Understanding the genetic control of adaptation and yield for Kazakh bread wheat

Kanat Yermekbayev

A thesis submitted to the University of East Anglia for the degree of Doctor of Philosophy

John Innes Centre

April 2022

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there from must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution. ©

DEDICATION

This work is dedicated to my Grandparents, Yermekbayev Serikbay Yermekbayuly and Shalimbetov Kuanyshbek Shalimbetuly

Abstract

Kazakhstan is a major wheat exporter in the region. However, an average wheat grain yield has not changed since 1960's in the country and plant breeding is mainly based on conventional breeding methods. Here, we developed the first segregating population for Kazakhstan with the cross of UK and KZ wheats. Thus, the crossing of two wheat varieties adapted and bred for diverse climatic conditions allows the discovery of variations that might be important for adaptation to Kazakh conditions. Several adaptation and yield related QTL were discovered among which are two plant height QTL with large additive effects. NILs and NIL combinations were developed for the loci. These new resources were used to study how height increasing QTL affect agronomic performance in Kazakhstan. The NIL analysis validated the effect of the two height QTL in the UK and showed that the chromosome 6A effect was the most stable. The height increases conferred by Pamyati Azieva (Kazakh wheat) on chromosomes 5A and 6A did not result in significant changes in yield or stress tolerance in Kazakh field trials. Recombinants from both QTL regions were used for fine-mapping leading to the identification of candidate genes and molecular markers for the loci. A large Central-Asian wheat panel was established. The panel was genetically fingerprinted with Axiom 35K genotyping platform revealing detailed population structure of Central-Asian germplasm and it's relationship with other global germplasm collections. The haplotype led approach in this population helped us to further refine the two target loci and haplotype analysis suggested that diversity for both loci is very low in Kazakh germplasm. Taking into account the yield data and the beneficial effect of reduced height for standing power it is proposed that these haplotypes are used to select against the common Pamyati Azieva alleles and introduce height reducing alleles, exemplifying a haplotype led pre-breeding strategy for Central-Asia. The new resources developed here can help with similar approaches to address the numerous trait targets of Central-Asian wheat breeders.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate ‘take down’ action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

Acknowledgments

First of all, I would like to express my full gratitude to Allah (SWT), most gracious and merciful, providing me with this great opportunity of conducting my PhD research at the John Innes Centre, the world-leading international centre of excellence in plant science, and submit my PhD thesis in the blessed month of Ramadan 2022.

Secondly, I would like to thank my primary supervisor, Simon Griffiths, for his endless support, guidance and enthusiasm and being always considered throughout the project and years of my life in the UK. I do not remember you saying "No" regardless of the case being project or life related! You supported me in everything. I learned so much from you in the field of genetics. So, I really appreciate every second of time I spent with you discussing the genetics. If I were asked to return the time and do my PhD research again, I would have certainly chosen you as my supervisor.

The next person who must receive "huge thanks" is my external supervisor, Yerlan Kenesbekovich Turuspekov, who I was inspired by initially. The same gratitude and thanks go to him. I still remember when entered your laboratory, I did not believe my eyes that such a well-equipped lab exits in Kazakhstan. Then I realised, it came with your enormous effort and time. I am also going through the same path you passed establishing my own lab and I am 100% sure that I have the full support of you! When I did my first DNA extraction and PCR in your lab, I had a dream to have a dream so that I could understand everything (theoretically and technically) what your lab was doing at that time. However, it needed time and proper education , so I had to go to the UK to study 😊. Anyway, it was a great honour for me being part of your team and that was one of the best times of my life.

I would like also to thank my secondary supervisor, Anthony Hall, for being extremely useful board who provided enthusiasm for the project from the very beginning till the very end. You always welcomed me in your lab. Your valuable thoughts made the project more successful than what it would have been.

The next thanks go to Simon Orford who has been a key and active member of my supervisory team. Your every support and advice are valued and appreciated so much. When I was recently thinking "O, I should have done this and that" to make my PhD thesis better, your encouragement "saying, actually Kanat this is a good sign, that means you have various ways to continue your research" was huge push for me to complete this thesis and think about brighter future.

I have been extremely lucky to be part of an excellent team not only in KZ, also in the UK. Thus, my huge thanks and appreciation go to each current and former member of the Griffiths' and Turuspekov' groups.

In Griffiths' team, I really want to thank Clare, Michelle and Sarah who helped me to settle not only in the lab, also in general life. I would say that's the most important part of PhD life abroad. Clare, also thanks for preparing my working bench before I joined to the group. I did take a photo of the note you left for me on my bench saying "Welcome Kanat! I cleaned up everything for you, this is your bench" and at that time I thought, I will include this photo into the "Acknowledgments" section of my thesis when I will be saying thanks to Clare, but unfortunately my phone was broken down and I couldn't recover the photo ☹. Additionally, many thanks to Luzie for providing me with "Lifesaving" codes. Luzie, you were always helpful when I stuck with stats and coding. Also, thanks for offering your home to stay during my last months of my PhD life. That was something I appreciated so much. I would also like to deliver a special thanks to Rajani, Charlie, Richard and Sophie for their endless and on-time helps. Thanks guys, I am really grateful what you have done for me! Also, thanks to Liz who is the former PhD student in the Griffiths lab.

In Yerlan Kenesbekovich's lab, firstly, I would like to thank Saule Izteleuovna who was always very considered about me before and during the PhD research, especial thanks for organising all the paperwork related to my field trails in Kazakhstan. Saule Izteleuovna, again, many thanks for arranging these all for me. I am sure nobody could have done it better than you! Also, my gratitude goes to Akerke and Alibek for sending and receiving the seeds from and to the UK, and for arranging field experiments in Almaty. I also thank other members of the lab for welcoming me warmly always when I am/was there.

