent resulting from the drought and the decrease in the area sown in favor of more profitable crops like canola and soybeans. For the world's largest producer of durum wheat (Canada), the drop in production is about 38% (4.5 million tons (Mt) compared to 7.8 Mt last year). During this season, a significant increase in durum production occurs in the Mediterranean basin in comparison with the previous season (Sauvage, 2017).

The largest durum producers in the Mediterranean basin are Syria, Turkey, and Italy followed by Morocco, Algeria, Spain, France, and Tunisia (Habash et al., 2009). In Morocco, the production was increased by 10.6% with 2.43 million tons, but it has recorded a significant decrease (44.74%) in the next season (2018/2019) registering a rate of 1.34 million tons. During the 2019/2020 season, the production has further decreased by 41.20%, thus dropped to 0.79 million tons according to the ONICL (Office national interprofessionnel des céréales et des légumineuses). A new record of 2.83 million tons was recorded for the 2020/2021 season followed by an important decrease of 67.4% with 0.81 million tons in 2021/2022. These fluctuations in the production are mainly due the fact that the cereal crops are largely cultivated in rainfed systems. Therefore, the major environmental constraints limiting the production of durum wheat in this region are drought and extreme temperatures. Changes in total seasonal precipitation and its pattern of variability are both important, and the occurrence of moisture stress during flowering, pollination, and grain-filling is harmful to wheat production.

Moreover, farming wheat in this Mediterranean environment is characterized by a combination of drought with terminal heat at the grain-filling period. This characteristic is ultimately detrimental to grain yield (Habash et al., 2009). These conditions are exacerbated with the advent of global warming, which mainly affects the Mediterranean region. Morocco, is among the most affected countries in this region. Habash et al., (2009) forecast a temperature increase of about 5°C for Morocco in 2099 as well as a reduction in precipitation of up to 30%.

These effects are already being felt during the last three decades in North Africa where an increase in drought frequency during the middle to the end of winter was registered. This phenomenon coincides with the tillering-booting developmental stages and has had a negative impact on durum wheat production (Belaid et al., 2005). In addition, crop cycle has been shortened by almost one month since 1970 in some areas of southern Morocco, making winters warmer and drier. Consequently, the effect of some diseases and insects on wheat was exacerbated. For example, *Puccinia tritici* (leaf rust), *Pyrenophora tritici-repentis* (tan spot), *Fusarium* species (*culmorum* and *graminearum*), and *Helminthosporium sativum* have become more prevalent and limiting durum production. Also, the ecological conditions become progressively favorable to the occurrence of stem rust caused by *Puccinia graminis* f.sp. *tritici*, *Septoria tritici*, and *Mayetiola destructor* ‘Hessian fly’, the major constraints to farming wheat in Morocco, North Africa and Europe. These parasites alone and along with the abiotic stresses (drought, heat, cold...) engendered by climate changes affect largely wheat production. Hence, durum breeders are working hard in order to develop tolerant cultivars to the different stresses. In this context, durum wheat breeding programs have been established and the cooperation between international centers has been enhanced.

In the durum breeding program for Mediterranean dryland established by CIMMYT and ICARDA, several varieties adapted to the Mediterranean drylands were released (Lahn, Cham1, Korifla, Omrabi, Faraj, Nachit etc ...) and have known a lot of success in this region. To benefit more from this progress and the results achieved, several durum populations were developed from crosses between the released varieties; of which the population Lahn/Cham1 that consisted of 112 Recombined Inbred Line (RIL) derived by single-seed descent selection from the cross ICD-MN91-015.

Several studies concerning this population and other durum populations were conducted to develop new varieties with enhance stresses tolerance, yield, and quality, along with other traits of interest. These studies are based essentially on the physiological, biochemical, morphological and phenotypical aspects and have greatly contributed to the improvement of grain yield. Although new tools and techniques like molecular markers were used to follow characters transmission in many studies, there are no studies that have investigated the genetic of the important characters like grain yield in its details. Understanding the genetics functionality and the genome organization using molecular markers and bioinformatic tools has a great scientific value, to improve further yield potential and other traits.

In this context, and to investigate and understand the genetics functionality of grain yield, we have used in this study the durum population cited above (Lahn/Cham1), and the study was aimed to: 1) construct a high-density genetic linkage map of Lahn/Cham1 population. 2) identify the Quantitative Trait Loci (QTL) linked to grain yield. 3) identify the SNPs mapped into QTL regions; 4) extend the SNPs sequence analysis in terms of functional and comparative analysis, and 5) identify the candidate genes constituting the specific QTLs harbored by 2A chromosome and linked to grain yield.

The use of high throughput genetic marker technique like SNP molecular markers to genotype this population allow us the construction of a high-density genetic linkage map. This map constitutes the essential step for understanding the grain yield functionality. Further, the QTLs linked to grain yield were identified, and through the Bioinformatics approaches, the QTL genes composition was established. The choice of parents to develop this population was studied. Lahn is an improved durum wheat variety, developed for favorable and high input environments in the Near East and North Africa. It is a high yielding cultivar with cold tolerance, resistant to yellow rust (*Puccinia striiformis*) and moderately resistant to leaf rust (*Puccinia recondita* Rob.), powdery mildew (*Erysiphe graminis* DC.), common bunt (*Tilletia caries* D.C.), and to *Septoria* leaf blotch (*Mycosphaerella graminicola*). Further, Lahn has a large kernel size with superior hectoliter weight, milling extraction index, and values for gluten strength. Similarly, the second parent Cham1 is also a variety from the CIMMYT/ICARDA durum program. It is adapted to the Mediterranean climate with high yield stability, and it exhibits a good resistance to drought and to yellow rust with some tolerance to Russian wheat aphid. Cham1 has also high values for osmotic adjustment and high yellow pigment content, but its gluten strength is low. Additionally, Cham1 is susceptible to leaf rust, stem rust, *Septoria tritici*, and powdery mildew. This variety has been released in several countries in the Mediterranean basin (Syria, Sudan, Portugal, Algeria, Cyprus, Turkey, Iraq, and Jordan) under different names.

Literature review

Chapter 1: Literature review

1. History and evolution of wheat in the grass family

1.1. Wheat domestication history

The wheat domestication constitutes a milestone in the history of human societies marking the beginning of the Neolithic era which will result in the adoption of a production economy based on agriculture and breeding. This is around 10000 to 15000 BC. JC when wheats were domesticated (Bozzini, 1988) in the south-west of Asia, and more precisely the mountainous region of the *fertile crescent* (Figure 1) between the Mediterranean coasts and the plains of the Tigris and the Euphrates (Harlan and Zohary, 1966; Feldman and Sears, 1981; Shewry, 2009). Archaeological studies showed that the end of the period in which hunting and gathering were the main means of subsistence, coincided with the beginning of wheat cultivation and the sedentary societies establishment (Diamond, 2002). The wheat cultivation history can be subdivided into three phases (Kislev, 1984).

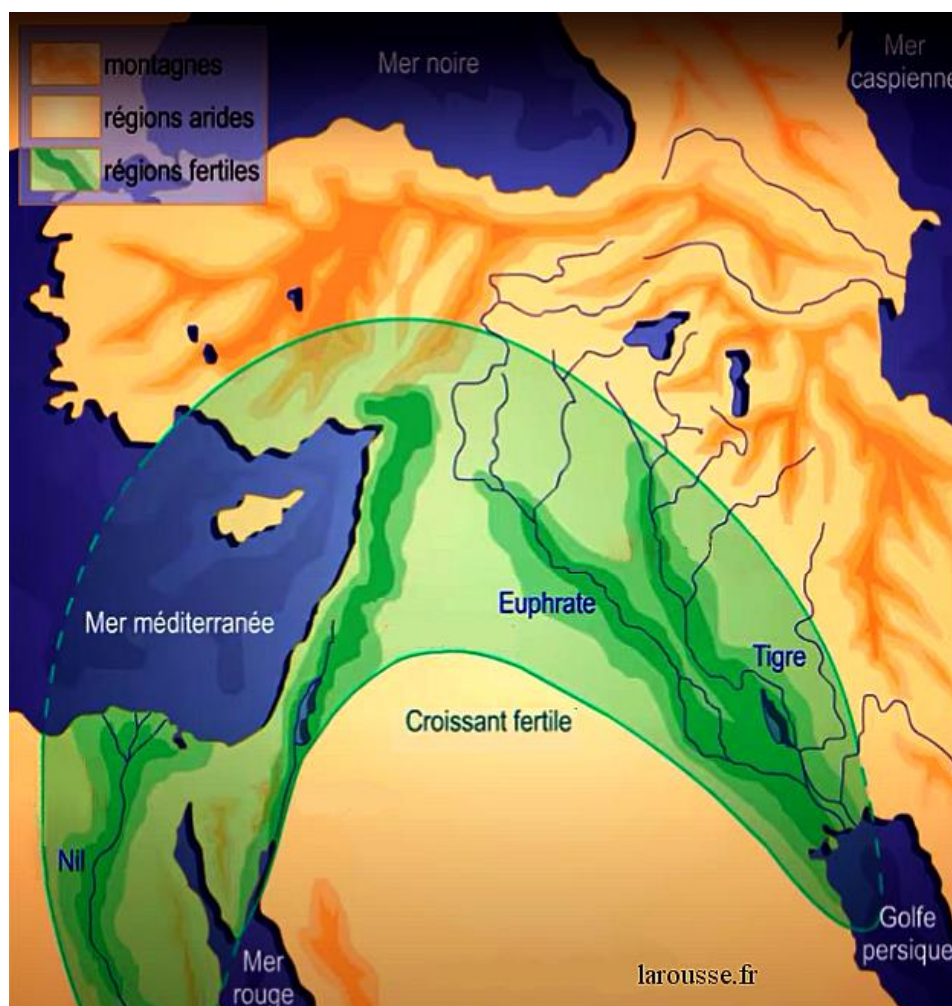


Figure 1: Geographical location of the fertile crescent

1.1.1. The agrotechnical revolution

This phase of field crop husbandry is now thought to have developed about the 11th millennium B.C. Many agricultural tools were unearthed from Natufian sites dated 10300-8500 B.C such as: glossy sickle blades (some of them still mounted on their bone handles), querns, pestles, and pounding stones. Moreover, in sedimentary layers of Tell Abu Hureyra dated from late Natufian in Nord of Syria, important quantities of wild einkorn and other cereals (wild barley and wild rye) as well as wild legumes were identified (Hillman, 1975). These Natufian findings indicate an economy based on wild wheats and barley seeds collection and consumption in addition to hunting and fishing (Garrod and Reckitt, 1957). The sickle invention and its use for harvesting constitutes a major component of the agrotechnical revolution. The employment of this tool reduced considerably physical exertion and it is two times more efficient than manual reaping (Kislev, 1984). Also, it reduced spikelet loss during collection. This feature allowed to the Natufian man to allow cereal grains to ripen more fully before the harvest, without loss of yield. At this stage, the Natufian people seems to be ready for intentional growing of cereals.

1.1.2. The domestication revolution

This process consists to a wild animal and plant material adaption to the conditions rearing and to the uses of this material (Diamond, 2002). At this stage, the farmers were looking for phenotypes well adapted to these agricultural practices. At this stage and in addition to the natural selection, the first farmers even unconsciously exerted selection pressure on wheat by looking for certain important alleles leading to the fixation of several alleles of certain gens called genes of domestication within the species. Such as the *Tg* genes “*Tenacious glumes*” and *Q* genes “free-threshing” illustrate this evolutionary mechanism (Faris et al., 2003; Jantasuriyarat et al., 2004; Azhaguvel et al., 2006; Nalam et al., 2006; Feuillet et al., 2008; Sood et al., 2009). Another example regarding the *Br* genes “*Brittle rachis*”: wheat opted for brittle rachis allele to ensure an easy seed dissemination, however farmers opted for the non-brittle rachis allele to limit yield losses thus exerting strong selection pressure (Nalam et al., 2006; Feuillet et al., 2008).

1.1.3. The agricultural expansion

The migration and trade progression promoted wheat crops spread to Europe and Asia simultaneously (Balfourier et al., 2007). The wheat crops appear in Greece, Cyprus, and India from 6500 BC, in Egypt around 6000 BC, reaches Germany and Spain around 5000 BC,

England and Scandinavia around 3000 BC and finally China around 2000 BC. In the American continent, wheat did not appear until the Christopher Columbus's arrival during the XVIIth century (Bogard, 2012).

Currently, the main cultivated species of the *Triticum* genera are *Triticum aestivum* L. (common wheat) and *Triticum durum* (durum wheat). They differ mainly in the appearance and in grain composition as well as in their agroecological distribution, durum wheat is found in the Mediterranean zones while bread wheat is cultivated in more temperate zones.

The cumulated technological advances during the wheat expansion: use of animal traction (from 3000 BC), crop rotation, use of organic manure, flail for grain threshing (from the Middle Ages), seed drill installation at the 18th century, mineral nitrogen fertilizers application, the arrival of combine harvesters and semi-dwarf varieties of the green revolution during the nineteenth century were greatly increased the productivity of wheat crops yield.

1.2. Taxonomy

The Grass (*Poaceae* or *Gramineae*) is a family with more than 600 genera and 10000 species (Clayton and Renvoize, 1986), this family is of ecological dominance, though growing in diverse latitudes and climates (Twiss, 1992). Many species of this family having great agronomic importance have been domesticated such as: rice (*Oryza* genera), Maize (*Zea* genera), sorghum (*Sorghum* genera), oats (*Avena* genera), rye (*Secale* genera), barley (*Hordeum* genera) and wheat (*Triticum* and *Aegilops* genera). The wheat group includes many species, which according to the hierarchical classification of species are angiosperms (flowering plants) monocotyledons of the *Poaceae* family, of the *Pooideae* subfamily and of the *Triticeae* tribe (Twiss, 1992). The *Triticum* genera includes species with various ploidy levels, we find diploid species such as the einkorn Wheat (*triticum monococcum*) or, tetraploid species such as emmer wheat (*Triticum dicoccum*) or durum wheat (*Triticum Durum*) and finally hexaploid species such as common wheat (*Triticum aestivum*).

1.3. Wheat evolution

The economic interest of the domesticated species has made of the *Poaceae* a family highly studied in genomics. Comparative studies among species of this family have shown good synteny (Moore et al., 1995; Bennetzen and Freeling, 1997; Singh et al., 2004; Colasuonno et al., 2013; Molnár et al., 2016). Other works showed that around 60 to 70 million years (Mya) ago, *Poaceae* species diverged from a common ancestor inhabiting diverse ecological and agricultural systems (Moore et al., 1995; Kellogg, 2001; Gaut, 2002; Prasad et al., 2005). The

divergence dates estimation of several *Poaceae* species were based on the sequence comparisons of the acetyl-CoA carboxylase (Acc) orthologous gene, generally present in a single copy in the *Poaceae* family (Huang et al., 2002; Chalupska et al., 2008; Nasernakhaei et al., 2015). Thus, the divergence of the *Panicoideae* to constitute a subfamily is estimated at 60 Mya, followed by the *Oryzoideae* called also *Ehrhartoideae* at 50 Mya, oats at 31 Mya, and from the *Triticeae*, barley diverged at 11,6 and rye at 6,7 Mya (Figure 2). It is generally accepted that the different species of wheat with the A, B and D genomes were diverged at 2,4-2.5 Mya (Huang et al., 2002; Devos et al., 2005; Chalupska et al., 2008).

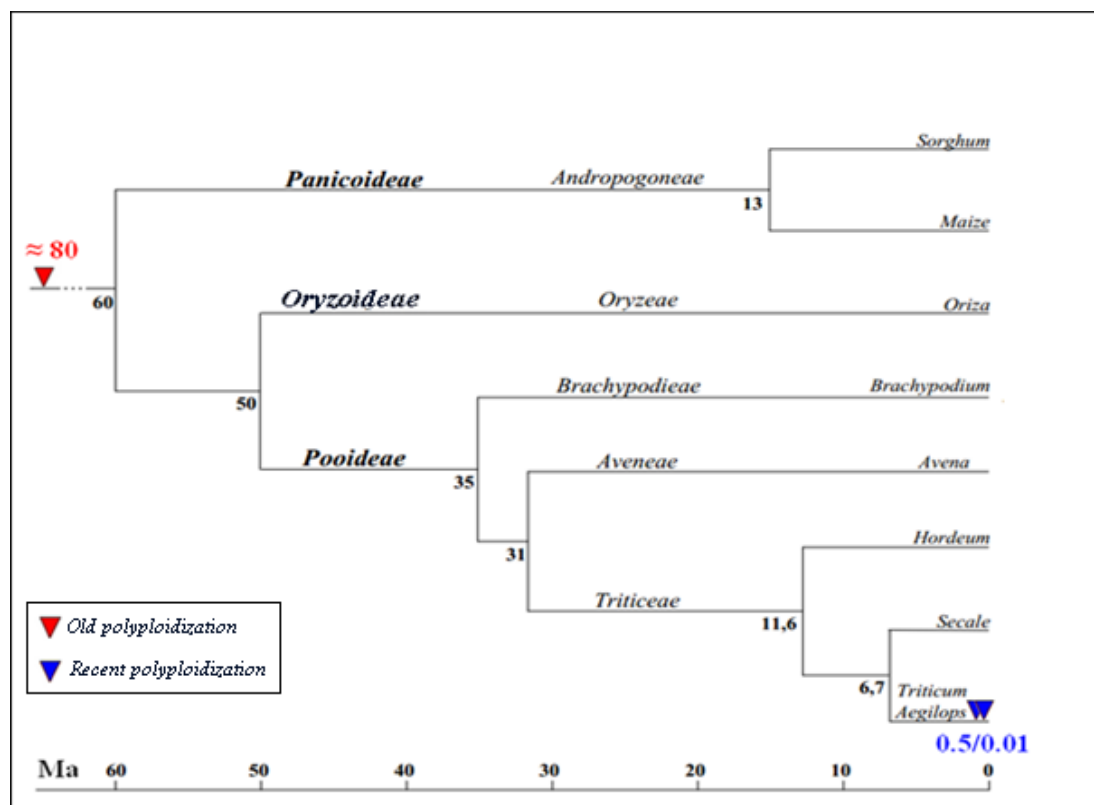


Figure 2: Divergence species of the *Poaceae* family and polyploidization events (Soreng et al., 2015)

The evolutionary process of eukaryotic genomes was marked by an important phenomenon “polyploidization” which was more common in plants, especially in angiosperms (Wolfe, 2001; Adams and Wendel, 2005; Soltis, 2005; Cui et al., 2006; Salse et al., 2008), this polyploidization can occur by the genome duplication “autopolyploidy” or by the fusion of two distinct parental closed species to form a hybrid “allopolyploidy” (Wolfe, 2001).

The *Poaceae* have undergone several polyploidizations during their evolution (Salse et al., 2008; Bolot et al., 2009). Some are old (about 80 Mya), it is called paleopolyploidy and

then it is common to all species of this family (Chalupska et al., 2008). Others are more recent, therefore restricted to certain tribes / species (Chalupska et al., 2008) see Figure 2. This is the case of the *Triticeae* tribe, where two major polyploidization events were detected leading to a series of species constituting this tribe with varied level of ploidy (Figure 2). The first polyploidization at 0.5 Mya ago before the domestication occurred via the crossing between two diploid species at $2n = 14$ *Triticum urartu* providing the A genome and a close species of *Aegilops speltoides* which has now disappeared providing the B genome, the progeny of this cross is represented by tetraploid wheats such as durum wheat "*Triticum turgidum ssp durum*" and the emmer wheat "*Triticum dicoccum*". The second event, which occurred 10000 years ago, is a cross between *Triticum tauschii* providing the D genome and a domesticated variety of tetraploid wheat (Lagudah et al., 1991; Boyko et al., 1999; Shewry, 2009) give birth to hexaploid wheats such as common wheat "*Triticum aestivum*" or dinkel/hulled wheat "*Triticum spelta*" (Figure 3).

This allopolyploidization has led to the formation of very large genomes. The durum wheat genome made up of 14 pairs of chromosomes grouped into seven homoeologous groups representing the genomes of each ancestor, whose chromosomal formula is: $2n = 4x = 28$, AABB. While the bread wheat genome is structured into 21 pairs of chromosomes also grouped into seven homeologous groups representing the genomes of each ancestor, its chromosomal formula is $2n = 6x = 42$, AABBDD (Figure 3).

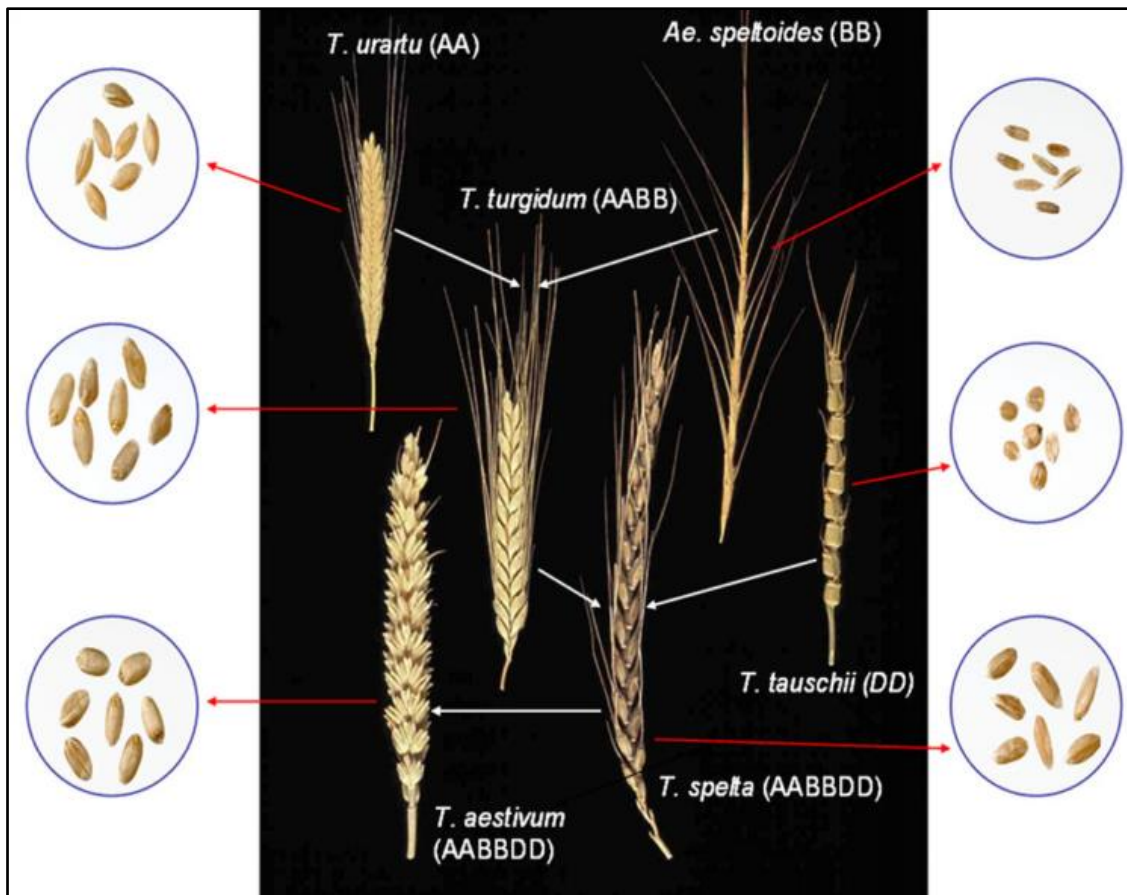


Figure 3: Evolution and genetic relationship between durum and bread wheat and related diploid grasses, with examples of spikes and seeds (Shewry, 2009)

2. Economic importance

Wheat and other cereals (maize and rice) constitutes the basis of the human food. Also, it is a privileged resource for animal feed and for multiple industrial applications. It is the most widely grown cereal in the world on an area of 216 million hectares, with 766 million tons (Mt) as harvest in 2019 (FAO, 2020).

In Morocco, the cereal sector is important socio-economically, it occupies very significant place in the Moroccan agriculture. Indeed, this sector alone accounts for nearly 71% of the total Useful Agricultural Area (UAA). It generates 20% of the overall agricultural turnover with fluctuations depending on the annual precipitation. It contributes with 19% to the people employment, also it contributes to human consumption and animal feeding, generating very marked interdependencies with the livestock sector (Ministry of agriculture, 2022).

Between 2003 and 2019, the grain yield per hectare in cereals was increased by 42%, from 12 to 17 quintals (Ministry of agriculture, 2022). As for the production, it was increased by 25%, rising from around 64 million quintals (between 2003 and 2007) to around 80 million

quintals (between 2015 and 2019). This increase was achieved despite a drop in the national area by -32%, decreasing from 5.35 Mha in 2008 to 3.65 Mha in 2019. This drop is part of the green Morocco plan's objective for the reduction of the area of cereals through the conversion to crops with high added values. With about 1.4 million national farms, Moroccan production generates a turnover of more than 15 billion dirhams. Since 2019 season, the mean of the national production in cereals has dropped again to 56 million quintals. The Figure 4 shows the national production over the two last decades. In general, every year Morocco invests for cereal imports nearly 6 billion dirhams of cereals.

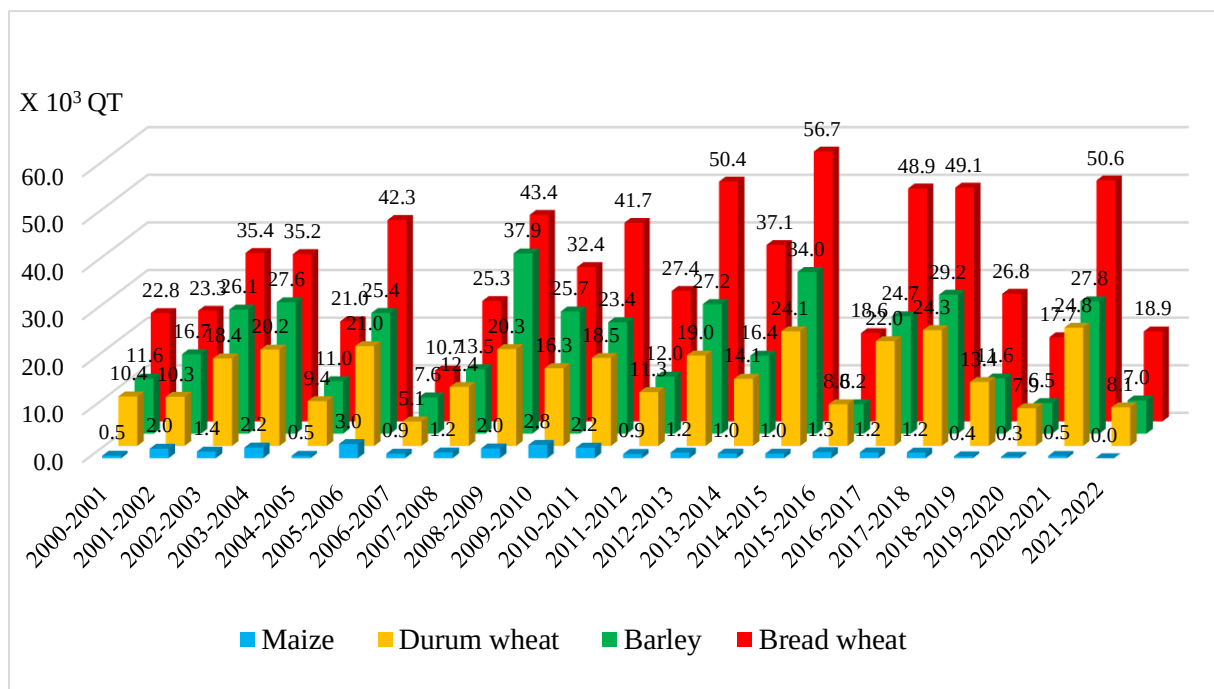


Figure 4: National production of cereals (Maize, durum wheat, barley, bread wheat) over the two last decades (ONICL, 2022)

Cereals are grown in the different agro-climatic zones of the country in rotation with other annual crops represented mainly by legumes, industrial crops, and fodder crops. The main production regions are located in the rainfed areas of the plains and plateaus of Zaer, Chaouia, Abda, Haouz, Tadla, Gharb and Saïs where the vast majority of farms practice cereal growing, whatever their size (Ministry of agriculture, 2022). The geographical distribution of cereal crops is summarized in the following map (Figure 5).

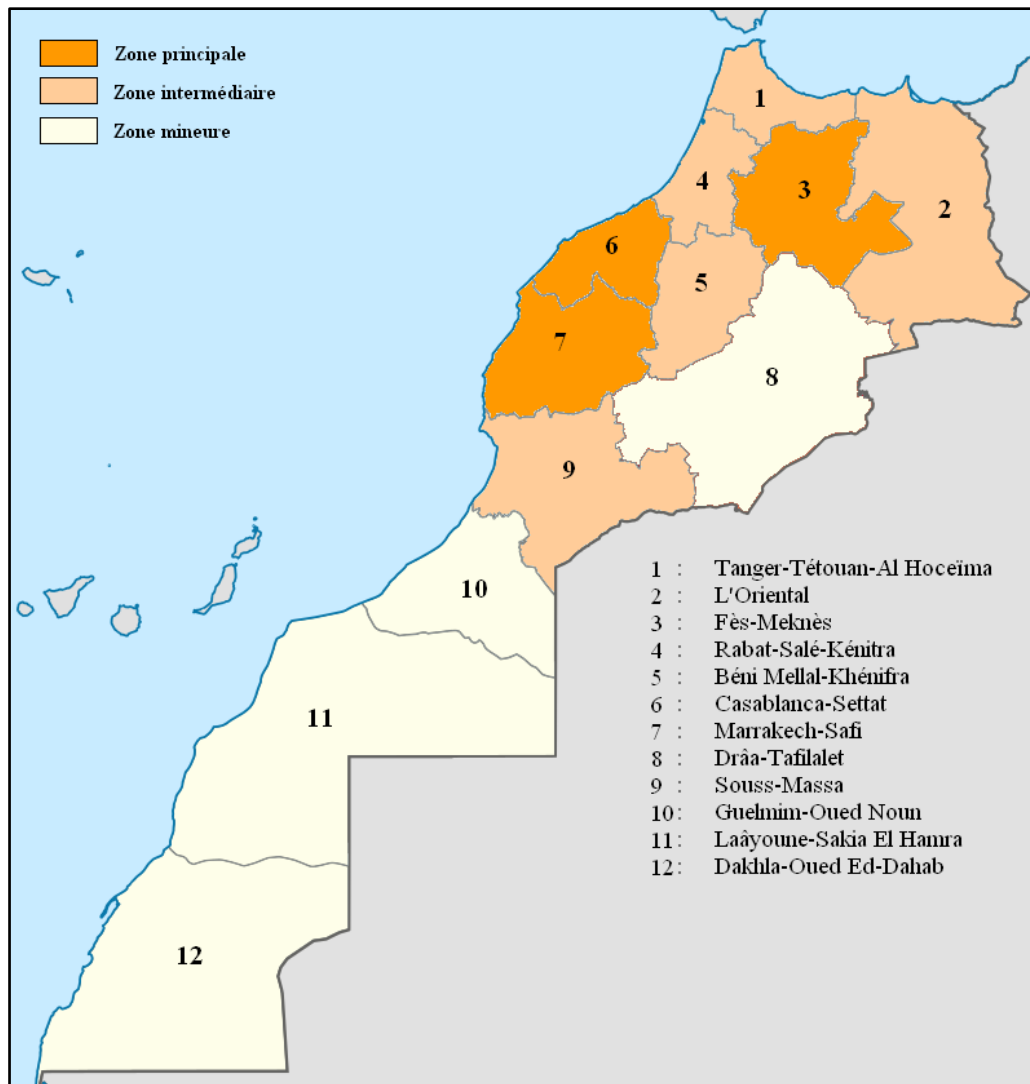


Figure 5: Geographical distribution of the cereal sector (fellah-trade, 2022)

Durum wheat occupies an important place in the cereal crops, it has a significant impact on the world economy. In addition to its use in bread production, it serves also to produce pasta and several Mediterranean end products such as Frike, couscous, bulgur (Nachit, 1992), also it is an important source of calories, protein, and vitamin A in this region (Nachit, 1992; Elouafi and Nachit, 2004; El Yadini et al., 2013).

3. Genetic and QTL mapping

3.1. Genetic linkage map

A genetic map is an inventory of genes or genetic markers of species or group of individuals. It defines the relative order of genetic markers and/or genes, and the distances separating them from each other on each chromosome. Concretely it is a representation of the genome species. The construction of a genetic map takes place in two main stages. Firstly,

linkage groups are formed (sets of linked markers). Then, a given marker is assigned to the group when it is linked to at least one of the markers of this group. To constitute these groups, the recombination frequency r is used to indicate the distance between two markers, by testing, using the most satisfactory and general approach for linkage analysis named maximum likelihood method (LOD) (Haldane and Smith, 1947; Morton, 1955), the hypothesis $H_0: r = 0.5$, i.e. that “there is independence of the markers” against the hypothesis $H_1: r < 0.5$ meaning that “the two markers are linked”. Thereafter, the markers are ordered within each linkage group according to the most probable order. To transform the recombination frequency between adjacent markers into genetic distance (d) in centimorgan (cM), several functions were available:

- Morgan’s function:

$$d = r$$

- Haldane’s function (Haldane, 1919) which considers the possible multiple crossing-over on a chromosome and assumes independence between them:

$$d = -\frac{1}{2} \ln(1-2r)$$

- The Kosambi function (Kosambi, 1943) which, more generally, considers the interference between the different crossing-over. Meaning that the presence of a crossing-over decreases the probability of another crossing-over at proximity:

$$d = \frac{1}{4} \ln \left(\frac{1+2r}{1-2r} \right)$$

3.2. QTLs mapping

3.2.1. Statistical tools and QTL detection

The traits of agronomic importance are quantitative due to the collective action of many genes termed quantitative trait loci (QTLs) interacting with environment and with other QTL (Geldermann, 1975; Johansson, 1909). The statistical approach for QTL detection takes various forms. The first methods consisted of the variance’s analysis testing the marker effect on the performances (Neimann-Sorensen and Robertson, 1961). Thereafter, the maximum likelihood approach (LOD) was proposed to find QTL in a linkage group (Lander and Botstein, 1989). Basing on these approaches several methods are available:

3.2.1.1 Single Marker Analysis

We investigate whether a QTL localized in the vicinity of the marker considered, would be involved in the variability of the quantitative character studied. Two approaches are considered:

- The ANOVA approach (Soller et al., 1976): If the marker and the QTL are considered to be confounded in the strict sense of the term, this means that the recombination r is equal to zero, the ANOVA approach consists in applying a linear ANOVA model to a single factor which was the genotype to the marker.
- The maximum likelihood approach (Luo and Kearsey, 1989; Rebai et al., 1995). The quantitative trait is assumed to be normally distributed in each genotype class. This method is advantageous for not requiring the existence of a complete genetic map, but it does not allow to differentiate between a QTL with relatively strong effect and far from the marker from another QTL very close to the marker with a weak effect.

3.2.1.2 Interval mapping

When a genetic linkage map is available, it is possible to envisage a QTL research approach across the entire genome. The analysis is done from a linkage group to other. The general idea is to test the hypothesis of the absence of QTL in each genomic position. Lander and Botstein (1989) have proposed the Interval Mapping for the backcross populations. This method considers two markers at the same time thus defining successive intervals over the whole genome. When two markers are linked with a recombination rate r , the QTL is tested if it is located between these two markers. In this technique which: considers two markers simultaneously, finds the most likely position, estimates its effect, and tests its significance, is more powerful and robust than single marker analysis (Haley et al., 1994). The interval mapping method has several advantages. But also has several problems, particularly in distinguishing the effects of multiple linked QTL. When there are two or more QTLs located on a chromosome, the QTLs mapping can be seriously biased, and mapped to wrong positions or obtaining a phantom QTL between two real QTLs on the same linkage group (Knott and Haley, 1992; Martínez and Curnow, 1992). To increase the reliability and accuracy of QTLs mapping, the effects of possible multiple linked QTLs on a chromosome should be adequately separated in testing (Zeng, 1994).

3.2.2. Improvements of QTLs detection methods and position accuracy

To overcome these problems of QTLs detection, the analysis is pushed towards the multiple QTLs detection, so Haley and Knott (1992), Martinez and Curnow (1992) proposed simultaneous detection of two QTL located in two different intervals, this method has been proposed in the context of multiple linear regression (of the phenotype on the markers). The coefficients of the regression are interpreted in terms of QTL locations and its effects. Moreover, the interaction effects between the two QTLs can be considered (Knott et al., 1996), in this analysis which greatly increases the number of parameters to be estimated, several methods are available:

3.2.2.1. Composite Interval Mapping (CIM)

The Multiple QTL Mapping (MQM) method (Jansen, 1993; Jansen, 1994; Jansen, 1996; Jansen and Stam, 1994) and the Composite interval mapping (CIM) method developed by Zeng (1994) are similar. They consider the position studied and the impact of other chromosomal regions on the variability of the trait. These methods have been developed for crosses between “inbred” lines. The CIM method consists of adding the marker effects in cofactors in the interval mapping model. Thus, in a multiple linear regression, the effects of other chromosomal regions are integrated as cofactors into the analysis using markers flanking these regions (Zeng, 1994).

3.2.2.2. Multiple Interval Mapping (MIM)

Zeng et al. (1999) and Kao et al. (1999) proposed the MIM method, which is a generalization of the CIM (Composite Interval Mapping) for any number of QTLs. It is an iterative method, i.e., we start with a model with total absence of QTL, and at each iteration the model can be increased or decreased by one or more QTLs at a time. The main flaw of this method is the relatively large number of parameters to estimate and therefore the calculation time. Indeed, it is necessary at each iteration estimate the position and the effect for each QTL, as well as the epistasis effects between these QTLs.

3.2.2.3. Inclusive Composite Interval Mapping (ICIM)

The CIM method is the most commonly used for QTLs mapping. However, the algorithm implemented in the popular QTL Cartographer software may not completely ensure all its advantageous properties. In addition, different background marker selection methods may give very different mapping results, and the nature of the preferred method is not clear. Li et al. (2007) proposed a modified algorithm (ICIM) in which the marker selection is conducted only once through stepwise regression by considering all marker information simultaneously, and

the phenotypic values are then adjusted by all markers retained in the regression equation except the two markers flanking the current mapping interval. The adjusted phenotypic values are finally used in interval mapping (IM). The modified algorithm has a simpler form than that used in CIM, but a faster. In addition, ICIM retains all advantages of CIM over IM and avoids its disadvantages (Li et al., 2007). Extensive simulations using two genomes and various genetic models indicated that ICIM has increased detection power, reduced false detection rate, and less biased estimates of QTL effect (Li et al., 2007). Subsequently this method was extended to cover the analysis and mapping of dominance, additivity, and epistasis interactions between QTLs (Li et al., 2008; Zhang et al., 2008; Li et al., 2010; Zhang et al., 2010; Li et al., 2012; Li et al., 2015)

In wheat, several QTLs studies were conducted for different complex traits like pasta quality (Zhang et al., 2008), crown rot resistance (Ma et al., 2010), resistance to Soil-borne cereal mosaic virus (Maccaferri et al., 2011; Russo et al., 2012), nitrogen use efficiency (Quraishi et al., 2011), rusts resistance (Haile et al., 2012; Zeng et al., 2019), yellow pigment content (Colasuonno et al., 2014). And for grain yield and its components (Maccaferri et al., 2008 ; McIntyre et al., 2010 ; Bogard et al., 2011 ; Heidari et al., 2011 ; Wang et al., 2011 ; Bennett et al., 2012; Blanco et al., 2012 ; Dura et al., 2013 ; Wang et al., 2013 ; Dura et al., 2014 ; Graziani et al., 2014; Li et al., 2014; Zhang et al., 2016; Marcotuli et al., 2017; Tura et al., 2019).

4. Bioinformatics

4.1. Outlines

Bioinformatics, also known as "computational biology" or "in silico biology" is a scientific discipline which encompasses the collection, the organization, and the analysis of biological data using bioinformatics tools. The term bioinformatics combines two approaches, the first approach has a computational orientation which consists of the elaboration of algorithms and the development of programs to extract biological information as well as the creation of databases. The second approach consists of bio-analysis, which focuses on the analysis of biological data and their significance in a biological context. These two approaches, bioanalytical and computer science, are complementary and are linked to each other in their evolution. Indeed, studying and understanding the biological systems is made possible through the development of adapted and efficient tools and programs. Bioinformatics is a young discipline that dates from the second half of the twentieth century. From the emergence of the

central dogma in molecular biology (Crick 1958) marking the foundations of molecular biology until the publication of the complete sequences of the genome of *Saccharomyces cerevisiae* and of Human towards the end of the XXth century, in addition to the significant development of sequencing materials, and the development of more powerful calculation programs, the data are generated with a higher throughput, hence the need to gather them in data banks. The analysis of this large amount of data is possible only through the bioinformatic.

4.2. Databases

Since the use of high-throughput technologies, particularly those of sequencing. The data generated has increased exponentially, this creates needs to rapid, easy, and organized storage / access to information by biologists. These needs become an engine for the development of bioinformatics. Therefore, the creation of databases. Most scientific information was kept in “hard” format: written records, articles in scientific journals, books, artifacts, and various sorts of images, eventually including photographs, x-ray pictures, and CT scans. The information content changed with new discoveries and interpretations, but the form of the information was stable and well understood (Esnayra et al., 2000). With the rapid growth of information generated, the transition to electronic support became self-evident to allow to scientists faster research and analysis. In the early 1970s, there were only banks of protein sequences and their three-dimensional structures. The first nucleic acid banks appeared towards the end of the 1970s. Depending on the type of information contained in the data banks, a distinction is made between: Primary sequence databases, in which are stored the sequences corresponding to basic objects in molecular biology (DNA, RNA and protein) from all organizations, and value-added banks that focus their efforts on particular points such as a particular organism or a given theme, providing additional information.

4.2.1. Primary sequence databases

The primary sequence databases are archival in nature. They contain raw sequence data (experimental results) with some interpretations and explanations. There are also redundancies in the primary databases, the same sequence might be submitted by different laboratories, sometimes under different names (Choudhuri, 2014). A great majority of protein sequences in the primary databases are derived from computational translation of the open reading frame (ORF); hence they have not been experimentally verified for the most part. There are three primary databases that contain all the sequences data so far generated. These are **GenBank**, **EMBL database** (also called the **EMBL-Bank**), and **DDBJ** (DNA Databank of Japan). These

databases are interconnected; so, the data submitted to any one of these databases are shared by the others, hence it can be retrieved from all three databases (Choudhuri, 2014).

The GenBank was created in 1979 at the Los Alamos National Laboratory and was called the Los Alamos Sequence Database. It was renamed GenBank in 1982 and became a public database. During 1989 to 1992, GenBank transitioned to the newly created NCBI, a division of the National Library of Medicine (NLM), located on the campus of the US National Institutes of Health (NIH) in Bethesda, MD. GenBank is built and distributed by the NCBI. NCBI began accepting direct submissions to GenBank in 1993. Since its creation, GenBank has grown at an exponential rate, doubling in size every 18 months (Choudhuri, 2014). The NCBI home page is <http://www.ncbi.nlm.nih.gov/>.

The EMBL (European Molecular Biology Laboratory) was founded in July 1974 based on an intergovernmental treaty of nine European countries plus Israel. Then, it has grown in membership; Luxembourg became the twentieth member in 2007, and Australia joined as an associate member in 2008. The EMBL is located in Heidelberg, Germany. An outstation of EMBL is the European Bioinformatics Institute (EBI), located at Hinxton, near Cambridge, UK. The EMBL database as a central depository of nucleotide sequence was created in 1981 and was known as the EMBL Data Library. The EMBL Data Library moved to the EBI in 1993 and became the precursor to the current EMBL-Bank, which is also maintained at the EBI. The expression “EMBL-Bank” is not frequently used. In the literature, the EMBL-Bank is mostly referred as EMBL nucleotide sequence database or EMBL database (Choudhuri, 2014). The EMBL-Bank is now part of the European Nucleotide Archive (ENA). The ENA is developed and maintained at the EMBL-EBI under the guidance of the International Nucleotide Sequence Database Consortium (INSDC). The home page of EMBL-EBI is <http://www.ebi.ac.uk/>. Various databases and tools maintained by EMBL-EBI and made freely available for use can be accessed using EMBL Services at <http://www.ebi.ac.uk/services>.

The DDBJ has been in operation since 1986 and it is maintained at the National Institutes of Genetics at Mishima, Japan. DDBJ is the sole nucleotide-sequence data bank in Asia. The DDBJ home page is <http://www.ddbj.nig.ac.jp/>. Many improvements and added features of DDBJ were discussed (Choudhuri, 2014).

The INSDC (<http://www.insdc.org/>), a collaborative consortium, was initiated between GenBank, EMBL (ENA), and DDBJ to connect these three databases. This collaboration created the International Nucleotide Sequence Database (INSD). For over 30 years, the INSDC

has maintained the primary nucleotide-sequence database (Brunak et al., 2002). The INSDC advisory board is composed of members of each of the databases' advisory bodies. The INSDC has a policy of providing free and unrestricted access to all the available data to scientists worldwide.

4.2.2. Secondary (value – added) databases

The secondary databases extract the information from primary (archival) databases. Multiple entries of the same sequence in primary databases are merged to create a single sequence in the secondary databases. Therefore, they are non-redundant databases. The sequence and all linked information are manually curated. The final sequence flat file has links to all the original entries about the sequence. The value-added database comes to answer to questions such as which genes are expressed under specific conditions and how gene expression vary under different conditions. The added value comes from: data processing, additional annotation, mapping to standardized vocabularies or ontologies (such as Experimental Factor Ontology) and analysis to extract gene expression profiles and other results from primary data.

4.2.2.1. NCBI RefSeq database

It is a secondary database providing a collection of curated, non-redundant, well-annotated sequences including genomic DNA, transcripts, and proteins. The RefSeq database also provides a lot of other information about these sequences, such as characterization, mutation, polymorphism analysis, expression studies, and comparative analyses. Although the RefSeq is non-redundant database, it does include transcripts which are alternatively spliced and encodes for the same protein or distinct protein isoforms, in addition to orthologs, paralogs, and alternative haplotypes (Choudhuri, 2014).

4.2.2.2. Suiss-Prot

It is one of the best secondary databases of proteins. It is non-redundant and curated database, and It is a part of the larger database system called the Universal Protein Resource Knowledgebase (UniProtKB), which was initiated in 2002 by the UniProt consortium. The UniProtKB consists of two parts: UniProtKB/Swiss-Prot (reviewed, manually annotated) and UniProtKB/TrEMBL (unreviewed, automatically annotated; TrEMBL = translated EMBL). UniProtKB/Swiss-Prot contains manually annotated records and information obtained from the literature and curator-evaluated computational analysis. Whereas UniProtKB/TrEMBL contains computationally analyzed records that still need full manual annotation. Differences found between various sequencing reports are analyzed and fully described in the feature table,

such as alternative splicing events and polymorphisms. UniProt currently contains four databases: UniProtKB, UniProt Reference Clusters (UniRef), UniProt Archive (UniParc), and UniProt Metagenomic and Environmental Sequences (UniMES) (Choudhuri, 2014).

4.2.2.3. Rice Annotation Project database (RAP-DB)

The IRGSP, formally established in 1998, pooled the resources of sequencing groups in ten nations to obtain a complete finished quality sequence of the rice genome (*Oryza sativa* L. ssp. *japonica* cv. Nipponbare). Finished quality sequence is defined as containing less than one error in 10,000 nucleotides, having resolved ambiguities, and having made all state-of-the-art attempts to close gaps. The IRGSP released a high-quality map-based draft sequence in December 2002 (Sasaki and International Rice Genome Sequencing, 2005). Upon completion of Rice genome sequencing in 2004, the Annotation of this genome was required, and the Rice Annotation Project (RAP) was organized in 2004 (Ohyanagi et al., 2006; Itoh et al., 2007; Sakai et al., 2013). This database contains IRGSP genome sequences, RAP-DB loci with their identifiers, transcriptomes with links to: Gene Ontology, domain and motif information, cDNA information and other information. It also contains tools like the BLAST program.

4.2.2.4. Unité de Recherches en Génomique Info (URGI)

The “Unité de Recherches en Génomique Info = URGI” founded in 2002, is a research unit in genomics and bioinformatics at French National Institute for Agriculture, Food and Environment (INRAE), dedicated to plants and crop parasites. The URGI research activity covers data integration as well as genome structure and dynamics. The research goal at URGI is to acquire a better understanding of genome structure, evolution, and function, also to help researchers to interpret their data. The URGI unit hosts a bioinformatics platform which supports research activity on plants of agronomic interest for INRA, it contains tools and software generally developed by INRA such as the BLAST program, a centralized information system, which stores genetics and genomics data (genomic annotation, genetic resources, genetic markers, QTL, GWAS "association studies", polymorphisms, and phenotypes ...). It allows to the scientists genomic and genetic data access to extract valuable information on genes of agronomic interest and on the structure of the genome and data from several species such as grapes, wheat, maize, *Arabidopsis*, tomato, poplar, and certain trees (oak, maritime pine).

4.2.2.5. Protein Data Bank (PDB)

The PDB at Brookhaven National Laboratory, is a database containing experimentally determined three-dimensional structures of proteins, nucleic acids, and other biological

macromolecules. It is the single global repository of experimentally determined 3D structures of biological macromolecules, it was established in 1971 and becoming the first open-access digital resource in the biological sciences. A significant increase in structure deposition throughout its existence is observed, e.g., going from 54,000 protein structures in 2008 to 160,000 in 2020 (Shinada et al., 2020). The most of the PDB inputs is Protein structures, but there are also structures of RNA and DNA molecules, nucleic acid-protein complexes. These structures are determined experimentally by X-ray crystallography, NMR, or electron microscopy (Sussman et al., 1998; Kouranov et al., 2006; Berman et al., 2007; Burley et al., 2017; Goodsell et al., 2020; Shinada et al., 2020).

4.2.2.6. Gene Ontology (GO)

The GO consortium funded by the National Institute of Health (NIH) started in 1998. Initially, GO contained only three model organisms. Then, it is extended to over 3200 (Ashburner et al., 2000; The Gene Ontology Consortium, 2019). The Gene Ontology (GO) resource is the world's most comprehensive source of information about the functions and products of genes (proteins and non-coding RNAs). This information can be reads by human and by machines. Therefore, it plays a critical role in the computational analysis of genomic and biomedical data (Carbon et al., 2021). It covers genes from different organisms from all kingdoms of living organisms, as well as viruses (Carbon et al., 2021). The Gene Ontology resource (GO; <http://geneontology.org>) is the most comprehensive and widely used knowledgebase concerning the functions of genes. In GO, all functional knowledge is structured and represented in a suitable form for computational analysis, which is essential to support modern biological research. The GO knowledgebase is structured using a formal ontology, by defining classes of gene functions (GO terms) that have specified relations to each other. Currently, The GO knowledgebase contains over 7 million annotations to genes/gene products from over 3,200 species (The Gene Ontology Consortium, 2019). The GO has become the standard database in genome annotation. In addition, it is incorporated as a reference in several other databases, notably Swiss-Prot and RAP-DB.

4.3. Data research

The important development of the Internet has allowed the easy access to databases and to tools made available by bioinformatic centers. The organization degree of these data varies from simple files to tables of relational systems (Ghosh, 1991). However, most of data is organized as sequential files or flat files. This situation arose for several reasons. The flat file

format can be manipulated and comprehended by both humans and machines. The almost intuitive structure of the flat file comparing to other formats, and it can be easily constructed. Specific software has been developed with the aim of making databases more easily exploitable and allowing users to extract information from them.

The data exploitation is conventionally based on the search of similarity, which consists in first time in searching for the closest sequences in the database to an unknown initial sequence named “Query”. In second times, the query is characterized according to the closest sequences founds in the database. The research of the closest sequences to the query is based on similarity research by pairwise alignment between the query and sequences in database in order to obtain as many matches (identities or similarities) as possible. Several algorithms have been developed in order to find the best alignment between the query and one of the sequences in the database (Needleman and Wunsch, 1970). These algorithms are extremely sensitive, but far too slow and expensive in memory. At the end of the 80s Two programs leading to a rapid and satisfactory result have been developed: FASTA (Pearson and Lipman, 1988) and BLAST (Altschul et al., 1990). they currently constitute the basic tools of the search for similarity in the banks of sequences. They offer both efficiency and speed to their users.

4.3.1. FASTA

This algorithm was originally a program developed by Pearson and Lipman (1988) serving to the alignment of DNA and protein sequence. It is the descendant of the FASTAP program (Lipman and Pearson, 1985). The research principle of FASTA program consist in considering small sequences (word) from the query and by using local sequence alignment to find matches in database sequences. Then, the research will be expanded for the length of the query using an adapted algorithm (Smith and Waterman, 1981). The FASTA file format used as input for this software is now largely used by other sequence database search tools (such as BLAST) and sequence alignment programs (Clustal, T-Coffee, etc.)

4.3.2. BLAST

Basic Local Alignment Search Tool (Altschul et al., 1990), is a similarity search program. It based on the same principle as for FASTA, that is, from the best alignments or very high scoring small segments the alignment will expanded. Although the initial BLAST algorithm is very fast and assigns a statistical value to the score obtained, however it did not allow insertions or deletions. In order to meet different needs, this algorithm has been modified several times. Thus, BLAST2 allows insertions and deletions, but the statistic is no longer exact. Subsequently

comes the PSI-BLAST (Altschul et al., 1997), in addition to filters like the programs SEG (Wootton and Federhen, 1996), the results were improved. The interpretation of these results comes down to the analysis of the significance of the alignments is statistically evaluated according to the length and composition of the sequence, the size of the library and the scoring matrix used. Thus, BLAST calculates a score (the E-value), which corresponds to the probability of randomly observing this alignment in the database used. then, more the E-value tends towards zero, more the alignment is significant. However, the final decision is also based on a visual inspection of the result by the experimenter.

It can be noted that for very short query sequences, the E-Value is very low, thus we cannot rely on these values to estimate the significance of the alignments. The sequences aligned and detected by BLAST are also called “hit”.

5. History of wheat genetic linkage map and the importance of its construction in durum breeding programs

5.1. Wheat genetic linkage map history

Genetic studies of any organism require the establishment of genetic linkage map made up of molecular markers covering the genome. Genetic linkage maps saturated with molecular markers are the basic and essential tool in advanced genetic studies such as: positional gene cloning, QTLs mapping, and marker-assisted selection (MAS). At the beginning, it has been possible to construct linkage maps only for the intensively studied organisms, such as bacteria, yeast, or fruit flies, in which some visible mutations were available as genetic markers. But since the 80s, the molecular markers advent allows showing the polymorphism at the DNA level has made possible the establishment of genetic maps based on molecular markers for many organisms including: human (White et al., 1985; Schumm et al., 1988; Dib et al., 1996; Kong et al., 2002), tomato (Helentjaris et al., 1986; Tanksley et al., 1992; DeVicente and Tanksley, 1993; Saliba-Colombani et al., 2000), maize (Helentjaris et al., 1986; Helentjaris, 1987; Hoisington and Coe, 1990), rice (McCouch et al., 1988; Causse et al., 1994; Harushima et al., 1998), diploid and hexaploid oat (O'Donoghue et al., 1992; Rayapati et al., 1994; O'Donoghue et al., 1995), barley (Graner et al., 1991; Heun et al., 1991; Kleinhofs et al., 1993; Backes et al., 1995), wheat related diploid (Gill et al., 1991; Lagudah et al., 1991; Boyko et al., 1999), hexaploid wheat (Chao et al., 1989; Devos et al., 1992; Xie et al., 1993; Deynze et al., 1995; Nelson et al., 1995a; Nelson et al., 1995b; Jia et al., 1996; Marino et al., 1996; Röder et

al., 1998; Korzun et al., 1999), tetraploid wheat (Nachit et al., 2001; Blanco et al., 1996; Blanco et al., 1998a; Korzun et al., 1999)...

The construction of genetic linkage maps relies on the choice of parental lines, segregating population, and markers to reveal polymorphism. Ease in saturating a molecular map is strongly dependent on the polymorphism revealed between the parental lines. The genetic maps can be constructed based on F₂, F₃ generations, backcrosses, Doubled Haploids or Recombined Inbred Lines (RILs). The use of RILs is more advantageous than the other kind of populations (Saliba-Colombani et al., 2000) due: 1) The RIL can be reproduced easily; 2) The experiments replicability over years and environments witch favoring the genetic analysis of quantitative traits; 3) low level of heterozygosity, dominant markers can be used with the same efficiency as the co-dominant ones; and 4) the several selfing generations increases the probability of recombination between genes. Then, the constructed maps will facilitate the fine mapping of regions around genes or QTL of interest. Moreover, the RIL populations can be used in an unlimited number of studies of different traits in different environments, allowing efficient comparative QTL mapping (Alonso-Blanco et al., 1998).

For durum wheat, the linkage map construction did not start until the end of the 90s. The first durum linkage map constructed with 113 molecular markers, essentially Restriction Fragment Length Polymorphisms (RFLP) markers was made by the segregation data collected from a population consisted of 65 Recombinant Inbred Lines (RILs) derived from an inter-specific cross and based on (Blanco et al., 1996), then enriched with more RFLP marker (Blanco et al., 1998a; Blanco et al., 1998b), Korzun et al., (1999) integrated Simple Sequence Repeats (SSR) into this genetic linkage map, later Lotti et al., (2000) integrates Amplified Fragment Length Polymorphisms (AFLP). This map was enriched with more RFLP, AFLP and SSR markers (Blanco et al., 2004). Other genetic linkage maps were developed by this team from other durum populations (Gadaleta et al., 2009; Marone et al., 2009; Verlotta et al., 2010; Blanco et al., 2011; Blanco et al., 2012; Russo et al., 2012). Later these genetic linkage maps were served to construct a high-density consensus map (Marone et al., 2012a). Another dense linkage maps consisted of 9040 of which 9007 were Single Nucleotide Polymorphisms (SNP) was developed to studying durum seed morphology (Russo et al., 2014). Also, the linkage map of Gadaleta et al., (2009) was enriched with SNPs and Diversity Arrays Technology (DArT) for yellow pigment content QTLs investigation in durum wheat (Colasuonno et al., 2014). Recently other dense genetic linkage map was developed from other population (Duilio × Avonlea) consisted of 5444 SNPs (Marcotuli et al., 2017) and 5134 SNPs (Mangini et al., 2020).

The second linkage map was established 5 years after the first one of Blanco et al., (1996), it was constructed by (Nachit et al., 2001). This linkage map was based on an F₉ durum population consisted of 110 RILs derived by single-seed descent from an intraspecific cross (ICD-MN91-0012) between Jennah Khetifa and Cham1 made in 1991 at the Tel Hadya research station (Aleppo province, Syria) by the CIMMYT/ICARDA durum breeding program for Mediterranean dryland, the linkage map was constructed with 306 molecular markers (138 RFLPs, 26 SSRs, 134 AFLPs, five SSPs, and three known genes “one pyruvate decarboxylase and two lipoxygenases”) and covered 3598 cM. Then, this linkage map was served to identify chromosomal position of genes controlling important traits: Lipoxygenase (Hessler et al., 2002) and polyphenol oxidase (Watanabe et al., 2006), also this linkage map was used for detection of QTLs controlling important agronomical trait like soil-penetrating ability of roots (Kubo et al., 2007). Later, 162 loci including 6 known genes, 37 dESTs, and 119 markers (53 RFLPs, 1 SSRs, 65 AFLPs) that were unlinked in the previous map (Nachit et al., 2001) have been merged with the 306 molecular markers previously mapped (Diab et al., 2007; Diab et al., 2008). During the same durum wheat-breeding program (CIMMYT/ICARDA durum breeding program for Mediterranean dryland), several other durum populations were developed (Belikh2/Omrabi5, Omrabi5/*Triticum dicoccoides* 600545, Lahn/Cham1, Haurani/Cham1...), some durum populations were served for genetic linkage map construction and to QTL of important complex traits detection (Elouafi and Nachit, 2004; Farouk et al., 2021), for detection of molecular markers associated with agronomically important traits (Dura et al., 2013; Dura et al., 2014), and for physiological and other studies (Kehel et al., 2010; Bort et al., 2014; Habash et al., 2014).

Later, other research teams developed new genetic linkage maps based on other durum populations (Maccaferri et al., 2008; Mantovani et al., 2008; Maccaferri et al., 2011; Maccaferri et al., 2012; Graziani et al., 2014) and enriched them (Van Poecke et al., 2013). Later, 13 genotyping data set from 13 different populations were assembled to construct a high-density consensus genetic linkage map (Maccaferri et al., 2015).

5.2. The importance of high-density genetic linkage in durum breeding programs

High-density genetic linkage maps constitute the crucial step in the QTL detection (Heidari et al., 2011). The introduction of Diversity Array Technology (DArT) markers (Mantovani et al., 2008; Peleg et al., 2008; Gadaleta et al., 2009; Verlotta et al., 2010; Blanco et al., 2011; Blanco et al., 2012; Marone et al., 2012a; Russo et al., 2012) and Single Nucleotide

Polymorphism (SNP) (Zhang et al., 2008; Trebbi et al., 2011; Van Poecke et al., 2013) allow the construction of saturated linkage maps. Recently, high-density linkage maps were constructed using DArT and SNP markers (Marcotuli et al., 2017; Mangini et al., 2020; Simeone et al., 2020; Ruan et al., 2020).

Several studies were aimed at the detection of QTL linked to grain yield (Quarrie et al., 2005; Marza et al., 2006; Kuchel et al., 2007; Krigwi et al., 2007; Maccaferri et al., 2008; Heidari et al., 2011; Bennett et al., 2012; Tura et al., 2019). When the QTLs are localized on the chromosome, this allows to the scientists to study deeper the genetic basis and to enhance their knowledge about the trait of interest.

In this way and with the aim of studying efficiently the grain yield underlying genetical and physiological mechanisms, we have conducted this study, it was aimed firstly at the construction of a high-density linkage map and at the identification of grain yield QTL using data from genotyping by SSRs and SNPs and phenotyping for grain yield of Lahn/Cham1 population. Then, the annotation of the SNPs localized in the regions of the QTLs identified and the dissection of the four QTLs detected on the chromosome 2A using the bioinformatic tools and public databases cited above.

Detection of grain yield QTLs
in the durum population
Lahn/Cham1 tested in
contrasting environments

Chapter II: Detection of grain yield QTLs in the durum population Lahn/Cham1 tested in contrasting environments

1. Introduction

The complex traits such as grain yield is linked to multiple quantitative trait loci (QTL) and to their interactions with the environment. For durum breeders, the phenotyping of grain yield, the genotyping, and the construction of a genetic map constitute a crucial step in the QTL detection (Heidari et al., 2011). In this context, the Lahn/Cham1 population was used in this work, it consists of 112 (RIL) derived from a single seed descent selection from the cross ICD-MN91-015 between the varieties Lahn and Cham1. The cross was made at CIMMYT/ICARDA durum breeding program for Mediterranean dryland, was phenotyped for grain yield across 11 contrasting environments and genotyped with mainly SSR and SNPs markers to construct the genetic linkage map and then, identify the QTLs linked to grain yield.

The choice of Lahn and Cham1 parent was due to their wide cultivation in the Mediterranean basin and to their high performance. Lahn is high productive in the favorable environments and Cham1 is wide adapted, yield stable, and drought tolerant (Nachit et al., 2001; Kehel et al., 2010; Habash et al., 2014). Moreover, Lahn is resistant to yellow rust (*Puccinia striiformis*) and moderately resistant to leaf rust (*Puccinia recondita*), powdery mildew (*Erysiphe graminis*), common bunt (*Tilletia caries*), and to Septoria leaf blotch (*Mycosphaerella graminicola*). Also, Lahn has a large kernel size with superior hectoliter weight, milling extraction index, and gluten strength values. As for Cham1, also originated from the CIMMYT/ICARDA durum program, adapted to the Mediterranean climate, and has high yield stability. It exhibits an excellent resistance to drought and yellow rust with some tolerance to Russian wheat aphid. In addition, Cham1 also has high values for osmotic adjustment and high yellow pigment content, but its gluten strength is weak. However, Cham1 is susceptible to leaf rust, stem rust, Septoria tritici, and powdery mildew. This variety has a wide adaptation.

2. Materials and methods

2.1. Experimental design

The trials were run in 11 contrasting environments using the Augmented Design (AD), according to Federer (1956) and Kehel et al. (2010), with 5 checks (Omrabi5, Haurani, Korifla, Waha, and Gidara2). The AD is based on randomized complete block design (RCBD) with 6 blocks. Each trial had 144 plots arranged as a grid layout of 12 rows by 12 columns. The GY is

harvested from 5 m² plot and estimated as the harvested grain's weight in kilogram (kg) per hectare (ha).

2.2. Population genotyping

The population genotyping was conducted at the ICARDA durum molecular breeding laboratory. The Lahn/Cham1 population was genotyped with SSR markers and 9400 SNP. Then, the genetic linkage map was constructed with 247 molecular markers: 216 SSR, 14 Expressed Sequence Repeats (EST-SSR), 10 AFLP, 6 Seed Storage Proteins (Gliadin (Gli) and glutenin (Glu)), and 1 Pseudogliadine gene. Thereafter, we have enriched the SSR map with 1425 polymorphic SNP (Su et al., 2018).

2.3. Linkage map construction and QTLs detection

We used the QTL IciMapping software V 4.1.0.0 (Wang et al., 2016) for mapping. Kosambi mapping function (Kosambi, 1944) was used to transform the recombination frequencies into centiMorgan (cM) distances.

The QTL IciMapping software V 4.1.0.0 was also served for the QTLs detection; the scanning was performed using the model “Inclusive Composite Interval Mapping Additive (ICIM-ADD)” through stepwise regression according to Li et al. (2007). The walking speed was 1.0 cM, and the stepwise regression probability 0.01. We used a LOD threshold of 3 with 1000 permutations and an error of 0.05. The QTL scan was performed using two approaches: the first one was a combined QTL analysis across all environments to identify QTLs with additive-by-environment (A by E) interaction effects named also Multi-Environment Trials (MET analysis), the second one was in single/specific environmental analysis named (Biparental analysis, the scan was performed for each one of the 11 environments separately.

3. Results

3.1. RIL performance

The results show that in the irrigated and favorable environments, the parent's grain yield performance is generally above the mean grain yields of the RILs mean. Whereas in the average yielding environments of 4000-5000 kg/ha (02RF, 02INC) and low yielding environments of 2000-2500 kg/ha (02BR), the parent's grain yield was approximately similar to that of the mean grain yields of the RILs. Table 2 (article 1) shows statistical parameters, including the trial heritability values. Further, in all environments, several RILs have out-yielded both parents. The highest RIL grain yield means were achieved in the irrigated environments (01IR), the

early sowing environment with supplementary irrigation (02EP), and the favorable and high input environment (02GH).

3.2. Linkage map construction

The genetic linkage map covered 6122.22 cM. The A-chromosomes had 2902.61 cM (47.41%) and harbor 117 SSR markers (47.36%) and 598 SNP (41.96%) with an average distance between adjacent loci of 4.06 cM. whereas the B chromosomes had 3219.61 cM (52.59%) and harbor 130 SSR markers (52.64%) and 827 SNP (58.04%) with an average distance between adjacent loci of 3.36 cM. The highest number of SNP mapped was on 2B with 205 SNP. Whereas the lowest molecular markers number mapped was on 7B with 23 SNP markers.

For the group of A chromosomes, 2A harbored the largest number of markers (152) with the highest SNP mapping number (133). It covered 502.84 cM and it had of 3.31 cM per marker. However, the 6A was with 746.24 cM and 96 markers is the largest and the least dense chromosome (7.77 cM/marker) in all durum genome. The densest chromosome in A chromosomes is 1A (3.17cM/marker). The least enriched chromosome in the A chromosomes with SNP markers was the chromosome 5A (24 SNP). The entire A-chromosomes covering had 715 markers with a map size of 2902.61 cM and an average density of 4.06 cM/marker.

As for the group of B chromosomes, 2B had the highest number of molecular markers (226). It also had the highest number of SNP (205) mapped not only in B genome but in the entire durum chromosomes. It was relatively saturated (3.07 cM/marker). The lowest coverage in B-chromosomes and in all durum chromosomes with 23 SNP was the 7B chromosome. The 6B was the most densely saturated chromosome (2.85 cM/marker) with 350.54 cM size. Further, the total number of molecular markers harbored by the B chromosomes was 957 with a size of 3219.61 cM. the average, the average density of 3.36 cM per marker. There were more SNP mapped in the B chromosomes (827) than in the A-chromosomes (598). The B chromosome map size was larger than that for the A-chromosomes: 3219.61 vs. 2902.61 cM. Similarly, for markers density, B chromosomes were denser (3.66 cM/marker) than A chromosomes (4.06 cM/marker).

3.3. QTL detection:

3.3.1. In specific environment analysis

Three QTLs were detected in the rainfall-favorable environment (01GH and 02GH), and two in the irrigated and hot environment (02RQ). Further, four QTLs were detected in the cold favorable (02EP) and in the dry environments (02INC and 02RF) see Figure 6.

3.3.2. In Multi-Environmental QTL mapping

A total of 126 QTLs was detected covering all chromosomes were detected for GY. The A-chromosomes contained 53.17% of the total QTLs detected. The chromosomes 2A, 2B, and 6A harbored the highest QTLs number: 13, 18, and 13, respectively. The GY-QTL distribution and their flanking markers are reported in the Figure 6. Further, the A chromosomes QTL contributions were higher than in B chromosomes 67 vs. 59. Moreover, the contribution of Lahn parent was 74 QTLs, of which 40 originated from A chromosomes and 34 from B chromosomes, while Cham1 contributed with 52 QTLs, 29 of which were originated from A chromosomes and 23 from B chromosomes (Figure 6).

1A chromosome

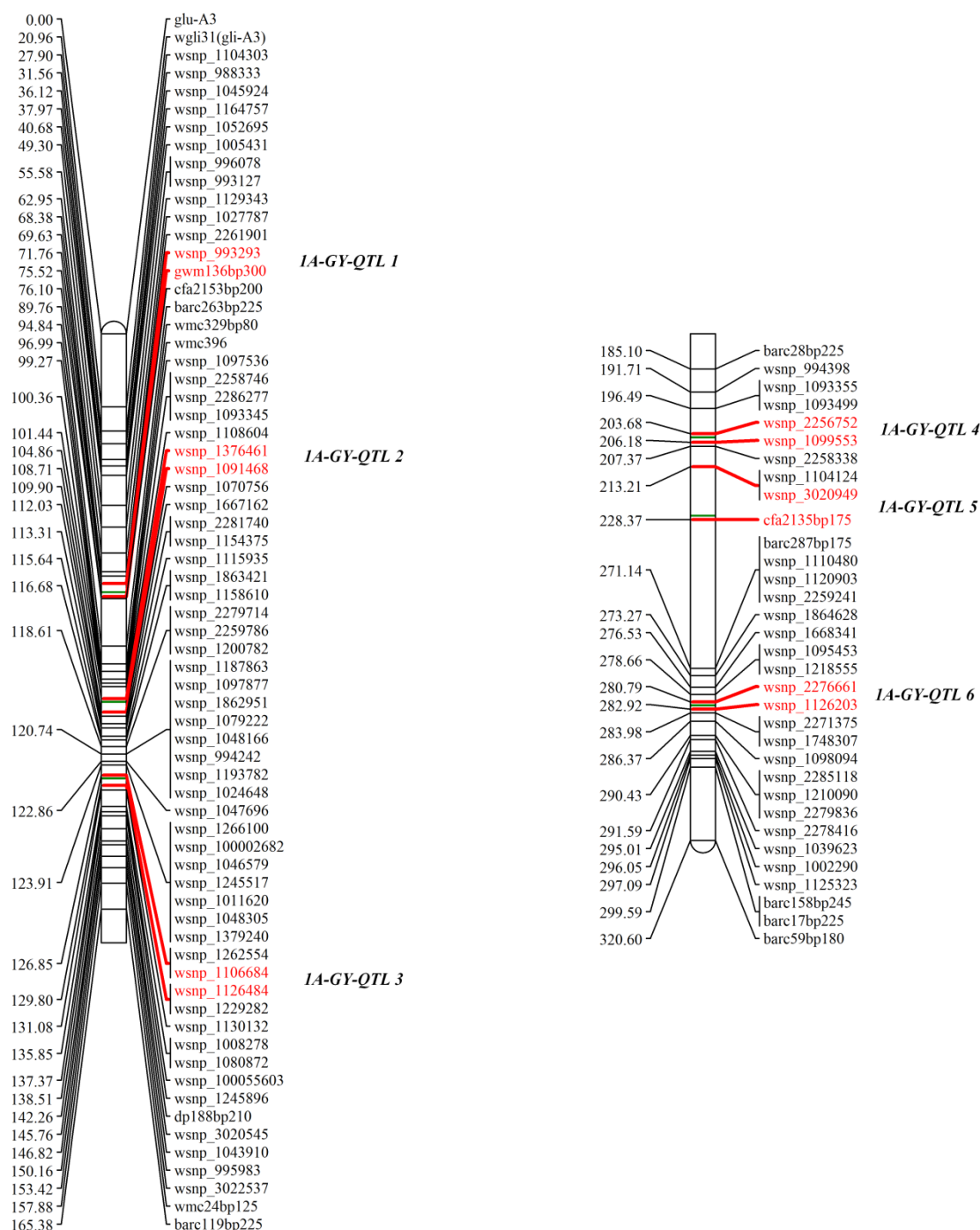


Figure 6: Durum Lahn/Cham1 genetic linkage map with the grain yield QTLs detected using MET (mentioned in black color) and BIP (mentioned in different colors) approaches

1B chromosome

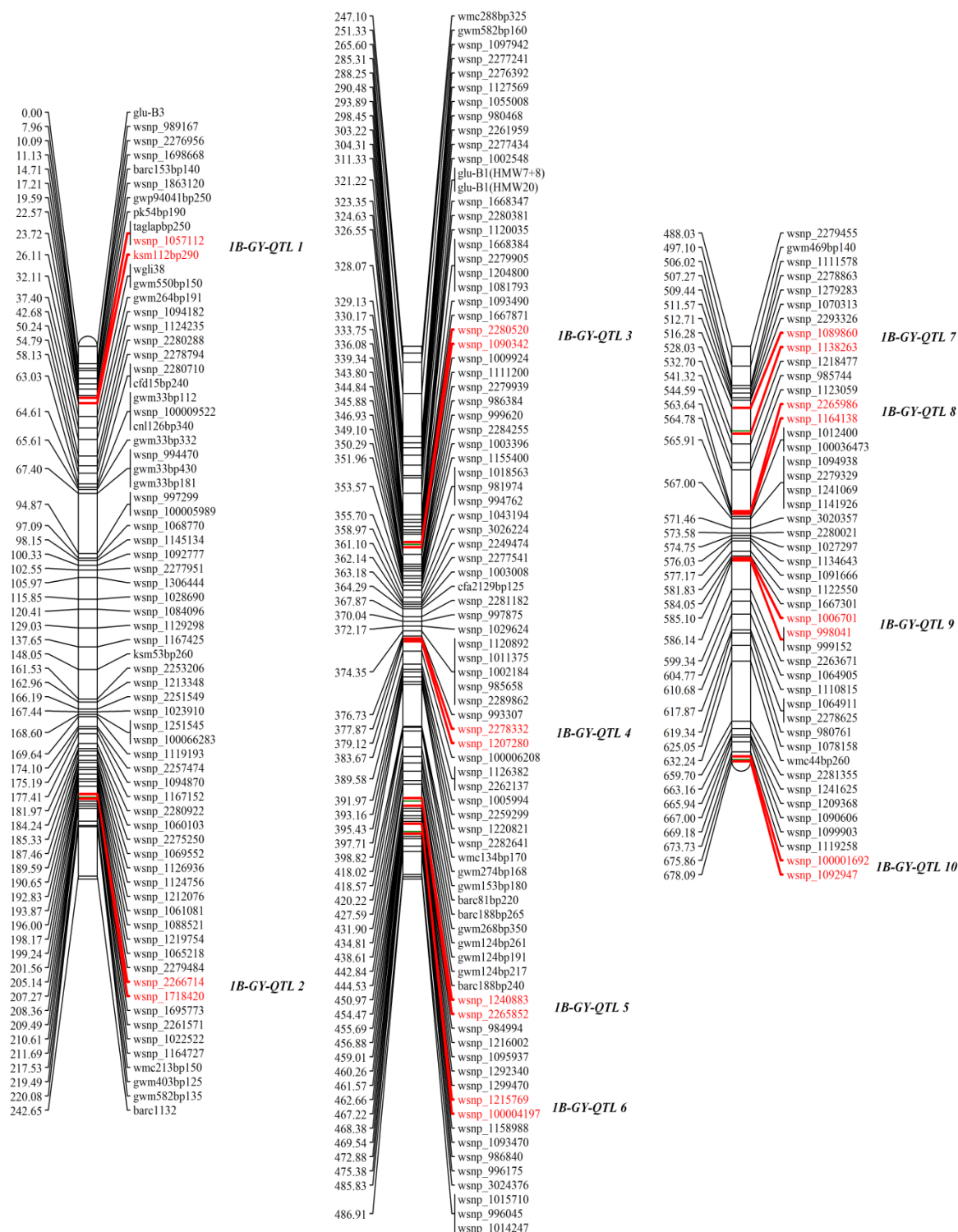


Figure 6: (Continued.)

2A chromosome

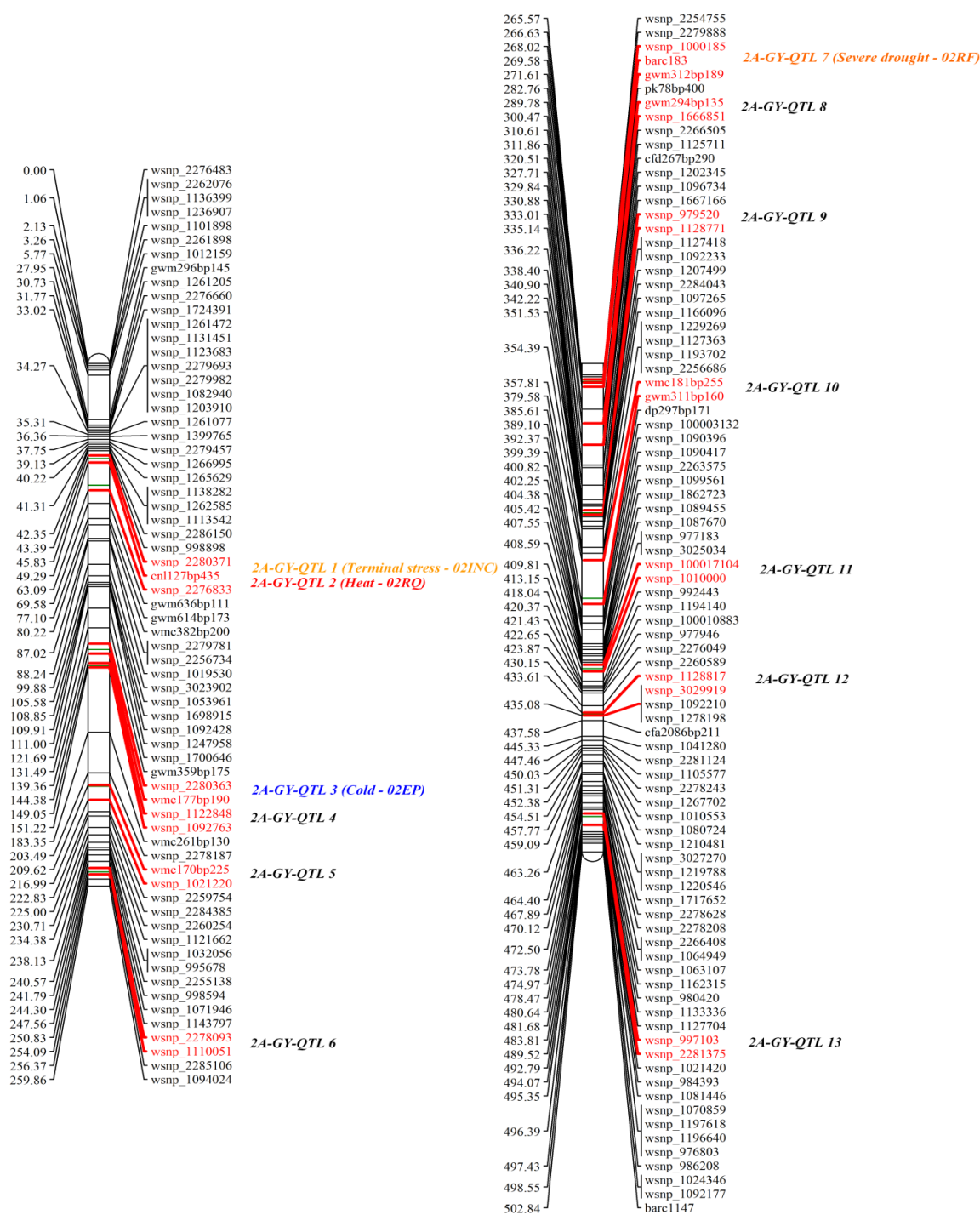


Figure 6: (Continued.)