The next, but important thanks go to my all-family members. I am profoundly grateful to my dad, Abay, and mum, Laziza, for their love and encouragement, without whom I would never have enjoyed so many opportunities. I am thankful to my brother, Askat and his family, and Almas, for the unceasing support and attention. Special thanks to my wife, Makpal and my lovely daughter Akniyet and wonderful sons Almat, Mansur and Akhmediyar without whom I would never achieved what I have achieved so far. I do not remember anytime that they distracted me from doing my PhD research. Instead, their warm welcome after a long day of work has always put me on the path.

I would like to express my sincere gratitude to "Bolashak" and the John Innes Centre for generously funding my studies. Special thanks to my adviser Elvira Myrzaglieva for always being responsive and arranging my flight tickets.

I would also like to thank all JIC Crop Genetics Department, Field Experimentation and Horticultural Services members, my friends and other people who supported me and offered deep insight into the study.

I would also like to thank GRU members, especially Liz for obtaining phytosanitary certificates on time when I needed them.

Table of Contents

Table of figures

Table of tables

List of Abbreviations

BAM - Binary Alignment Map

CA - Central Asian

CAC - Central Asian Countries

CAWBIN - Central Asian Wheat Breeding Initiative

CER - Controlled environment room

CNVs - Copy number variations

CS - Chinese Spring

DH - Doubled Haploids

DLTAA - Deviation from the LTAA

dNILs - Double NILs

DTEM - Days To Ear Emergence or Heading Date

EC - Exome capture

EPS - Earliness Per Se

EST - Estimated difference between the means

ETN - Effective tiller number

FTD - First Ten Days

GH – Glasshouse

GWAS - Genome Wide Association Study

GWMS - Grain weight per main spike

GWP - Grain weight per plant

HI - Harvest index

HTB - Haplotype Block

HTC - Hydrothermal coefficient

InDels - Insertion–deletion mutations

Int1 - Internode 1

Int2 - Internode 2

Int3 - Internode 3

Int4 - Internode 4

Int5 - Internode 5

KASP - Kompetitive Allele Specific PCR

LLP - Limited Liability Partnership

LTAA - Long Term Annual Average

LTD - Last Ten Days

MAS - Marker Assisted Selection

MSW - Main spike weight

NDVI - Normalized difference vegetation index

NILs - Near Isogenic Lines

NKP - Number of kernels per plant

NKS - Number of kernels per spike

nNILs - Non-NILs

NP_1m2 - Number of plants per square meter

NPR - Number of plants per row

NS_1m2 - Number of spikes per square meter

NSMS - Number of spikelets per main spike

PCR - Polymerase Chain Reaction

PH - Plant height

PHCs - Plant height components

PI - Photoperiod insensitive

PL - Peduncle length

PPD - Photoperiod response locus

QTL analysis - Quantitative Trait Loci Analysis

RAG - Resource allocation to grains

RC - Recombinant classes

Rht - Reduced height

RILs - Recombinant Inbred Lines

SAM - Sequence Alignment Map

SL - Spike length

SLE - Seed length

SN_1m2 - Number of seedlings per sq.meter

sNILs - Single NILs

SNP - Single Nucleotide Polymorphism

SSD - Single Seed Descent

SSM - Seed per square meter

STD - Second Ten Days

SWI - Seed Width

TAGB - Total Above Ground Biomass

TD - Ten days

TGW - Thousand kernel weight

VCF - Variant Call Format

VRN – Vernalisation

WGS - Whole Genome Sequencing

YP - Yield per plot (10m2)

YR - Yield per row

YSM - Yield per square meter

1. Chapter 1: Introduction

1.1 General introduction to wheat

1.1.1 Origin, domestication, and importance

Wheat is a cereal crop. The wheat classification, constructed after many years of concerns, mis/assumptions, hypothesis, archaeological records and scientific studies, of the cultivated diploid, tetraploid and hexaploid wheats and their possible wild progenitors shows that there are 6 main species and 18 subspecies. *Triticum urartu* Tum. ex Gand. and *Triticum zhukovskyi* Men. & Er. do not have any subspecies. Bread or common wheat belongs to *Triticum aestivum* L. at the species rank, *aestevium* at the subspecies rank. There are three early assumptions of wheat domestication i) de Candolle's (1886) assumption with an indication of Euphrates basin and vicinity as a first area where wheat domestication took place, ii) Solms-Laubach (1899) offered the Gobi Desert as an initial place of wheat cultivation, iii) Much (1908) assumed that wheat was domesticated in Europe. The idea of emergence of hexaploid wheat as a result of hybridizations of tetra and wild diploid wheats (McFadden and Sears, 1946) is widely acknowledged among today's scientific community. In spite of the fact that there is much speculation on its origin, the Fertile Crescent is mainly considered to be the place where wheat originated (Smith, 1995; Nesbitt and Samuel, 1998; Gustafson et al., 2009). The genome of hexaploid wheat (*Triticum aestivum* L. ssp. *aestevium*) consists of 7 homeologous groups of chromosomes, each of which has three - AA, BB and DD – homoeologous subgenomes derived from diploid (DD) *Aegilops tauschii* and tetraploid (BBAA) wild emmer (*Triticum turgidum* L. (Tell), ssp. *dicoccoides*) (Figure 1.1) which in turn inherited its AA genome from diploid *Triticum urartu* and BB from either diploid relative of *Aegilops speltoides* or some other extinct species (Feldman et al., 1997; Wang et al., 2013). Since its first cultivation bread wheat has been a crop of importance which supplies 1/5 of the total world population's calories (Bushuk, 1997). Therefore, hexaploid (2n=6x=42) bread wheat with one of the most complex genomes and origin of history is a vital food and feed crop in the world. Particularly, its importance in Kazakhstan is high because most of the local dietary commodities consisted of wheat flour (Eken et al., 2016).