2B chromosome

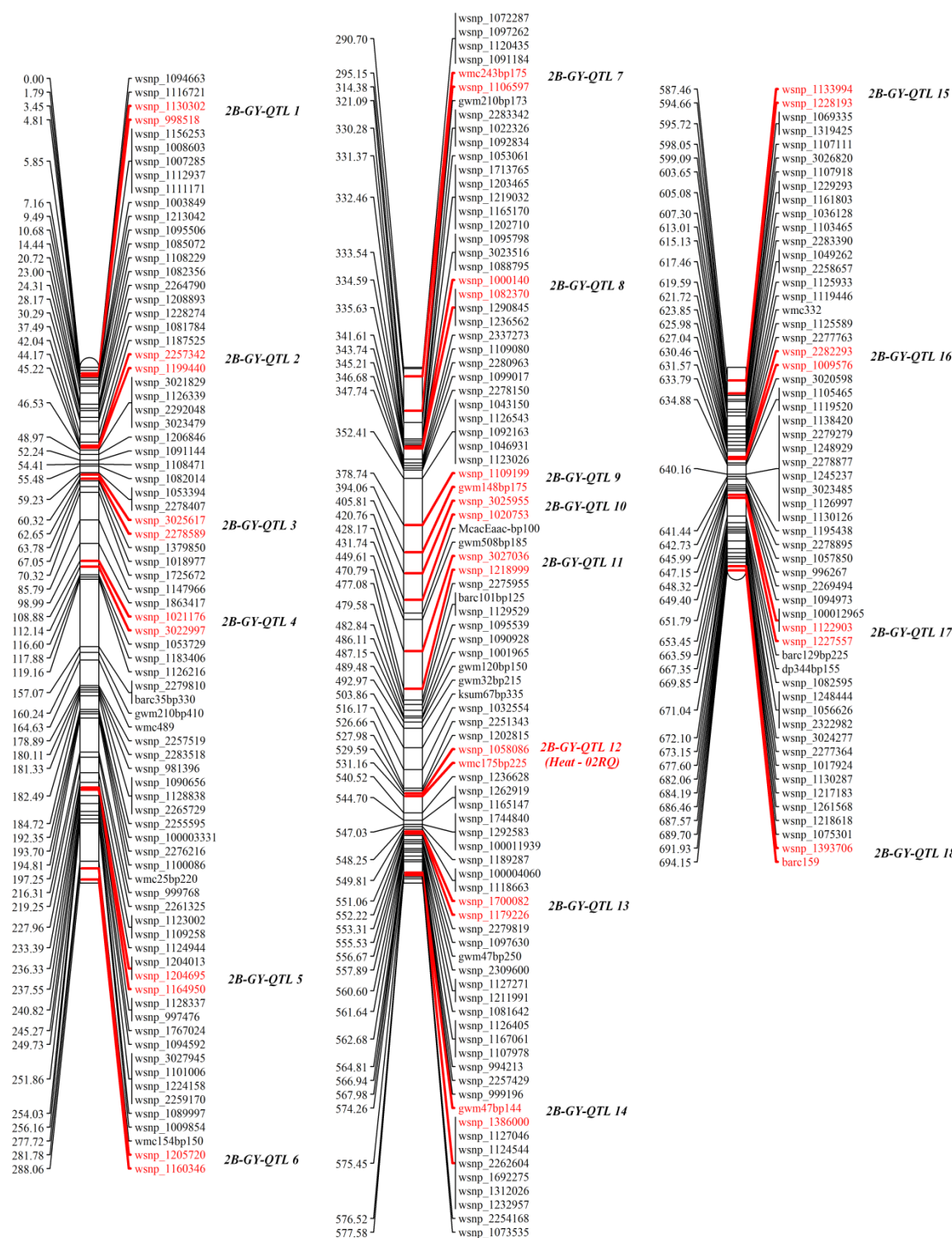


Figure 6: (Continued.)

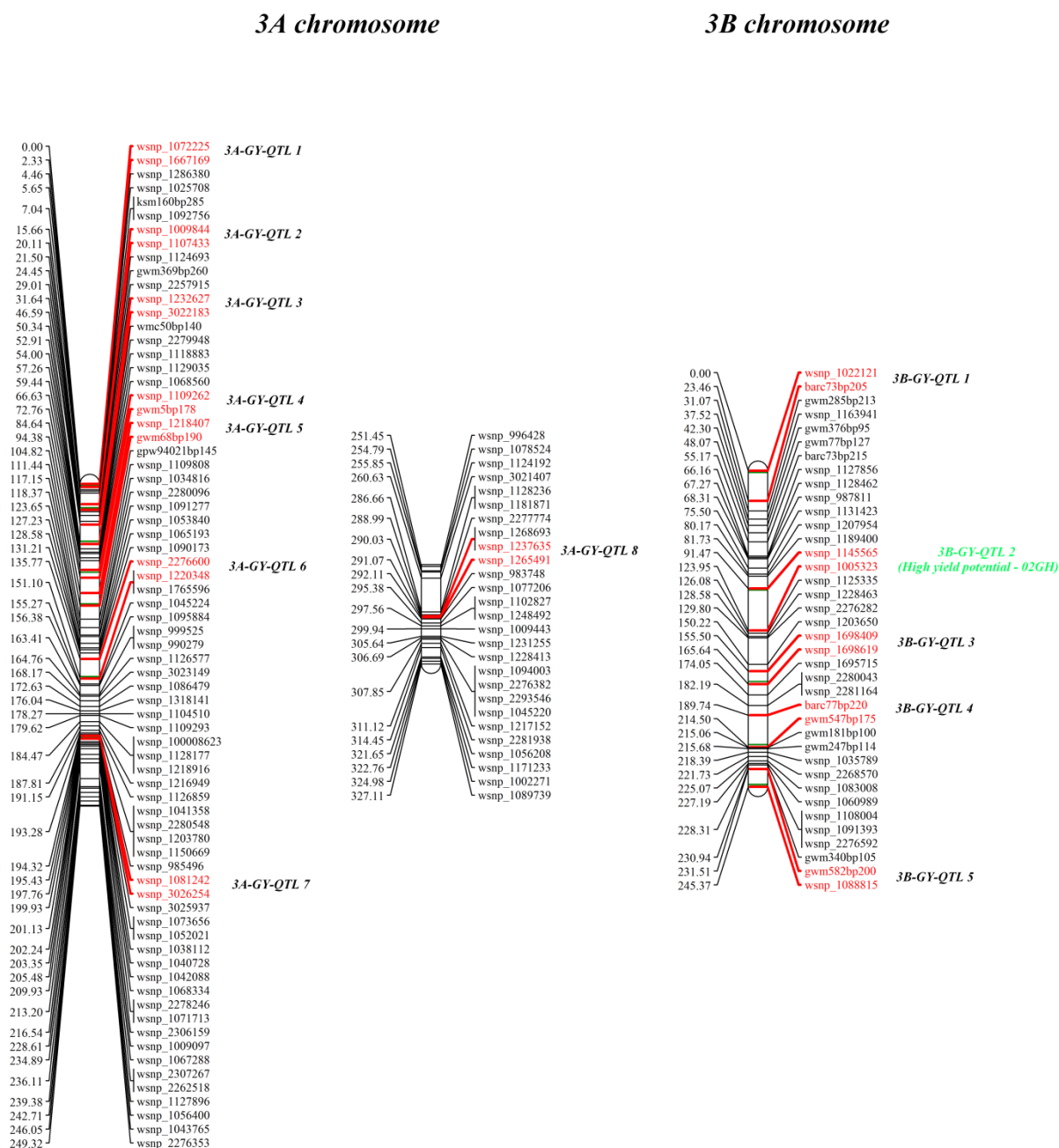


Figure 6: (Continued.)

4A chromosome

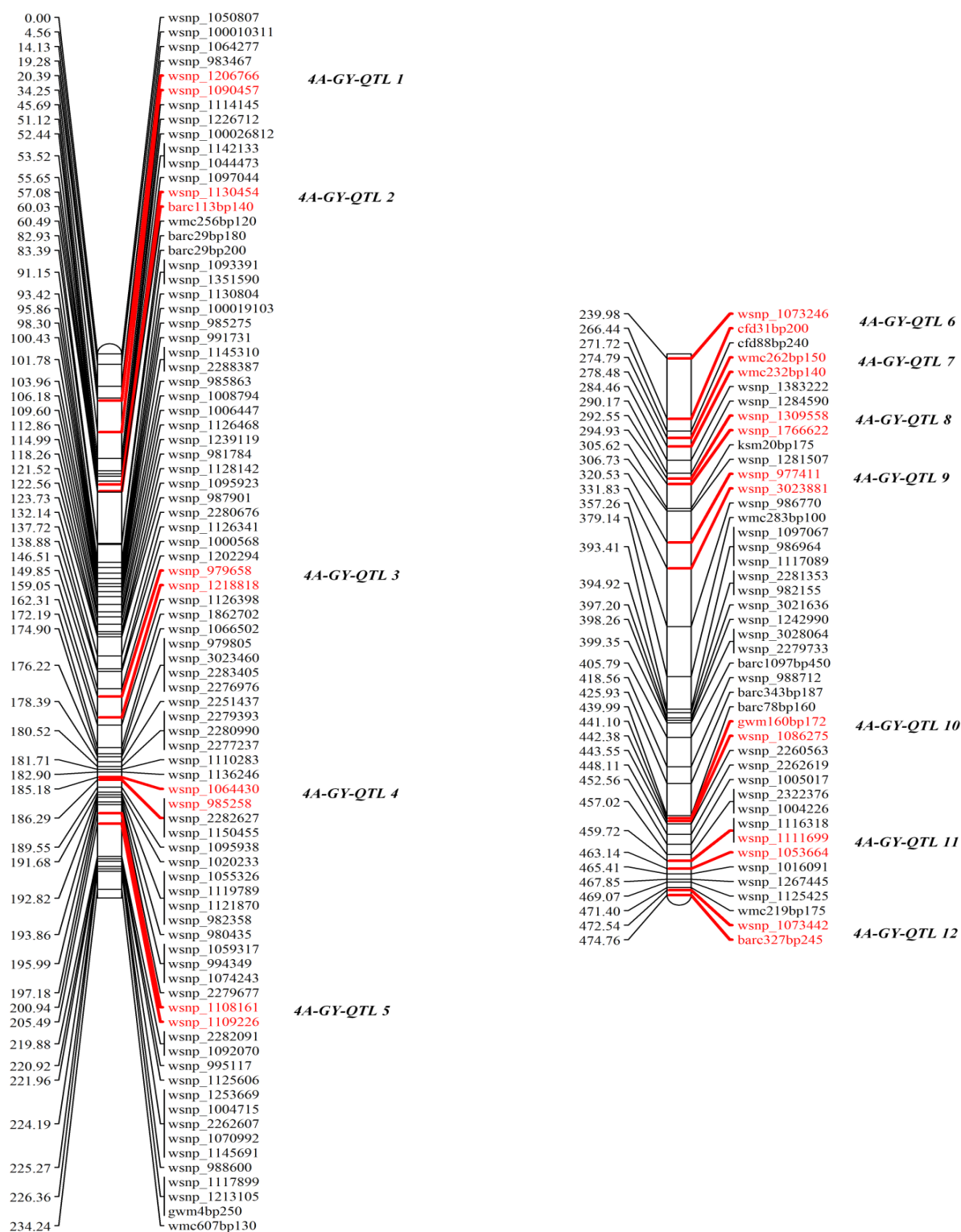


Figure 6: (Continued.)

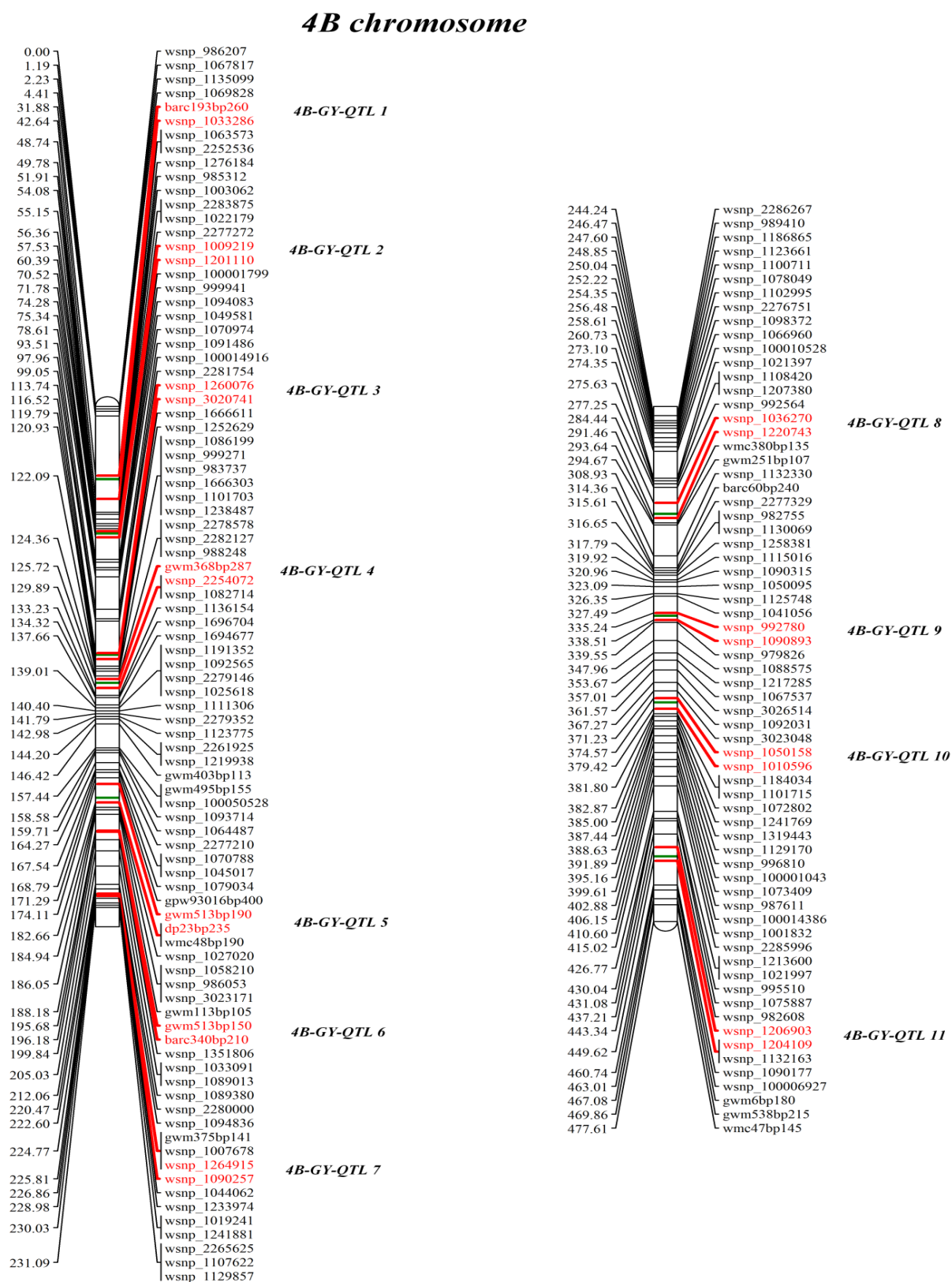


Figure 6: (Continued.)

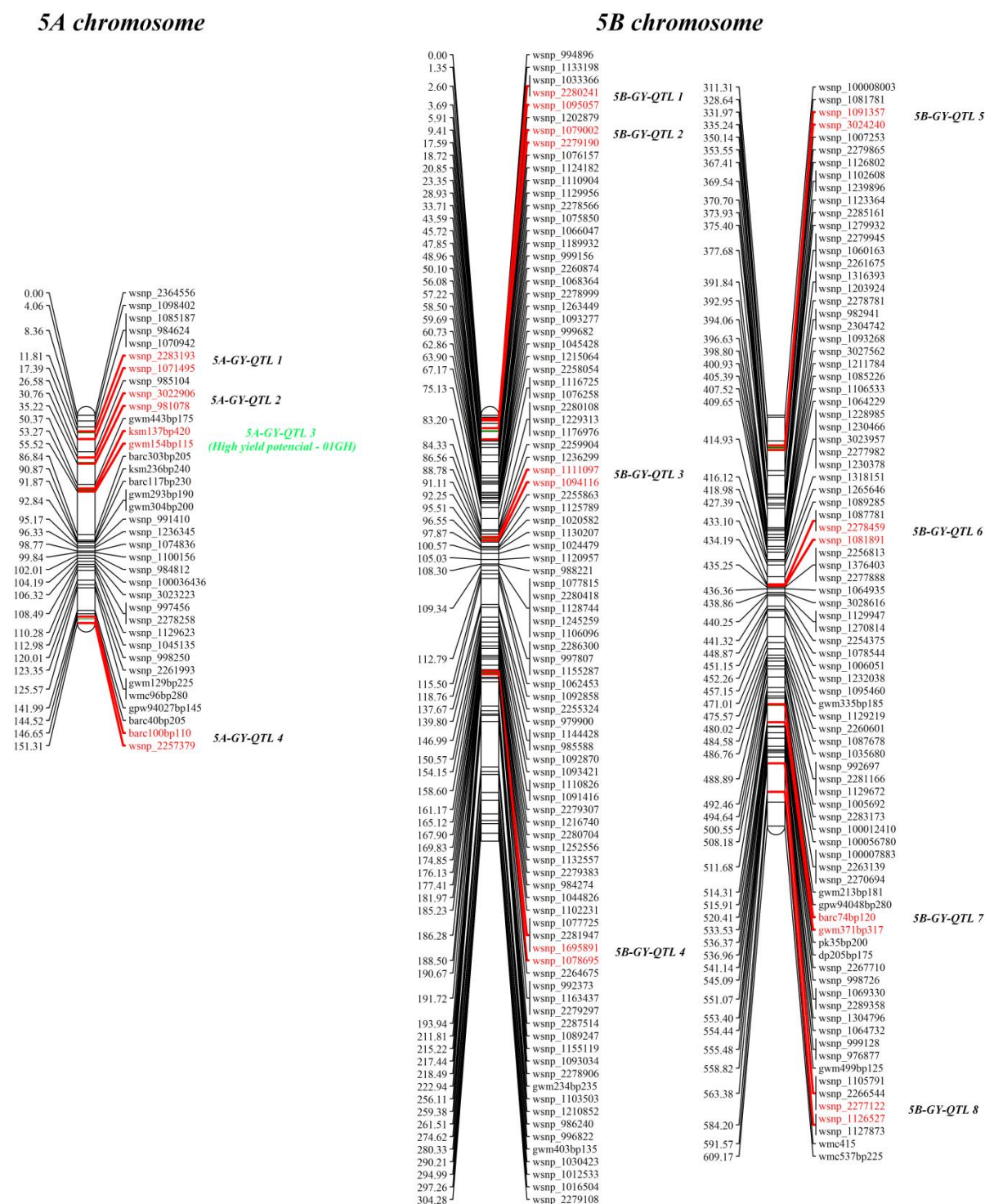


Figure 6: (Continued.)

6A chromosome

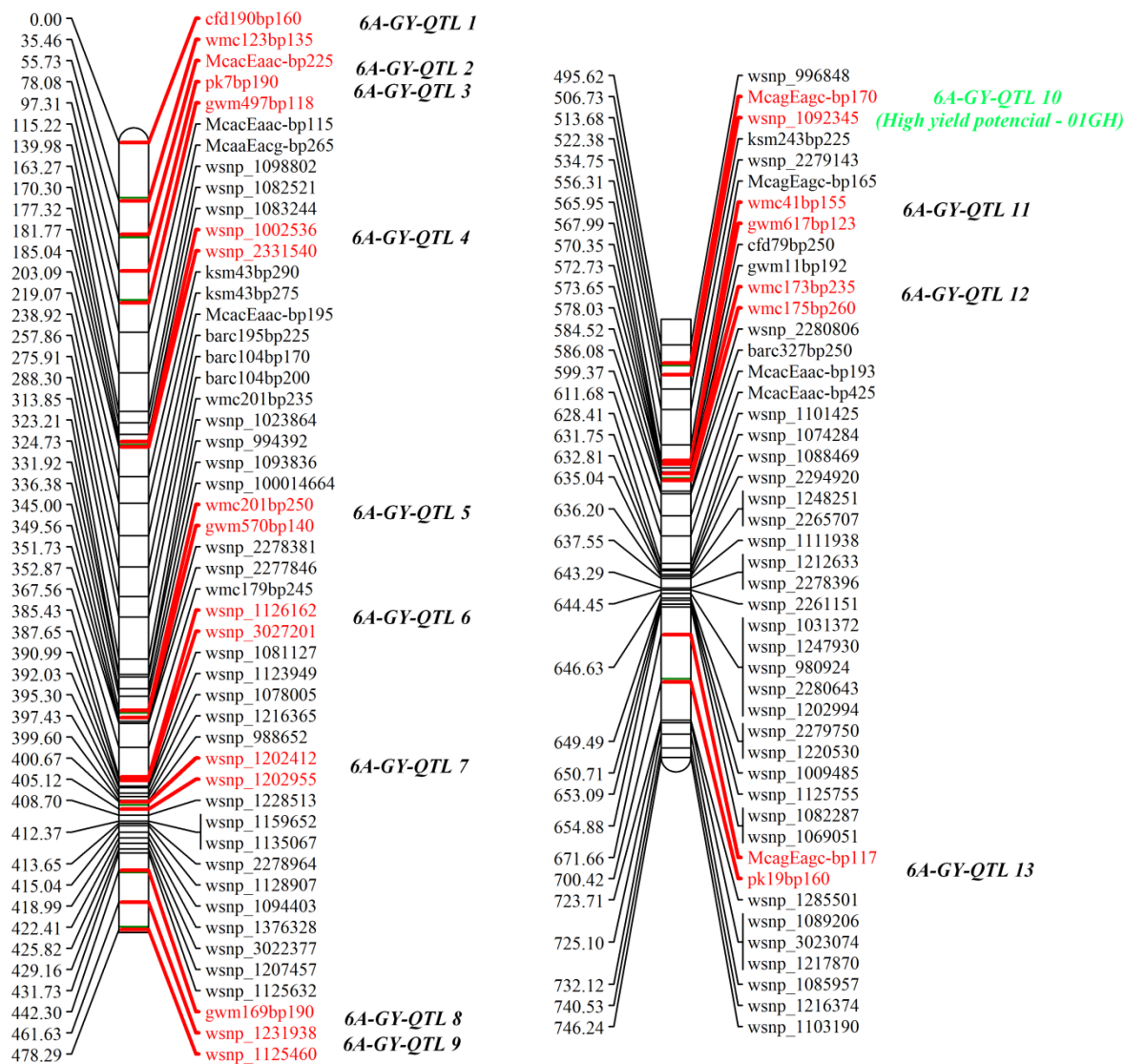


Figure 6: (Continued.)

6B chromosome

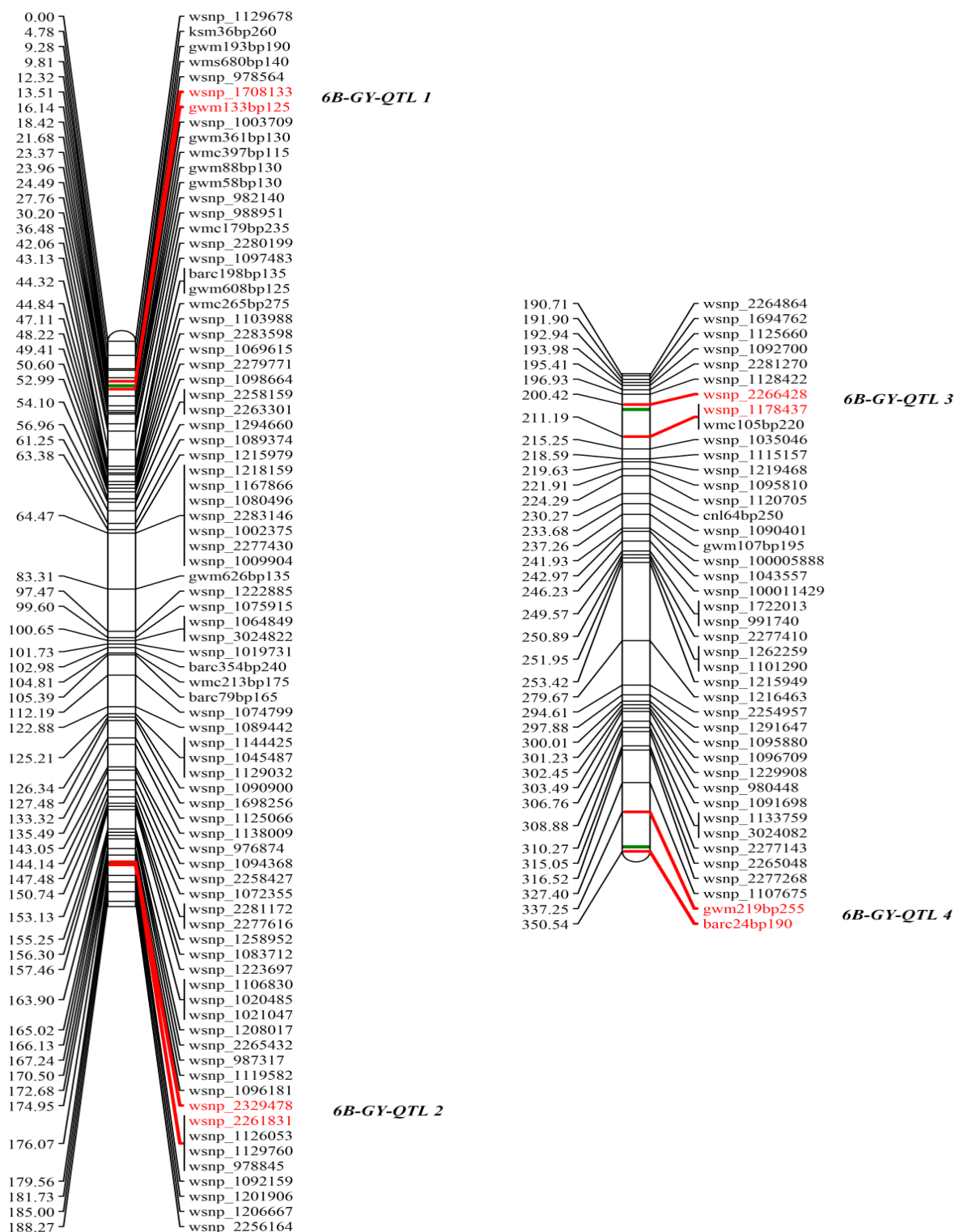


Figure 6: (Continued.)

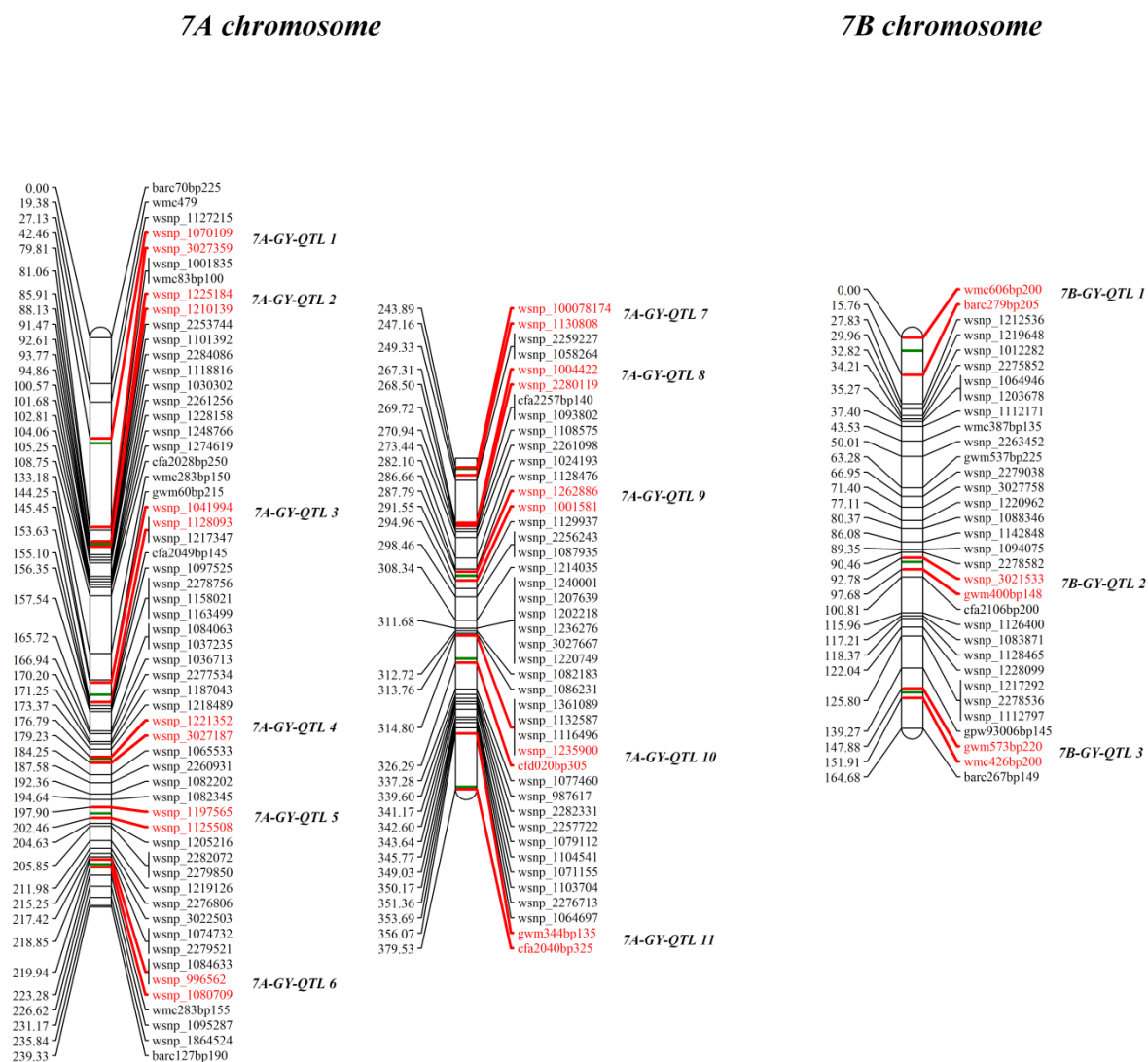


Figure 6: (Continued.)



Detection of grain yield QTLs in the durum population Lahn/Cham1 tested in contrasting environments

Issame FAROUK¹, Ahmad ALSALEH², Jihan MOTOWAJ³, Fatima GABOUN⁴, Bouchra BELKADI¹, Abdelkarim FILALI MALTOUF¹, Zakaria KEHEL³, Ismahane ELOUAFI⁵, Nasserelhaq NSARELLAH⁴, M. Miloudi NACHIT^{3,*}

¹Laboratory of Microbiology and Molecular Biology, Department of Biology, Med V University, Rabat, Morocco

²Department of Science and Technology Bozok University, Yozgat, Turkey

³ICARDA, The International Center for Agricultural Research in the Dry Areas, Rabat, Morocco

⁴INRA, National Institute of Agronomical Research, Unity of Biotechnology Research, Rabat, Morocco

⁵ICBA, International Center for Biosaline Agriculture, Dubai, United Arab Emirates

Received: 21.08.2020 • Accepted/Published Online: 19.11.2020 • Final Version: 09.02.2021

Abstract: Durum wheat (*Triticum turgidum* L. var durum) is tetraploid wheat (AABB); it is the main source of semolina and other pasta products. Grain yield in wheat is quantitatively inherited and influenced by the environment. The genetic map construction constitutes the essential step in identifying quantitative trait loci (QTLs) linked to complex traits, such as grain yield. The study aimed to construct a genetic linkage map of two parents that are widely grown durum cultivars (Lahn and Cham1) in the Mediterranean basin, which is characterized by varying climate changes. The genetic linkage map of Lahn/Cham1 population consisted of 112 recombinant inbred lines (RILs) and was used to determine QTLs linked to the grain yield in 11 contrasting environments (favorable, cold, dry, and hot). Simple sequence repeat (SSR) molecular markers were used to construct an anchor map, which was later enriched with single nucleotide polymorphisms (SNPs). The map was constructed with 247 SSRs and enriched with 1425 SNPs. The map covered 6122.22 cM. One hundred and twenty-six QTLs were detected on different chromosomes. Chromosomes 2A and 4B harbored the most significant grain yield QTLs. Furthermore, by comparison with several wheat mapping populations, all the A and B chromosomes of Lahn/Cham1 QTLs contributed to grain yield. The results showed that the detected QTLs can be used as a potential candidate for marker-assisted selection in durum breeding programs.

Key words: Molecular markers, grain yield, genetic linkage map, SNP, QTL

1. Introduction

Durum wheat originated through merging of two genomes (AABB) through natural interspecific hybridization and genome duplication (Zhang et al., 2012). Durum wheat (*Triticum turgidum* L. var durum) is the primary source of semolina for the production of pasta and couscous in the Mediterranean basin, and its cracked grain for burghul in the Middle East (Elouafi and Nachit, 2004). Moreover, durum wheat is grown mainly in drylands and is affected mainly by year-to-year environmental variation, drought, cold, and heat stresses (Nachit, 1992; El Hassouni et al., 2019).

Grain yield is quantitatively inherited and linked to multiple quantitative trait loci (QTLs) and their interactions with the environment. QTLs are DNA regions associated with a specific trait (Oduola et al., 2003). The

phenotyping of grain yield and the use of molecular markers for the genotyping and the construction of a genetic map constitute a crucial step in QTL detection (Heidari et al., 2011). Molecular linkage maps have been constructed for durum wheat by several geneticists (Blanco et al., 1996, 1998; Korzun et al., 1999; Nachit et al., 2001; Elouafi and Nachit, 2004; Hail et al., 2012; Nagel et al., 2014; Alsaleh et al., 2015). The first durum map was made by the interspecific cross and was based on 65 recombinant inbred lines (RILs) and restriction fragment length polymorphism (RFLP) markers (Blanco et al., 1996, 1998), and later simple sequence repeats (SSR) were integrated into this genetic map (Korzun et al., 1999). The second durum map was based on an intraspecific cross Jennah Khetifa/Cham1 with 110 RILs using RFLP, SSR, amplified fragment length polymorphism (AFLP), seed

* Correspondence: mmnachit@gmail.com



storage proteins, and genes (Nachit et al., 2001). Several other studies on durum genetic mappings were conducted (Maccaferri et al., 2008, 2010; Haile et al., 2012; Dura et al., 2013). Moreover, single nucleotide polymorphisms (SNPs) have been included in durum wheat genetic maps (Poecke et al., 2013). Some researchers have integrated a significant number of SNP in bread and durum wheat genetic linkage maps (Maccaferri et al., 2015; Su et al., 2018; Zeng et al., 2019). Many QTL studies for different traits have been conducted for wheat (Tixier et al., 1998; Varshney et al., 2000; Sourdille et al., 2003; Elouafi and Nachit, 2004; Li et al., 2007; Pozniak et al., 2007; Maccaferri et al., 2008; Zhang et al., 2008; Bennett et al., 2012; Haile et al., 2012; Zhang et al., 2016; Tura et al., 2019; Xu et al., 2019; Zaim et al., 2020).

In this study, the Lahn/Cham1 population was phenotyped for grain yield across 11 contrasting environments and was genotyped with mainly SSR and SNP markers to construct the genetic linkage map and to identify the QTLs linked to grain yield. The choice of Lahn and Cham1 as parents was due to their wide cultivation in the Mediterranean basin. Moreover, Lahn is highly productive in favorable environments and Cham1 is widely adapted, yield-stable, and drought-tolerant (Nachit et al., 2001; Kehel et al., 2010; Habash et al., 2014). We have provided the Lahn/Cham1 population to Barcelona University and Tunisia to be phenotyped for stable carbon isotope discrimination and other physiological traits (Bort et al., 2014). Moreover, the current study is the first on the genetic linkage mapping and QTL detection for Lahn/Cham1. (Lahn “Improved durum variety” was wrongly quoted as Jennah Khetifa, landrace in the paper published by Bort et al. (2014)).

The study's aims were: 1) to construct and enrich the durum genetic linkage map of the Lahn/Cham1 population with SNP markers, and 2) to determine the QTLs linked to the grain yield in different contrasting environments.

2. Materials and methods

2.1. Plant materials

The studied durum population consisted of 112 F_{12} lines derived from a single-seed descent selection from the cross ICD-MN91-015 between the varieties Lahn and Cham1. The cross was made at CIMMYT/ICARDA durum breeding program for Mediterranean drylands. Lahn is an improved durum wheat variety developed for favorable and high-input environments with cold tolerance. It is also resistant to yellow rust (*Puccinia striiformis*) and moderately resistant to leaf rust (*Puccinia recondita*), powdery mildew (*Erysiphe graminis*), common bunt (*Tilletia caries*), and to Septoria leaf blotch (*Mycosphaerella graminicola*). Furthermore, Lahn has a large kernel size with

superior hectoliter weight, milling extraction index, and gluten strength values. As for Cham1, the second parent of the population, it is also originated from the CIMMYT/ICARDA durum program, adapted to the Mediterranean climate, and has high yield stability. It exhibits an excellent resistance to drought and yellow rust with some tolerance to Russian wheat aphid. Cham1 also has high values for osmotic adjustment and high yellow pigment content, but its gluten strength is weak. However, Cham1 is susceptible to leaf rust, stem rust, *Septoria tritici*, and powdery mildew. This variety has a wide adaptation; it has been released in several countries in the Mediterranean basin and outside (Turkey, Syria, Cyprus, Iraq, Algeria, Portugal, Sudan, etc.) under different names.

2.2. Experimental field design

The Lahn/Cham1 population trial consisted of 112 recombinant inbred lines (RILs) and the two parents. For the experimental design, we used the augmented design (AD), according to Federer (1956) and Kehel et al. (2010), with 5 checks (Omrabi5, Haurani, Korifla, Waha, and Gidara2). The AD is based on randomized complete block design (RCBD) with 6 blocks. Each trial had 144 plots arranged as a grid layout of 12 rows by 12 columns. The Lahn/Cham1 population was grown in the field for grain yield (GY) evaluation in contrasting environments (Nachit and Elouafi, 2004; Nachit et al., 2016). The GY is harvested from 5 m² plot and estimated as the harvested grain's weight in kilogram per hectare. The testing sites description is shown in Table 1 with their locations, growing seasons, crop rotations, supplementary irrigation, and temperatures during the trial's tillering and flowering stages.

2.3. Population genotyping

The genotyping of the population was conducted at the ICARDA durum molecular breeding laboratory. The Lahn/Cham1 population was genotyped with SSR markers and 1500 polymorphic SNPs selected from 9400 SNP platform. The protocol for DNA extraction, molecular assay for microsatellites (SSR), AFLP, and seed storage proteins was conducted as described by Nachit et al. (2001). The genotyping by sequencing of RILs was performed using a whole-genome profiling service for SNP and diversity array technology sequencing (DarTseq **markers**). One hundred microliters of 50 ng μL^{-1} was analyzed using Diversity Array Technology¹, as described by Akbari et al. (2006).

2.4. Genetic linkage map construction

The genetic linkage map was constructed with 247 molecular markers: 216 SSR, 14 expressed sequence repeats (EST-SSR), 10 AFLP, 6 seed storage proteins

¹<http://www.diversityarrays.com>

Table 1. Testing sites, environments (crop rotation) with growing seasons, sowing dates, annual rainfall, added irrigations, and mean temperatures at tillering and flowering stages for the field trials of Lahn/Cham1 population.