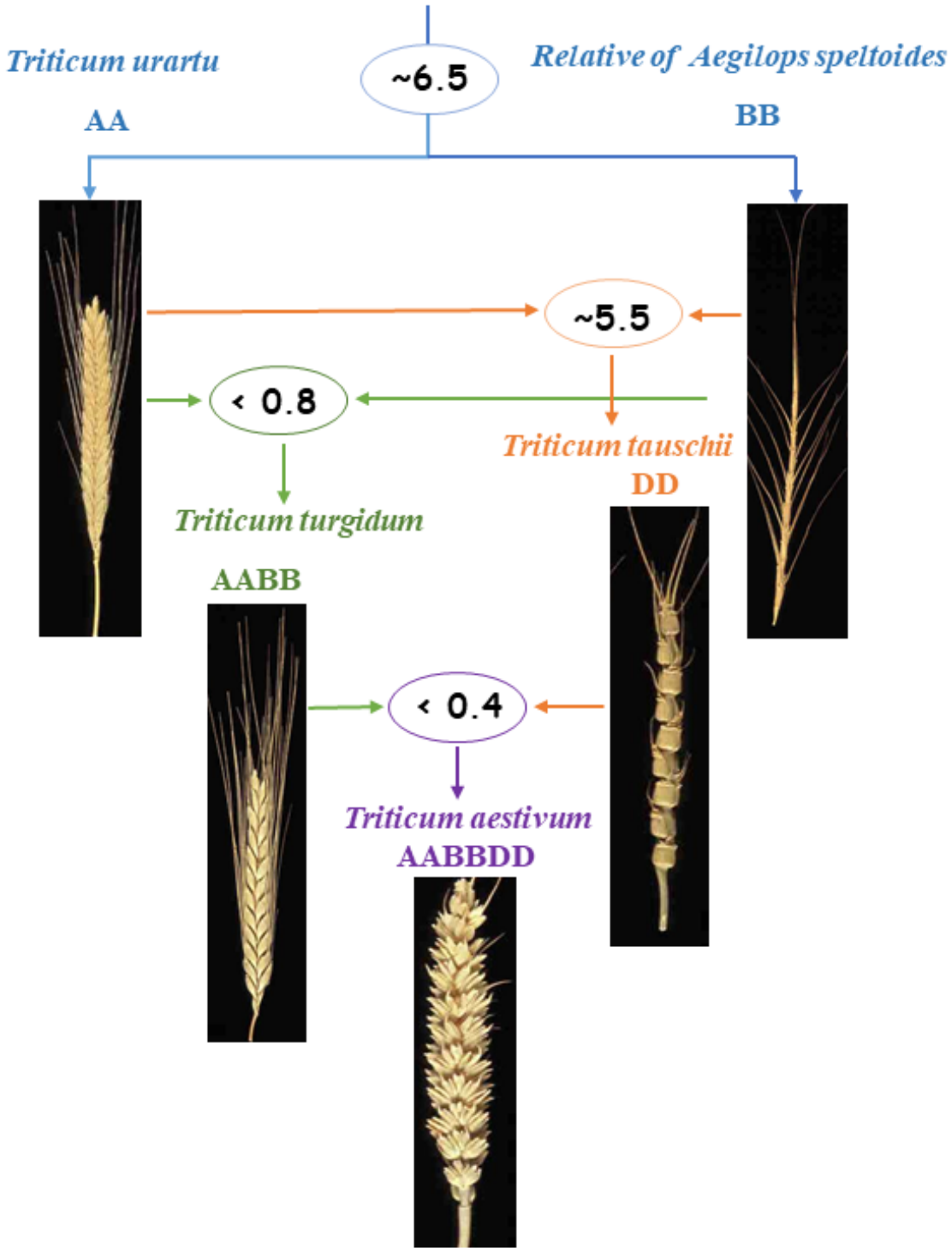


Figure 1.1 Wheat genome evolution

The figure provides visual demonstration of the wheat genome evolution. The genome of hexaploid wheat (AABBDD)Triticum aestivum L. ssp. aestevium derived from diploid (DD) Aegilops tauschii and tetraploid (BBAA) wild emmer (Triticum turgidum L. (Tell), ssp. dicoccoides). The figure adapted from Shewry (2009) and Mayer et al., (2014).

1.1.2 Wheat yield and production around the globe and in Kazakhstan

Bread wheat (*Triticum aestivum*) is one of the most adaptable and widely grown cereal crops across the world for food as well as feed. Global wheat production increased, as an average yield did, significantly in the last decade despite the decline of its grown area (Table 1.1 and Table 1.2).

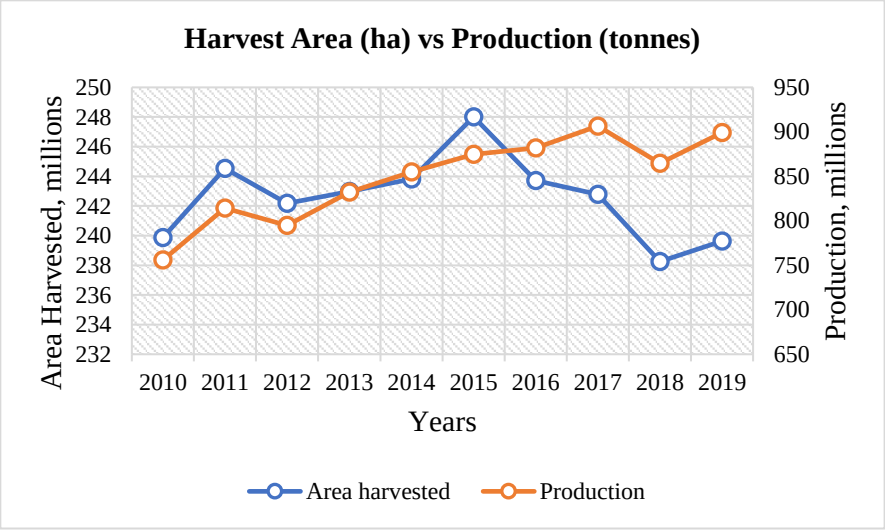


Table 1.1 Global Wheat Harvested area and Production

The table demonstrates that although global wheat harvested area has shrunken, its overall production has increased.

Data source: FAOSTAT

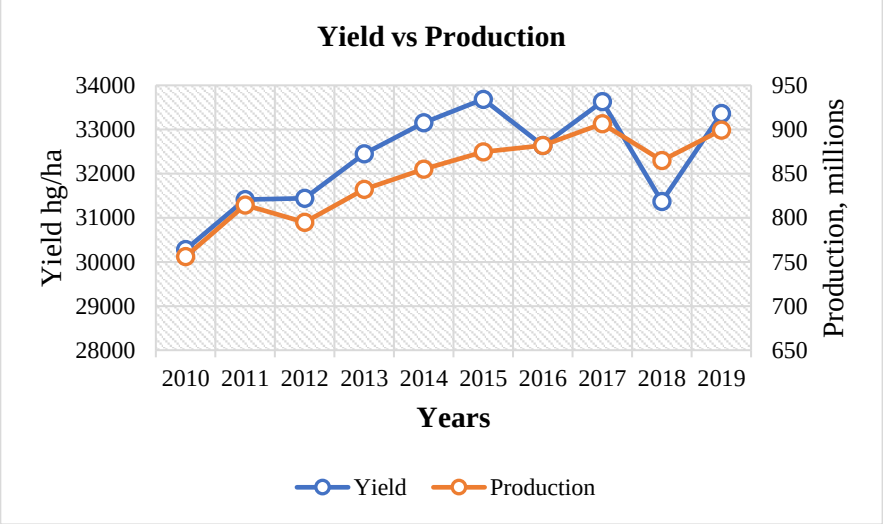


Table 1.2 An average Global Wheat Grain Yield and Production

The global production of wheat increased mainly due to increased grain yield per hectare.

Data source: FAOSTAT

However, the trend has not changed in Kazakhstan since 1990s (Table 1.3 and Table 1.4) when the country declared independence from the USSR. The country has not seen a significant yield increase since the reclamation of fallow land which took place in 1960s (data not shown). At best an average wheat yield in the country could reach 1.6 t/ha, at worst it may be as low as 0.5 t/ha compared with global average which in turn has effects on the country's total wheat production (Table 1.4). Despite a significant correlation between sowing area and total production in the country, Pearson's correlation coefficient was not that strong (Figure 1.2). This indicates yield dependency on environmental changes. Conversely, there is a high consistency between the average yield and total production (Figure 1.3). As Kazakhstan is the main wheat grain provider in

Central Asia (CA), volatile wheat production poses a significant threat to the global food security. Thus, long-term systematic wheat breeding programs with the use of current advanced technologies and techniques may assist to overcome the existing problems.

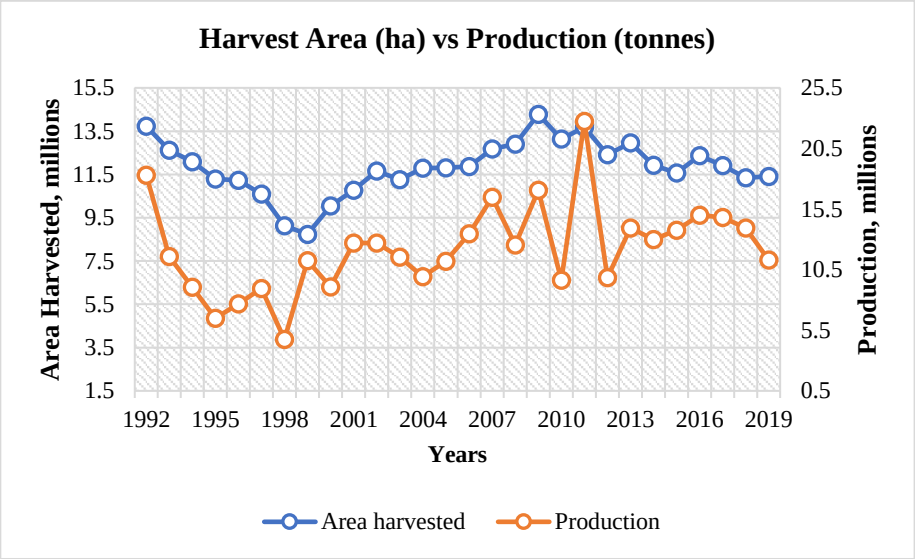


Table 1.3 Wheat Harvested area and Production in Kazakhstan

The table shows that an increase in wheat showing area in Kazakhstan does not always correlate with an increased grain yield proving strong environmental effects on an average and overall grain yield.