Sites	Environment	Season	Sowing date	Rainfall (mm)	Irrigation (mm)	Mean temperature (°C) at	
						Tillering (Zadoks scale 29)	Flowering (Zadoks scale 69)
Tel Hadya/Syria (ICARDA main station) (36°0.1'N ^l 36°56'E ^m , 284m)	02EP ^a (lentil)	01/02	Mid-October	330	70 at sowing	2	20
	01IR ^b (vetch above biomass harvested)	00/01	Mid-November	330	100 (50 at early tillering & 50 at booting)	8	24
	01INC ^c (vetch biomass incorporated in soil)	00/01	Idem	330	None	8	24
	02INC ^d (vetch biomass incorporated in soil)	01/02	Idem	330	None	8	24
	02RF ^e (vetch above biomass harvested)	01/02	Idem	330	None	8	24
Breda/Syria (dry) (35°0'N ^l 38°0'E ^m , 300m)	02BR ^f (fallow)	01/02	Idem	260	None	6	27
Ghab/Syria (high rainfall) (35°38'N 36°14'E ^m , 280m)	01GH ^g (faba beans)	00/01	Idem	650	None	5	20
	02GH ^h (faba beans)	01/02	Idem	614	None	5	20
Raqqqa/Syria (irrigated) (35°57' N ^l 39°0'E ^m , 295m)	01RQ ⁱ (chickpea)	00/01	Idem	150	450 (150 at sowing, 150 at tillering & 150 at flowering)	10	30
	02RQ ^j (chickpea)	01/02	Idem	150	Idem	10	30
Terbol/Lebanon (high & favorable rainfall) (33°49'N ^l 35°58'E ^m , 890m)	01TR ^k (faba beans)	00/01	Idem	550	None	2	20

^a02EP: Early planting in 2001/2002; ^b01IR: Irrigated in 2000/2001; ^c01INC: Incorporated in the field in 2000/2001; ^d02INC: Incorporated in the field in 2001/2002; ^e02RF: Rainfed in 2001/2002; ^f02BR: Breda station in 2001/2002; ^g01GH: Ghab valley in 2000/2001; ^h02GH: Ghab valley in 2001/2002; ⁱ01RQ: Raqqqa station in 2000/2001; ^j02RQ: Raqqqa station in 2001/2002; ^k01TR: Terbol station in 2000/2001; ^lN: latitude in North; ^mE: longitude in Est and m: altitude in meters.

(gliadin (Gli) and glutenin (Glu)), and 1 Pseudogliadin gene. Furthermore, we enriched the SSR map with 1425 polymorphic SNPs (Su et al., 2018); the remaining 75 SNPs were unlinked. We used the QTL IciMapping software V 4.1.0.0 (Wang et al., 2016) for mapping. Kosambi mapping function (Kosambi, 1944) was used to transform the recombination frequencies into centiMorgan (cM) distances.

2.5. Grain yield QTL detection

QTL IciMapping software V 4.1.0.0 (Wang et al., 2016) was used to detect the QTLs; the scanning was performed by the model “Inclusive Composite Interval Mapping Additive (ICIM-ADD)” through stepwise regression according to Li et al. (2007). The walking speed for all QTLs was 1.0 cM, and the stepwise regression probability was 0.01. We used a LOD threshold of 3 with 1000 permutations and an error

of 0.05. Combined QTL analysis across all environments was conducted to identify QTLs with additive-by-environment (A by E) interaction effects, using the same parameters and the threshold of LOD 3 (Li et al., 2007) for the ICIM method of QTL analysis.

3. Results

3.1. Trial performance across the contrasting environments

Table 2 shows RIL mean, maximum, and minimum yields along the population parents Lahn (parent 1) and Cham1 (parent 2). In the irrigated and favorable environments, the parent's grain yield performance is above the mean grain yields of the RILs mean. In the average yielding environments of 4000–5000 kg/ha (02RF, 02INC) and low yielding environments of 2000–2500 kg/ha (02BR),

Table 2. Statistical parameters of grain yield (kg/ha) for RIL mean, maximum, and minimum yielders (kg/ha), and parents across testing environment.

Parameter	02RF	01IR	02EP	01INC	02INC	02BR	01GH	02GH	01RQ	02RQ	01TR
RIL ^a mean	4301	8590	8014	6010	5033	2311	6669	7947	6644	7517	6221
Maximum	5257	11217	13753	8457	6563	3373	11033	11933	10133	11083	8667
Minimum	2697	5033	3827	3433	2865	1273	2333	3100	1557	4277	1333
P1 ^b (Lahn)	2873	9825	8787	6363	4821	2470	7517	8600	7867	8100	6000
P2 ^c (Cham1)	3137	9617	10340	6433	5385	2170	8317	9867	7667	8433	7333
LSD ^d (5%)	436	758	1263	600	424	324	1306	2254	1897	1273	1620
CV ^e %	8.28	8.15	14.22	8.69	7.2	11.15	16.9	22.85	20.87	13.69	23.54
Trial H2 ^f	0.36	0.43	0.54	0.46	0.32	0.44	0.29	0	0.09	0.31	0
EGA ^g (%)	5.27	27.39	28.27	19.95	34.1	13.6	40.44	12.19	-6.55	27.59	16.77

^aRIL: Recombinant inbred lines; ^bP1: Parent 1 (Lahn); ^cP2: Parent 1(Cham1); ^dLSD: Least significant difference;

^eCV: Coefficient of variability; ^fH2: Heritability; ^gEGA: Expected genetic advance.

the parent's grain yield was approximately similar to that of the mean grain yields of the RILs. Table 2 shows statistical parameters, including the trial heritability values. Furthermore, in all environments, several RILs have outyielded both parents (Table 2) significantly.

The highest RIL grain yield means were achieved in the irrigated environments (01IR), the early sowing environment with supplementary irrigation (02EP), and the favorable and high-input environment (02GH). The maximum yields were observed in the early sowing environment with 13,753 kg/ha (02EP) followed by the favorable/high-input environment with fertile soil (02GH) 11,933 kg/ha, the irrigated environment (01IR) with 11,217 kg/ha, the favorable and high-input environment with fertile soil (01GH) 11,033 kg/ha, and the hot irrigated environment with 11,083 and 10,133 kg/ha (02RQ and 01RQ). The Ghab site with high soil fertility is located in a valley in north-western Syria, where soil is loamy and consists of sand, clay, silt, and black topsoil. Its fertility provides consistent high crops yields across years.

3.2. Map construction

The SNPs were mapped to all 14 chromosomes of durum wheat. The A chromosomes had 2902.61 cM (47.41%), whereas the B chromosomes had 3219.61 cM (52.59%). Furthermore, the A chromosomes were covered by 117 SSR markers (47.36%) and 598 SNPs (41.96%) with an average distance between adjacent loci of 4.06 cM. B chromosomes were covered by 130 SSR markers (52.64%) and 827 SNPs (58.04%) with an average distance between adjacent loci of 3.36 cM. The highest number of SNPs mapped was on 2B with 205 SNPs, whereas the lowest number of molecular markers mapped was on 7B with 23 SNP markers (Table 3).

For the group of A chromosomes, 2A harbored the largest number of markers (152) with the highest SNP mapping number (133), followed by 4A (128) with 110 SNPs and by 1A and 7A (101 each) with 85 and 88 SNPs, respectively (Table 3). The 2A showed the highest markers enrichment in the group of A chromosomes with a size of 502.84 cM and a density of 3.31 cM per marker. However, the 6A was the largest chromosome with 746.24 cM, but it was covered only with 96 markers; consequently, it has the lowest cM density per marker (7.77 cM/marker). The least enriched chromosome in the A chromosomes with SNP markers was the chromosome 5A (24 SNP). The entire A chromosomes had 715 markers with a map size of 2902.61 cM and an average density of 4.06 cM/marker (Table 3).

As for the group of B chromosomes, 2B had the highest number of molecular markers (226). It also had the highest number of SNPs (205) mapped; 1B and 5B followed it with 203 and 178, respectively. In both 1B and 5B, 166 SNPs each were mapped (Table 3). The lowest coverage in B chromosomes was the 7B chromosome with SNPs (23). The total number of molecular markers harbored by the B chromosomes was 957.

In the B chromosomes, 2B covered 694.15 cM, and it was relatively saturated: 3.07 cM per marker. Moreover, it was the highest enriched genome with SNP markers of the entire genome. The 6B was the most densely saturated chromosome: 2.85 cM/marker and 350.54 cM size. Furthermore, the whole B chromosome was covered by 957 markers with a size of 32,19.61 cM and an average density of 3.36 cM per marker. There were more SNPs mapped in the B chromosomes than in the A chromosomes: 827 for B vs. 598 for A chromosomes; and for the total markers mapped: 957 for B vs. 715 for

Table 3. Assignment and distribution of molecular markers, %, cM size, and coverage across the 14 A and B genomes for SSR and SNP enriched maps in the durum Lahn/Cham1 population.

Chromosome	Markers				Size (cM ^b) in SSR ^c / SNP ^a enriched map	cM ^b /marker in SSR ^c / SNP ^a enriched map
	SSR ^c	SNP ^a	All markers	%		
1A	16	85	101	6.04	184.29/320.6	11.52/3.17
2A	19	133	152	9.09	227.73/502.84	11.98/3.31
3A	6	94	100	5.98	85.75/327.11	14.29/3.27
4A	18	110	128	7.66	158.80/474.76	8.22/3.71
5A	13	24	37	2.21	65.18/151.31	5.01/4.09
6A	32	64	96	5.74	481.43/746.24	13.07/7.77
7A	13	88	101	6.04	224.69/379.53	17.28/3.76
A chromosomes	117	598	715	42.76	1427.87/2902.61	12.20/4.06
1B	37	166	203	12.14	252.27/678.09	6.82/3.34
2B	21	205	226	13.52	246.29/694.15	11.73/3.07
3B	11	27	38	2.27	81.13/245.37	7.37/6.46
4B	18	138	156	9.33	111.63/477.61	6.20/3.06
5B	12	166	178	10.65	101.55/609.17	8.46/3.42
6B	21	102	123	7.36	165.68/350.54	7.89/2.85
7B	10	23	33	1.97	105.32/164.68	10.53/4.99
B chromosomes	130	827	957	57.24	1063.87/3219.61	8.18/3.36
Total	247	1425	1672	100	2491.74/6122.22	10.09/3.66

^aSNP: Single nucleotide polymorphism; ^bcM: CentiMorgan; ^cSSR: Simple sequence repeat.

A chromosomes. The B chromosome map size was larger than that of the A chromosomes: 3219.61 vs. 2902.61 cM. Similarly, for marker density (Table 3), B chromosomes were denser (3.66 cM/marker) than A chromosomes (4.06 cM/marker).

3.3. QTL mapping for grain yield in specific environments

In the specific environments, three QTLs were detected in the rainfall-favorable environment (01GH and 02GH), and two in the irrigated and hot environment (02RQ). Furthermore, four QTLs were detected in the cold favorable environment (02EP) and the dry environments (02INC and 02RF).

In 01GH, the first QTL was located at 5A chromosome and flanked on the left by ksm137bp420 and on the right by gwm154bp115 (Table 4). The second QTL was located at 6A and flanked on the left by McagEagc-bp170 and on the right by wsnp_1092345. The last QTL was detected at 6B chromosome, and its left flanking marker was wsnp_1708133, and the right flanking marker was gwm133bp125 (Table 4; Figure).

In the hot irrigated environment (02RQ), a QTL was detected at 2A with the flanking markers cnl127bp435 and wsnp_2276833 (Table 4; Figure). The analysis also showed

another QTL at 2B chromosome, which was flanked by wsnp_1058086 and wmc175p225 (Table 4; Figure).

In addition, four other QTLs were detected, three of which at 2A chromosome in 02EP, 02INC, and in 02RF, whereas the fourth QTL was detected at 3B chromosome in 02GH environment (Table 4; Figure).

The expression of the QTLs occurred in all environments: moderately stressed (02INC and 02RF), favorable and cold (01GH, 02GH, and 02EP), and hot irrigated (02RQ). The differences of QTL detection between the high-input favorable sites (01GH and 02GH) are probably due to the different seasonal rainfall distribution: 02GH was affected by early drought.

3.4. Mapping in multienvironment trials

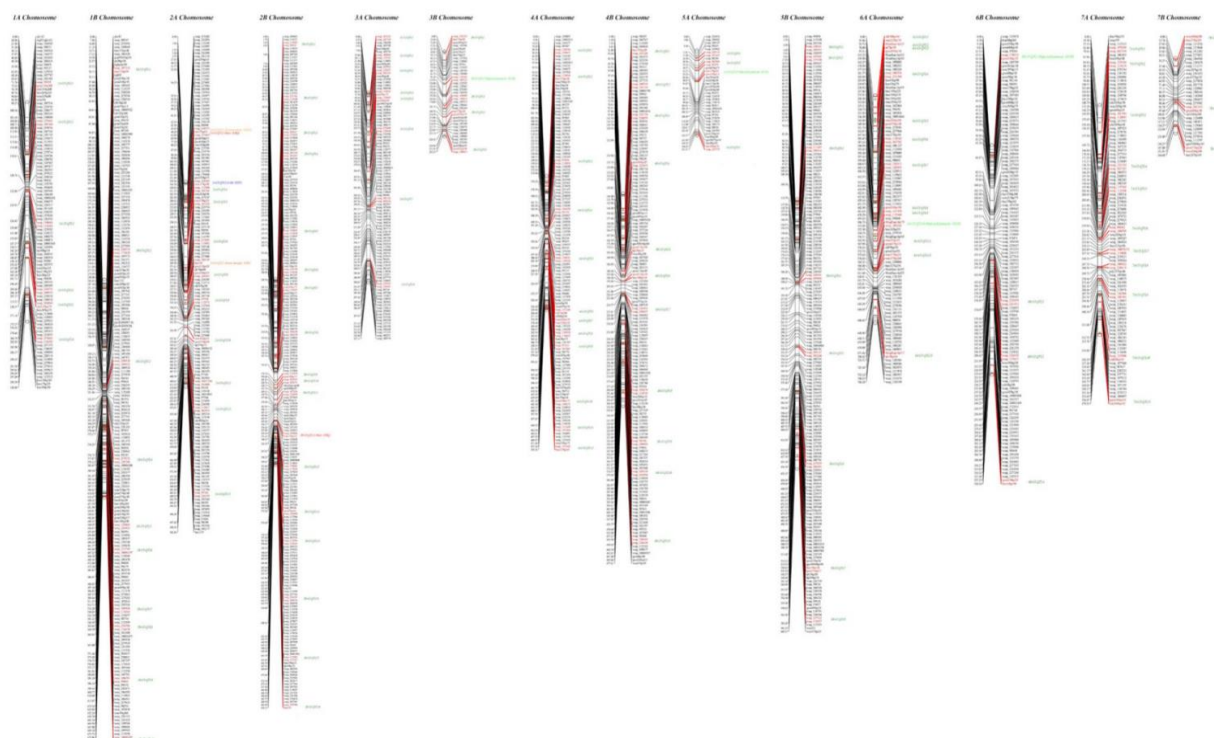
In multienvironmental mapping, 126 QTLs were detected for GY. The A chromosomes contained 53.17% of the total QTLs detected. The chromosomes 2A, 2B, and 6A harbored the highest QTL numbers: 13, 18, and 13, respectively. The GY-QTL distribution and their flanking markers are reported in Table 5 and Figure.

The number of QTL contributions to grain yield was 126. The QTL detected covered all chromosomes. Furthermore, the A chromosomes QTL contributions

Table 4. Significant QTL regions in contrasting environments for Lahn/Cham1 population with their chromosomal location, position, flanking markers, LOD scores, additive effect, and phenotypic variations.

Environment	Chromosome	Position (cM)	Left marker	Right marker	LOD ^b	PVE ⁱ (%)	Add ^j
01GH	5A	54	<i>ksm^a137bp420</i>	<i>gwm^c154bp115</i>	3.9248	11.4299	450.9197
01GH	6A	508	<i>McagEagc^f-bp170</i>	<i>wsnp^g_1092345</i>	3.4404	12.1341	520.5310
01GH	6B	14	<i>wsnp^g_1708133</i>	<i>gwm^c133bp125</i>	3.0045	8.8522	397.7209
02RQ	2A	63	<i>cnl^b127bp435</i>	<i>wsnp^g_2276833</i>	3.3358	9.435	-518.7356
02RQ	2B	531	<i>wsnp^g_1058086</i>	<i>wmc^e175bp225</i>	4.4356	13.0559	-610.2649
02EP	2A	142	<i>wsnp^g_2280363</i>	<i>wmc^e177bp190</i>	3.2202	12.5648	614.8357
02GH	3B	95	<i>wsnp^g_1145565</i>	<i>wsnp^g_1005323</i>	3.1525	11.4755	-295.1018
02INC	2A	46	<i>wsnp^g_2280371</i>	<i>cnl^b127bp435</i>	3.9061	15.0246	-180.9878
02RF	2A	270	<i>barc^d183</i>	<i>gwm^c312bp189</i>	3.6291	14.1871	-159.9529

^aksm: Kansas State University (EST-SSR); ^bcnl: Cornell University (EST-SSR); ^cgwm: Gatersleben Wheat Microsatellite (SSR); ^dbarc: Beltsville Agriculture Research Center (SSR); ^ewmc: Wheat Microsatellite Consortium (SSR); ^fMcagEagc: MseI+cag and EcoRI+agc (AFLP); ^gwsnp: Wheat SNP (SNP); ^hLOD: logarithm of odds; ⁱPVE: Phenotypic variation explained by QTL at the current scanning position; ^jAdd: Additive effect at the current scanning position; positive numbers of Add shows the allelic effect from Lahn, negative numbers from Cham1.

**Figure.** Chromosomes and grain yield QTLs of Lahn/Cham1 genetic linkage map.

were higher than in B chromosomes 67 vs. 59. The chromosomes with the highest number of QTL were 2B (18) followed by 2A and 6A (13) for each, whereas

the lowest chromosome with QTL was 7B (Table 5; Figure). The comparison with several published wheat mapping populations (Table 6) showed that Lahn/Cham1

Table 5. Significant QTL regions for multienvironment trials with their chromosomes, position, flanking markers, LOD scores, additive effect, and phenotypic variations.

Chromosome	Position (cM)	Left marker	Right marker	LOD ^m	LOD (A) ⁿ	LOD (AbyE) ^o	PVE ^p	PVE (A) ^q	PVE (AbyE) ^r	Add ^s
1A	75	wsnp_993293 ^l	gwm136bp300 ^e	4.25	1.01	3.24	1.91	0.78	1.13	-61.09
1A	105	wsnp_1376461 ^l	wsnp_1091468 ^l	3.45	0.10	3.35	0.91	0.08	0.83	19.37
1A	127	wsnp_1106684 ^l	wsnp_1126484 ^l	5.27	0.36	4.91	2.11	0.29	1.82	37.24
1A	203	wsnp_1093499 ^l	wsnp_2256752 ^l	5.39	0.41	4.98	2.89	0.31	2.57	-38.66
1A	228	wsnp_3020949 ^l	cfa2135bp175 ^h	3.72	0.40	3.32	3.63	0.31	3.33	-38.29
1A	281	wsnp_2276661 ^l	wsnp_1126203 ^l	3.10	0.39	2.71	1.04	0.31	0.74	38.28
1B	26	wsnp_1057112 ^l	ksm112bp290a	3.20	0.24	2.95	0.90	0.19	0.71	30.38
1B	207	wsnp_2266714 ^l	wsnp_1718420 ^l	3.05	0.98	2.07	2.58	0.76	1.82	-60.09
1B	334	wsnp_2280520 ^l	wsnp_1090342 ^l	4.40	0.98	3.43	3.83	0.76	3.07	60.27
1B	379	wsnp_2278332 ^l	wsnp_1207280 ^l	3.04	0.24	2.80	1.72	0.19	1.53	30.22
1B	451	wsnp_1240883 ^l	wsnp_2265852 ^l	6.59	1.84	4.75	3.51	1.45	2.07	84.25
1B	467	wsnp_1215769 ^l	wsnp_100004197 ^l	3.83	0.17	3.66	3.98	0.13	3.85	25.29
1B	528	wsnp_1089860 ^l	wsnp_1138263 ^l	3.09	0.28	2.81	3.49	0.22	3.27	32.32
1B	564	wsnp_2265986 ^l	wsnp_1164138 ^l	3.16	0.03	3.12	1.90	0.02	1.88	10.86
1B	586	wsnp_1006701 ^l	wsnp_998041 ^l	5.27	0.18	5.09	3.36	0.14	3.23	-26.09
1B	677	wsnp_10001692 ^l	wsnp_1092947 ^l	3.46	0.93	2.54	2.59	0.71	1.88	-58.38
2A	46	wsnp_2280371 ^l	cnl127bp435 ^b	8.02	1.73	6.29	4.09	1.35	2.74	-80.09
2A	60	cnl127bp435 ^b	wsnp_2276833 ^l	7.08	4.27	2.81	7.08	2.99	4.09	-119.35
2A	142	wsnp_2280363 ^l	wmc177bp190 ^g	8.38	0.05	8.33	10.59	0.04	10.55	14.45
2A	151	wsnp_1122848 ^l	wsnp_1092763 ^l	7.66	1.33	6.33	7.01	1.03	5.98	-72.61
2A	210	wmc170bp225 ^g	wsnp_1021220 ^l	3.60	1.32	2.28	3.14	1.02	2.12	-69.75
2A	253	wsnp_2278093 ^l	wsnp_1110051 ^l	6.17	2.16	4.02	5.16	1.60	3.56	-87.19
2A	269	wsnp_1000185 ^l	barc183 ^f	7.52	1.84	5.68	3.54	1.41	2.13	-81.99
2A	290	gwm294bp135 ^e	wsnp_1666851 ^l	5.02	1.61	3.40	4.05	1.25	2.81	-77.04
2A	335	wsnp_979520 ^l	wsnp_1128771 ^l	4.24	1.72	2.52	3.07	1.34	1.72	-79.95
2A	377	wmc181bp255 ^g	gwm311bp160 ^e	5.17	2.86	2.31	3.52	1.96	1.56	-97.15
2A	413	wsnp_10017104 ^l	wsnp_1010000 ^l	4.07	0.89	3.18	3.75	0.69	3.06	-57.45
2A	434	wsnp_1128817 ^l	wsnp_3029919 ^l	5.24	2.52	2.72	3.94	1.92	2.02	-95.52
2A	484	wsnp_997103 ^l	wsnp_2281375 ^l	5.24	2.56	2.67	4.83	2.00	2.84	-97.50
2B	4	wsnp_1130302 ^l	wsnp_998518 ^l	3.21	0.16	3.05	3.35	0.12	3.23	23.86
2B	45	wsnp_2257342 ^l	wsnp_1199440 ^l	3.61	0.23	3.38	2.37	0.18	2.19	28.93
2B	61	wsnp_3025617 ^l	wsnp_2278589 ^l	4.08	0.59	3.49	3.73	0.40	3.34	43.44
2B	112	wsnp_1021176 ^l	wsnp_3022997 ^l	5.15	0.16	5.00	4.44	0.12	4.32	24.25
2B	237	wsnp_1204695 ^l	wsnp_1164950 ^l	3.24	0.71	2.53	3.07	0.51	2.56	49.24
2B	282	wsnp_1205720 ^l	wsnp_1160346 ^l	3.63	0.24	3.38	3.49	0.19	3.30	30.38
2B	314	wmc243bp175 ^g	wsnp_1106597 ^l	3.20	0.52	2.68	1.56	0.41	1.16	-45.31
2B	335	wsnp_100140 ^l	wsnp_1082370 ^l	3.09	0.82	2.28	3.79	0.62	3.16	54.40
2B	379	wsnp_1109199 ^l	gwm148bp175 ^e	4.72	0.12	4.59	1.87	0.09	1.78	-21.16
2B	406	wsnp_3025955 ^l	wsnp_1020753 ^l	4.66	0.17	4.49	2.66	0.13	2.53	-24.79

Table 5. (Continued.)

Chromosome	Position (cM)	Left marker	Right marker	LOD ^m	LOD (A) ⁿ	LOD (AbyE) ^o	PVE ^p	PVE (A) ^q	PVE (AbyE) ^r	Add ^s
2B	468	wsnp_3027036 ^l	wsnp_1218999 ^l	4.98	1.19	3.79	2.22	0.81	1.41	-63.13
2B	531	wsnp_1058086 ^l	wmc175bp225 ^s	8.08	2.15	5.94	8.17	1.67	6.50	-89.23
2B	552	wsnp_1700082 ^l	wsnp_1179226 ^l	4.80	0.20	4.59	2.27	0.15	2.12	-26.95
2B	575	gwm47bp144 ^e	wsnp_1386000 ^l	4.23	0.45	3.78	2.00	0.36	1.64	-42.25
2B	594	wsnp_1133994 ^l	wsnp_1228193 ^l	3.73	1.00	2.73	2.11	0.75	1.35	-60.56
2B	631	wsnp_2282293 ^l	wsnp_1009576 ^l	5.87	0.62	5.26	3.18	0.49	2.69	-52.58
2B	653	wsnp_1122903 ^l	wsnp_1227557 ^l	3.15	0.10	3.04	1.79	0.08	1.70	-20.49
2B	694	wsnp_1393706 ^l	barc159 ^f	3.63	0.01	3.62	3.17	0.01	3.16	-4.94
3A	0	wsnp_1072225 ^l	wsnp_1667169 ^l	6.41	0.00	6.41	5.50	0.00	5.50	1.63
3A	20	wsnp_1009844 ^l	wsnp_1107433 ^l	3.79	0.21	3.58	3.18	0.16	3.02	27.88
3A	46	wsnp_1232627 ^l	wsnp_3022183 ^l	3.72	0.13	3.60	3.37	0.10	3.28	21.62
3A	67	wsnp_1109262 ^l	gwm5bp178 ^e	4.22	0.03	4.19	2.02	0.03	1.99	-10.92
3A	94	wsnp_1218407 ^l	gwm68bp190 ^e	4.65	1.85	2.79	3.80	1.42	2.38	82.33
3A	151	wsnp_2276600 ^l	wsnp_1220348 ^l	4.13	0.67	3.46	3.09	0.52	2.57	49.95
3A	196	wsnp_1081242 ^l	wsnp_3026254 ^l	5.35	1.26	4.10	4.79	0.97	3.82	69.19
3A	291	wsnp_1237635 ^l	wsnp_1265491 ^l	3.10	0.60	2.50	2.97	0.47	2.50	47.71
3B	1	wsnp_1022121 ^l	barc73bp205 ^f	3.69	0.04	3.65	3.61	0.03	3.58	11.79
3B	92	wsnp_1145565 ^l	wsnp_1005323 ^l	4.48	0.31	4.17	2.56	0.23	2.33	-33.34
3B	164	wsnp_1698409 ^l	wsnp_1698619 ^l	5.16	1.63	3.53	2.49	1.17	1.31	76.47
3B	213	barc77bp220 ^f	gwm547bp175 ^e	3.66	0.92	2.74	2.32	0.67	1.65	57.24
3B	245	gwm582bp200 ^e	wsnp_1088815 ^l	4.15	0.87	3.28	3.07	0.67	2.40	56.77
4A	21	wsnp_1206766 ^l	wsnp_1090457 ^l	3.30	0.00	3.29	2.35	0.00	2.35	4.06
4A	59	wsnp_1130454 ^l	barc113bp140 ^f	7.38	1.17	6.21	3.27	0.86	2.41	66.30
4A	150	wsnp_979658 ^l	wsnp_1218818 ^l	3.48	0.02	3.46	3.21	0.01	3.20	-8.17
4A	186	wsnp_1064430 ^l	wsnp_985258 ^l	3.25	0.21	3.04	4.00	0.16	3.83	-28.01
4A	204	wsnp_1108161 ^l	wsnp_1109226 ^l	3.33	0.45	2.87	1.84	0.33	1.51	-40.10
4A	266	wsnp_1073246 ^l	cfd31bp200 ^l	4.93	0.01	4.92	4.40	0.01	4.39	7.27
4A	276	wmc262bp150 ^s	wmc232bp140 ^s	4.75	0.01	4.74	4.63	0.01	4.62	-7.78
4A	293	wsnp_1309558 ^l	wsnp_1766622 ^l	4.95	0.00	4.95	2.49	0.00	2.49	1.74
4A	321	wsnp_977411 ^l	wsnp_3023881 ^l	4.99	0.00	4.99	6.62	0.00	6.62	3.57
4A	442	gwm160bp172 ^e	wsnp_1086275 ^l	6.34	0.04	6.31	7.30	0.02	7.28	10.36
4A	460	wsnp_1111699 ^l	wsnp_1053664 ^l	5.07	0.00	5.07	5.04	0.00	5.04	0.30
4A	474	wsnp_1073442 ^l	barc327bp245 ^f	5.62	0.19	5.43	7.89	0.13	7.75	25.31
4B	32	barc193bp260	wsnp_1033286	6.48	1.68	4.80	2.70	1.32	1.39	79.35
4B	58	wsnp_1009219	wsnp_1201110	6.33	1.75	4.58	3.99	1.34	2.65	80.30
4B	113	wsnp_2281754	wsnp_1260076	5.22	1.00	4.22	2.03	0.74	1.29	59.50
4B	126	gwm368bp287	wsnp_2254072	6.04	0.15	5.89	3.63	0.12	3.51	23.89
4B	182	gwm513bp190	dp23bp235	4.86	0.11	4.75	2.51	0.08	2.43	20.00
4B	196	gwm513bp150	barc340bp210	4.05	0.06	3.99	2.81	0.05	2.76	15.13
4B	225	wsnp_1264915	wsnp_1090257	3.09	0.02	3.06	1.93	0.02	1.91	-9.34

Table 5. (Continued.)

Chromosome	Position (cM)	Left marker	Right marker	LOD ^m	LOD (A) ⁿ	LOD (AbyE) ^o	PVE ^p	PVE (A) ^q	PVE (AbyE) ^r	Add ^s
4B	290	wsnp_1036270	wsnp_1220743	3.85	0.24	3.61	3.13	0.17	2.96	-28.40
4B	336	wsnp_992780	wsnp_1090893	3.80	0.15	3.65	1.68	0.11	1.57	22.84
4B	375	wsnp_1050158	wsnp_1010596	3.01	1.45	1.55	1.67	1.10	0.57	72.61
4B	449	wsnp_1206903	wsnp_1204109	3.21	0.82	2.40	2.25	0.62	1.62	-55.05
5A	12	wsnp_2283193 ^l	wsnp_1071495 ^l	3.80	0.02	3.78	2.38	0.02	2.36	9.03
5A	35	wsnp_3022906 ^l	wsnp_981078 ^l	4.25	0.00	4.25	1.48	0.00	1.48	1.08
5A	55	ksm137bp420a	gwm154bp115 ^e	8.60	3.33	5.27	6.37	2.62	3.75	111.84
5A	148	barc100bp110 ^f	wsnp_2257379 ^l	3.70	2.10	1.60	3.42	1.56	1.87	86.39
5B	3	wsnp_2280241 ^l	wsnp_1095057 ^l	5.21	2.65	2.56	3.63	2.06	1.57	-99.43
5B	10	wsnp_1079002 ^l	wsnp_2279190 ^l	6.93	3.72	3.21	4.00	2.83	1.16	-116.00
5B	89	wsnp_1111097 ^l	wsnp_1094116 ^l	3.42	0.03	3.40	3.50	0.02	3.48	-9.82
5B	187	wsnp_1695891 ^l	wsnp_1078695 ^l	4.23	0.19	4.04	4.90	0.14	4.76	-26.03
5B	332	wsnp_1091357 ^l	wsnp_3024240 ^l	3.41	0.08	3.33	2.83	0.06	2.77	-17.13
5B	434	wsnp_2278459 ^l	wsnp_1081891 ^l	3.39	0.59	2.79	3.61	0.47	3.14	-47.05
5B	521	barc74bp120 ^f	gwm371bp317 ^e	3.03	0.58	2.45	2.12	0.44	1.68	-45.84
5B	584	wsnp_2277122 ^l	wsnp_1126527 ^l	3.67	0.54	3.13	3.41	0.42	3.00	44.54
6A	35	cfd190bp160 ⁱ	wmc123bp135 ^s	3.70	0.39	3.30	2.58	0.29	2.29	45.50
6A	56	McacEaac-bp225d	pk7bp190 ^k	4.66	2.51	2.15	4.28	1.93	2.35	108.11
6A	97	pk7bp190 ^k	gwm497bp118 ^e	3.05	0.76	2.29	1.42	0.58	0.84	-58.32
6A	185	wsnp_1002536 ^l	wsnp_2331540 ^l	3.65	0.30	3.35	3.00	0.23	2.77	34.26
6A	345	wmc201bp250 ^s	gwm570bp140 ^e	4.32	0.50	3.82	1.75	0.40	1.35	43.80
6A	387	wsnp_1126162 ^l	wsnp_30272001 ^l	3.98	0.24	3.74	3.84	0.18	3.66	29.25
6A	401	wsnp_1202412 ^l	wsnp_1202955 ^l	3.81	0.02	3.79	4.07	0.02	4.05	-9.48
6A	443	gwm169bp190 ^e	wsnp_1231938 ^l	5.39	0.67	4.72	3.29	0.49	2.80	48.34
6A	477	wsnp_1231938 ^l	wsnp_1125460 ^l	4.09	0.06	4.03	2.25	0.05	2.20	-14.68
6A	507	McagEagc-bp170c	wsnp_1092345 ^l	6.00	0.23	5.76	4.33	0.18	4.15	33.96
6A	566	wmc41bp155 ^s	gwm617bp123 ^e	5.25	2.01	3.23	2.55	1.40	1.15	-112.97
6A	578	wmc173bp235 ^s	wmc175bp260 ^s	4.80	1.26	3.54	3.75	0.90	2.85	-122.39
6A	700	McagEagc-bp117c	pk19bp160 ^k	3.66	0.05	3.61	2.66	0.04	2.62	14.80
6B	15	wsnp_1708133 ^l	gwm133bp125 ^e	7.26	0.69	6.57	5.59	0.49	5.09	48.87
6B	175	wsnp_2329478 ^l	wsnp_2261831 ^l	4.85	0.00	4.84	1.95	0.00	1.95	4.27
6B	201	wsnp_2266428 ^l	wsnp_1178437 ^l	3.84	0.19	3.66	2.45	0.14	2.31	25.77
6B	350	gwm219bp255 ^e	barc24bp190 ^f	3.35	0.43	2.93	1.94	0.32	1.62	39.68
7A	43	wsnp_1070109 ^l	wsnp_3027359 ^l	4.09	0.37	3.72	3.69	0.28	3.41	36.34
7A	88	wsnp_1225184 ^l	wsnp_1210139 ^l	3.81	0.20	3.61	2.69	0.15	2.53	27.06
7A	150	wsnp_1041994 ^l	wsnp_1128093 ^l	5.42	0.54	4.88	4.18	0.34	3.84	40.19
7A	177	wsnp_1221352 ^l	wsnp_3027187 ^l	6.82	1.03	5.79	6.02	0.76	5.25	60.38
7A	202	wsnp_1197565 ^l	wsnp_1125508 ^l	3.78	1.48	2.30	2.73	1.13	1.60	73.22
7A	223	wsnp_996562 ^l	wsnp_1080709 ^l	4.23	1.35	2.88	4.23	1.06	3.17	71.05
7A	244	wsnp_100078174 ^l	wsnp_1130808 ^l	3.31	1.13	2.18	3.55	0.88	2.67	64.83

Table 5. (Continued.)

Chromosome	Position (cM)	Left marker	Right marker	LOD ^m	LOD (A) ⁿ	LOD (AbyE) ^o	PVE ^p	PVE (A) ^q	PVE (AbyE) ^r	Add ^s
7A	268	wsnp_1004422 ^l	wsnp_2280119 ^l	4.34	0.03	4.31	5.06	0.02	5.04	-10.06
7A	289	wsnp_1262886 ^l	wsnp_1001581 ^l	3.80	0.00	3.80	2.17	0.01	2.16	7.00
7A	326	wsnp_1235900 ^l	cfd020bp305 ⁱ	3.14	0.01	3.13	2.68	0.01	2.68	5.71
7A	379	gwm344bp135 ^e	cfa2040bp325 ^h	4.90	0.23	4.67	5.82	0.17	5.65	-28.56
7B	5	wmc606bp200 ⁸	barc279bp205 ^f	4.34	0.32	4.02	3.13	0.20	2.94	30.70
7B	93	wsnp_3021533 ^l	gwm400bp148 ^e	4.40	0.03	4.37	2.80	0.02	2.77	10.66
7B	150	gwm573bp220 ^e	wmc426bp200 ⁸	5.56	0.96	4.60	3.03	0.69	2.35	57.47

^aksm: Kansas State University (EST-SSR); ^bcnl: Cornell University (EST-SSR); ^cMcagEagc: MseI+cag and EcoRI+agc (AFLP); ^dMcacEaac: MseI+cac and EcoRI+aac (AFLP); ^egwm: Gatersleben Wheat Microsatellite (SSR); ^fbarc: Beltsville Agriculture Research Center (SSR); ^gwmc: Wheat Microsatellite Consortium; ^hcfa: Clermont-Ferrand A-genome (SSR); ⁱcfd: Clermont-Ferrand D-genome (SSR); ^jdp: Dehydrin primers "functional genes" (SSR); ^kpk: PK Gupta /india (SSR); ^lwsnp: Wheat SNP (SNP); ^mLOD: Logarithm of odds; ⁿLOD(A): LOD score for additive and dominance effects; ^oLOD(AbyE): LOD score for additive and dominance by environment effects; ^pPVE: Phenotypic variation explained by QTL at the current scanning position; ^qPVE(A): Phenotypic variation explained by additive and dominance effect at the current scanning position; ^rPVE(AbyE): Phenotypic variation explained by additive and dominance by environment effect at the current scanning position; ^sAdd: Additive effect at the current scanning position; positive numbers shows allelic effect from Lahn, negative numbers from Cham1.

population harbored more QTLs contributing to grain yield and in all A and B chromosomes (LOD used for QTL detection was ≥ 3). Some chromosomes (Table 5; Figure) were harboring a large number of QTLs, as in the case of 1A, 1B, 2A, 2B, 4A, 6A, and 7A.

Moreover, the QTLs with an important QTL*Environment effects were on the chromosomes 4A, where the additive effects ranged from 25.31 to 66.3 and from -40.1 to -28.0. On 5B, the additive effect ranged from -116 to -26.03, and one QTL with an additive effect of 44.54. On 7A, several QTLs with a positive additive effect ranged from 27.06 to 73.22 and one QTL with a negative additive effect (-28.56). The contribution of the Lahn parent was 74 QTLs, of which 40 originated from A chromosomes and 34 from B chromosomes, while Cham1 contributed with 52 QTLs, 29 of which were originated from A chromosomes and 23 from B chromosomes.

3. Discussion

The Lahn/Cham1 genetic linkage map has been shown to share several SSR markers with other linkage maps (Nachit et al., 2001; Elouafi and Nachit, 2004; Haile et al., 2012; Zhang et al., 2012; Zhang et al., 2012; Nagel et al., 2014; Alsaleh et al., 2015). It shared an important number of SSR markers with the high-density microsatellite consensus maps for bread wheat (Somers et al., 2004) and durum wheat (Maccaferri et al., 2015). Moreover, Lahn/Cham1 map was concordant with the marker orders of the consensus maps.

Su et al. (2018) adopted a similar approach to include SNP markers. Similarly, Zeng et al. (2019) constructed

a bread wheat linkage map with SNP markers using the Wheat 35K SNP array for genotyping the RILs of the population MX169 X CW86.

Kirigwi et al. (2007) reported that molecular markers on chromosome 4A linked to grain yield, explaining 20% of the phenotypic variation of grain yield. Concerning QTL detection, grain yield QTLs in wheat have been described in various studies (Maccaferri et al., 2010; McIntyre et al., 2010; Soriano et al., 2017). In Kuchel et al. (2007), bread wheat linkage map, which was mainly based on microsatellite molecular markers, the QTLs linked with grain yield were detected on chromosomes 1B, 2D, 3B, 5A, 6B, 7A, 7B, 7D. Multiple QTL detection on similar chromosomes were also reported by Bogard et al. (2011). Additionally, Maccaferri et al. (2008) also reported among the 16 QTLs that affected grain yield, two major QTLs on chromosomes 2BL and 3BS showed significant effects across several environments. Kato et al. (2000) also recovered five regions in wheat chromosome 5A that contributed to grain yield; and according to Heidari et al. (2011), QTLs located on chromosomes 6A, 6B, and 6D accounted mostly for total grain yield variation in wheat. Furthermore, McIntyre et al. (2010) established that several chromosomes harbored QTLs for grain yield on chromosomes 1B, 1D, 4A, 4D, 6A-a, 6B, and 7A. In addition, Quarrie et al. (2005) found that the strongest grain yield QTL effects were located on chromosomes 7AL and 7BL. Under drought conditions, the microsatellite wmc89 located on 4A chromosome was suggested as a marker to enhance drought tolerance in wheat (Kirigwi et al., 2007). Furthermore, Bennett et al. (2012) found that

Table 6. Comparison of detected grain yield QTLs harbored by A and B chromosomes in different wheat mapping populations.