Data source: FAOSTAT

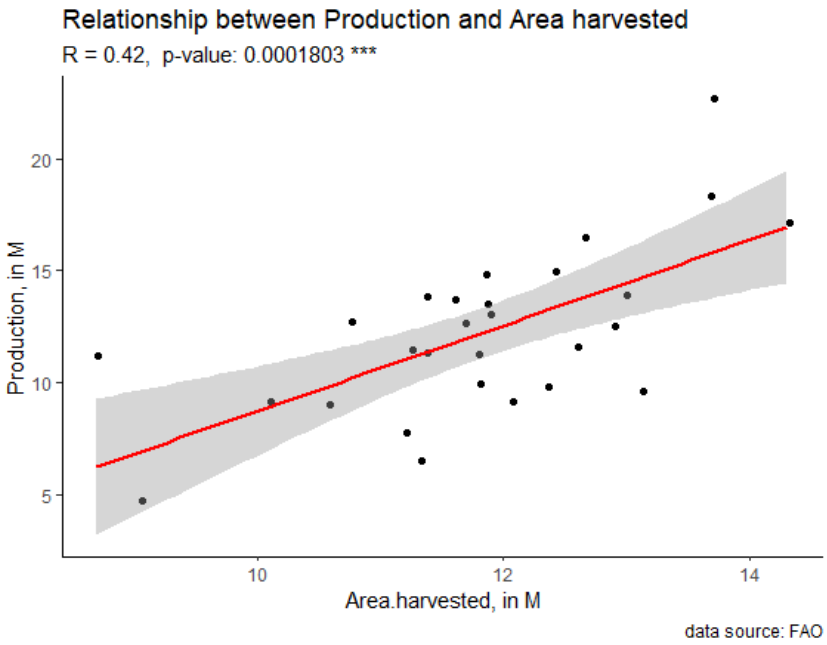


Figure 1.2 Wheat Total Production vs Area Harvested in Kazakhstan

Although correlation between sowing area and total production in the country was significant, Pearson's correlation coefficient was not that strong.

Data source: FAOSTAT

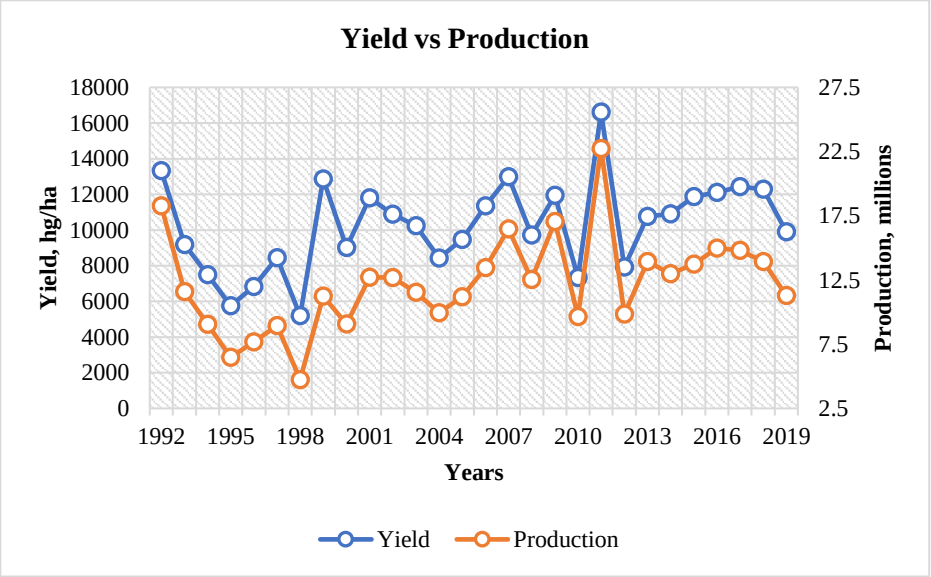


Table 1.4 An average Wheat Grain Yield and Production in Kazakhstan

The table shows that overall wheat production in Kazakhstan is mainly determined by an average yield.

Data source: FAOSTAT

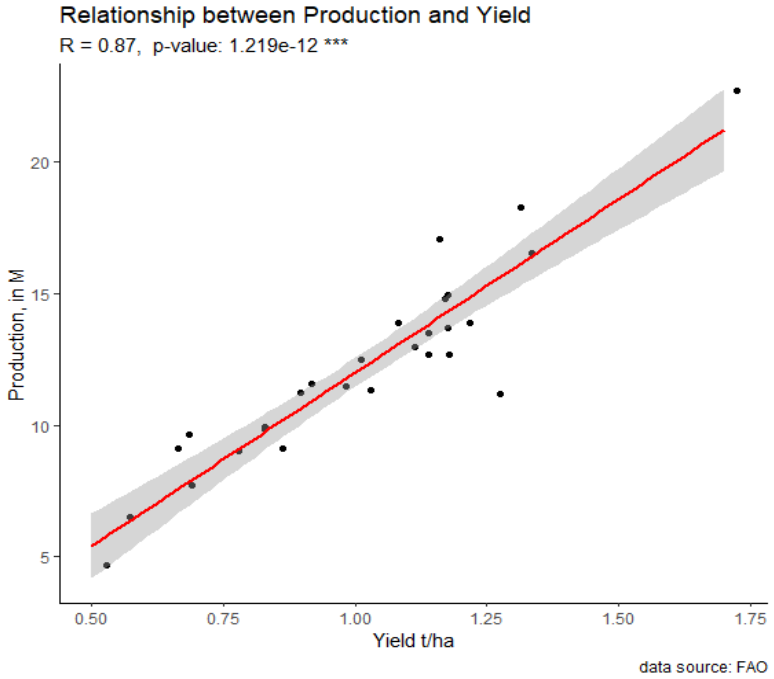


Figure 1.3 Wheat Total Production vs Average Yield in Kazakhstan

Pearson's correlation coefficient between the total production and average yield in Kazakhstan was strong compared to the correlation between total production and wheat harvested area.

1.1.3 The role of Kazakhstan in a global market

Wheat is in 85th place among world's most traded commodities, with a total trade of $44.1B, which makes it one of the important staple crops. In a global rank of the wheat exports and imports, Russia became the top origin of wheat in recent years and Egypt remains as the top importer respectively. Kazakhstan, despite unstable grain production, plays a significant role in contributing

to current and future global food security and is known as one of the major wheat suppliers in the world. Kazakhstan's share in global wheat exports among such most top exporters (Figure 1.4) accounted for about 1.5% at the minimum and 3.2% at the maximum with the mean of ~2.35%, depending on the season, in a last decade (www.oec.world).

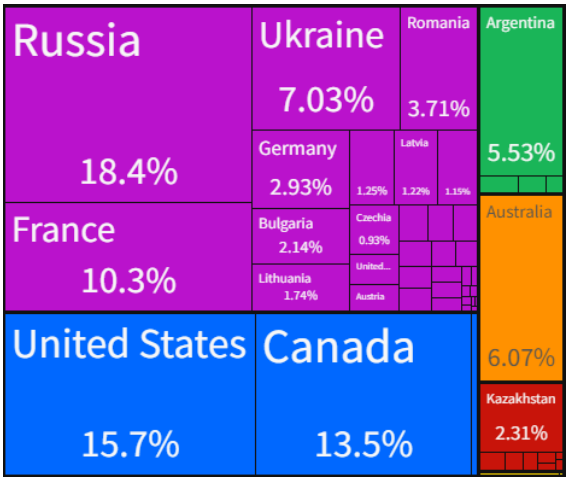


Figure 1.4 The world's topmost wheat exporters

Each country's contribution is shown as a percentage of the total.

Data source: (www.oec.world). The purple, blue, green, orange, red and yellow colour coding represents European, South American, North American, Australia, Asian and African countries respectively. Data for the year 2019.

Most of the time, one half of its harvested wheat is exported for the external use. The topmost consumers of the Kazakh bread wheat are Central Asian states. However, the high quality of Kazakhstani wheat attracts European countries as well (Figure 1.5).

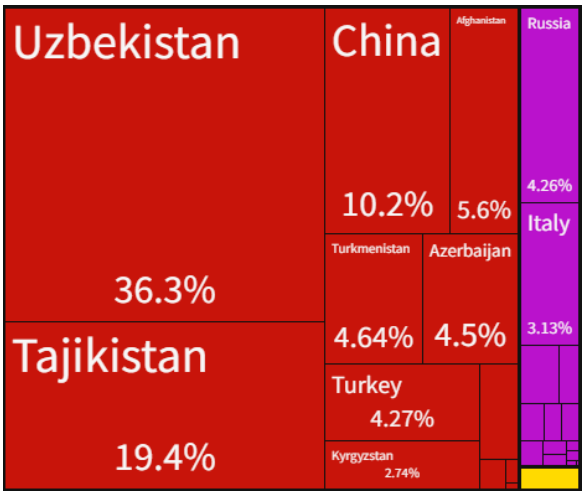


Figure 1.5 The list of top customers of the Kazakh bread wheat

The figure presents top Kazakh wheat consumers. Data source: (www.oec.world). The red, purple, and yellow colour coding represent Asian, European and African countries respectively. Data for the year 2019.

1.1.3.1 The attractive feature of Kazakh bread wheat in a global market

Kazakhstan is on the top among high-quality wheat producing countries. According to the quality classifications of the country, compatible with international requirements, for grain to meet the quality grade “superior” and “medium”, the cultivars must possess not less than 14 and 11 percent of protein content respectively. All varieties having less than 11 percent of protein are classified as “poor or weak” (Barayev et al., 1978). Therefore, during the breeding processes of new varieties the quality is highly controlled by breeders for the main two reasons:

I. Because of high consumption of wheat flour for pasta and tandyr naan (the traditional bread in Central Asian) industry within the country
II. To be competitive in domestic as well as international wheat export market

All these requirements contributed to make Kazakh wheat cultivars possess highly ranked bread baking quality (Omirbekova et al., 2016). Thus, high protein (14-16%) and gluten (21-40%) content (Abugalieva and Peña-Bautista, 2010), which are paramount aspects in bread-making performance as they contribute to the ability of dough to rise and maintain its shape, makes Kazakh wheat attractive in the world wheat trade. Moreover, wheat from Kazakhstan is suitable for steamed (Liu, 2021) and Chorleywood bread making processes. High concentrations of microelements such as Fe and Zn, with higher accumulation of Fe in spring and of Zn in winter wheats (Morgounov et al., 2007a; Gómez-Becerra et al., 2010) are also important characteristics of wheat from Kazakhstan.

1.2 General introduction to Kazakhstan

1.2.1 Economy, territory and location

Kazakhstan is the ninth largest country in the globe and located in the heart of Central Asia (Figure 1.6). It borders with countries such as Turkmenistan, Uzbekistan, Kyrgyzstan, China, Russia and through Caspian Sea with Azerbaijan. The country’s fast developing economy, comparing to its neighbours, mainly rests on the income from its rich mineral reserves. However, local people residing in rural villages mainly profit from agriculture and livestock farming. The country currently is striving to move towards green economy (www.qazaqstan.green).