RIL/ Marker	Kuchel et al. (2007)	Maccaferri et al. (2008)	Bennett et al. (2012)	Kirigwi et al. (2007)	Heidari et al. (2011)	Quarrie et al. (2005)	Bogard et al. (2011)	McIntyre et al. (2010)	Marza et al. (2006)	Tura et al. (2019)	Lahn/Cham1 population		
	182/251	249/254	250/456	127/ 136	107/371	96/567	140/499	192/587	132/410	155/3502	112/1672	Total QTL	Parent QTLs contribution per chromosome
1A									+	+	+	6	Lahn 3 Cham1 3
1B	+							+	+	+	+	10	Lahn 7 Cham1 3
2A										+	+	13	Lahn 1 Cham1 12
2B		+							+	+	+	18	Lahn 7 Cham1 11
3A			+							+	+	8	Lahn 7 Cham1 1
3B		+	+				+			+	+	5	Lahn 4 Cham1 1
4A			+	+				+		+	+	12	Lahn 8 Cham1 4
4B									+	+	+	11	Lahn 8 Cham1 3
5A							+		+	+	+	4	Lahn 4 Cham1 0
5B			+							+	+	8	Lahn 1 Cham1 7
6A					+			+	+		+	13	Lahn 8 Cham1 5
6B					+		+	+	+	+	+	4	Lahn 4 Cham1 0
7A			+			+	+	+	+	+	+	11	Lahn 9 Cham1 2
7B						+	+			+	+	3	Lahn 3 Cham1 0

+: detected QTL in a given chromosome

two of the detected QTLs linked to heat tolerance were located on chromosome 3B.

Marza et al. (2006) showed that the QTL contribution effects ranged from 7% to 23% in Ning7840xClark wheat population, where the marker alleles from the parent Clark were associated with a positive effect for the majority of QTLs for yield and its components on 1AL, 1B, 4B, 5A, 6A, and 7A, and specific QTLs on 2BL, 2BS, 2DL, and 6B. Tura et al. (2019) identified 38 grain yield QTLs spread over the whole genome of the Excalibur/Kuri population of bread wheat with the exception of 5D and 6A chromosomes. The grain yield QTLs with the highest QTL*Environment effects were observed on chromosomes 4A, 5B, and 7A. In the Lahn/Chm1 population, the QTLs linked to grain yield were harbored in all 14 chromosomes. These results agree with those of other studies (Heidari et al., 2011; Zhang et al., 2012; Farré et al., 2016). These results were also similar to those of the recently published work by Tura et al. (2019). In contrast, most of the populations showed a few QTLs (Table 6). It appears that the choice of genetically diverse population parents is of paramount importance to detect all involved QTLs linked with wheat grain yield (Table 6).

Furthermore, the analysis of the multienvironment trial showed the overall QTLs present on the 4B chromosome has an important additive effect with parent Lahn contribution ranging from 15.13 to 80.30 and the parent Cham1 contribution from -55.05 to -9.34. In light of this, it appears that the 4B chromosome contributes largely to grain yield in the dry areas.

Our results under stressed environments showed agreement with those of Ribaut et al. (1997) and Almeida

et al. (2013). It is expected that using marker-assisted selection for grain yield under drought and heat stresses will improve breeding under stress conditions.

Several studies have indicated the presence of QTLs and candidate genes involved in response to drought in the 4B chromosome (Nachit and Elouafi, 2004; Habash et al., 2009). In addition, in Mediterranean terminal stress, drought is usually combined with heat stress; both stresses are a major factor limiting grain yield in the Mediterranean region (Nachit, 1992; Nachit and Elouafi, 2004). Other findings on 4B have revealed the presence of QTLs linked to total biomass yield, grain yield, and straw yield (Li et al., 2014). Recently, the International Wheat Genome Sequence (IWGSC) released a wheat genome reference sequence along with annotated genes (called RefSeq v.1.0, IWGSC 2018)². This information, along the whole genome sequence dataset of 16 wheat varieties (Edwards et al., 2012), allows identifying new SNP markers for QTL fine-mapping and identifying genes responsible for abiotic and biotic stress tolerance in wheat; as consequence, SNP markers linked to detected QTLs need to be annotated for genes involved in traits tolerance/resistance and grain yield performance in diverse and contrasting environments.

Acknowledgments

We would like to thank the International Center for Agricultural Research in Dry Area (ICARDA) for the financial support of this study and the support staff of the CIMMYT/ICARDA durum breeding program.

²Wheat@URGI. Seq Repository [Online]. Website <http://wheat-urgi.versailles.inra.fr/Seq-Repository> [Accessed 22 Jan 2021]

References

- Akbari M, Wenz P, Craig V, Carling J, Xia L et al. (2006). Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. *Theoretical and Applied Genetics* 113: 1409-1420. doi: 10.1007/s00122-006-0365-4.
- Almeida G, Dan Macomb D, Moroboshi C, Nair S, Bore A et al. (2013). QTL mapping in three tropical maize populations reveals a set of constitutive and adaptive genomic regions for drought tolerance. *Theoretical and Applied Genetics* 126: 583-600. doi: 10.1007/s00122-012-2003-7.
- Alsaleh A, Baloch FS, Derma M, Zara M, Kilian B et al. (2015). Genetic linkage map of anatolian durum wheat derived from a cross of Kunduru-1149x Cham1. *Plant Molecular Biology Reporter* 33: 209-220. doi: 10.1007/s11105-014-0749-6.
- Bennett D, Reynolds M, Mullan D, Izanloo A, Kuchel H et al. (2012). Detection of two major grain yield QTL in bread wheat (*Triticum aestivum* L.) under heat, drought and high yield potential. *Theoretical and Applied Genetics* 125: 1473-1485. doi: 10.1007/s00122-012-1927-2.
- Blanco A, Bello MP, Cenci A, De Giovanni C, D'Ovidien R et al. (1998). A genetic linkage map of durum wheat. *Theoretical and Applied Genetics* 97: 721-728. doi: 10.1007/s001220050948.
- Blanco A, De Giovanni C, Laddomada B, Sciancalepore A, Simeone R et al. (1996). Quantitative trait loci influencing grain protein content in tetraploid wheats. *Plant Breeding* 115: 310-316. doi: 10.1111/j.1439-0523.1996.tb00925.x.
- Bogard M, Jourdan M, Jourdan M, Allard V, Maitre P et al. (2011). Anthesis date mainly explained correlations between post-anthesis leaf senescence, grain yield, and grain protein concentration in a winter wheat population segregating for flowering time QTLs. *Journal of Experimental Botany* 62 (10): 3621-3636. doi: 10.1093/jxb/err061.
- Bort J, Belhaj M, Latiri K, Kehel Z, Araus JL (2014). Comparative performance of the stable isotope signatures of carbon, nitrogen and oxygen in assessing early vigour and grain yield in durum wheat. *The Journal of Agricultural Science* 152 (3): 408-426. doi: 10.1017/S0021859613000269.

- Dura S, Dower M, Nachit MM, Al Shehab F (2013). Detection of molecular markers associated with yield and yield components in durum wheat (*Triticum turgidum* L. var. durum) under saline conditions. *Crop and Pasture Science* 64 (10): 957-964. doi: 10.1071/CP13287.
- Edwards D, Wilcox S, Barrera RA, Fleury D, Cavanagh CR et al. (2012). Bread matters: a national initiative to profile the genetic diversity of Australian wheat. *Plant Biotechnology Journal* 10: 703-708. doi: 10.1111/j.1467-7652.2012.00717.x.
- El Hassouni K, Berkadia B, Filali-Maltouf A, Tidiane-Sall A, Al-Abdellat A et al. (2019). Loci Controlling Adaptation to Heat Stress Occurring at the Reproductive Stage in Durum Wheat. *Agronomy* 9 (8): 414. doi: 10.3390/agronomy9080414.
- Elouafi I, Nachit MM (2004). A genetic linkage map of the *Durum* x *Triticum dicoccoides* backcross population based on SSRs and AFLP markers, and QTL analysis for milling traits. *Theoretical and Applied Genetics* 108: 401-413. doi: 10.1007/s00122-003-1440-8.
- Farré A, Sayers L, Leverington-Waite M, Goram R, Orford S et al. (2016). Application of a library of near isogenic lines to understand context dependent expression of QTL for grain yield and adaptive traits in bread wheat. *BMC Plant Biology* 16: 161. doi: 10.1186/s12870-016-0849-6.
- Federer WT (1956). Augmented (or hoonuiaku) designs. *Hawaiian Planters. Record* 55 (2): 191-208.
- Habash D, Kehel Z, Nachit MM (2009). Genomic approaches for designing durum wheat ready for climate change with a focus on drought. *Journal of Experimental Botany* 60 (10): 2805-2815. doi: 10.1093/jxb/erp211.
- Habash DZ, Baudo M, Hindle M, Powers SJ, Defoin-Platel M et al. (2014). Systems responses to progressive water stress in durum wheat. *PLOS ONE* 9: e108431. doi: 10.1371/journal.pone.0108431.
- Haile K J, Nachit MM, Hammer K, Badebo A, Röder M S (2012). QTL mapping of resistance to race Ug99 of *Puccinia graminis* f. sp. *tritici* in durum wheat (*Triticum durum* Desf.). *Molecular Breeding* 30: 1479-1493. doi: 10.1007/s11032-012-9734-7.
- Heidari B, Sayed-Tabatabaei BE, Saeidi G, Kearsy M, Suenaga K (2011). Mapping QTL for grain yield, yield components, and spike features in a doubled haploid population of bread wheat. *Genome* 54 (6): 517-527. doi: 10.1139/g11-017.
- Kato K, Miura H, Sawada S (2000). Mapping QTLs controlling grain yield and its components on chromosome 5A of wheat. *Theoretical and Applied Genetics* 101: 1114-1121. doi: 10.1007/s001220051587.
- Kehel Z, Habash DZ, Gezan SA, Welham SJ, Nachit MM (2010). Estimation of spatial trend and automatic model selection in augmented designs. *Agronomy Journal* 102: 1542-1555. doi: 10.2134/agronj2010.0175.
- Kirigwi FM, Van Ginkel M, Brown-Guedira G, Gill BS, Paulsen GM, et al. (2007). Markers associated with a QTL for grain yield in wheat under drought. *Molecular Breeding* 20: 401-413. doi: 10.1007/s11032-007-9100-3.
- Korzun V, Röder MS, Wendehake K, Pasqualone A, Lotti C et al. (1999). Integration of dinucleotide microsatellites from hexaploid bread wheat into a genetic linkage map of durum wheat. *Theoretical and Applied Genetics* 98: 1202-1207. doi: 10.1007/s001220051185.
- Kosambi DD (1944). The estimation of map distances from recombination values. *Annual Eugenics* 12: 172-175.
- Kuchel H, Williams KJ, Langridge P, Eagles HA, Jefferies SP (2007). Genetic dissection of grain yield in bread wheat. I. QTL analysis. *Theoretical and Applied Genetics* 115: 1029-1041. doi: 10.1007/s00122-007-0629-7.
- Li H, Ye G, Wang J (2007). A modified algorithm for the improvement of composite interval mapping. *Genetics* 175: 361-374. doi: 10.1534/genetics.106.066811.
- Li S, Jia J, Wei X, Zhang X, Li L et al. (2007). A intervarietal genetic map and QTL analysis for yield traits in wheat. *Molecular Breeding* 20: 167-178. doi: 10.1007/s11032-007-9080-3.
- Li ZK, Jiang XL, Peng T, Shi CL, Han SX et al. (2014). Mapping quantitative trait loci with additive effects and additive x additive epistatic interactions for biomass yield, grain yield, and straw yield using a doubled haploid population of wheat (*Triticum aestivum* L.). *Genetics and Molecular Research* 13 (1): 1412-1424. doi: 10.4238/2014.February.28.14.
- Maccaferri M, Ricci A, Salvi S, Milner SG, Noli E et al. (2015). High-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnology Journal* 13: 648-663. doi: 10.1111/pbi.12288.
- Maccaferri M, Sanguineti MC, Corneti S, Araus O, Ben Salem M et al. (2008). Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.). across a wide range of water availability. *Genetics* 178 (1): 489-511. doi: 10.1534/genetics.107.077297.
- Maccaferri M, Sanguineti MC, Demontis, El-Ahmed A, Del Moral LFG et al. (2010). Association mapping in durum wheat grown across a broad range of water regimes. *Journal of Experimental Botany* 62 (2): 409-438. doi: 10.1093/jxb/erq287.
- Marza F, Bai GH, Carver FB, Zhou WC (2006). Quantitative trait loci for yield and related traits in the wheat population Ning7840x Clark. *Theoretical and Applied Genetics* 112: 688-698. doi: 10.1007/s00122-005-0172-3.
- McIntyre CL, Mathews KL, Rattey A, Chapman SC, Drenth J et al. (2010). Molecular detection of genomic regions associated with grain yield and yield-related components in an elite bread wheat cross evaluated under irrigated and rainfed conditions. *Theoretical and Applied Genetics* 120: 527-541. doi: 10.1007/s00122-009-1173-4.
- Nachit MM (1992). Durum wheat breeding for Mediterranean dryland of North Africa and West Asia. In: *Durum Wheat International Workshop, Discussion on Durum Wheat: Challenges and Opportunity*. CIMMYT; Obregon, Mexico. p. 14-27.

- Nachit MM, Elouafi I (2004). Durum wheat adaptation in the Mediterranean dryland: Breeding, stress physiology, and molecular markers. In: Rao SC, Ryan J (editors). Challenges and Strategies of Dryland Agriculture. CSSA Special Publication no. 32, pp. 203-218. doi: 10.2135/cssaspecpub32.c13.
- Nachit MM, Elouafi I, Pagnotta A, El Saleh A, Iacono E et al. (2001). Molecular linkage map for an intraspecific recombinant inbred population of durum wheat (*Triticum turgidum* L. var. durum). Theoretical and Applied Genetics 102: 177-186. doi: 10.1007/s001220051633.
- Nachit MM, Motawaj J, Kehel Z, Habash D, Elouafi I et al. (2016). Gene flow as a source of adaptation of durum wheat to changing climate conditions: double gradient selection technique. In: Bari A, Damania AB, Mackay M, Dayanandan S (editors). Applied Mathematics and Omics to Assess Crop Genetic Resources for Climate Change Adaptive Traits. CRC Press. pp. 259-268.
- Nagel M, Navakode S, Scheibal V, Baum M, Nachit MM et al. (2014). The genetic basis of durum wheat germination and seedling growth under osmotic stress. Biologia Plantarum 58: 681-688. doi: 10.1007/s10535-014-0436-3.
- Oduola A, Joe M, Philip A, Alexander AB, John K et al. (2003). The nature and identification of quantitative trait loci: a community's view. Nature Reviews Genetics 4 (11): 911-916. doi: 10.1038/nrg1206.
- Poecke RM, Maccaferri M, Tang J, Truong HT, Janssen et al. (2013). Sequence-based SNP genotyping in durum wheat. Plant Biotechnology Journal 11: 809-817. doi: 10.1111/pbi.12072.
- Pozniak CJ, Knox RE, Clarke FR, Clarke JM (2007). Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. Theoretical and Applied Genetics 114: 525-537. doi: 10.1007/s00122-006-0453-5.
- Quarrie SA, Steed A, Calestani C, Semikhodskii A, Lebreton C et al. (2005). A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.). from the cross Chinese Spring $\times$ SQ1 and its use to compare QTLs for grain yield across a range of environments. Theoretical and Applied Genetics 110: 865-880. doi: 10.1007/s00122-004-1902-7.
- Ribaut JM, Jiang C, Gonzalez-de-Leon D, Edmeades GO, Hoisington D A (1997). Identification of quantitative trait loci under drought conditions in tropical maize. 2. Yield components and marker-assisted selection strategies. Theoretical and Applied Genetics 94: 887-896. doi: 10.1007/s001220050492.
- Röder MS, Korzun V, Wendehake K, Plaschke MHT, Leroy P et al. (1998). A microsatellite map of wheat. Genetics 149 (4): 2007-2023.
- Somers DJ, Isaac P, Edwards K (2004). A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 109: 1105-1114. doi: 10.1007/s00122-004-1740-7.
- Soriano JM, Malosetti M, Roselló M, Sorrells ME, Royo C (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: e0178290. doi: 10.1371/journal.pone.0178290.
- Sourdille P, Cadalen T, Guyomarch H, Snape JW, Perretant MR et al. (2003). An update of the courtot $\times$ Chinese spring intervarietal molecular marker linkage map for the QTL detection of agronomic traits in wheat. Theoretical and Applied Genetics 106: 530-538. doi: 10.1007/s00122-002-1044-8.
- Su Q, Zhang X, Zhang W, Zhang N, Song L et al. (2018). QTL detection for kernel size and weight in bread wheat (*Triticum aestivum* L.) using a high-density SNP and SSR-Based linkage map. Frontiers in Plant Science 9: 1484-1484. doi: 10.3389/fpls.2018.01484.
- Tixier MH, Sourdille P, Charmet G, Gay G, Jaby C et al. (1998). Detection of QTLs for crossability in wheat using a doubled-haploid population. Theoretical and Applied Genetics 97: 1076-1082. doi: 10.1007/s001220050994.
- Tura H, Edwards J, Gahlaut V, Garcia M, Sznajder B et al. (2019). QTL analysis and fine mapping of a QTL for yield-related traits in wheat grown in dry and hot environments. Theoretical and Applied Genetics 133: 239-257. doi: 10.1007/s00122-019-03454-6.
- Varshney RK, Prasad M, Roy JK, Kumar N, Harjit-Singh Dhaliwal HS et al. (2000). Identification of eight chromosomes and a microsatellite marker on 1AS associated with QTL for grain weight in bread wheat. Theoretical and Applied Genetics 100: 1290-1294. doi: 10.1007/s001220051437.
- Wang J, Zhang HL, Meng L (2016). QTL IciMapping software. 4.1.0.0. <http://www.isbreeding.net>
- Xu D, Wen W, Fu L, Li F, Li J et al. (2019). Genetic dissection of a major QTL for kernel weight spanning the Rht-B1 locus in bread wheat. Theoretical and Applied Genetics 132: 3191-3200. doi: 10.1007/s00122-019-03418-w.
- Zaïm M, Kabbaj H, Kehel Z, Gorjanc G, Filali-Maltouf A et al. (2020). Combining QTL analysis and genomic predictions for four durum wheat populations under drought conditions. Frontiers in Genetics 11: 316-330. doi: 10.3389/fgene.2020.00316.
- Zeng Q, Wu J, Huang S, Yuan F, Liu S et al. (2019). SNP-based linkage mapping for validation of adult plant stripe rust resistance QTL in common wheat cultivar Chakwal 86. The Crop Journal 7 (2): 176-186. doi: 10.1016/j.cj.2018.12.002.
- Zhang D, Hao C, Wang L, Zhang X (2012). Identifying loci influencing grain number by microsatellite screening in bread wheat (*Triticum aestivum* L.). Planta 236: 1507-1517. doi: 10.1007/s00425-012-1708-9.
- Zhang H, Chen J, Li R, Deng Z, Zhang K et al. (2016). Conditional QTL mapping of three yield components in common wheat (*Triticum aestivum* L.). The Crop Journal 4 (3): 220-228. doi: 10.1016/j.cj.2016.01.007.
- Zhang L, Luo JT, Hao M, Zhang LQ, Yuan ZW et al. (2012). Genetic map of *Triticum turgidum* based on a hexaploid wheat population without genetic recombination for D genome. BMC Genetics 13: 69. doi: 10.1186/1471-2156-13-69.
- Zhang W, Chao S, Manthey F, Chicaiza O, Brevis JC et al. (2008). QTL analysis of pasta quality using a composite microsatellite and SNP map of durum wheat. Theoretical and Applied Genetics. 117: 1361-1377. doi: 10.1007/s00122-008-0869-1.

Dissection of Quantitative Trait
Loci (QTL), annotation of Single
Nucleotide Polymorphism (SNP),
and Identification of Candidate
Genes for Grain Yield in
Triticum turgidum L. var *durum*

Chapter III: Dissection of Quantitative Trait Loci (QTL), annotation of Single Nucleotide Polymorphism (SNP), and Identification of Candidate Genes for Grain Yield in *Triticum turgidum* L. var *durum*

1. Introduction

The construction of genetic map sufficiently saturated with molecular markers, essentially SNP and DArT markers constitute the start point to identify candidate genes contributing in quantitative traits with an importance agronomic (Mantovani et al., 2008; Peleg et al., 2008; Zhang et al., 2008; Gadaleta et al., 2009; Verlotta et al., 2010; Blanco et al., 2011; Trebbi et al., 2011; Blanco et al., 2012; Marone et al., 2012a; Marone et al., 2012b; Russo et al., 2012; Van Poecke et al., 2013). The SNP and DArT markers are small sequence of DNA. This feature allowed the possibility of comparing these DNA sequences with the public database (Mérida-García et al., 2019; Tura et al., 2019; Xu et al., 2019; Mangini et al., 2020; Pradhan et al., 2020; Rahman et al., 2020). In this study, the public databases used for the identification of the candidate genes included in QTLs controlling grain yield and for the SNPs annotation were: URGI (Unité de Recherche Génomique Info) Blast of the Institut National de la Recherche Agronomique (INRA, France) and NCBI (National Center for Biotechnology Information, the USA).

2. Materials and methods

2.1. QTLs dissection and SNPs annotation

2.1.1. Dissection of QTLs identified on 2A chromosome and candidate genes identification

In this part, we focused on the chromosome 2A due to its strong contribution to grain yield (Li et al., 2007; Sukumaran et al., 2018; Mangini et al., 2021) and because it harbors 4 QTLs among 9 specific QTLs found. The full-length DNA sequence of 2A chromosome was downloaded from URGI and the 4 DNA sequences corresponding to the 4 QTL were isolated based on the flanked markers of these QTL regions using the GeneStudio software. More information about the environments, DNA sequences size, environmental conditions, QTL intervals, and flanking markers are reported in (Farouk et al., 2020) and summarized in Figure 7.

2A chromosome

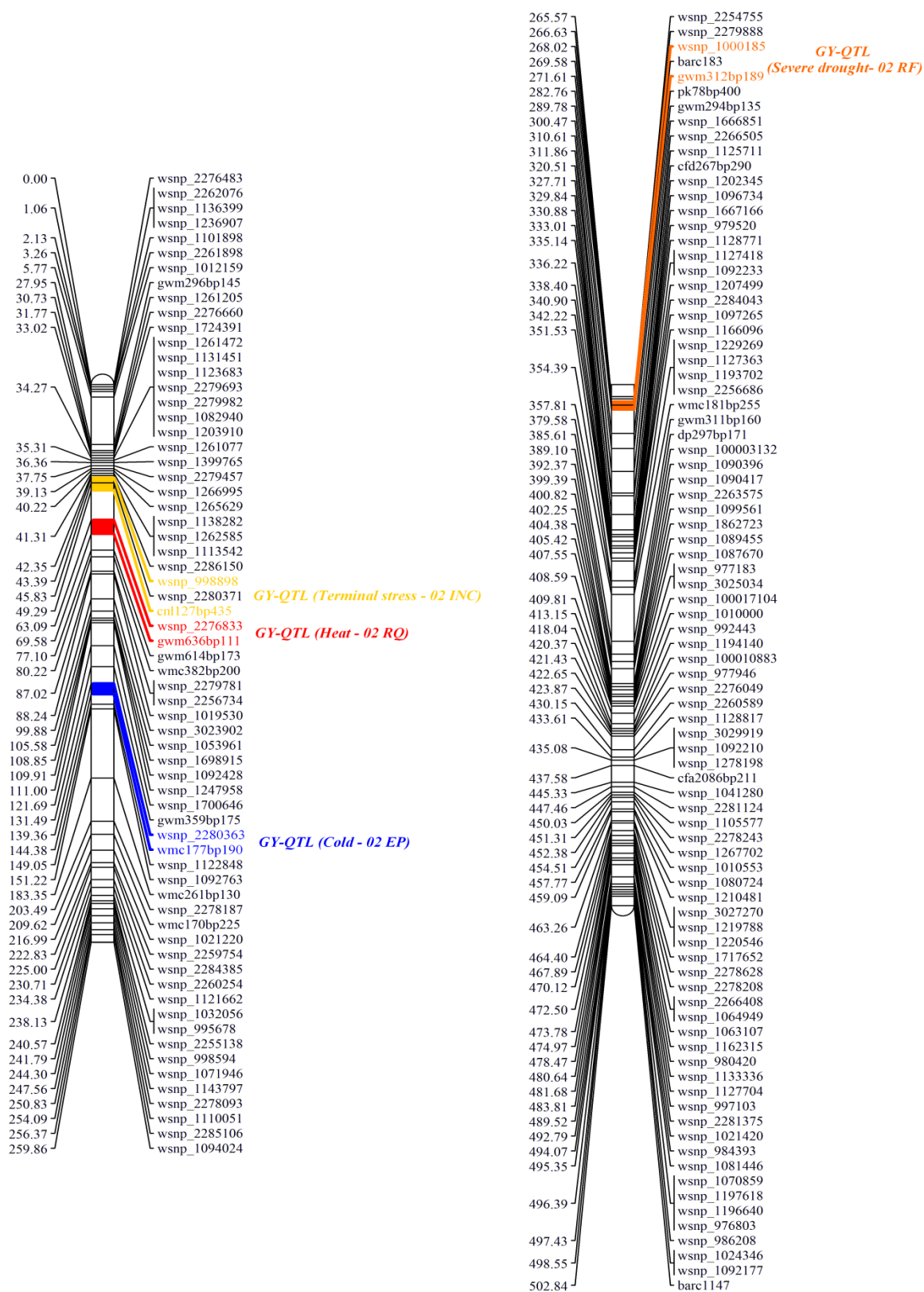


Figure 7: Genetic linkage map 2A chromosome QTLs mapping for grain yield of tolerance to drought, heat, and cold

Thereafter, the sequences isolated were served as queries for Blast search in NCBI against Rice and *Brachypodium* genomes to identify the candidate genes composing these QTLs (Figure 1 in article 2).

The NCBI blast search of DNA sequences of the QTLs specific to the dry, cold, and heat (INC+RF, EP, and RQ) environments show the presence of 407 candidate genes in both Rice and *Brachypodium* databases. Moreover, a set of 139 genes identified in Rice have showed no correspondence in *Brachypodium*, this set of genes is represented by retrotransposons and transposons.

In the dry environments (RF and INC), the QTLs detected were spanned on 4.57 Mb for RF and 2.64 Mb for INC. In RF (dryland), 147 candidate genes were identified and in INC (dryland with incorporation of vetch crop), 376 candidate genes, including genes involved in stress tolerance (biotic and abiotic stress) and in other cellular processes. The retrotransposons and transposons were also present in the two QTLs with a high frequency especially in the RF environment. Also, it was noticed that the incorporation of legume vetch in the soil increased the expression of candidate genes.

In the cold environment (EP) characterized by low temperature during the tillering developmental stage, the QTL detected was covered 1.33 Mb, it contained 17 candidate genes essentially involved in plant development and abiotic/biotic stress tolerance: as Germin-like protein 8-14, protein far1-related sequence 5, putative FBD-associated F-box protein, Photosystem II 44 kDa protein and others. The retrotransposons and transposons are also present in this QTL.

As for the QTL detected in the hot environment (RQ) characterized by high temperatures during the anthesis developmental stage compared with the other environments, the QTL region was spanned on a small interval of 0.1Mb that contained 6 genes, the first one is an LRR receptor-like serine/threonine-protein kinase gene involved in the biotic stress, the others candidate genes are a retrotransposons and transposons.

2.1.2. SNPs annotation

As mentioned above, 126 QTLs with additive-by-environment (A by E) interaction effects covering all chromosomes were detected for GY. A total of 583 SNPs were selected from the QTLs regions and analyzed through Blast (Basic Local Alignment Search Tool) homology search using the URGI Blast of INRA France

(<https://urgi.versailles.inra.fr/blast/blast.php>) with the aim to expand the SNP sequences in terms of length (Colasuonno et al., 2013). The derived genomic sequences (280bp) were served as query using NCBI Blast homology search to identify the candidate genes associated to the selected SNPs see Figure 2 in article 2.

3. Results

3.1. QTLs dissection and candidate genes identification

The dissection of the QTLs identified at 2A chromosome in dry, cold, and heat environments show the presence of 407 candidate genes in Rice as in *Brachypodium* databases. The candidate genes number found for both organisms and the QTL sizes in (Mb) are summarized in the diagram bellow. Moreover, in the Rice database another set of 139 genes identified have showed no correspondence in *Brachypodium* database were retrotransposons and transposons. The results of the candidate genes identification are summarized in the Figure 8.

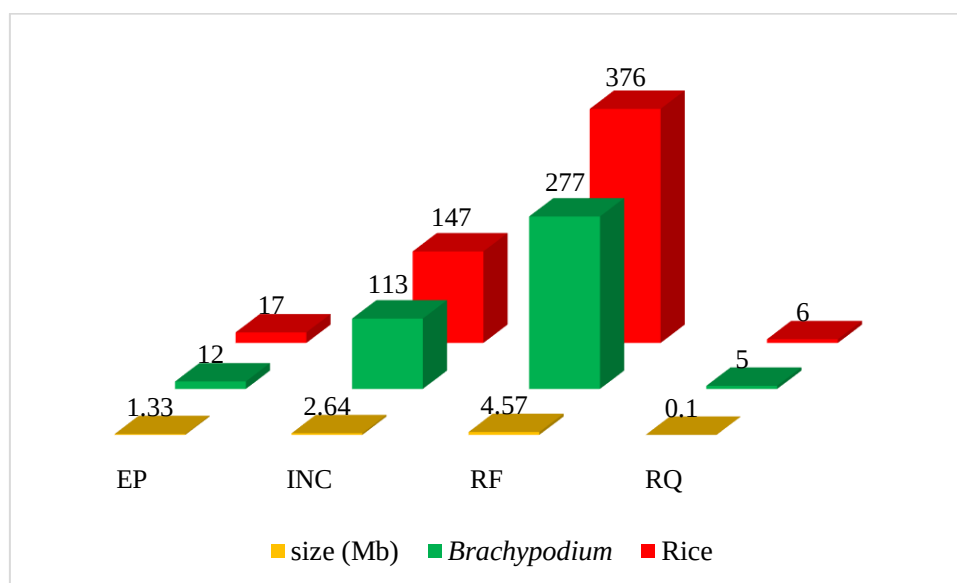


Figure 8: Number of candidate genes identified in the cold (EP), dry (INC and RF), and hot (RQ) environments in comparison with Rice and *Brachypodium* with the DNA sequence of the QTL sizes in Mega base (Mb)

The functional analysis shows that a part of candidate genes identified are involved essentially in tolerance to biotic and abiotic stresses, in plant growth and development, and a large part belong to the family of retrotransposon and transposon.

3.2. SNPs annotation

The results show that 265 present similarities in one or in both databases used (167 in URGI, 25 in NCBI and 73 in both data bases); however, these similarities do not correspond to any gene candidate genes (Figure 3 in article 2). For the 122 sequences annotated where we have identified candidate genes involved in different processes, 77 are enlarged using URGI database and annotated using NCBI, the remaining 45 SNPs sequences were annotated directly using NCBI, due to no-similarity presence in URGI for SNP enlargement, see Figure 4 in article 2. 196 did not show any similarity nor in URGI nor in NCBI databases, these sequences are not annotated.

Moreover, the annotation results show that 53% of candidate genes found were related to stresses tolerance, with 44.6% of them to biotic stress [powdery mildew, stem rust (sr32), leaf rust (lr10, lr34)], 32.3% to abiotic stress (cold heat, drought, salinity) and 23.1% to both stresses; 29.5% to plant development and growth; 3.3% to cell transport, and 1.6% to retrotransposon and transposon; and only 2.4% were with unknown function.

Further, 9.8% were involved in other cellular processes (secondary metabolism, pigments, herbicide detoxification, wall formation, etc.), see Figure 4 in article 2.

Some candidate genes are involved in several processes, such as the basic helix-loop-helix (bHLH) proteins pertaining to the superfamily of transcription factors (TFs) and linked with wsnp_3026224. These proteins have been linked with the regulation of diverse biological processes such as growth, development, and response to diverse stresses.

Moreover, in the 4BS, 57.1% of candidate genes found, were linked to stresses tolerance and 42.9% in plant development and growth (Figure 9).

4B chromosome

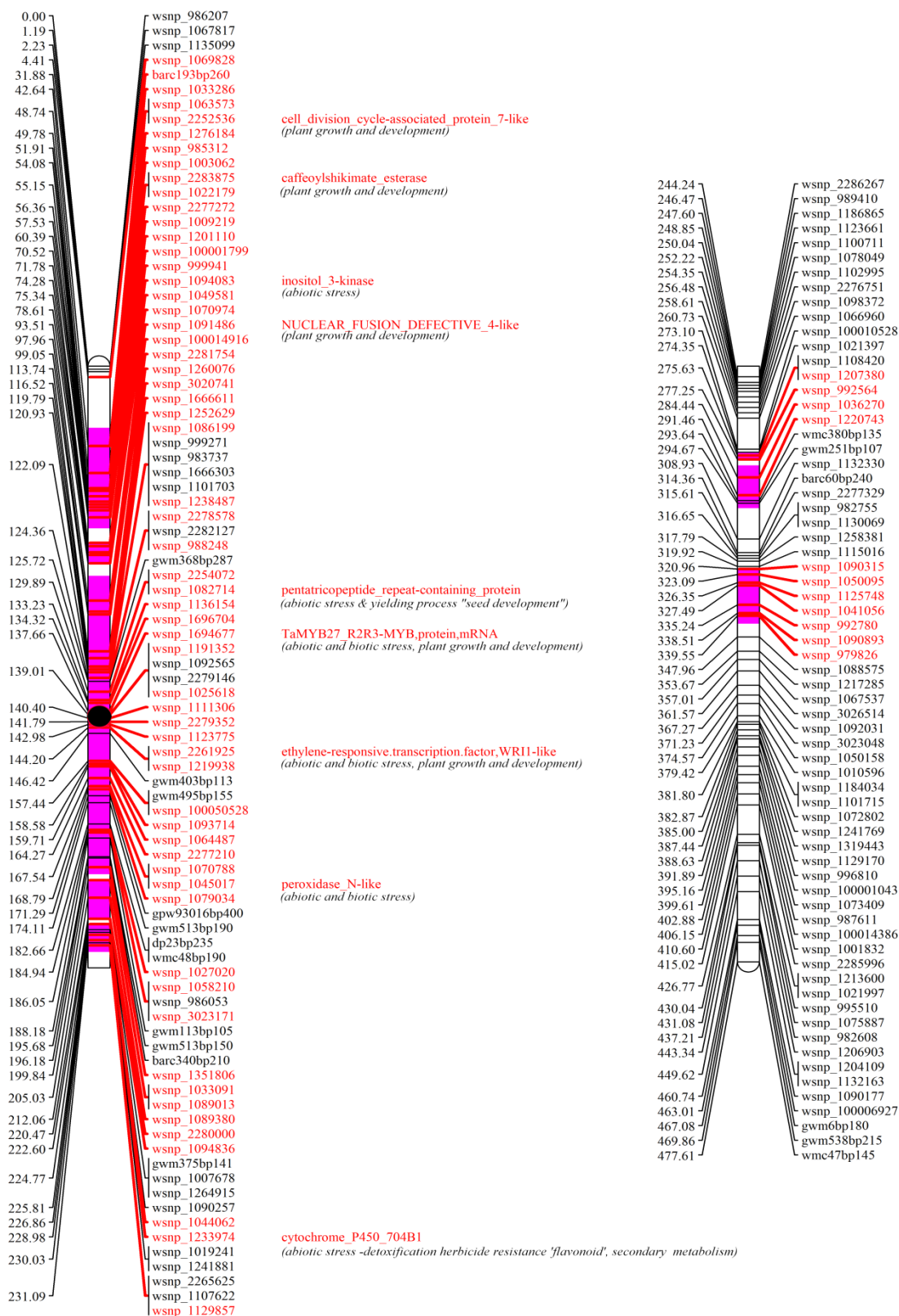


Figure 9: Genetic linkage map of 4B chromosome with QTLs and SNPs annotation

Dissection of Quantitative Trait Loci (QTL), annotation of Single Nucleotide Polymorphism (SNP), and Identification of Candidate Genes for Grain Yield in *Triticum turgidum* L. var durum

Issame Farouk^{*1}, Ahmad Alsaleh², Jihan Motowaj³, Fatima Gaboun⁴, Bouchra Belkadi¹, Abdelkarim Filali Maltouf¹, Zakaria Kehel³, Ismahane Elouafi⁵, Nasserelhaq Nsarellah⁴, Dimah Habash⁶, M. Miloudi Nachit³

¹Laboratory of Microbiology and Molecular Biology, Department of Biology, Faculty of Sciences, Med V University, Rabat, 10000, Morocco

²Dept. Science and Technology Bozok University, Yozgat, 66100, Turkey

³ICARDA. Av. Med Belarbi Alaoui, BP6299, Al Irfane, Rabat, 10000, Morocco

⁴INRA. Institut National de Recherche Agronomique Av. Med Belarbi Alaoui, Al Irfane, Rabat, 10000, Morocco

⁵ICBA. Dubai, 14660, United Arab Emirates

⁶Securewheat, AL3, St. Albans, UK

ARTICLE INFO

Article history:

Received: 29 June, 2020

Accepted: 04 October, 2020

Online: 20 October, 2020

Keywords:

Durum wheat

SNP

Candidate genes

Grain yield

QTL

Linkage map

ABSTRACT

Durum wheat (*Triticum turgidum* L. var durum) is among the most important crops in the world. High and stable grain yield in diverse environments is the major objective in durum breeding programs. This trait is linked to the quantitative trait loci (QTL). For the detection of QTL linked to the grain yield, it is necessary to construct a high-density genetic linkage map. The aims of this study were to detect the candidate genes comprised in the QTLs on 2A chromosome linked to grain yield and annotate the single nucleotide polymorphisms (SNPs). The linkage map of Lahn/Cham1 population was used to identify QTLs. In multi-environmental analysis and employing bioinformatic approaches, 583 sequences corresponding to the SNPs markers selected from the detected QTL regions were analyzed, 122 SNP sequences were annotated of which 53% of the candidate genes were involved in stresses tolerance, 29.5% in plant development and growth, and 3.3% in cell transport. Moreover, 1.6% of candidate genes were retrotransposon and transposon 2.4% with unknown function. Further 9.8% were related in other cellular processes. The results also showed that 66.7% of the candidate genes harbored on 4B chromosome, were involved in stresses tolerance and 33.3% in plant development and growth. Additionally, in the specific and stressed environments analysis, the DNA sequences of the four QTL detected on 2A chromosome were used for homology search, 546 candidate genes were identified of which some were present in several QTL (F-box gene family, hydroxyproline-rich glycoprotein-like), retrotransposons and transposons and others. This study provided information on employing SNPs markers to detect candidate genes linked with grain yield trait in durum wheat in contrasting environments (dry, cold, hot).

1. Introduction

Durum wheat (*Triticum turgidum* L. var durum) is a tetraploid wheat, its genome is constituted by seven homoeologous groups of chromosomes A and B. Its grain use is reported in [1, 2]. The

nutritional values for durum grain were published in earlier research works [1-3]. Grain yield is related to the QTL additive effects and to their interactions with the environment. Several wheat genetic linkage maps were constructed with various molecular markers, such as Restriction Fragment Length Polymorphism (RFLP), Amplified Fragment Length

^{*}Corresponding Author: Issame Farouk, Email: issame.farouk@gmail.com

www.astesj.com

<https://dx.doi.org/10.25046/aj0505125>

Polymorphism (AFLP), Simple Sequence Repeats (SSR) and other markers [4-13]. The construction of saturated maps with large number of molecular markers of Single Nucleotide Polymorphism (SNP) [14-16] and Diversity Array Technology (DArT) [10, 17-23] are fundamental to detect candidate genes for desirable traits in crops, such as durum wheat. Both markers (SNP and DArT) are small sequence of DNA. These sequences can be compared with the public databases [URGI (Unité de Recherche Génomique Info) Blast of the Institut National de la Recherche Agronomique (INRA, France) and NCBI (National Center for Biotechnology Information, the USA)] to identify the candidate genes involved in a QTL controlling a specific trait. Recently the markers (SNP and DArT) were used for QTLs regions detection and for the identification of the candidate genes involved in these QTLs of wheat [24-29]. The high and stable grain yield across diverse environments constitutes the principal aim for durum breeding programs.

The aims of this work were to: 1) identify the SNPs mapped into QTL regions; 2) extend the SNPs sequence analysis in terms of functional and comparative analysis; and 3) identify the candidate genes constituting the QTLs harbored by 2A chromosome and linked to grain yield.

2. Materials and methods

2.1. Plant materials

The genetic population used in this study (Lahn/Cham1) was published in [8, 30] and presented in the conference [31]. The population was phenotyped for grain yield in 11 contrasting environments (cold / drought / heat). The characteristics of the parents: Lahn and Cham1 were described in earlier paper [8, 30].

The Lahn/Cham1 genetic map that consisted of 1425 SNPs, 216 SSR and 31 other molecular markers was constructed using IciMapping program v 4.1.0.0 was presented in [31].

The QTLs linked to grain yield detected using this map were used for candidate genes identification and SNPs annotation.

2.2. SNPs annotation and candidate genes identification

The methodology used for the SNPs annotation and the candidate genes identification is as flowing:

For the QTLs detection in specific and stressed environments, 9 QTLs were detected, 4 of which were localized on 2A, due to its strong contribution to grain yield [32, 33], the full DNA sequence of 2A chromosome were downloaded from URGI and the 4 DNA sequences corresponding to the 4 QTL were isolated based on markers flanking of these QTL regions [24, 25, 29] using the GeneStudio software. More informations about the environments, DNA sequences size, environmental conditions, QTL intervals and flanking markers are summarized in Table 1. Thereafter, the sequences isolated were served as queries for Blast search in NCBI against Rice and *Brachypodium* genomes to identify the candidate genes composing these QTLs detected in the four environments (EP, INC, RF, and RQ), see Figure 1

In multi- environmental analysis, the 583 SNPs selected from the QTLs regions and harbored by durum chromosomes [31] were analyzed through Blast (Basic Local Alignment Search Tool)

search homology using the URGI Blast of (INRA), France (<https://urgi.versailles.inra.fr/blast/blast.php>) with the aim to expand the SNP sequences in terms of length [34] (Figure 2).

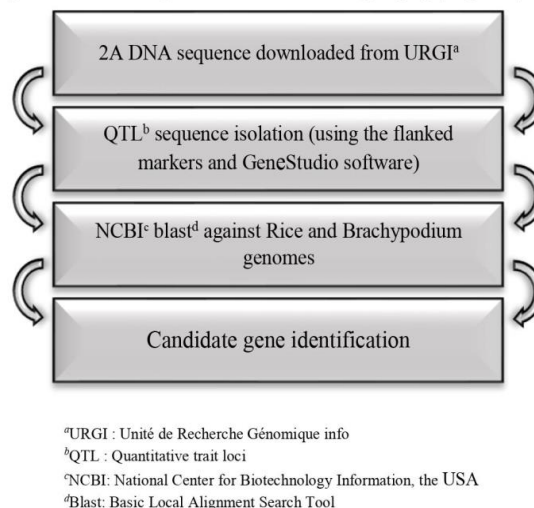


Figure 1: Candidate genes identification in specific and stressed environments

Subsequently, the derived genomic sequences (280bp) were served as query to identify the candidate genes using Blast search homology in NCBI (National Center for Biotechnology Information, the USA), see Figure 2. The Venn diagram was used to present the results.

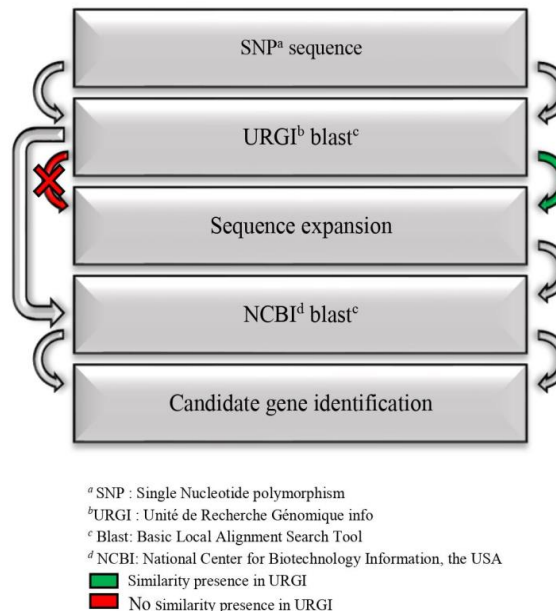


Figure2: Candidate genes identification in multi-environmental analysis

More details about the candidate genes identified in combined and in specific environments QTLs analysis are shown in supplementary materials S1 and S2, respectively.

I. Farouk et al. /Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 5, 1020-1027 (2020)

Table 1: QTL detected on 2A chromosome with the environmental conditions, QTL peak positions in cM, QTL regions interval size in centimorgan (cM) and in Mega bases (Mb) and with QTLs region flanking markers.

Environment	Environmental Conditions	Peak position (cM)	QTL ^a interval size		Flanking markers	
			cM	Mb	Left marker	Right marker
EP ^b	Cold	142	5.02	1.33	wsnp ^f _2280363	wmc ^g 177bp190
INC ^c	Drought (vetch incorporation)	046	5.90	2.64	wsnp ^f _998898	cnl ^h 127bp435
RF ^d	Drought	270	3.59	4.57	wsnp ^f _1000185	gwm ⁱ 312bp189
RQ ^e	Heat	063	6.49	0.10	wsnp ^f _2276833	gwm ⁱ 636bp111

^aQTL: Quantitative Trait Loci

^bEP: Early planting at Tel Hadya / Syria (ICARDA main station), 36°0.1'N 36°56'E, 284m, the sowing is in mid-October, 330mm of rainfall and 70mm of irrigation at sowing, the temperatures were 2 °C at Tillering and 20 °C at flowering

^cINC: Incorporated in field (vetch biomass incorporation in soil) at same station, the sowing date in Mid-November with 330mm as rainfall, the temperatures were 8 °C at tillering and 24 °C at flowering

^dRF: Rainfed (vetch above biomass harvested, no incorporation in soil) at the same station, sowing in Mid-November, the rainfall temperatures are also the same
^eRQ: Raqqa station (35°57' N 39°0'E, 295m) with chickpea rotation, sowing in Mid-November, the rainfall was 150mm and 450 mm of irrigation, were 10 °C at tillering and 30 °C at flowering

^fSNP: Single Nucleotide Polymorphism

^gwmc: Wheat Microsatellite Consortium (SSR)

^hcnl: Cornell University (EST-SSR)

ⁱgwm: Gatersleben Wheat Microsatellite (SSR)

3. Results

3.1. Candidate genes detection in specific and stressed environments

The search for candidate genes in the QTLs corresponding to the environments (dry / cold / heat) show the presence of 407 candidate genes in Rice as in *Brachypodium* databases. Moreover, in the Rice database 139 genes identified were retrotransposons and transposons that have showed no correspondence in *Brachypodium* database.

3.1.1. Dry environment

As for the dry environments (RF and INC), the QTLs detected in these environments (Table 2) spanned on 4.57 Mb for RF and 2.64 Mb for INC. In RF (dryland) were identified 147 candidate genes and in INC (dryland with incorporation of vetch crop) 376 candidate genes, including genes involved in stress tolerance (biotic and abiotic stress) and in other cellular processes, the retrotransposons and transposons were also present in the two QTLs with a high frequency especially in the RF environment

(Supplementary: Table S1). Also, it was noticed that the incorporation of legume vetch in the soil increased the expression of candidate genes. This expression is higher in dryland with vetch incorporation in soil (INC) than in dryland without vetch incorporation in the soil (RF). This may be due to increased availability of nitrogen and water through the organic matter incorporation in the soil.

3.1.2. Cold environments

In the cold environment (EP) characterized by low temperature during the tillering developmental stage (Table 2), the QTL detected was spanned on 1.33 Mb, it contained 17 candidate genes among those involved in plant development and abiotic/biotic stress tolerance: as Germin-like protein 8-14, protein far1-related sequence 5, putative FBD-associated F-box protein, Photosystem II 44 kDa protein and others. The retrotransposons and transposons are also present in this QTL (Supplementary: Table S1).

Table 2: The environmental conditions, the number of candidate genes identified for the four QTL linked to grain yield of the four stressed environments (EP, INC, RF and RQ) with the QTL interval size, and the number of candidate genes identified for Rice and *Brachypodium*

Environment	Environmental Conditions	QTL ^a Interval size (Mb)	Candidates genes number	
			Rice	<i>Brachypodium</i>
EP ^b	Cold	1.33	017	012
INC ^c	Drought (vetch incorporation)	2.64	147	113
RF ^d	Drought	4.57	376	277
RQ ^e	Heat	0.10	006	005

^aQTL: Quantitative Trait Loci

^bEP: Early planting at Tel Hadya / Syria (ICARDA main station), 36°0.1'N 36°56'E, 284m, the sowing is in mid-October, 330mm of rainfall and 70mm of irrigation at sowing, the temperatures were 2 °C at Tillering and 20 °C at flowering

^cINC: Incorporated in field (vetch biomass incorporation in soil) at same station, the sowing date in Mid-November with 330mm as rainfall, the temperatures were 8 °C at tillering and 24 °C at flowering

^dRF: Rainfed (vetch above biomass harvested, no incorporation in soil) at the same station, sowing in Mid-November, the rainfall temperatures are also the same

^eRQ: Raqqa station (35°57' N 39°0'E, 295m) with chickpea rotation, sowing in Mid-November, the rainfall was 150mm and 450 mm of irrigation, were 10 °C at tillering and 30 °C at flowering

Table 3: SNPs annotated and not annotated in A and B genomes

Genome	SNPs			
	Annotated	Not annotated	Total	Percentage of SNPs annotated (%)
1A	15	24	39	38.46
2A	17	63	80	21.25
3A	6	33	39	15.38
4A	8	31	39	20.51
5A	1	11	12	08.33
6A	5	14	19	26.31
7A	7	46	53	13.20
A genome	59	222	281	20.99
1B	17	35	52	32.69
2B	16	63	79	20.25
3B	2	10	12	16.66
4B	9	64	73	12.32
5B	5	28	33	15.15
6B	13	29	42	30.95
7B	1	10	11	09.10
B genome	63	239	302	20.86

3.1.3. Hot environments

Further, the QTL detected in the hot environment (RQ) which is characterized by high temperatures during the anthesis developmental stage compared with the other environments (Table 2), the QTL region was spanned on a small interval of 0.1Mb that contained 6 genes, the first one is an LRR receptor-like serine/threonine-protein kinase gene involved in plant resistance to pathogens and the others are a retrotransposons and transposons (Supplementary: Table S1).

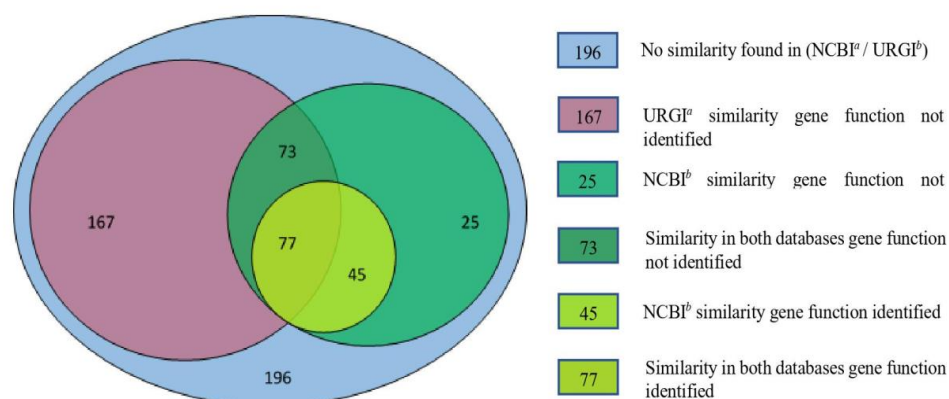
3.2. Mapping of SNPs to identify candidate genes

Of the mapped SNPs, 80 were mapped on 2A. Additional information for other chromosomes is shown in Table 3.

As for the candidate genes number, the 2A along 1B harbored the highest number (17), while in terms of annotation percentage for other chromosomes see Table 3.

The SNPs sequences availability provides the possibility of identifying gene loci contributing to grain yield.

Among the 583 SNP sequences selected, 122 sequences were annotated. the candidate genes involved in various growth, development processes, and stress tolerances have been identified, of which 77 sequences were enlarged with URGI database and annotated with NCBI. Whereas the remaining 45 SNPs sequences were annotated directly with NCBI, as no-similarity presence was available to enlarge in URGI database (Figure 3). Several sequences did not show any similarity in URGI and NCBI databases



^aURGI : Unité de Recherche Génomique info

^bNCBI: National Center for Biotechnology Information. The USA

Figure 3: Annotation and similarity study on Venn diagram of SNP.

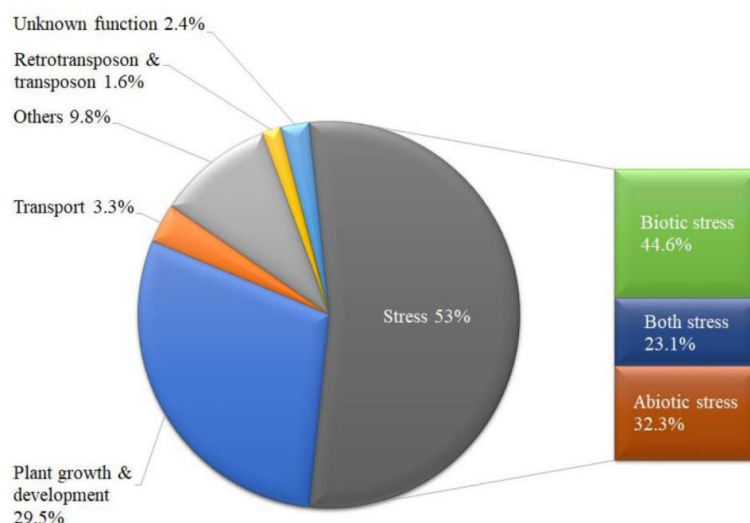


Figure 4: Functional classification of candidate genes co-located with QTLs linked to grain yield.

The 583 SNPs marked on Venn diagram (Figure 3) showed that 77 sequences were annotated (candidate genes identified) with similarity presence with both databases; while 45 sequences were annotated with similarity presence only in NCBI. Further, 73 were not annotated but had similarity in both databases. Also 167 SNP had similarity only in URGI and 25 only in NCBI, consequently both sets of SNPs were not annotated. The remaining 196 sequences were not annotated, showing no similarity in the two databases.

The analysis with bioinformatic tools showed that 53% of candidate genes found were related to stresses tolerance, with 44.6% of them to biotic stress [powdery mildew, stem rust (sr32), leaf rust (lr10, lr34)], 32.3% to abiotic stress (cold heat, drought, salinity) and 23.1% to both stresses; 29.5% to plant development and growth; 3.3% to cell transport, and 1.6% to retrotransposon and transposon; and only 2.4% were with unknown function.

Further, 9.8% were involved in other cellular processes (secondary metabolism, pigments, herbicide detoxification, wall formation, etc.), see Figure 4. More informations about the SNPs annotation and candidate gene functions were shown in the supplementary table (S2).

Some candidate genes are involved in several processes, such as the basic helix-loop-helix (bHLH) proteins pertaining to the superfamily of transcription factors (TFs) and linked with wsnp_3026224. These proteins have been linked with the regulation of diverse biological processes as growth, development, and response to diverse stresses.

Moreover, in the 4BS, 57.1% of candidate genes found, were linked to stresses tolerance and 42.9% in plant development and growth.

www.astesj.com

4. Discussion

4.1. Candidate genes detection in specific and stressed environments

We found among 9 QTLs related to grain yield, four were located on 2A chromosome, these QTLs were detected in diverse stressed environments: EP characterized by cold, INC and RF by drought, and RQ by heat. These results are in accordance with several studies on grain yield QTLs located on 2A chromosome [24, 32, 33].

4.1.1. Dry environment

In the dry environment (INC), among the candidate genes were serpin-Z5, Probable glucuronosyltransferase Os04g0103100, Chalcone synthase 1, Fatty acid amide hydrolase isoform X1, Amino acid permease 3 (Supplementary: Table S1). This is in the agreement with the study of [30] where the transcriptomic pattern under progressive water stress of the two parents: Lahn and Cham1 (parents of the current studied population) and one recombined inbred line where the candidate genes cited above showed an overexpression. Such as the candidate gene coding for the mitochondrial metalloendopeptidase OMA1 identified in the QTL of the dry environment RF (Supplementary: Table S1). This was also confirmed by [30]. Similarly, the F-box and associated protein family were overexpressed in the dry environments RF and INC (Supplementary Table S1), but under cold stress was downexpressed [35]. Some genes of this large family under abiotic stress may be over or underexpressed depending on the stress [36]. As for the candidate gene coding for the Hydroxyproline-rich glycoprotein-like identified in the QTLs detected in the dry and cold environments: RF, INC, and EP

1024

(Supplementary: Table S1), this gene was associated with drought tolerance in barley [37] and with the resistance to bacterial pathogen in rice [38].

Further, other candidate genes identified in the dry environment (RF): GDSE esterase/lipase EXL3; the cold environment (EP): UDP-glycosyltransferase 79-like isoform, and the glutathione S-transferase identified in both environments (Supplementary: Table S1). These genes play role in plant development and growth through the action on ABA homeostasis [39]; germination and growth [40]; detoxification and plant pathogen resistance [41].

4.1.2. Cold environment

In the cold environment (EP), among the candidate genes composing the QTL detected: 5'-3' exoribonuclease 3 isoform X1, NADH dehydrogenase; as for the QTL of dry environment (RF), it contains the genes: starch synthase3, chloroplastic/amyloplastic and a photosystem I P700 chlorophyll an apoprotein A1 candidate gene. While the gene coding for the Photosystem II 44 kDa protein were identified in the QTLs for the cold (EP) and dry (RF) environment, as for the ankyrin repeat family protein-like it was found in the INC environment (Supplementary Table S1). All these candidate genes expression was altered under water stress [30]. Some candidate genes identified in QTLs for INC and RF environments like zinc finger family genes play an important role in physiological traits of yield potential under heat stress [27]. Moreover, in the QTL for the RF environment, the gene coding for BTB/POZ and MATH domain-containing protein 2 was highly repeated in this QTL, this gene is involved in heat tolerance and in plant growth and development [42].

4.1.3. Hot environment

In the hot environment RQ, six candidate genes were identified, one coding for LRR receptor-like serine/threonine-protein kinase RCH1 involved in resistance to pathogens [43, 44]. The remaining candidate genes were belonging to the family of retrotransposon/transposase. This gene family was also present in the other QTLs especially in RF (dry environment) with a high frequency. The abundance of this gene family presence can be explained by their involvement in stress conditions, [30] reported that the activity of these genes was altered under water stress and by the fact that the retrotransposon/transposon may represents 80 to 90% of wheat genome [45].

4.2. Detection of candidate genes in other chromosomes

Additionally, the presence of the wsnp_993293 SNP marker in the chromosome 1A and its linkage with the candidate gene implicated in powdery mildew resistance, confirms earlier mention on one parent of the population (Lahn) that carries moderate resistance to powdery mildew. Additionally, the wsnp_2265986, located in the 1B chromosome and linked to 3-ketoacyl-CoA synthase 6-like, this candidate gene plays also a role in abiotic (drought, heat and cold) and in biotic stress resistance. This reflects the genes carried by the second parent Cham1 for

good drought and heat tolerance, osmotic adjustment, and proline content.

The candidate genes covered the entire genome and were linked to different biotic (powdery mildew, rust, etc.) and abiotic (drought, cold, heat, etc.) stress resistance / tolerance (Fig 4).

In the current work, a large part of candidate genes harbored by the 2A, in addition to 4B chromosome, were related to stress tolerance. They were also involved in plant growth and development. The 4BS results corroborate earlier studies [25, 46-50] on the importance of the 4BS and its contribution to grain yield in stressed environments.

5. Conclusion

The candidate genes found were linked to yield performance under drought and temperature extreme conditions. Also, this study showed the importance of 2A chromosomes for contrasting environments (drought, cold, and heat).

In this study specific and general candidate genes were detected. For the specific environments, the candidate genes in cold environment (EP) involved the genes of 5'-3' exoribonuclease 3 isoform X1 and NADH dehydrogenase. As for the dry environment (RF) the genes included were starch synthase3, chloroplastic/amyloplastic, and a photosystem I P700 chlorophyll. Further, in the dry environment (RF), the gene coding for BTB/POZ and MATH domain-containing protein2 was highly repeated in this environment. This gene is also effectuates heat tolerance, plant development and growth. Similarly, the ankyrin repeat family protein-like was found in the dry environment (INC), it contains the genes synthase3, chloroplastic/amyloplastic and photosystem I P700 chlorophyll. For the general candidate genes involved in diverse stressed environments, the following genes were identified: The Photosystem II 44 kDa protein were identified in the cold (EP) and the dry (RF) environments. Also, some of the candidate genes are operating in all 3 stresses, such as the zinc finger family genes. This gene also plays a significant role in physiology of yield potential under heat stress frequently.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgement

We would like to thank the International Center for Agricultural Research in Dry Area (ICARDA) for the financial support of this study and the support staff of the CIMMYT/ICARDA durum-breeding program.

References

- [1] I. Elouafi, MM. Nachit, "A genetic linkage map of the Durum x *Triticum dicoccoides* backcross population based on SSRs and AFLP markers, and QTL analysis for milling traits" Theor Appl Genet, **108**(3), 401-413, 2004. <https://doi.org/10.1007/s00122-003-1440-8>
- [2] MM. Nachit, "Durum wheat breeding for Mediterranean drylands of North Africa and West Asia. In: Rajram S, Saari EE, Hetel GP (eds) Durum Wheats" Challenges and Opportunities. CIMMYT, Ciudad Obregon, Mexico, **14**, 14-27, 1992.

- [3] El Yadini, I. Kahama, M. Labhili, A. El Yadini, M. Azeqour, "La technique TILLING comme nouvelle biotechnologie pour la protection et l'amélioration du blé dur (TILLING : A new biotechnological method for protection and improvement of durum wheat)" J Mater Env Sci, **4**, 931-934, 2013.
- [4] M. Maccaferri, M. Sanguineti, S. Corneti, J. Ortega, M. Salem, J. Bort, E. DeAmbrogio, L. del Moral, A. Demontis, A. El-Ahmed, "Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.) across a wide range of water availability" Genetics, **178**(1), 489-511, 2008. <https://doi.org/10.1534/genetics.107.077297>
- [5] B. Heidari, BE. Sayed-Tabatabaei, G. Saeidi, M. Kearsy, K. Suenaga, P. Gulick, "Mapping QTL for grain yield, yield components, and spike features in a doubled haploid population of bread wheat" Genome, **54**(6), 517-527, 2011. <https://doi.org/10.1139/g11-017>
- [6] A. Blanco, M. Bellomo, A. Cenci, C. De Giovanni, R. D'ovidio, E. Iacono, B. Laddomada, M. Pagnotta, E. Porceddu, A. Sciancalepore, "A genetic linkage map of durum wheat" Theor App Genet **97**(5-6), 721-728, 1998. <https://doi.org/10.1007/s001220050948>
- [7] SA. Dura, MA. Duwayri, MM. Nachit, "Detection of molecular markers associated with yield and yield components in durum wheat (*Triticum turgidum* L. var. durum Desf.) under drought conditions" Afr J Agr, **8**(19), 2118-2128, 2013. <https://doi.org/10.5897/AJAR11.2031>
- [8] MM. Nachit, I. Elouafi, A. Pagnotta, A. El saleh, E. acono, M. Labhili, A. Asbati, M. Azrak, H. azzam, D. Bensch, M. Khairallah, JM. Ribaut, OA. Tanzarella, E. Porceddu, ME. Sorrells, "Molecular linkage map for an intraspecific recombinant inbred population of durum wheat (*Triticum turgidum* L. var. durum)". Theor App Genet, **102**, 177-186, 2001. <https://doi.org/10.1007/s001220051633>
- [9] A. Alsaleh, FS. Baloch, M. Derya, M. Azrak, B. Kilian, H. Özkan, MM. Nachit, "Genetic Linkage Map of Anatolian Durum Wheat Derived from a Cross of Kunduru-1149× Chaml". Plant Mol Biol Rep, **33**, 209-220, 2014. <https://doi.org/10.1007/s11105-014-0749-6>
- [10] Mantovani, M. Maccaferri, MC. Sanguineti, R. Tuberosa, I. Catizone, P. Wenzl, B. Thomson, J. Carling, E. Huttner, E. DeAmbrogio, "An integrated DArT-SSR linkage map of durum wheat" Mol Breed, **22**(4), 629-648, 2008. <https://doi.org/10.1007/s11032-008-9205-3>
- [11] M. Nagel, S. Navakode, V. Scheibal, M. Baum, MM. Nachit, M. Röder, A. Bömer, "The genetic basis of durum wheat germination and seedling growth under osmotic stress" Biol Plantarum, **58**, 681-688, 2014. <https://doi.org/10.1007/s10535-014-0436-3>
- [12] KJ. Haile, MM. Nachit, H. Karl, B. Ayele, SM. Röder, "QTL mapping of resistance to race Ug99 of Puccinia graminis f. sp. tritici in durum wheat (*Triticum durum* Desf.)" Mol Breed, **30**, 1479-1493, 2012. <https://doi.org/10.1007/s11032-012-9734-7>
- [13] S. Dura, M. Duwayri, MM. Nachit, F. Al Sheyab, "Detection of molecular markers associated with yield and yield components in durum wheat (*Triticum turgidum* L. var. durum) under saline conditions" Crop Past Sci, **64**(10), 957-964, 2013. <https://doi.org/10.5897/AJAR11.2031>
- [14] RM. Poেকে, M. Maccaferri, J. Tang, HT. Truong, A. Janssen, NJ. Orsouw, S. Salvi, MC. Sanguineti, R. Tuberosa, EA Vossen, "Sequence based SNP genotyping in durum wheat" Plant Biotech J, **11**(7), 809-817, 2013. <https://doi.org/10.1111/pbi.12072>
- [15] D. Trebbi, M. Maccaferri, P. de Heer, A. Sørensen, S. Giuliani, S. Salvi, M. Sanguineti, A. Massi, E. van der Vossen, R. Tuberosa, "High-throughput SNP discovery and genotyping in durum wheat (*Triticum durum* Desf.)" Theor App Genet, **123**(4), 555-569, 2011. <https://doi.org/10.1007/s00122-011-1607-7>
- [16] W. Zhang, S. Chao, F. Manthey, O. Chicaiza, JC. Brevis, V. Echenique, J. Dubcovsky, "QTL analysis of pasta quality using a composite microsatellite and SNP map of durum wheat" Theor App Genet, **117**(8) 1361-1377, 2008. <https://doi.org/10.1007/s00122-008-0869-1>
- [17] A. Blanco, P. Colasuonno, A. Gadaleta, G. Mangini, A. Schiavulli, R. Simeone, AM. Digesù, P. De Vita, AM. Mastrangelo, L. Cattivelli, "Quantitative trait loci for yellow pigment concentration and individual carotenoid compounds in durum wheat" J Cereal Sci, **54**(2), 255-264, 2011. <https://doi.org/10.1016/j.jcs.2011.07.002>
- [18] A. Blanco, G. Mangini, A. Giancaspro, S. Giove, P. Colasuonno, R. Simeone, A. Signorile, P. Vita, AM. Mastrangelo, L. Cattivelli, "Relationships between grain protein content and grain yield components through quantitative trait locus analyses in a recombinant inbred line population derived from two elite durum wheat cultivars" Mol Breed, **30**(1), 79-92, 2011. <https://doi.org/10.1007/s11032-011-9600-z>
- [19] A. Gadaleta, A. Giancaspro, S. Giove, S. Zacheo, G. Mangini, R. Simeone, A. Signorile, A. Blanco, "Genetic and physical mapping of new EST-derived SSRs on the A and B genome chromosomes of wheat" Theor App Genet, **118**(5), 1015-1025, 2009. <https://doi.org/10.1007/s00122-008-0958-1>
- [20] D. Marone, G. Laidò, A. Gadaleta, P. Colasuonno, D. Ficco, A. Giancaspro, S. Giove, G. Panio, M. Russo, P. De Vita, L. Cattivelli, R. Papa, A. Blanco, "A high-density consensus map of A and B wheat genomes" Theor App Genet, **125**(8), 1619-1638, 2012. <https://doi.org/10.1007/s00122-012-1939-y>
- [21] MA. Russo, DBM. Ficco, D. Marone, P. De Vita, V. Vallega, C. Rubies-Autonell, C. Ratti, P. Ferragionio, V. Giovannelli, N. Pecchioni, "A major QTL for resistance to soil-borne cereal mosaic virus derived from an old Italian durum wheat cultivar" J Plant Interact, **7**(4), 290-300, 2012. <https://doi.org/10.1080/17429145.2011.640437>
- [22] A. Verlotta, V. De Simone, AM. Mastrangelo, L. Cattivelli, R. Papa, D. Trono, "Insight into durum wheat Lpx-B1: a small gene family coding for the lipoxigenase responsible for carotenoid bleaching in mature grains" BMC plant biol, **10**, number (263), 1-18, 2010. <https://doi.org/10.1186/1471-2229-10-263>
- [23] Z. Peleg, Y. Saranga, T. Suprunova, Y. Ronin, MS. Röder, A. Kilian, AB. Korol, T. Fahima, "High-density genetic map of durum wheat x wild emmer wheat based on SSR and DArT markers" Theor App Genet, **117**, 103-115, 2008. <https://doi.org/10.1007/s00122-008-0756-9>
- [24] H. Tura, J. Edwards, V. Gahlaut, M. Garcia, Sznajder, U. Baumann, F. Shahinnia, M. Reynolds, P. Langridge, HS. Balyan, "QTL analysis and fine mapping of a QTL for yield-related traits in wheat grown in dry and hot environments" Theor App Genet, **1-19**, 2019. <https://doi.org/10.1007/s00122-019-03454-6>
- [25] D. Xu, W. Wen, L. Fu, F. Li, J. Li, L. Xie, X. Xia, Z. Ni, Z. He, S. Cao, "Genetic dissection of a major QTL for kernel weight spanning the Rht-B1 locus in bread wheat" Theor App Genet, **132**(11), 3191-3200, 2019. <https://doi.org/10.1007/s00122-019-03418-w>
- [26] R. Mérida-García, G. Liu, S. He, V. Gonzalez-Dugo, G. Dorado, S. Gálvez, I. Solís, P. Zarco-Tejada, J. Reif, P. Hernandez P, "Genetic dissection of agronomic and quality traits based on association mapping and genomic selection approaches in durum wheat grown in Southern Spain" PloS one **14**(2), e0211718, 1-24, 2019. <https://doi.org/10.1371/journal.pone.0211718>
- [27] S. Pradhan, MA. Babar, G. Bai, J. Khan, D. Shahi, M. Avci, J. Guo, J. McBreen, S. Asseng, S. Gezan, et al, "Genetic dissection of heat-responsive physiological traits to improve adaptation and increase yield potential in soft winter wheat" BMC Genomics **21**, 315, 1-15, 2020. <https://doi.org/10.1186/s12864-020-6717-7>
- [28] M.S. Rahman, K.J. Linsell, J.D. Taylor, M.J. Hayden, N.C. Collins, K.H. Oldach, "Fine mapping of root lesion nematode (*Pratylenchus thornei*) resistance loci on chromosomes 6D and 2B of wheat" Theor App Genet **133**, 635-652, 2020. <https://doi.org/10.1007/s00122-019-03495-x>
- [29] J. Yu, Y. Miao, S. Yang, Z. Shi, N. Miao, M. Ding, H. Zhang, Y. Jiang, J. Rong, "Identification and mapping of a photoperiod response gene (QPPd.zafu-4A) on wild emmer wheat (*Triticum turgidum* L.) chromosome 4AL" Euphytica **215**, 146, 1-12, 2019. <https://doi.org/10.1007/s10681-019-2469-3>
- [30] D. Habash, M. Baudo, M. Hindle, S. Powers, M. Defoin-Platel, R. Mitchell, M. Saqi, C. Rawlings, K. Latini, J. Araus, et al, "Systems responses to progressive water stress in durum wheat" PloS one, **9**, e108431, 1-21, 2014. <https://doi.org/10.1371/journal.pone.0108431>
- [31] I. Farouk, F. Gaboun, Z. Kehel, A. Alsaleh, B. Belkadi, I. Elouafi, J. Motowaj, A. Filali maltouf, MM. Nachit, "Dissection of QTL linked to grain yield and identification of candidate genes involved in grain yield formation using comparative SNP sequences analysis," in 2020 IEEE the first International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco, pp. 1-6, 2020. <https://doi.org/10.1109/IRASET48871.2020.9091989>
- [32] S. Li, J. Jia, X. Wei, X. Zhang, L. Li, H. Chen, Y. Fan, H. Sun, X. Zhao, T. Lei, et al, "A intervarietal genetic map and QTL analysis for yield traits in wheat" Mol Breeding **20**, 167-178, 2007. <https://doi.org/10.1186/s12870-018-1308-3>
- [33] S. Sukumaran, M.P. Reynolds, C. Sansaloni, "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, 1-16, 2018. <https://doi.org/10.3389/fpls.2018.00081>
- [34] P. Colasuonno, MA. Maria, A. Blanco, Gadaleta, "Description of durum wheat linkage map and comparative sequence analysis of wheat mapped DArT markers with rice and *Brachypodium* genomes" BMC Genet, **14**, number (114), 1-9, 2013. <https://doi.org/10.1186/1471-2156-14-114>
- [35] C. Xin, R. Hou, F. Wu, Y. Zhao, H. Xiao, W. Si, ME. Ali, L. Cai, J. Guo, "Analysis of cytosine methylation status in potato by methylation-sensitive amplified polymorphisms under low-temperature stress" J Plant Biol, **58**, 383-390, 2015.

- <https://doi.org/10.1007/s12374-015-0316-1>
- [36] M. Jain, A. Nijhawan, R. Arora, P. Agarwal, S. Ray, P. Sharma, S. Kapoor, AK. Tyagi, JP. Khurana, "F-Box Proteins in Rice. Genome-Wide Analysis, Classification, Temporal and Spatial Gene Expression during Panicle and Seed Development, and Regulation by Light and Abiotic Stress" *Plant Physiol*, **143**, 1467-1483, 2007.
<https://doi.org/10.1104/pp.106.091900>
- [37] SF. Abou-Elwafa, "Identification of genes associated with drought tolerance in barley" *Biol Plantarum* **62**, 299-306, 2018.
<https://doi.org/10.1007/s10535-017-0765-0>
- [38] IS. Kumar, N. Zaharin, K. Nadarajah, "In silico identification of resistance and defense related genes for bacterial leaf blight (BLB) in rice" *J Pure Appl Microbiol*, **12**, 1867-1877, 2018.
<https://dx.doi.org/10.22207/JPAAM.12.4.22>
- [39] Z. Liu, J-P. Yan, D-K. Li, Q. Luo, Q. Yan, Z-B. Liu, L-M. Ye, J-M. Wang, X-F. Li, Y. Yang, "UDP-Glucosyltransferase71C5, a Major Glucosyltransferase, Mediates Abscissic Acid Homeostasis in Arabidopsis" *Plant Physiol*, **167**, 1659-1670, 2015.
<https://doi.org/10.1104/pp.15.00053>
- [40] C-P. Lai, L-M. Huang, L-FQ. Chen, M-T. Chan, J-F. Shaw, "Genome-wide analysis of GDSL-type esterases/lipases in Arabidopsis" *Plant Mol Biol*, **95**, 181-197, 2017.
<https://doi.org/10.1007/s11103-017-0648-y>
- [41] G. Gullner, T. Komives, L. Király L, P. Schröder, "Glutathione S-Transferase Enzymes in Plant-Pathogen Interactions" *Front Plant Sci* **9**, 1836, 1-19, 2018.
<https://doi.org/10.3389/fpls.2018.01836>
- [42] K. Morimoto, N. Ohama, S. Kidokoro, J. Mizoi, F. Takahashi, D. Todaka, J. Mogami, H. Sato, F. Qin, J-S. Kim, et al, "BPM-CUL3 E3 ligase modulates thermotolerance by facilitating negative regulatory domain-mediated degradation of DREB2A in Arabidopsis" *Proceedings of the National Academy of Sciences* **114**(40), E8528-E8536, 2017.
<https://doi.org/10.1073/pnas.1704189114>
- [43] S. Padmarasu, DJ. Sargent, A. Patocchi, M. Troggo, P. Baldi, G. Linsmith, L. Poles, M. Jansch, M. Kellerhals, S. Tartarini, R. Velasco, "Identification of a leucine-rich repeat receptor-like serine/threonine-protein kinase as a candidate gene for Rvi12 (Vb)-based apple scab resistance" *Mol Breeding*, **38**, 73, 1-14, 2018.
<https://doi.org/10.1007/s11032-018-0825-y>
- [44] D. Pradhan, D. Mathew, SK. Mathew, PA. Nazeem, "Identifying the markers and tagging a leucine-rich repeat receptor-like kinase gene for resistance to anthracnose disease in vegetable cowpea [*Vigna unguiculata* (L.) Walp.]" *J Hort Sci Biotech*, **93**, 225-231, 2018.
<https://doi.org/10.1080/14620316.2017.1362962>
- [45] V. Jamilloux, J. Daron, F. Choulet, H. Quesneville, "De novo annotation of transposable elements: tackling the fat genome issue" *Proceedings of the IEEE*, **105**, 474-481, 2016.
<https://doi.org/10.1109/JPROC.2016.2590833>
- [46] AA. Diab, R. Kantety, N. Ozturk, D. Benscher, MM. Nachit, M. Sorrells, "Drought-inducible genes and differentially expressed sequence tags associated with components of drought tolerance in durum wheat" *Sci Res Essay*, **3**, 009-026, 2008.
- [47] D. Habash, Z. Kehel, M. Nachit, "Genomic approaches for designing durum wheat ready for climate change with a focus on drought" *J Exp Bot*, **60**(10), 2805-2815, 2009.
<https://doi.org/10.1093/jxb/erp211>
- [48] MM. Nachit, "Association of grain yield in dryland and carbon isotope discrimination with molecular markers in durum (*Triticum turgidum* L. var. *durum*)" In *Proceedings of 9th International Wheat Genetics Symposium*, Saskatoon, ed. A.E. Slinkard, Canada, 218-223, 1998.
- [49] MM. Nachit, I.Elouafi, "Durum wheat adaptation in the Mediterranean dryland: Breeding, stress physiology, and molecular markers" In *Challenges and strategies for dry-land agriculture*, eds. S.C. Rao and J. Ryan, CSSA, Madison, WI: Crop Science of America and American Society of Agronomy, **32**, 203-218, 2004.
<https://doi.org/10.2135/cssaspecpub32.c13>
- [50] J. Li, S. Wen, C. Fan, M. Zhang, S. Tian, W. Kang, W. Zhao, C. Bi, Q. Wang, S. Lu, et al, "Characterization of a major quantitative trait locus on the short arm of chromosome 4B for spike number per unit area in common wheat (*Triticum aestivum* L.)" *Theor Appl Genet*, **133**, 2259-2269, 2020.
<https://doi.org/10.1007/s00122-020-03595-z>

Discussion

Discussion

In the irrigated and favorable environments, the grain yield's mean of the RILs used in this study are above the parent's grain yield performance and is approximatively similar to the parents' performance in the average yielding environments of 4000–5000 kg/ha (02RF, 02 and 01INC) and in the low yielding environment of 2000–2500 kg/ha (02BR). Although, in all environments, several RILs have largely outyielded both parents and a record was registered in Tel Hadya in early sowing environment with 13,753 kg/ha (02EP) followed by the favorable/high-input environment with fertile soil (02GH) 11,933 kg/ha, the irrigated environment (01IR) with 11,217 kg/ha, the favorable and high-input environment with fertile soil (01GH) 11,033 kg/ha, and the hot irrigated environment with 11,083 and 10,133 kg/ha (02RQ and 01RQ). These records of grain yield performance may be explained by the accumulation in the RILs of genetic heritage from the parents favoring the obtention of high grain yield records. It should be noted that the Lahn parent is high productive in the favorable environments, resistant to yellow rust and moderately resistant to other parasites (leaf rust "*Puccinia recondite*", powdery mildew "*Erysiphe graminis*", common bunt "*Tilletia caries*", and to Septoria leaf blotch "*Mycosphaerella graminicola*"). Also, Lahn has a large kernel size with superior hectoliter weight, milling extraction index, and gluten strength values. As for Cham1, it has high yield stability. It exhibits an excellent resistance to drought and yellow rust with some tolerance to Russian wheat aphid. In addition, Cham1 also has high values for osmotic adjustment and high yellow pigment content, but its gluten strength is weak.