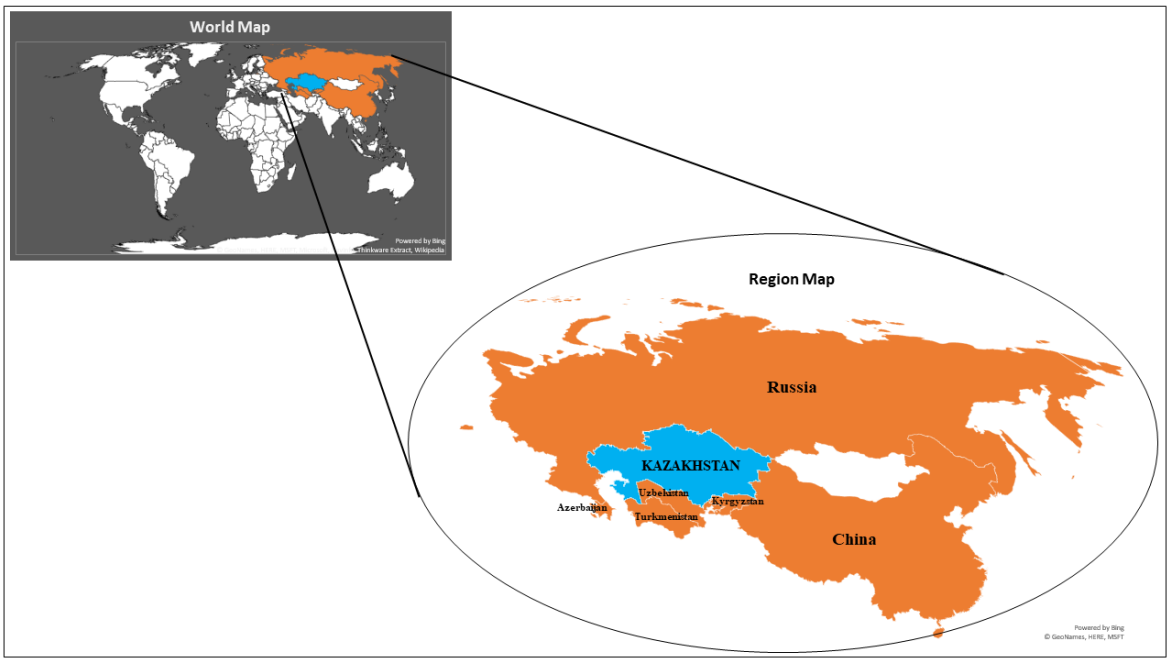


Figure 1.6 The map of Kazakhstan and surrounding neighbours

1.2.2 Climate zones and soil types

Kazakhstan geographically extends from 40.6˚ to 54.9˚ N and from 46.8˚ to 87.8˚ E. More than 86% of the territory is plain and lowlands, and only 3.6% covered by forests. The rest is occupied by high mountain ranges which are in the east and the south-east. Kazakhstan's nature occupies five landscape diversity zones: forest-steppe, steppe, semidesert, desert and mountains and foothills (Pilifosova et al., 1997). These zones were categorised based on soil type. Deserts or semi deserts can be found in the South-West, mountains are natural attractions of the Central, East and South-East Kazakhstan, and North of the republic is mainly in the steppe and forest steppe zones. Semi-arid steppes are important for wheat production in Kazakhstan (Mizina et al., 1996). The soil formation is based on the principles of latitudinal (horizontal) and vertical zonalities (Pachikin et al., 2014). There are four main types of soil in Kazakhstan: chernozem (black soil), multi coloured chestnut, fulvous and mountainous soils. Chernozem occupies 25.5 million hectares, or 9.5% of the territory of the republic, while multi coloured chestnut and fulvous soils take up 90.6 Mha, or 34% and 120 Mha, or 44% respectively. Further, **chernozem** can be categorised into two types; *standard* and *southern chernozems*, the latter possesses lower humus content of 6-4% comparing to 8-6% of the standard: **multi coloured chestnut** soils with humus content of 4.5-2.0%, into three; *dark-chestnut, ordinary chestnut* and *light chestnut soils*: **fulvous** soils with only 2.0-1.0% humus, into two; *brown desert* and *gray-brown desert soils*. The **mountainous soils** can only be found in mountain ranges of the country (Figure 1.7).

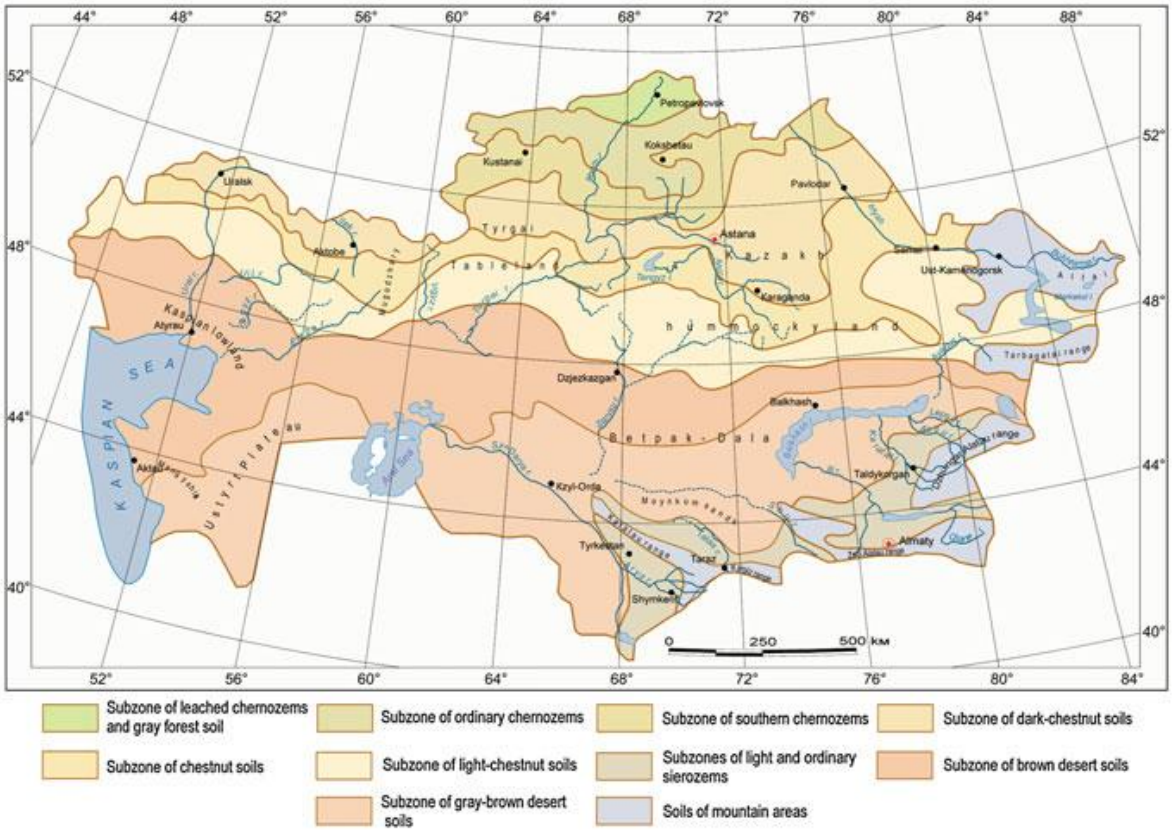


Figure 1.7 The soil types of Kazakhstan

The map adapted from Pachikin et al., (2014).