As for the genetic linkage map. The SSR molecular markers used for genotyping the Lahn/Cham1 population used in this study and for constructing consensus linkage map (Somers et al., 2004), were used as anchors and their order was kept as in the Somers's consensus map, then the SNPs markers were freely assigned to the durum chromosomes. The Lahn/Cham1 genetic linkage map share several SSR markers with other linkage maps (Nachit et al., 2001; Elouafi and Nachit, 2004; Haile et al., 2012; Zhang et al., 2012; Nagel et al., 2014; Alsaleh et al., 2015). Also, it shared an important number of SSR markers with the high-density microsatellite consensus maps for bread wheat (Somers et al., 2004) and for durum wheat (Maccaferri et al., 2015). Moreover, the Lahn/ Cham1 genetic linkage map was concordant with the marker orders of other linkage maps. Other researchers adopted same approach to construct genetic linkage maps (Su et al., 2018; Zeng et al., 2019). The Shixin828/Kenong2007 RIL population of Su et al. (2018) was genotyped for SNPs using the illumina iSelect 90K SNP

Array containing 81,587 SNP markers, this may explain the high number of SNP (6130 SNP vs 1425 for Lahn/Cham1 map) successfully assigned to the Shixin828/Kenong2007 map. On the other hand, the MX169/ CW86 RIL population of Zeng et al. (2019) was genotyped using the Wheat 35K SNP array, Zeng have successfully assigned a number of SNPs similar to that in our map (1480 for MX169/CW86 population vs 1425 for Lahn/Cham1 population). The comparison of the ratios of SNPs assigned to the maps / SNPs used for the population genotyping show an important ratio for our population in comparison with the others (7.51% for Shixin828/Kenong2007, 4.21% for MX169/ CW86, and 15.16% for Lahn/Cham1), the ratio obtained for Lahn/Cham1 population show that the number of SNPs assigned to the map is very important in comparison with the other maps of Su et al., (2018) and Zeng et al., (2019).

Grain yield is a typical quantitative trait controlled by a complex genetic system (QTL) and strongly influenced by both environmental factors and agronomic management. In this study, a grain yield QTL detection in contrasting environments was conducted based on the inclusive composite interval mapping (ICIM) using IciMapping 4.1 software. Moreover, a Multienvironment Traits (MET) analysis was employed for combined QTL analysis across environments to verify the QTL identified in the individual environment and evaluate the QTL $\times$ environment interactions. The results show the presence of 126 QTL over all chromosomes, whereas in the study of Tura et al., (2019) that was conducted in 32 environments, 38 QTLs for grain yield were detected. In addition, (Li et al., 2015) has detected only eight QTL for grain yield in a MET analysis conducted on the “Ning7840 / Clark” population that consisted of 127 RIL, the trials were performed in 9 environments (3 location for three years). The high frequency of QTL $\times$ environment interaction (126) detected in this work return to the response of the strong genotype of Lahn and Cham1 parents (cultivars adapted to the Mediterranean region that's know important changes in term of temperature and precipitations during the agricultural Saison) to the experiment fields (in Syria and Lebanon). Moreover, several works (Nachit, 1998; Nachit and Elouafi, 2004; Diab et al., 2007; Diab et al., 2008; Habash et al., 2009; Haile et al., 2012; Dura et al., 2013; Wang et al., 2018; Tura et al., 2019; Xu et al., 2019) affirms that the drought is one of the most important abiotic stresses that severely affects plant growth and development. Furthermore, the damages caused can be increased if drought occurs during critical periods of development cycle or in presence with other abiotic stresses like temperature extremes, salinity, or biotic stresses as septoria, tan spot, and rusts. Consequently, durum wheat can sense and respond to these different conditions and adapt by expressing genes responsible for the tolerance to the different stresses, for growth and development, and for other

processes (Habash et al., 2009). This well explaining the high frequency of QTLs detection in this work.

The investigation of these QTL allows us the identify the SNPs markers linked to specific candidate genes. Then, the SNPs markers contained in the QTL regions were selected and their DNA sequences were expanded and aligned in NCBI Blast to annotate them. Habash et al., (2009) reported that durum wheat can sense and respond to these different external conditions and adapt by expressing genes responsible for the tolerance to the different stresses, for growth and development, and for other processes. In accordance with this affirmation, the results shows that almost half of the SNPs annotated are linked to candidate genes involved in the tolerance to abiotic stresses with its different types (drought, cold, heat, etc). Some candidate genes play dual role in both biotic and abiotic stresses at the same time. Moreover, the results shows that the SNP wsnp_993293 located in 1A chromosome at the 82.02 cM seems to be linked to a candidate gene involved in resistance to powdery mildew. This result agrees with the Lahn parent feature which is the resistance to powdery mildew (*Erysiphe graminis* DC.). Lahn is also resistant to yellow rust (*Puccinia striiformis*) and moderately resistant to leaf rust (*Puccinia recondita* Rob.), common bunt (*Tilletia caries* D.C.), and to *Septoria* leaf blotch (*Mycosphaerella graminicola*), the investigation of denser genetic linkage map with SNP or DArT can identify molecular markers linked to the genes responsible for tolerance to these pathogens useful for Marker Assisted Selection (MAS).

In addition, the MET analysis and the annotation proved the importance of 4B chromosome specially de short arm, the results show that the candidates genes found in 4B are essentially involved in stress tolerance (Figure 9). Diab *et al.*, (2008) reported that under drought conditions, loci have been linked to yield traits and several candidate genes and differentially expressed sequence tags (dESTs) were also statistically linked with QTLs for traits related to drought stress (Diab et al., 2008). Furthermore, on 4B chromosome these candidate genes and dESTs were co-located with QTLs related to traits responsible for tolerance to drought on the short arm of chromosome (Li et al., 2007; Habash et al., 2009; Yu et al., 2019). In this work, the SNPs annotation for the short arm of 4B chromosome show that 57.14% of candidate genes found are involved in stresses tolerance and 42.86% in plant growth and reproduction. These results confirm earlier findings (Nachit, 1998; Nachit et al., 2004; Diab et al., 2008; Habash et al., 2009; Xu et al., 2019) on the contribution of the short arm of 4B chromosome to grain yield. Note also that 1.6% of candidate genes found are transposons and retrotransposons, this will be discussed below.

In another way, a QTLs detection in specific environments was performed. All QTLs detected in this way were detected in MET analysis (Figure 6). The expression of the QTLs occurred in all environments: moderately stressed (02INC and 02RF), favorable and cold (01GH, 02GH, and 02EP), and hot irrigated (02RQ). The differences of QTLs detection between the high input favorable sites (01GH and 02GH) are probably due to the different seasonal rainfall distribution: 02GH was affected by early drought. Furthermore, we noticed the total absence of QTLs detection in the severe drought stressed environment (02BR), this may be explained by the fact that the severe stress adversely affects the gene expressions and the genetic variance. Therefore, the possibility to detect QTLs was reduced (Ribaut et al., 1997; Almeida et al., 2013). In addition, the absence of QTL in the 01IR and 01TR environments may be due to the optimal environmental conditions (important irrigation) which affecting adversely the QTL detection, Almeida et al., (2013) reported that for consistent detection of QTL ensuring an optimal drought stress is very critical. The QTL detection results show that 4 QTLs among 9 detected were harbored by the 2A chromosome, this means that this chromosome has a strong contribution in grain yield, these results are in agreement with other studies (Li et al., 2007; Sukumaran et al., 2018; Mangini et al., 2021). Moreover, the QTLs detected on 2A chromosome in the moderately stressed environments (02INC, 02RF) and in the favorable and the hot irrigated environment (02RQ) have a negative additive effect e.g., the genetic material comes from Cham1 parent, this agrees with the characteristics of Cham1 cited above (adapted to the Mediterranean climate and it exhibits an excellent resistance to drought). However, the remaining QTL detected on 2A chromosome at the cold favorable environment (02EP) have an additive effect positive e.g., the genetic material comes from Lahn parent which is an improved durum wheat variety developed for favorable with cold tolerance. The same thing was observed for the 5 grain yield QTLs detected on the other durum chromosomes. The four QTLs of grain yield detected on the 2A chromosome were spanned in different chromosomal regions (Table 2, article 2), meaning that depending on environmental conditions, several groups of genes in different chromosomal regions may be expressed to respond to the stress exerted by the environment. For example:

- ✓ In the dry environments (INC and RF):

The alignment of the chromosomal region (QTL) expressed in this environment against Rice genome has allowed the identification of 147 candidate genes vs 113 for *Brachypodium* in the QTL identified in the INC environment, among which: serpin-Z5, Probable glucuronosyltransferase Os04g0103100, Chalcone synthase 1, Fatty acid amide

hydrolase isoform X1, Amino acid permease 3. This set of genes were over expressed under water progressive stress (Habash et al., 2014), noted that the work of Habash et al., (2014) was conducted on the parent generating the RIL population used in our work (Lahn, Cham1), and one RIL of this population. The results of Habash et al., (2014) concord perfectly and confirms the results of the annotation established in this work.

As for the RF environment, the alignment of the DNA sequence in NCBI against Rice genome allowed us to identify 376 candidate genes vs 277 for *Brachypodium*, of which the F-Box and associated protein family that has also showed an over expression in the work of Habash et al., (2014). The F-Box gene family is involved in tolerance to abiotic stress according to (Moon et al., 2004; Hua and Vierstra, 2011), in another work (Li et al., 2018) conducted on transgenic tobacco under heat stress showed that this gene family is involved in this stress, in this work, the F-Box gene were identified in RF and INC environments (dry environment) but not in the heat environment (RQ), this is may be due to small interval spanning of the QTL detected in this environment on the 2A chromosome (Table 2, Article 2) and probably can be identified in the other QTL detected for this environment on 2B chromosome (future expected study).

✓ In the cold environment (EP)

We have identified 17 candidate genes for Rice genome against 12 for *Brachypodium*. Several candidate genes were identified especially in this environment like: UDP-glycosyltransferase 79-like isoform X2, involved in biotic and abiotic stress (Li et al., 2001; Yu et al., 2017), and in detoxification (Huang et al., 2021; Wang et al., 2021). Recently, the study of (Subramanyam and Nemacheck, 2021) has demonstrated that this gene can play an important role in Hessian fly resistance for wheat, the detection of this gene in the grain yield QTL at the EP environment on the 2A chromosome may be due to coincidence of the first annual generation of the Hessian fly which begins with fall emergence of adults from volunteer infested wheat with the tillering stage of Lahn/Cham1 durum wheat population at early planting environment. The acyl-desaturase 6, chloroplastic isoform X1 “acyl-carrier-protein” candidate gene identified in the same environment (EP) also seems involved in insect tolerance (Kazaz et al., 2021), this suggest the genes expression in synergy between the UDP-glycosyltransferase 79-like isoform X2 and the acyl-desaturase 6, chloroplastic isoform X1 “acyl-carrier-protein” and other genes for an optimal resistance to the Hessian fly, this in accordance with the Cham1 parent feature that It exhibits some tolerance to Russian wheat aphid (insect affecting

wheat). Moreover, Nsarellah et al., (2003) reported that most of the gene resistance to Hessian fly are located on A chromosomes, this is in accordance with the funding of this study which showed the localization of the gene coding for the acyl-carrier-protein on 2A chromosome.

✓ In the hot environment (RQ)

The QTL detected on this environment was spanned on a small chromosomal region (0.1 Mb) contained a small number of candidates genes (six). Among them we have identified the gene coding for the LRR receptor-like serine/threonine-protein kinase RCH1, involved in resistance to fungal pathogens for apple and cowpea (Padmarasu et al., 2018; Pradhan et al., 2018). The five candidate genes found are belonging to the retrotransposon / transposon gene family, this family of genes is discussed below.

We have noticed that same candidate genes are represented in different conditional environments like: the Hydroxyproline-rich glycoprotein-like identified in the QTLs detected in the dry (RF, INC) and cold (EP) environments, this gene was associated with tolerance to diverse stresses, this result corroborate the studies (Kumar and Nadarajah, 2020; Kumar et al., 2018) indicate that this gene is linked with the resistance to bacterial pathogen in rice, and the study of (Abou-Elwafa, 2018) indicating the involvement in tolerance to drought stress in barley. Several studies on durum wheat have been reported the involvement of the proline and soluble saccharides accumulation in drouth tolerance (Dib et al., 1994; Rekika et al., 1998; Nachit et al., 2004). The absence of the Hydroxyproline-rich glycoprotein-like detection in the heat environment (RQ) may be explained as for the F-Box genes family by the small QTL size detected on 2A chromosome and may be present or duplicated in the QTL detected on 2B or in other chromosomal regions, knowing that more than 65% of annotated genes in plants have a duplicated copy (Panchy et al., 2016). Moreover, durum is an allotetrapolyploid wheat originated through merging of two genomes (A and B genomes) through natural interspecific hybridization and genome duplication (Zhang et al., 2012), in duo the polyploidization witch is a major factor of genome evolution (Kondrashov, 2012; Magadum et al., 2013, Huo et al., 2018) this is in accordance with the results of this work where several candidate genes were found duplicated on same chromosome (2A) like: BTB/POZ and MATH involved in biotic and abiotic stress (He et al., 2019; Zhang et al., 2019) was duplicated 56 times in the QTLs detected in the dry environment (02RF). Other duplicated candidate genes were detected in multiple QTL (02INC and 02RF) like the gene coding for Zink finger proteins which was reported acting as positive regulators of plant responses to biotic and heat stresses, and as negative regulators of

cold and drought stresses (Nakai et al., 2013a; Nakai et al., 2013b; Ganie et al., 2020) was also duplicated in both environments.

Furthermore, in this work we have found that in the hot environment, the candidate genes identified are in majority belong to the transposons and retrotransposons genes family, further these genes were highly frequent in the other QTLs detected in the other environments on 2A chromosome and especially in the RF environment (dry environment). This is in the agreement with an early work (Jamilloux et al., 2016) where is reported that the transposable elements may represents about 80 to 90% of wheat genome, The abundance of this genes family suggest that the transposable elements plays an important role in stress tolerance, in adaptation to the environmental conditions, and in plants evolution (Grandbastien, 1998; Mansour, 2007; Lisch, 2013; Jamilloux et al., 2016; Negi et al., 2016; Dubin et al., 2018). The involvement of these elements in stress tolerance was confirmed by Habash et al., (2014) where their activity was altered under progressive water stress.

Conclusion and perspectives

Conclusion

Since the wheats domestication in the Neolithic era around 10000 to 15000 BC. JC, they have constituted the basis of human nutrition which resulted in the adoption of a production economy based on agriculture and natural and farmed selection. Durum wheat has an important place in the cereal crops with a significant impact globally on the world economy, and especially in the West Asia and North Africa (WANA) region and in the other Mediterranean countries, which produce 50% of the durum wheat production. In these regions, durum is not only used for bread production, but also to produce pasta and several Mediterranean end products such as Frikke and couscous. Moreover, durum is known for its nutritional value and its composition in protein, and vitamin A. The climatic condition in the Mediterranean region is becoming more and more stressing due the global warming. The cereal crops in Morocco and the other Mediterranean countries are largely under rainfed conditions. Furthermore, durum crop is adversely affected by these abiotic stresses in relation with global warming (heat, temperature, and precipitation fluctuations...) and by biotic stresses (parasites and pathogen organisms). The breeders have to develop new cultivars with high yield and quality, well adapted and tolerant to these stressed conditions. Therefore, breeders have to understand the genomic of these complex traits to bring them together in a new improved durum varieties for farmers.

Our study dealt with understanding the genomics of the grain yield trait in durum. Firstly, we have to mention that some RIL of the population used (Lahn/Cham1) showed great performances and have outyielded both parents, these RIL may constitute promising cultivars and may be released for farmers.

Moreover, the high-density genetic linkage map constructed basing on the population studied in this work allowed us to identify the QTLs linked to the grain yield. The important number of QTL detected in the Multi-Environmental traits analysis (MET) covering all chromosomes, prove the importance of this population in term of genetic resources for grain yield. This method of QTL analysis allowed us the selection of SNP markers contained in these QTLs and the alignment of their DNA sequences with public databases (URGI and NCBI). The result of this alignment shows that more than half of these SNPs (53%) are linked to candidate genes involved in stresses tolerance. This method of QTL analysis (MET) confirms the importance of the 4B chromosome reported in several earlier studies in stress tolerance.

As for the second method of QTLs detection (QTL scan in specific environments “Biparental analysis”). The 2A chromosome shows great importance and it may contribute to grain yield under different climatic conditions. The annotation of de four QTLs detected on this chromosome show the presence of several candidate genes involved essentially in stresses tolerance.

The results of this work have enhanced our knowledge about the genomic of the grain yield in terms of genes that may contribute to this complex trait and provided to us raw material to develop markers for tracking these genes over durum generations in durum breeding programs.

Perspectives

The SNPs selected (using MET approach) are of great interest and may be used in Marker Assisted Selection (MAS) and will constitute powerful tools for tracking the linked genes in durum breeding programs.

The DNA sequences corresponding to the QTLs identified on 2A chromosome may constitute the basis for the development of molecular marker (SSR, SNP, DArT...) linked to the candidate genes of interest. These molecular markers are useful for breeders to follow the transferability of the linked genes to the progeny in the durum breeding programs.

The QTLs dissection may be generalized for the other QTL detected by the BIP analysis over all durum genome.

Also, the methodologies of the QTLs detection and their annotation may be used for other complex traits of economic importance like quality by using this population or other durum populations.

References

References

- Abou-Elwafa, S. F. (2018). Identification of genes associated with drought tolerance in barley. *Biologia Plantarum* **62**, 299-306.
- Adams, K. L., and Wendel, J. F. (2005). Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology* **8**, 135-141.
- Almeida, G. D., Makumbi, D., Magorokosho, C., Nair, S., Borém, A., Ribaut, J.-M., Bänziger, M., Prasanna, B. M., Crossa, J., and Babu, R. (2013). QTL mapping in three tropical maize populations reveals a set of constitutive and adaptive genomic regions for drought tolerance. *Theoretical and Applied Genetics* **126**, 583-600.
- Alonso-Blanco, C., Koornneef, M., and Stam, P. (1998). The use of recombinant inbred lines (RILs) for genetic mapping. In "Arabidopsis Protocols", pp. 137-146. Springer.
- Alsaleh, A., Baloch, F. S., Derya, M., Azrak, M., Kilian, B., Özkan, H., and Nachit, M. (2015). Genetic linkage map of Anatolian durum wheat derived from a cross of Kunduru-1149× Cham1. *Plant Molecular Biology Reporter* **33**, 209-220.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* **25**, 3389-3402.
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics* **25**, 25-29.
- Azhaguvel, P., Vidya-Saraswathi, D., and Komatsuda, T. (2006). High-resolution linkage mapping for the non-brittle rachis locus btr1 in cultivated×wild barley (*Hordeum vulgare*). *Plant Science* **170**, 1087-1094.
- Backes, G., Graner, A., Foroughi-Wehr, B., Fischbeck, G., Wenzel, G., and Jahoor, A. (1995). Localization of quantitative trait loci (QTL) for agronomic important characters by the

- use of a RFLP map in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* **90**, 294-302.
- Balfourier, F., Roussel, V., Strelchenko, P., Exbrayat-Vinson, F., Sourdille, P., Boutet, G., Koenig, J., Ravel, C., Mitrofanova, O., Beckert, M., and Charmet, G. (2007). A worldwide bread wheat core collection arrayed in a 384-well plate. *Theoretical and Applied Genetics* **114**, 1265-1275.
- Belaid, A., Nsarellah, N., Laamari, A., Nachit, M., and Amri, A. (2005). Assessing the economic impact of durum wheat research in Morocco. Aleppo, Syria: International Centre for Agricultural Research in the Dry Areas (ICARDA).
- Bennett, D., Reynolds, M., Mullan, D., Izanloo, A., Kuchel, H., Langridge, P., and Schnurbusch, T. (2012). Detection of two major grain yield QTL in bread wheat (*Triticum aestivum* L.) under heat, drought and high yield potential environments. *Theoretical and Applied Genetics* **125**, 1473-1485.
- Bennetzen, J. L., and Freeling, M. (1997). The unified grass genome: synergy in synteny. *Genome research* **7**, 301-306.
- Berman, H., Henrick, K., Nakamura, H., and Markley, J. L. (2007). The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucleic Acids Research* **35**, D301-D303.
- Blanco, A., Bellomo, M. P., Cenci, A., De Giovanni, C., D'Ovidio, R., Iacono, E., Laddomada, B., Pagnotta, M. A., Porceddu, E., Sciancalepore, A., Simeone, R., and Tanzarella, O. A. (1998a). A genetic linkage map of durum wheat. *Theor Appl Genet* **97**, 721-728.
- Blanco, A., Bellomo, M. P., Lotti, C., Maniglio, T., Pasqualone, A., Simeone, R., Troccoli, A., and Di Fonzo, N. (1998b). Genetic mapping of sedimentation volume across environments using recombinant inbred lines of durum wheat. *Plant Breeding* **117**, 413-417.
- Blanco, A., Colasuonno, P., Gadaleta, A., Mangini, G., Schiavulli, A., Simeone, R., Digesù, A. M., De Vita, P., Mastrangelo, A. M., and Cattivelli, L. (2011). Quantitative trait loci for yellow pigment concentration and individual carotenoid compounds in durum wheat. *Journal of Cereal Science* **54**, 255-264.

- Blanco, A., Giovanni, C. d., Laddomada, B., Sciancalepore, A., Simeone, R., Devos, K. M., and Gale, M. D. (1996). Quantitative trait loci influencing grain protein content in tetraploid wheats. *Plant Breeding* **115**, 310-316.
- Blanco, A., Mangini, G., Giancaspro, A., Giove, S., Colasuonno, P., Simeone, R., Signorile, A., De Vita, P., Mastrangelo, A. M., Cattivelli, L., and Gadaleta, A. (2012). Relationships between grain protein content and grain yield components through quantitative trait locus analyses in a recombinant inbred line population derived from two elite durum wheat cultivars. *Molecular Breeding* **30**, 79-92.
- Blanco, A., Simeone, R., Cenci, A., Gadaleta, A., Tanzarella, O., Porceddu, E., Salvi, S., Tuberosa, R., Figliuolo, G., Spagnoletti Zeuli, P., Röder, M., and Korzun, V. (2004). Extension of the *Messapia* x *dicoccoides* linkage map of *Triticum turgidum* (L.) Thell. *Cellular & molecular biology letters* **9**, 529-41.
- Bogard, M. (2012). Analyse génétique et écophysiologique de l'écart à la relation teneur en protéines - rendement en grains chez le blé tendre (*Triticum aestivum* L.). *thèse*.
- Bogard, M., Jourdan, M., Allard, V., Martre, P., Perretant, M. R., Ravel, C., Heumez, E., Orford, S., Snape, J., Griffiths, S., Gaju, O., Foulkes, J., and Le Gouis, J. (2011). Anthesis date mainly explained correlations between post-anthesis leaf senescence, grain yield, and grain protein concentration in a winter wheat population segregating for flowering time QTLs. *Journal of Experimental Botany* **62**, 3621-3636.
- Bolot, S., Abrouk, M., Masood-Quraishi, U., Stein, N., Messing, J., Feuillet, C., and Salse, J. (2009). The 'inner circle' of the cereal genomes. *Current Opinion in Plant Biology* **12**, 119-125.
- Bort, J., Belhaj, M., Latiri, K., Kehel, Z., and Araus, J. L. (2014). Comparative performance of the stable isotope signatures of carbon, nitrogen and oxygen in assessing early vigour and grain yield in durum wheat. *The Journal of Agricultural Science* **152**, 408-426.
- Boyko, E. V., Gill, K. S., Mickelson-Young, L., Nasuda, S., Raupp, W. J., Ziegler, J. N., Singh, S., Hassawi, D. S., Fritz, A. K., Namuth, D., Lapitan, N. L. V., and Gill, B. S. (1999). A high-density genetic linkage map of *Aegilops tauschii*, the D-genome progenitor of bread wheat. *Theoretical and Applied Genetics* **99**, 16-26.

- Bozzini, A. (1988). Origin, distribution, and production of durum wheat in the world. In: *Durum wheat: Chemistry and Technology*, G. Fabriani and C. Lintas (ed.). AACC, St. Paul, MN., 1-16.
- Brunak, S., Danchin, A., Hattori, M., Nakamura, H., Shinozaki, K., Matise, T., and Preuss, D. (2002). Nucleotide Sequence Database Policies. *Science* **298**, 1333-1333.
- Burley, S. K., Berman, H. M., Kleywegt, G. J., Markley, J. L., Nakamura, H., and Velankar, S. (2017). Protein Data Bank (PDB): The Single Global Macromolecular Structure Archive. In "Protein Crystallography: Methods and Protocols" (A. Wlodawer, Z. Dauter and M. Jaskolski, eds.), pp. 627-641. Springer New York, New York, NY.
- Carbon, S., Douglass, E., Good, B. M., Unni, D. R., Harris, N. L., Mungall, C. J., Basu, S., Chisholm, R. L., Dodson, R. J., and Hartline, E. (2021). The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Research* **49**, D325-D334.
- Causse, M. A., Fulton, T. M., Cho, Y. G., Ahn, S. N., Chunwongse, J., Wu, K., Xiao, J., Yu, Z., Ronald, P. C., and Harrington, S. E. (1994). Saturated molecular map of the rice genome based on an interspecific backcross population. *Genetics* **138**, 1251-1274.
- Chalupska, D., Lee, H. Y., Faris, J. D., Evrard, A., Chalhoub, B., Haselkorn, R., and Gornicki, P. (2008). *Acc* homoeoloci and the evolution of wheat genomes. *Proceedings of the National Academy of Sciences* **105**, 9691-9696.
- Chao, S., Sharp, P. J., Worland, A. J., Warham, E. J., and Koebner, R. M. D. (1989). RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. *Theor Appl Genet* **78**, 495–504.
- Choudhuri, S. (2014). Data, Databases, Data Format, Database Search, Data Retrieval Systems, and Genome Browsers Chapter 5. In "Bioinformatics for Beginners" (S. Choudhuri, ed.), pp. 77-131. Academic Press, Oxford.
- Clayton, W., and Renvoize, S. (1986). *Genera Graminum*. pp. 389. Her Majesty's Stationery Office, London.
- Colasuonno, P., Gadaleta, A., Giancaspro, A., Nigro, D., Giove, S., Incerti, O., Mangini, G., Signorile, A., Simeone, R., and Blanco, A. (2014). Development of a high-density SNP-based linkage map and detection of yellow pigment content QTLs in durum wheat. *Molecular Breeding* **34**, 1563-1578.

- Colasuonno, P., Maria, M. A., Blanco, A., and Gadaleta, A. (2013). Description of durum wheat linkage map and comparative sequence analysis of wheat mapped DArT markers with rice and *Brachypodium* genomes. *BMC Genetics* **14**, 114.
- Cui, L., Wall, P. K., Leebens-Mack, J. H., Lindsay, B. G., Soltis, D. E., Doyle, J. J., Soltis, P. S., Carlson, J. E., Arumuganathan, K., Barakat, A., Albert, V. A., Ma, H., and dePamphilis, C. W. (2006). Widespread genome duplications throughout the history of flowering plants. *Genome Research* **16**, 738-749.
- DeVicente, M., and Tanksley, S. (1993). QTL analysis of transgressive segregation in an interspecific tomato cross. *Genetics* **134**, 585-596.
- Devos, K. M., Beales, J., Ogihara, Y., and Doust, A. N. (2005). Comparative sequence analysis of the phytochrome C gene and its upstream region in allohexaploid wheat reveals new data on the evolution of its three constituent genomes. *Plant molecular biology* **58**, 625-641.
- Devos, K., Atkinson, M., Chinoy, C., Liu, C., and Gale, M. (1992). RFLP-based genetic map of the homoeologous group 3 chromosomes of wheat and rye. *Theoretical and Applied Genetics* **83**, 931-939.
- Deynze, A. V., Dubcovsky, J., Gill, K., Nelson, J., Sorrells, M., Dvořák, J., Gill, B., Lagudah, E., McCouch, S., and Appels, R. (1995). Molecular-genetic maps for group 1 chromosomes of *Triticeae* species and their relation to chromosomes in rice and oat. *Genome* **38**, 45-59.
- Diab, A. A., Fahmy, A. H., Hassan, O. S., Nachit, M., and Momtaz, O. A. (2007). Identification of chromosomal regions and genetic contributions of genes controlling yield and other agronomic traits in durum wheat grown under different Egyptian environmental conditions. *World Journal of Agricultural Sciences* **3**, 401-422.
- Diab, A. A., Kantety, R., Ozturk, N., Benscher, D., Nachit, M., and Sorrells, M. (2008). Drought-inducible genes and differentially expressed sequence tags associated with components of drought tolerance in durum wheat. *Scientific Research and Essay* **3**, 9-26.
- Diamond, J. (2002). Evolution, consequences and future of plant and animal domestication. *Nature* **418**, 700-707.

- Dib, C., Fauré, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Kazan, J., and Seboun, E. (1996). A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* **380**, 152-154.
- Dib, T. A., Monneveux, P., Acevedo, E., and Nachit, M. M. (1994). Evaluation of proline analysis and chlorophyll fluorescence quenching measurements as drought tolerance indicators in durum wheat (*Triticum turgidum* L. var. *durum*). *Euphytica* **79**, 65-73.
- Dubin, M. J., Mittelsten Scheid, O., and Becker, C. (2018). Transposons: a blessing curse. *Current Opinion in Plant Biology* **42**, 23-29.
- Dura, S. A., Duwayri, M. A., and Nachit, M. M. (2013). Detection of molecular markers associated with yield and yield components in durum wheat (*Triticum turgidum* L. var. *durum* Desf.) under drought conditions. *African Journal of Agricultural Research* **8**, 2113-2117.
- Dura, S., Duwayri, M., Nachit, M., and Al Sheyab, F. (2014). Detection of molecular markers associated with yield and yield components in durum wheat (*Triticum turgidum* L. var. *durum*) under saline conditions. *Crop and Pasture Science* **64**, 957-964.
- El Yadini, M., Kahama, I., Labhilili, M., El Yadini, A., and Azeqour, M. (2013). La technique TILLING comme nouvelle biotechnologie pour la protection et l'amélioration du blé dur (TILLING: A new biotechnological method for protection and improvement of durum wheat). *J Mater Environ Sci* **4**, 931-934.
- Elouafi, I., and Nachit, M. M. (2004). A genetic linkage map of the Durum × *Triticum dicoccoides* backcross population based on SSRs and AFLP markers, and QTL analysis for milling traits. *Theoretical and Applied Genetics* **108**, 401-413.
- Esnayra, J., Pool, R., and Council, N. R. (2000). Bioinformatics: Converting data to knowledge.
- FAO. (2020). <https://www.fao.org/> (visité le : 02/03/2021).
- Faris, J. D., Fellers, J. P., Brooks, S. A., and Gill, B. S. (2003). A Bacterial Artificial Chromosome Contig Spanning the Major Domestication Locus Q in Wheat and Identification of a Candidate Gene. *Genetics* **164**, 311-321.
- Farouk, I., Alsaleh, A., Motowaj, J., Gaboun, F., Belkadi, B., Filali maltouf, A., Kehel, Z., Elouafi, I., Nsarellah, N., Habash, D., and Nachit, M. M. (2020). Dissection of Quantitative Trait Loci (QTL), annotation of Single Nucleotide Polymorphism (SNP),

- and Identification of Candidate Genes for Grain Yield in *Triticum turgidum* L. var durum. *Advances in Science, Technology and Engineering Systems Journal* **5**, 1020-1027.
- Farouk, I., Alsaleh, A., Motowaj, J., Gaboun, F., Belkadi, B., Maltouf, A. F., Kehel, Z., Elouafi, I., Nsarellah, N., and Nachit, M. M. (2021). Detection of grain yield QTLs in the durum population Lahn/Cham1 tested in contrasting environments. *Turkish Journal of Biology* **45**, 65-78.
- Fellahtrad. (2021). <https://www.fellah-trade.com> (visit   le: 01/03/2021)
- Feldman, M., and Sears, E. R. (1981). The Wild Gene Resources of Wheat. *Scientific American* **244**, 102-113.
- Feuillet, C., Langridge, P., and Waugh, R. (2008). Cereal breeding takes a walk on the wild side. *Trends in Genetics* **24**, 24-32.
- Gadaleta, A., Giancaspro, A., Giove, S. L., Zacheo, S., Mangini, G., Simeone, R., Signorile, A., and Blanco, A. (2009). Genetic and physical mapping of new EST-derived SSRs on the A and B genome chromosomes of wheat. *Theoretical and Applied Genetics* **118**, 1015-1025.
- Ganie, S. A., Ahammed, G. J., and Wani, S. H. (2020). Vascular plant one zinc-finger (VOZ) transcription factors: novel regulators of abiotic stress tolerance in rice (*Oryza sativa* L.). *Genetic Resources and Crop Evolution* **67**, 799-807.
- Garrod, D. A., and Reckitt, A. (1957). The Natufian culture: the life and economy of a Mesolithic people in the Near East. *Proc. Brit. Acad.* **43**, 211-227.
- Gaut, B. S. (2002). Evolutionary dynamics of grass genomes. *New Phytologist* **154**, 15-28.
- Geldermann, H. (1975). Investigations on inheritance of quantitative characters in animals by gene markers I. Methods. *Theoretical and Applied Genetics* **46**, 319-330.
- Ghosh, D. (1991). New developments of a transcription factors database. *Trends in biochemical sciences* **16**, 445-447.
- Gill, K. S., Lubbers, E. L., Gill, B. S., Raupp, W. J., and Cox, T. S. (1991). A genetic linkage map of *Triticum tauschii* (DD) and its relationship to the D genome of bread wheat (AABBDD). *genome* **34**, 362-374.

- Goodsell, D. S., Zardecki, C., Di Costanzo, L., Duarte, J. M., Hudson, B. P., Persikova, I., Segura, J., Shao, C., Voigt, M., Westbrook, J. D., Young, J. Y., and Burley, S. K. (2020). RCSB Protein Data Bank: Enabling biomedical research and drug discovery. *Protein Science* **29**, 52-65.
- Grandbastien, M.-A. (1998). Activation of plant retrotransposons under stress conditions. *Trends in Plant Science* **3**, 181-187.
- Graner, A., Jahoor, A., Schondelmaier, J., Siedler, H., Pillen, K., Fischbeck, G., Wenzel, G., and Herrmann, R. (1991). Construction of an RFLP map of barley. *Theoretical and Applied Genetics* **83**, 250-256.
- Graziani, M., Maccaferri, M., Royo, C., Salvatorelli, F., and Tuberosa, R. (2014). QTL dissection of yield components and morpho-physiological traits in a durum wheat elite population tested in contrasting thermo-pluviometric conditions *Crop and Pasture Science* **65**, 80-95.
- Habash, D. Z., Baudo, M., Hindle, M., Powers, S. J., Defoin-Platel, M., Mitchell, R., Saqi, M., Rawlings, C., Latiri, K., Araus, J. L., Abdulkader, A., Tuberosa, R., Lawlor, D. W., and Nachit, M. M. (2014). Systems responses to progressive water stress in durum wheat. *PloS one* **9**, e108431-e108431.
- Habash, D. Z., Kehel, Z., and Nachit, M. (2009). Genomic approaches for designing durum wheat ready for climate change with a focus on drought. *Journal of Experimental Botany* **60**, 2805-2815.
- Haile, J. K., Nachit, M. M., Hammer, K., Badebo, A., and Röder, M. S. (2012). QTL mapping of resistance to race Ug99 of *Puccinia graminis* f. sp. *tritici* in durum wheat (*Triticum durum* Desf.). *Molecular Breeding* **30**, 1479-1493.
- Haldane, J. (1919). The probable errors of calculated linkage values, and the most accurate method of determining gametic from certain zygotic series. *Journal of genetics* **8**, 4, 291-297.
- Haldane, J. B., and Smith, C. A. (1947). A new estimate of the linkage between the genes for colour-blindness and *haemophilia* in man. *Annals of Eugenics* **14**, 10-31.
- Haley, C. S., Knott, S. A., and Elsen, J. M. (1994). Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics* **136**, 1195-1207.

- Harlan, J. R., and Zohary, D. (1966). Distribution of Wild Wheats and Barley. **153**, 1074-1080.
- Harushima, Y., Yano, M., Shomura, A., Sato, M., Shimano, T., Kuboki, Y., Yamamoto, T., Lin, S. Y., Antonio, B. A., and Parco, A. (1998). A high-density rice genetic linkage map with 2275 markers using a single F2 population. *Genetics* **148**, 479-494.
- He, Y.-M., Liu, K.-K., Zhang, H.-X., Cheng, G.-X., Ali, M., Ul Haq, S., Wei, A.-M., and Gong, Z.-H. (2019). Contribution of CaBPM4, a BTB Domain-Containing Gene, to the Response of Pepper to *Phytophthora capsici* Infection and Abiotic Stresses. *Molecular Plant Pathology* **9**, 417.
- Heidari, B., Sayed-Tabatabaei, B. E., Saeidi, G., Kearsey, M., Suenaga, K., and Gulick, P. (2011). Mapping QTL for grain yield, yield components, and spike features in a doubled haploid population of bread wheat. *Genome* **54**, 517-527.
- Helentjaris, T. (1987). A genetic linkage map for maize based on RFLPs. *Trends in Genetics* **3**, 217-221.
- Helentjaris, T., Slocum, M., Wright, S., Schaefer, A., and Nienhuis, J. (1986). Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. *Theoretical and Applied Genetics* **72**, 761-769.
- Hessler, T., Thomson, M. J., Benscher, D., Nachit, M., and Sorrells, M. (2002). Association of a lipoxygenase locus, Lpx-B1, with variation in lipoxygenase activity in durum wheat seeds. *Crop Science - CROP SCI* **42**, 1695-1700.
- Heun, M., Kennedy, A., Anderson, J., Lapitan, N., Sorrells, M., and Tanksley, S. (1991). Construction of a restriction fragment length polymorphism map for barley (*Hordeum vulgare*). *Genome* **34**, 437-447.
- Hillman, G. (1975). The plant remains from Tell Abu Hureyra: a preliminary report. In "Proceedings of the Prehistoric Society", Vol. **41**, pp. 70-73.
- Hoisington, D., and Coe, E. (1990). Mapping in maize using RFLPs. In "Gene manipulation in plant improvement II", pp. 331-352. Springer.
- Hua, Z., and Vierstra, R. D. (2011). The cullin-RING ubiquitin-protein ligases. *Annual review of plant biology* **62**, 299-334.
- Huang, S., Sirikhachornkit, A., Su, X., Faris, J., Gill, B., Haselkorn, R., and Gornicki, P. (2002). Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the

- Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. *Proceedings of the National Academy of Sciences*. **99**, 8133-8138.
- Huang, X.-x., Zhao, S.-m., Zhang, Y.-y., Li, Y.-j., Shen, H.-n., Li, X., and Hou, B.-k. (2021). A novel UDP-glycosyltransferase 91C1 confers specific herbicide resistance through detoxification reaction in *Arabidopsis*. *Plant Physiology Biochemistry* **159**, 226-233.
- Huo, N., Zhang, S., Zhu, T., Dong, L., Wang, Y., Mohr, T., Hu, T., Liu, Z., Dvorak, J., Luo, M.-C., Wang, D., Lee, J.-Y., Altenbach, S., and Gu, Y. Q. (2018). Gene Duplication and Evolution Dynamics in the Homeologous Regions Harboring Multiple Prolamin and Resistance Gene Families in Hexaploid Wheat. *frontiers in Plant Science* **9**.
- Itoh, T., Tanaka, T., Barrero, R. A., Yamasaki, C., Fujii, Y., Hilton, P. B., Antonio, B. A., Aono, H., Apweiler, R., Bruskiewich, R., Bureau, T., Burr, F., Costa de Oliveira, A., Fuks, G., Habara, T., Haberer, G., Han, B., Harada, E., Hiraki, A. T., Hirochika, H., Hoen, D., Hokari, H., Hosokawa, S., Hsing, Y., Ikawa, H., Ikeo, K., Imanishi, T., Ito, Y., Jaiswal, P., Kanno, M., Kawahara, Y., Kawamura, T., Kawashima, H., Khurana, J. P., Kikuchi, S., Komatsu, S., Koyanagi, K. O., Kubooka, H., Lieberherr, D., Lin, Y.-C., Lonsdale, D., Matsumoto, T., Matsuya, A., McCombie, W. R., Messing, J., Miyao, A., Mulder, N., Nagamura, Y., Nam, J., Namiki, N., Numa, H., Nurimoto, S., O'Donovan, C., Ohyanagi, H., Okido, T., Oota, S., Osato, N., Palmer, L. E., Quetier, F., Raghuvanshi, S., Saichi, N., Sakai, H., Sakai, Y., Sakata, K., Sakurai, T., Sato, F., Sato, Y., Schoof, H., Seki, M., Shibata, M., Shimizu, Y., Shinozaki, K., Shinso, Y., Singh, N. K., Smith-White, B., Takeda, J.-i., Tanino, M., Tatusova, T., Thongjuea, S., Todokoro, F., Tsugane, M., Tyagi, A. K., Vanavichit, A., Wang, A., Wing, R. A., Yamaguchi, K., Yamamoto, M., Yamamoto, N., Yu, Y., Zhang, H., Zhao, Q., Higo, K., Burr, B., Gojobori, T., Sasaki, T., and Project, f. t. R. A. (2007). Curated genome annotation of *Oryza sativa* ssp. *japonica* and comparative genome analysis with *Arabidopsis thaliana*. *Genome Research* **17**, 175-183.
- Jamilloux, V., Daron, J., Choulet, F., and Quesneville, H. (2016). De novo annotation of transposable elements: tackling the fat genome issue. *Proceedings of the IEEE* **105**, 474-481.
- Jansen, R. C. (1993). Interval mapping of multiple quantitative trait loci. *Genetics* **135**, 205-211.

- Jansen, R. C. (1994). Controlling the type I and type II errors in mapping quantitative trait loci. *Genetics* **138**, 871-881.
- Jansen, R. C. (1996). A General Monte Carlo Method for Mapping Multiple Quantitative Trait Loci. *Genetics* **142**, 305-311.
- Jansen, R. C., and Stam, P. (1994). High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* **136**, 1447-1455.
- Jantasuriyarat, C., Vales, M. I., Watson, C. J. W., and Riera-Lizarazu, O. (2004). Identification and mapping of genetic loci affecting the free-threshing habit and spike compactness in wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* **108**, 261-273.
- Jia, J., Devos, K., Chao, S., Miller, T., Reader, S., and Gale, M. (1996). RFLP-based maps of the homoeologous group-6 chromosomes of wheat and their application in the tagging of Pm12, a powdery mildew resistance gene transferred from *Aegilops speltoides* to wheat. *Theoretical and Applied Genetics* **92**, 559-565.
- Johannsen, W. (1909). Elemente der Exakten Erblchkeitslehre. *Science* **30**, 851-853.
- Kao, C.-H., Zeng, Z.-B., and Teasdale, R. D. (1999). Multiple Interval Mapping for Quantitative Trait Loci. *Genetics* **152**, 1203-1216.
- Kazaz, S., Miray, R., and Baud, S. (2021). Acyl–Acyl Carrier Protein Desaturases and Plant Biotic Interactions. *Cells* **10**, 674.
- Kehel, Z., Habash, D. Z., Gezan, S. A., Welham, S. J., and Nachit, M. M. (2010). Estimation of Spatial Trend and Automatic Model Selection in Augmented Designs. *Agron. J* **102**, 1542-1552.
- Kellogg, E. A. (2001). Evolutionary History of the Grasses. *Plant Physiology* **125**, 1198-1205.
- Kislev, M. E. (1984). Emergence of wheat agriculture. *Paléorient* **10**, 61-70.
- Kleinhofs, A., Kilian, A., Maroof, M. S., Biyashev, R. M., Hayes, P., Chen, F., Lapitan, N., Fenwick, A., Blake, T., and Kanazin, V. (1993). A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. *Theoretical and Applied Genetics* **86**, 705-712.
- Knott, S. A., and Haley, C. S. (1992). Aspects of maximum likelihood methods for the mapping of quantitative trait loci in line crosses. *Genetical Research* **60**, 139-151.

- Knott, S. A., Elsen, J. M., and Haley, C. S. (1996). Methods for multiple-marker mapping of quantitative trait loci in half-sib populations. *Theoretical and Applied Genetics* **93**, 71-80.
- Kondrashov, F. A. (2012). Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proceedings of the Royal Society B: Biological Sciences* **279**, 5048-5057.
- Kong, A., Gudbjartsson, D. F., Sainz, J., Jonsdottir, G. M., Gudjonsson, S. A., Richardsson, B., Sigurdardottir, S., Barnard, J., Hallbeck, B., Masson, G., Shlien, A., Palsson, S. T., Frigge, M. L., Thorgeirsson, T. E., Gulcher, J. R., and Stefansson, K. (2002). A high-resolution recombination map of the human genome. *Nature Genetics* **31**, 241-247.
- Korzun, V., Röder, M., Wendehake, K., Pasqualone, A., Lotti, C., Ganai, M., and Blanco, A. (1999). Integration of dinucleotide microsatellites from hexaploid bread wheat into a genetic linkage map of durum wheat. *Theoretical and Applied Genetics* **98**, 1202-1207.
- Kosambi, D. D. (1943). The estimation of map distances from recombination values. *Annals of Eugenics* **12**, 172-175.
- Kouranov, A., Xie, L., de la Cruz, J., Chen, L., Westbrook, J., Bourne, P. E., and Berman, H. M. (2006). The RCSB PDB information portal for structural genomics. *Nucleic Acids Research* **34**, D302-D305.
- Kubo, K., Elouafi, I., Watanabe, N., Nachit, M. M., Inagaki, M. N., Iwama, K., and Jitsuyama, Y. (2007). Quantitative trait loci for soil-penetrating ability of roots in durum wheat. *Plant Breeding* **126**, 375-378.
- Kumar, I. S., and Nadarajah, K. (2020). A meta-analysis of quantitative trait loci associated with multiple disease resistance in rice (*Oryza sativa* L.). *Plants* **9**, 1491.
- Kumar, I. S., Zaharin, N., and Nadarajah, K. (2018). In silico identification of resistance and defense related genes for bacterial leaf blight (BLB) in rice. *J. Pure Appl. Microbiol* **12**, 1867-1877.
- Lagudah, E. S., Appels, R., Brown, A. H. D., and McNeil, D. (1991). The molecular-genetic analysis of *Triticum tauschii*, the D-genome donor to hexaploid wheat. *Genome* **34**, 375-386.

- Lander, E. S., and Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**, 185-199.
- Li, C., Bai, G., Carver, B. F., Chao, S., and Wang, Z. (2015). Single nucleotide polymorphism markers linked to QTL for wheat yield traits. *Euphytica* **206**, 89-101.
- Li, H., Hearne, S., Bänziger, M., Li, Z., and Wang, J. (2010). Statistical properties of QTL linkage mapping in biparental genetic populations. *Heredity* **105**, 257-267.
- Li, H., Ribaut, J.-M., Li, Z., and Wang, J. (2008). Inclusive composite interval mapping (ICIM) for digenic epistasis of quantitative traits in biparental populations. *Theoretical and applied genetics* **116**, 243-60.
- Li, H., Ye, G., and Wang, J. (2007). A Modified Algorithm for the Improvement of Composite Interval Mapping. *Genetics* **175**, 361-374.
- Li, H., Zhang, L., and Wang, J. (2012). Estimation of statistical power and false discovery rate of QTL mapping methods through computer simulation. *Chinese Science Bulletin* **57**, 2701-2710.
- Li, Q., Wang, W., Wang, W., Zhang, G., Liu, Y., Wang, Y., and Wang, W. (2018). Wheat F-Box Protein Gene TaFBA1 Is Involved in Plant Tolerance to Heat Stress. *Frontiers*
- Li, S., Jia, J., Wei, X., Zhang, X., Li, L., Chen, H., Fan, Y., Sun, H., Zhao, X., Lei, T., Xu, Y., Jiang, F., Wang, H., and Li, L. (2007). A intervarietal genetic map and QTL analysis for yield traits in wheat. *Molecular Breeding* **20**, 167-178.
- Li, S., Wang, J., and Zhang, L. (2015). Inclusive Composite Interval Mapping of QTL by Environment Interactions in Biparental Populations. *PLOS ONE* **10**, e0132414.
- Li, Y., Baldauf, S., Lim, E.-K., and Bowles, D. J. (2001). Phylogenetic Analysis of the UDP-glycosyltransferase Multigene Family of *Arabidopsis thaliana*. *Journal of Biological Chemistry* **276**, 4338-4343.
- Li, Z., Jiang, X., Peng, T., Shi, C., Han, S., Tian, B., Zhu, Z., and Tian, J. (2014). Mapping quantitative trait loci with additive effects and additive x additive epistatic interactions for biomass yield, grain yield, and straw yield using a doubled haploid population of wheat (*Triticum aestivum* L.). *Genet. Mol. Res* **13**, 1412-1424.
- Lipman, D., and Pearson, W. (1985). Rapid and sensitive protein similarity searches. *Science* **227**, 1435-1441.

- Lisch, D. (2013). How important are transposons for plant evolution? *Nature Reviews Genetics* **14**, 49-61.
- Lotti, C., Salvi, S., Pasqualone, A., Tuberosa, R., and Blanco, A. (2000). Integration of AFLP markers into an RFLP-based map of durum wheat. *Plant Breeding* **119**, 393-401.
- Luo, Z. W., and Kearsey, M. J. (1989). Maximum likelihood estimation of linkage between a marker gene and a quantitative locus. *Heredity* **63**, 401-408.
- Ma, J., Li, H. B., Zhang, C. Y., Yang, X. M., Liu, Y. X., Yan, G. J., and Liu, C. J. (2010). Identification and validation of a major QTL conferring crown rot resistance in hexaploid wheat. *Theor Appl Genet* **120**.
- Maccaferri, M., Francia, R., Ratti, C., Rubies-Autonell, C., Colalongo, C., Ferrazzano, G., Tuberosa, R., and Sanguineti, M. C. (2012). Genetic analysis of Soil-Borne Cereal Mosaic Virus response in durum wheat: evidence for the role of the major quantitative trait locus QSbm.ubo-2BS and of minor quantitative trait loci. *Molecular Breeding* **29**, 973-988.
- Maccaferri, M., Ratti, C., Rubies-Autonell, C., Vallega, V., Demontis, A., Stefanelli, S., Tuberosa, R., and Sanguineti, M. C. (2011). Resistance to Soil-borne cereal mosaic virus in durum wheat is controlled by a major QTL on chromosome arm 2BS and minor loci. *Theoretical and Applied Genetics* **123**, 527-544.
- Maccaferri, M., Ricci, A., Salvi, S., Milner, S. G., Noli, E., Martelli, P. L., Casadio, R., Akhunov, E., Scalabrin, S., Vendramin, V., Ammar, K., Blanco, A., Desiderio, F., Distelfeld, A., Dubcovsky, J., Fahima, T., Faris, J., Korol, A., Massi, A., Mastrangelo, A. M., Morgante, M., Pozniak, C., N'Diaye, A., Xu, S., and Tuberosa, R. (2015). A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnology Journal* **13**, 648-663.
- Maccaferri, M., Sanguineti, M. C., Corneti, S., Ortega, J. L. A., Salem, M. B., Bort, J., DeAmbrogio, E., del Moral, L. F. G., Demontis, A., El-Ahmed, A., Maalouf, F., Machlab, H., Martos, V., Moragues, M., Motawaj, J., Nachit, M., Nserallah, N., Ouabbou, H., Royo, C., Slama, A., and Tuberosa, R. (2008). Quantitative Trait Loci for Grain Yield and Adaptation of Durum Wheat (*Triticum durum* Desf.) Across a Wide Range of Water Availability. *GENETICS* **178**, 489-511.

- Magadum, S., Banerjee, U., Murugan, P., Gangapur, D., and Ravikesavan, R. (2013). Gene duplication as a major force in evolution. *Journal of genetics* **92**, 155-161.
- Mangini, G., Blanco, A., Nigro, D., Signorile, M. A., and Simeone, R. (2020). Candidate Genes and Stable QTL for Grain Yield and Seed Size in Durum Wheat. Research Square.
- Mangini, G., Blanco, A., Nigro, D., Signorile, M. A., and Simeone, R. (2021). Candidate Genes and Quantitative Trait Loci for Grain Yield and Seed Size in Durum Wheat. *Plants* **10**.
- Mansour, A. (2007). Epigenetic activation of genomic retrotransposons. *Cell Molecular Biology* **6**, 99-107.
- Mantovani, P., Maccaferri, M., Sanguineti, M., Tuberosa, R., Catizone, I., Wenzl, P., Thomson, B., Carling, J., Huttner, E., DeAmbrogio, E., and Kilian, A. (2008). An integrated DArT-SSR linkage map of durum wheat. *Mol Breed* **22**, 629–648.
- Marcotuli, I., Gadaleta, A., Mangini, G., Signorile, A. M., Zacheo, S. A., Blanco, A., Simeone, R., and Colasuonno, P. (2017). Development of a High-Density SNP-Based Linkage Map and Detection of QTL for β -Glucans, Protein Content, Grain Yield per Spike and Heading Time in Durum Wheat. *International Journal of Molecular Sciences* **18**.
- Marino, C. L., Tuleen, N. A., Hart, G. E., Nelson, J. C., Sorrells, M. E., Lu, Y. H., Leroy, P., and Lopes, C. R. (1996). Molecular genetic maps of the group 6 chromosomes of hexaploid wheat (*Triticum aestivum* L. em. Thell.). *Genome* **39**, 359-366.
- Marone, D., Del Olmo, A. I., Laidò, G., Sillero, J. C., Emeran, A. A., Russo, M. A., Ferragonio, P., Giovanniello, V., Mazzucotelli, E., De Leonardis, A. M., De Vita, P., Blanco, A., Cattivelli, L., Rubiales, D., and Mastrangelo, A. M. (2009). Genetic analysis of durable resistance against leaf rust in durum wheat. *Molecular Breeding* **24**, 25-39.
- Marone, D., Laidò, G., Gadaleta, A., Colasuonno, P., Ficco, D. B. M., Giancaspro, A., Giove, S., Panio, G., Russo, M. A., De Vita, P., Cattivelli, L., Papa, R., Blanco, A., and Mastrangelo, A. M. (2012a). A high-density consensus map of A and B wheat genomes. In "Theor Appl Genet", Vol. 125, pp. 1619–1638.
- Marone, D., Panio, G., Ficco, D. B. M., Russo, M. A., De Vita, P., Papa, R., Rubiales, D., Cattivelli, L., and Mastrangelo, A. M. (2012b). Characterization of wheat DArT markers: genetic and functional features. *Mol Genet Genom* **287**, 741–753.

- Martínez, O., and Curnow, R. N. (1992). Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. *Theoretical and Applied Genetics* **85**, 480-488.
- McCouch, S., Kochert, G., Yu, Z., Wang, Z., Khush, G., Coffman, W., and Tanksley, S. (1988). Molecular mapping of rice chromosomes. *Theoretical and Applied Genetics* **76**, 815-829.
- McIntyre, C. L., Mathews, K. L., Rattey, A., Chapman, S. C., Drenth, J., Ghaderi, M., Reynolds, M., and Shorter, R. (2010). Molecular detection of genomic regions associated with grain yield and yield-related components in an elite bread wheat cross evaluated under irrigated and rainfed conditions. *Theoretical and Applied Genetics* **120**, 527-541.
- Mérida-García, R., Liu, G., He, S., Gonzalez-Dugo, V., Dorado, G., Gálvez, S., Solís, I., Zarco-Tejada, P. J., Reif, J. C., and Hernandez, P. (2019). Genetic dissection of agronomic and quality traits based on association mapping and genomic selection approaches in durum wheat grown in Southern Spain. *PloS one* **14**, e0211718-e0211718.
- Ministry of agriculture. (2022). <https://www.agriculture.gov.ma/fr/filiere/Cerealiere> (visité le: 02/03/2022).
- Molnár, I., Vrána, J., Burešová, V., Cápál, P., Farkas, A., Darkó, É., Cseh, A., Kubaláková, M., Molnár-Láng, M., and Doležel, J. (2016). Dissecting the U, M, S and C genomes of wild relatives of bread wheat (*Aegilops* spp.) into chromosomes and exploring their synteny with wheat. *The Plant Journal* **88**, 452-467.
- Moon, J., Parry, G., and Estelle, M. (2004). The ubiquitin-proteasome pathway and plant development. *The Plant Cell* **16**, 3181-3195.
- Moore, G., Devos, K. M., Wang, Z., and Gale, M. D. (1995). Cereal Genome Evolution: Grasses, line up and form a circle. *Current Biology* **5**, 737-739.
- Morton, N. E. (1955). Sequential tests for the detection of linkage. *American journal of human genetics* **7**, 277-318.
- Nachit, M. (1992). Durum wheat breeding for Mediterranean drylands of North Africa and West Asia. Durum Wheats: Challenges and Opportunities. International Workshop; Cd. Obregon, Son. (Mexico); 23-25 Mar 1992.

- Nachit, M. (1998). Associations of grain yield in dryland and carbon isotope discrimination with molecular markers in durum (*Triticum turgidum* L. var *durum*). International Wheat Genetics Conference. Proceedings of the International Wheat Genetics Symposium, 9; Saskatoon (Canada); 2-7 Aug 1998. In "International Wheat Genetics Conference. Proceedings of the International Wheat Genetics Symposium", Vol. 9, Saskatoon (Canada)
- Nachit, M. M., and Elouafi, I. (2004). Durum Wheat Adaptation in the Mediterranean Dryland: Breeding, Stress Physiology, and Molecular Markers. In "Challenges and Strategies of Dryland Agriculture" (S. C. Rao and J. Ryan, eds.), pp. 203-218. Crop Science Society of America and American Society of Agronomy, Madison, WI.
- Nachit, M. M., Elouafi, I., Pagnotta, A., El Saleh, A., Iacono, E., Labhilili, M., Asbati, A., Azrak, M., Hazzam, H., Benscher, D., Khairallah, M., Ribaut, J. M., Tanzarella, O. A., Porceddu, E., and Sorrells, M. E. (2001). Molecular linkage map for an intraspecific recombinant inbred population of durum wheat (*Triticum turgidum* L. var. *durum*). *Theoretical and Applied Genetics* **102**, 177-186.
- Nachit, M. M., Elouafi, I., Rao, S., and Ryan, J. (2004). "Durum wheat adaptation in the Mediterranean dryland: Breeding, stress physiology, and molecular markers," Crop Science of America and American Society of Agronomy Madison, WI.
- Nachit, M., Baum, M., Impiglia, A., and Ketata, H. (1995). Studies on some grain quality traits in durum wheat grown in Mediterranean environments. In "Durum wheat quality in the Mediterranean region". Vol. 22, pp. 181-187. Zaragoza: CIHEAM.
- Nagel, M., Navakode, S., Scheibal, V., Baum, M., Nachit, M., Röder, M., and Börner, A. (2014). The genetic basis of durum wheat germination and seedling growth under osmotic stress. *Biologia plantarum* **58**, 681-688.
- Nakai, Y., Fujiwara, S., Kubo, Y., Sato, M. H., and behavior (2013a). Overexpression of VOZ2 confers biotic stress tolerance but decreases abiotic stress resistance in *Arabidopsis*. *Plant signaling* **8**, e23358.
- Nakai, Y., Nakahira, Y., Sumida, H., Takebayashi, K., Nagasawa, Y., Yamasaki, K., Akiyama, M., Ohme-Takagi, M., Fujiwara, S., and Shiina, T. (2013b). Vascular plant one-zinc-finger protein 1/2 transcription factors regulate abiotic and biotic stress responses in *Arabidopsis*. *Plant Journal* **73**, 761-775.

- Nalam, V. J., Vales, M. I., Watson, C. J. W., Kianian, S. F., and Riera-Lizarazu, O. (2006). Map-based analysis of genes affecting the brittle rachis character in tetraploid wheat (*Triticum turgidum* L.). *Theoretical and Applied Genetics* **112**, 373-381.
- Nasernakhaei, F., Rahiminejad, M. R., Saeidi, H., and Tavassoli, M. (2015). Phylogenetic relationships among the Iranian *Triticum* diploid gene pool as inferred from the loci Acc1 and Pgk1. *Phytotaxa* **201**, 111-121.
- Needleman, S. B., and Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology* **48**, 443-453.
- Negi, P., Rai, A. N., and Suprasanna, P. (2016). Moving through the Stressed Genome: Emerging Regulatory Roles for Transposons in Plant Stress Response. *frontiers in Plant Science* **7**.
- Neimann-Sorensen, A., and Robertson, A. (1961). The Association between Blood Groups and Several Production Characteristics in Three Danish Cattle Breeds. *Acta Agriculturae Scandinavica* **11**, 163-196.
- Nelson, J. C., Deynze, A. E. V., Sorrells, M. E., Autrique, E., Lu, Y. H., Merlino, M., Atkinson, M., and Leroy, P. (1995a). Molecular mapping of wheat. Homoeologous group 2. *Genome* **38**, 516-524.
- Nelson, J. C., Deynze, A. E. V., Sorrells, M. E., Autrique, E., Lu, Y. H., Negre, S., Bernard, M., and Leroy, P. (1995b). Molecular mapping of wheat. Homoeologous group 3. *Genome* **38**, 525-533.
- Nsarellah, N., Amri, A., Nachit, M., Bouhssini, M. E., Lhaloui, S., and Lorenzoni (2003). New durum wheat with Hessian fly resistance from *Triticum araraticum* and *T. carthlicum* in Morocco. *Plant breeding* **122**, 435-437.
- O'Donoghue, L., Sorrells, M., Tanksley, S., Autrique, E., Deynze, A. V., Kianian, S., Phillips, R., Wu, B., Rines, H., and Rayapati, P. (1995). A molecular linkage map of cultivated oat. *Genome* **38**, 368-380.
- O'Donoghue, L., Wang, Z., Röder, M., Kneen, B., Leggett, M., Sorrells, M., and Tanksley, S. (1992). An RFLP-based linkage map of oats based on a cross between two diploid taxa (*Avena atlantica* × *A. hirtula*). *Genome* **35**, 765-771.

- Ohyanagi, H., Tanaka, T., Sakai, H., Shigemoto, Y., Yamaguchi, K., Habara, T., Fujii, Y., Antonio, B. A., Nagamura, Y., Imanishi, T., Ikeo, K., Itoh, T., Gojobori, T., and Sasaki, T. (2006). The Rice Annotation Project Database (RAP-DB): hub for *Oryza sativa* ssp. *japonica* genome information. *Nucleic Acids Research* **34**, D741-D744.
- ONICL. (2022). <https://www.onicl.org.ma/> (visit  le : 05/10/2022)
- Padmarasu, S., Sargent, D. J., Patocchi, A., Troggio, M., Baldi, P., Linsmith, G., Poles, L., J nsch, M., Kellerhals, M., Tartarini, S., and Velasco, R. (2018). Identification of a leucine-rich repeat receptor-like serine/threonine-protein kinase as a candidate gene for Rvi12 (Vb)-based apple scab resistance. *Molecular Breeding* **38**, 73.
- Panchy, N., Lehti-Shiu, M., and Shiu, S.-H. (2016). Evolution of gene duplication in plants. *Plant physiology* **171**, 2294-2316.
- Pearson, W. R., and Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proceedings of the National Academy of Sciences* **85**, 2444-2448.
- Peleg, Z., Saranga, Y., Suprunova, T., Ronin, Y., R der, M. S., Kilian, A., Korol, A. B., and Fahima, T. (2008). High-density genetic map of durum wheat x wild emmer wheat based on SSR and DArT markers. *Theor Appl Genet* **117**, 103–115.
- Pradhan, D., Mathew, D., Mathew, S. K., and Nazeem, P. A. (2018). Identifying the markers and tagging a leucine-rich repeat receptor-like kinase gene for resistance to anthracnose disease in vegetable cowpea [*Vigna unguiculata* (L.) Walp.]. *The Journal of Horticultural Science and Biotechnology* **93**, 225-231.
- Pradhan, S., Babar, M. A., Bai, G., Khan, J., Shahi, D., Avci, M., Guo, J., McBreen, J., Asseng, S., Gezan, S., Baik, B.-K., Blount, A., Harrison, S., Sapkota, S., St. Amand, P., and Kunwar, S. (2020). Genetic dissection of heat-responsive physiological traits to improve adaptation and increase yield potential in soft winter wheat. *BMC Genomics* **21**, 315.
- Prasad, V., Str mberg, C. A. E., Alimohammadian, H., and Sahni, A. (2005). Dinosaur Coprolites and the Early Evolution of Grasses and Grazers. *Science* **310**, 1177-1180.
- Quraishi, U. M., Abrouk, M., Murat, F., Pont, C., Foucrier, S., Desmaizieres, G., Charmet, G., Paux, E., Murigneux, A., Guerreiro, L., Lafarge, S., Le Gouis, J., Feuillet, C., and Salse, J. (2011). Cross-genome map based dissection of a nitrogen use efficiency ortho-metaQTL in bread wheat unravels concerted cereal genome evolution. *Plant J* **65**.

- Rahman, M. S., Linsell, K. J., Taylor, J. D., Hayden, M. J., Collins, N. C., and Oldach, K. H. (2020). Fine mapping of root lesion nematode (*Pratylenchus thornei*) resistance loci on chromosomes 6D and 2B of wheat. *Theoretical and Applied Genetics* **133**, 635-652.
- Rayapati, P., Gregory, J., Lee, M., and Wise, R. (1994). A linkage map of diploid *Avena* based on RFLP loci and a locus conferring resistance to nine isolates of *Puccinia coronata* var. 'avenae'. *Theoretical and Applied Genetics* **89**, 831-837.
- Rebai, A., Goffinet, B., and Mangin, B. (1995). Comparing Power of Different Methods for QTL Detection. *Biometrics* **51**, 87-99.
- Rekika, D., Nachit, M., Araus, J., and Monneveux, P. (1998). Effects of water deficit on photosynthetic rate and osmotic adjustment in tetraploid wheats. *Photosynthetica* **35**, 129-138.
- Ribaut, J. M., Jiang, C., Gonzalez-de-Leon, D., Edmeades, G. O., and Hoisington, D. A. (1997). Identification of quantitative trait loci under drought conditions in tropical maize. 2. Yield components and marker-assisted selection strategies. *Theoretical and Applied Genetics* **94**, 887-896.
- Röder, M. S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M.-H., Leroy, P., and Ganal, M. W. (1998). A microsatellite map of wheat. *Genetics* **149**, 2007-2023.
- Ruan, Y., Yu, B., Knox, R. E., Singh, A. K., DePauw, R., Cuthbert, R., Zhang, W., Piche, I., Gao, P., Sharpe, A., and Fobert, P. (2020). High Density Mapping of Quantitative Trait Loci Conferring Gluten Strength in Canadian Durum Wheat. *Frontiers in Plant Science* **11**.
- Russo, M. A., Ficco, D. B. M., Laidò, G., Marone, D., Papa, R., Blanco, A., Gadaleta, A., De Vita, P., and Mastrangelo, A. M. (2014). A dense durum wheat $\times$ *T. dicoccum* linkage map based on SNP markers for the study of seed morphology. *Molecular Breeding* **34**, 1579-1597.
- Russo, M. A., Ficco, D. B. M., Marone, D., De Vita, P., Vallega, V., Rubies-Autonell, C., Ratti, C., Ferragonio, P., Giovanniello, V., Pecchioni, N., Cattivelli, L., and Mastrangelo, A. M. (2012). A major QTL for resistance to soil-borne cereal mosaic virus derived from an old Italian durum wheat cultivar. *Journal of Plant Interactions* **7**, 290-300.
- Sakai, H., Lee, S. S., Tanaka, T., Numa, H., Kim, J., Kawahara, Y., Wakimoto, H., Yang, C.-c., Iwamoto, M., Abe, T., Yamada, Y., Muto, A., Inokuchi, H., Ikemura, T., Matsumoto,

- T., Sasaki, T., and Itoh, T. (2013). Rice Annotation Project Database (RAP-DB): An Integrative and Interactive Database for Rice Genomics. *Plant and Cell Physiology* **54**, e6-e6.
- Saliba-Colombani, V., Causse, M., Gervais, L., and Philouze, J. (2000). Efficiency of RFLP, RAPD, and AFLP markers for the construction of an intraspecific map of the tomato genome. *Genome* **43**, 29-40.
- Salse, J., Bolot, S., Throude, M., Jouffe, V., Piegu, B., Quraishi, U. M., Calcagno, T., Cooke, R., Delseny, M., and Feuillet, C. (2008). Identification and Characterization of Shared Duplications between Rice and Wheat Provide New Insight into Grass Genome Evolution. *The Plant Cell journal* **20**, 11-24.
- Sasaki, T., and International Rice Genome Sequencing, P. (2005). The map-based sequence of the rice genome. *Nature* **436**, 793-800.
- Sauvage, H., (2017). Article published in Cultivar-magazine (<https://www.cultivar.fr/sinformer/des-stocks-mondiaux-au-plus-bas>).
- Schumm, J. W., Knowlton, R. G., Braman, J. C., Barker, D. F., Botstein, D., Akots, G., Brown, V. A., Gravius, T. C., Helms, C., Hsiao, K., Rediker, K., Thurston, J. G., and Donis-Keller (1988). Identification of more than 500 RFLPs by screening random genomic clones. *American journal of human genetics* **42**, 143-159.
- Shewry, P. R. (2009). Wheat. *Journal of Experimental Botany* **60**, 1537-1553.
- Shinada, N. K., Schmidtke, P., and de Brevern, A. G. (2020). Accurate Representation of Protein-Ligand Structural Diversity in the Protein Data Bank (PDB). *International Journal of Molecular Sciences* **21**.
- Simeone, R., Piarulli, L., Nigro, D., Signorile, M. A., Blanco, E., Mangini, G., and Blanco, A. (2020). Mapping Powdery Mildew (*Blumeria graminis* f. sp. *tritici*) Resistance in Wild and Cultivated Tetraploid Wheats. *International Journal of Molecular Sciences* **21**.
- Singh, N. K., Raghuvanshi, S., Srivastava, S. K., Gaur, A., Pal, A. K., Dalal, V., Singh, A., Ghazi, I. A., Bhargav, A., Yadav, M., Dixit, A., Batra, K., Gaikwad, K., Sharma, T. R., Mohanty, A., Bharti, A. K., Kapur, A., Gupta, V., Kumar, D., Vij, S., Vydianathan, R., Khurana, P., Sharma, S., McCombie, W. R., Messing, J., Wing, R., Sasaki, T., Khurana, P., Mohapatra, T., Khurana, J. P., and Tyagi, A. K. (2004). Sequence analysis of the

- long arm of rice chromosome 11 for rice–wheat synteny. *Functional & Integrative Genomics* **4**, 102-117.
- Smith, T. F., and Waterman, M. S. (1981). Identification of common molecular subsequences. *Journal of Molecular Biology* **147**, 195-197.
- Soller, M., Brody, T., and Genizi, A. (1976). On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *Theoretical and Applied Genetics* **47**, 35-39.
- Soltis, P. S. (2005). Ancient and Recent Polyploidy in Angiosperms. *The New Phytologist* **166**, 5-8.
- Somers, D. J., Isaac, P., and Edwards, K. (2004). A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* **109**, 1105-1114.
- Sood, S., Kuraparthi, V., Bai, G., and Gill, B. S. (2009). The major threshability genes soft glume (sog) and tenacious glume (Tg), of diploid and polyploid wheat, trace their origin to independent mutations at non-orthologous loci. *Theoretical and Applied Genetics* **119**, 341-351.
- Soreng, r. j., Peterson, p. m., Romaschenko, k., Davidse, g., Zuloaga, f. o., Judziewicz, e. j., Filgueiras, t. s., Davis, j. i., and Morrone, o. (2015). A worldwide phylogenetic classification of the *Poaceae* (*Gramineae*). *Journal of Systematics and Evolution* **53**, 117-137.
- Su, Q., Zhang, X., Zhang, W., Zhang, N., Song, L., Liu, L., Xue, X., Liu, G., Liu, J., Meng, D., Zhi, L., Ji, J., Zhao, X., Yang, C., Tong, Y., Liu, Z., and Li, J. (2018). QTL Detection for Kernel Size and Weight in Bread Wheat (*Triticum aestivum* L.) Using a High-Density SNP and SSR-Based Linkage Map. *Frontiers in plant science* **9**, 1484-1484.
- Subramanyam, S., and Nemacheck, J. A. (2021). Initiation of compatible wheat-Hessian fly interactions triggers the expression of a novel UDP-glycosyltransferase, MdesUGT1, in virulent Hessian fly larvae. *Arthropod-Plant Interactions* **15**, 363-374.
- Sukumaran, S., Reynolds, M. P., and 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. *Frontiers in Plant Science* **9**.

- Sussman, J. L., Lin, D., Jiang, J., Manning, N. O., Prilusky, J., Ritter, O., and Abola, E. E. (1998). Protein Data Bank (PDB): Database of Three-Dimensional Structural Information of Biological Macromolecules. *Acta Crystallographica Section D* **54**, 1078-1084.
- Tanksley, S., Ganai, M., Prince, J., De Vicente, M., Bonierbale, M., Broun, P., Fulton, T., Giovannoni, J., Grandillo, S., and Martin, G. (1992). High density molecular linkage maps of the tomato and potato genomes. *Genetics* **132**, 1141-1160.
- The Gene Ontology Consortium (2019). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research* **47**, D330-D338.
- Trebbi, D., Maccaferri, M., de Heer, P., Sørensen, A., Giuliani, S., Salvi, S., Sanguineti, M. C., Massi, A., van der Vossen, E. A. G., and Tuberosa, R. (2011). High-throughput SNP discovery and genotyping in durum wheat (*Triticum durum* Desf.). *Theoretical and Applied Genetics* **123**, 555-569.
- Tura, H., Edwards, J., Gahlaut, V., Garcia, M., Sznajder, B., Baumann, U., Shahinnia, F., Reynolds, M., Langridge, P., and Balyan, H. S. (2019). QTL analysis and fine mapping of a QTL for yield-related traits in wheat grown in dry and hot environments. *Theoretical and Applied Genetics* **133**, 239–257.
- Twiss, P. C. (1992). Predicted World Distribution of C3 and C4 Grass Phytoliths. In "Phytolith Systematics: Emerging Issues" (G. Rapp and S. C. Mulholland, eds.), pp. 113-128. Springer US, Boston, MA.
- Van Poecke, R. M. P., Maccaferri, M., Tang, J., Truong, H. T., Janssen, A., van Orsouw, N. J., Salvi, S., Sanguineti, M. C., Tuberosa, R., and van der Vossen, E. A. G. (2013). Sequence-based SNP genotyping in durum wheat. *Plant Biotechnology Journal* **11**, 809-817.
- Verlotta, A., De Simone, V., Mastrangelo, A. M., Cattivelli, L., Papa, R., and Trono, D. (2010). Insight into durum wheat Lpx-B1: a small gene family coding for the lipoxygenase responsible for carotenoid bleaching in mature grains. *BMC Plant Biology* **10**, 263.
- Wang, H., Zhao, S., Mao, K., Dong, Q., Liang, B., Li, C., Wei, Z., Li, M., and Ma, F. (2018). Mapping QTLs for water-use efficiency reveals the potential candidate genes involved in regulating the trait in apple under drought stress. *BMC Plant Biology* **18**, 136.

- Wang, J., Liu, W., Wang, H., Li, L., Wu, J., Yang, X., Li, X., and Gao, A. (2011). QTL mapping of yield-related traits in the wheat germplasm 3228. *Euphytica* **177**, 277-292.
- Wang, X., Yu, Y., Sang, J., Wu, Q., Zhang, X., and Lin, Z. (2013). Intraspecific linkage map construction and QTL mapping of yield and fiber quality of *Gossypium babardense*. *Australian Journal of Crop Science* **7**, 1252.
- Wang, Z., Li, H., Wei, Z., Sun, H., He, Y., Gao, J., Yang, Z., and You, J. (2021). Overexpression of UDP-glycosyltransferase genes enhanced aluminum tolerance through disrupting cell wall polysaccharide components in soybean. *Plant Soil*, 1-13.
- Watanabe, N., Akond, A. S. M. G. M., and Nachit, M. M. (2006). Genetic mapping of the gene affecting polyphenol oxidase activity in tetraploid durum wheat. *Journal of Applied Genetics* **47**, 201-205.
- White, R., Leppert, M., Bishop, D. T., Barker, D., Berkowitz, J., Brown, C., Callahan, P., Holm, T., and Jerominski, L. (1985). Construction of linkage maps with DNA markers for human chromosomes. *Nature* **313**, 101-105.
- Wolfe, K. H. (2001). Yesterday's polyploids and the mystery of diploidization. *Nature Reviews Genetics* **2**, 333-341.
- Wootton, J. C., and Federhen, S. (1996). Analysis of compositionally biased regions in sequence databases. *Methods Enzymol* **266**, 554-71.
- Xie, D., Devos, K., Moore, G., and Gale, M. (1993). RFLP-based genetic maps of the homoeologous group 5 chromosomes of bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* **87**, 70-74.
- Xu, D., Wen, W., Fu, L., Li, F., Li, J., Xie, L., Xia, X., Ni, Z., He, Z., and Cao, S. (2019). Genetic dissection of a major QTL for kernel weight spanning the Rht-B1 locus in bread wheat. *Theoretical and Applied Genetics* **132**, 3191-3200.
- Yu, J., Hu, F., Dossa, K., Wang, Z., and Ke, T. (2017). Genome-wide analysis of UDP-glycosyltransferase super family in *Brassica rapa* and *Brassica oleracea* reveals its evolutionary history and functional characterization. *BMC genomics* **18**, 1-18.
- Yu, J., Miao, Y., Yang, S., Shi, Z., Miao, N., Ding, M., Zhang, H., Jiang, Y., and Rong, J. (2019). Identification and mapping of a photoperiod response gene (QPpd.zafu-4A) on wild emmer wheat (*Triticum turgidum* L.) chromosome 4AL. *Euphytica* **215**, 146.

- Zeng, Q., Wu, J., Huang, S., Yuan, F., Liu, S., Wang, Q., Mu, J., Yu, S., Chen, L., Han, D., and Kang, Z. (2019). SNP-based linkage mapping for validation of adult plant stripe rust resistance QTL in common wheat cultivar Chakwal 86. *The Crop Journal*.
- Zeng, Z. B. (1994). Precision mapping of quantitative trait loci. *Genetics* **136**, 1457-1468.
- Zeng, Z.-B., Kao, C.-H., and Basten, C. J. (1999). Estimating the genetic architecture of quantitative traits. *Genetical Research* **74**, 279-289.
- Zhang, C., Gao, H., Li, R., Han, D., Wang, L., Wu, J., Xu, P., and Zhang, S. (2019). GmBTB/POZ, a novel BTB/POZ domain-containing nuclear protein, positively regulates the response of soybean to *Phytophthora sojae* infection. *Molecular Plant Pathology* **20**, 78-91.
- Zhang, H., Chen, J., Li, R., Deng, Z., Zhang, K., Liu, B., and Tian, J. (2016). Conditional QTL mapping of three yield components in common wheat (*Triticum aestivum* L.). *The Crop Journal* **4**, 220-228.
- Zhang, L., Li, H., Li, Z., and Wang, J. (2008). Interactions between markers can be caused by the dominance effect of quantitative trait loci. *Genetics* **180**, 1177-90.
- Zhang, L., Luo, J.-T., Hao, M., Zhang, L.-Q., Yuan, Z.-W., Yan, Z.-H., Liu, Y.-X., Zhang, B., Liu, B.-L., Liu, C.-J., Zhang, H.-G., Zheng, Y.-L., and Liu, D.-C. (2012). Genetic map of *Triticum turgidum* based on a hexaploid wheat population without genetic recombination for D genome. *BMC Genetics* **13**, 69.
- Zhang, L., Wang, S., Li, H., Deng, Q., Zheng, A., Li, S., Li, P., Li, Z., and Wang, J. (2010). Effects of missing marker and segregation distortion on QTL mapping in F2 populations. *Theoretical and Applied Genetics* **121**, 1071-1082.
- Zhang, W., Chao, S., Manthey, F., Chicaiza, O., Brevis, J. C., Echenique, V., and Dubcovsky, J. (2008). QTL analysis of pasta quality using a composite microsatellite and SNP map of durum wheat. *Theor Appl Genet* **117**, 1361–1377.