The soil content of the semi-arid wheat growing steppe zone is ordinary/typical/standard chernozem and southern chernozem, and the mix of two, and includes the north of Kazakhstan, most parts of Kokshetau and Kostanay Region and northern part of Akmola and Pavlodar Regions. The steppe zone includes dark chestnut, ordinary chestnut and light chestnut soils, and occupies mainly Northwest part of the country, Torgai and Aktobe Region, Kostanay, Akmola, Pavlodar and Karaganda Regions (Mizina et al., 1996; Sommer et al., 2013). Agricultural cereal crops grown in typical chernozem are feed and malting barley, bread wheat, rape, peas, oats, winter rye, buckwheat, soybean, the mixtures of grain legumes and forage crops, combined sowings of perennial legumes and cereal grasses, perennial legumes, perennial grasses and maize; in southern chernozem: strong and valuable bread wheat, durum wheat, barley, peas, chickpea, mustard, oats, buckwheat, millet, sunflower, pea vine, perennial grasses, combined sowings of perennial grasses and legumes; crops mainly cultivated in dark chestnut are strong bread wheat, chick pea, mustard, millet, Sudan grass, foxtail millet (Kurishbayev, 2003). The loss of humus content in soils of the northern Kazakhstan runs at 20–30%, while this indicator in southern Kazakhstan is even worse at 30–40% (Zubairov, 2002). Soil salinization is another issue where croplands are irrigated which is typical for southern Kazakhstan. This is mostly caused by the excessive water application to a level higher than needed by the crop (Ramazanov, 2006). Therefore, a drip irrigation technology was suggested which significantly reduces the amount of water being used and helps maintain soil quality (Karimov et al., 2009).

1.2.3 The wheat growing environments

Kazakhstan has enormous agronomic potential due to the fact that it occupies a huge area of 2,724,900 km^2. The total arable land under cultivation reached a maximum of 30 million hectares when the country was part of the USSR, then dropped to 18 million hectares in 2004 with gradual increase to 22 million hectares in 2017 (http://stat.gov.kz). Wheat occupied more than half of this area at about 14 million hectares. However, the wheat growing area has been shrinking gradually since 2009 from initial 14 to about 11 million hectares currently giving way to such economic crops as soybean. Of those 11 million hectares, 78-80% is occupied by spring wheat. Winter and durum wheats are cultivated more or less in 0,5-million-hectare area. The cropland of durum wheat decreased dramatically, twice as much as compared to Soviet Union times which at that time was more than 1 million hectares (Dorofeev et al., 1987). Nonetheless, durum wheat production is predicted by local authorities to increase in the near future. The spring wheat in Kazakhstan is mainly grown in the northern part, that is Kostanay with wheat cultivation area of 3.7 Mha (yield 1.1 t/ha), Petropavl - 2.2 Mha (1.7 t/ha), Pavlodar - 0.45 Mha (1 t/ha) and Akmola Regions – 3.7 Mha (1.1 t/ha). All environments are rainfed. Other regions of Kazakhstan also grow wheat in small quantities with the exception of Atyrau Region which stopped wheat cultivation since 2008 due to very low yield of ~0.09 t/ha. Likewise, Central Kazakhstan Region can be considered as a main wheat growing area as well, because land under wheat cultivation accounts for 0.63 Mha with an average yield of 1 t/ha. Durum is also cultivated in northern cities, Kostanay, Petropavl and Akmola with exception of Pavlodar, of the country. Because of extreme cold in winter, there is a high risk of sowing winter wheat in these areas. Therefore, these regions mainly focus on growing the spring wheat. The winter wheat can be grown alongside spring wheat in southern parts of Kazakhstan, which is comprised of three Regions: Turkestan, Jambul and Almaty with wheat growing areas of 0.2 Mha, 0.1 Mha and 0.13 Mha respectively. Grain yield is much higher in southern parts of Kazakhstan, where mostly winter wheats dominated, comparing to northern regions which grows spring wheats.

1.2.4 The wheat growing conditions of Kazakhstan

Depending on weather conditions, the total vegetation period of spring wheat in Kazakhstan varies between 90 and 105 days from seed to seed and is the shortest in the world. While for winter type, the duration from drilling to maturity takes about 210 and 270 days which is also weather dependent (Sommer et al., 2013). An average temperature in the northern part of Kazakhstan from November to March is mainly below 0˚C and from the end of May to the mid of September at around 18.5˚C, with an annual mean temperature of 1.6˚C (Yanai et al., 2005). The July - February average temperatures in the South and North of Kazakhstan are at around (33˚C; -5˚C) and (20˚C; -22˚C) respectively (www.worldweatheronline.com). However, for both regions there is an occasional potential increase to more than 48˚C and 39˚C in summer and decrease of more than -20˚C and -45˚C. The annual average precipitation ranges from 250 mm to 320 mm in the North,

while the amount of rainfall varies between 450 – 600 mm in the South Kazakhstan (De Beurs and Henebry, 2004; Abugaliyeva and Morgounov, 2016). Day lengths across the country also vary considerably (Figure 1.8) (Turuspekov et al., 2013).

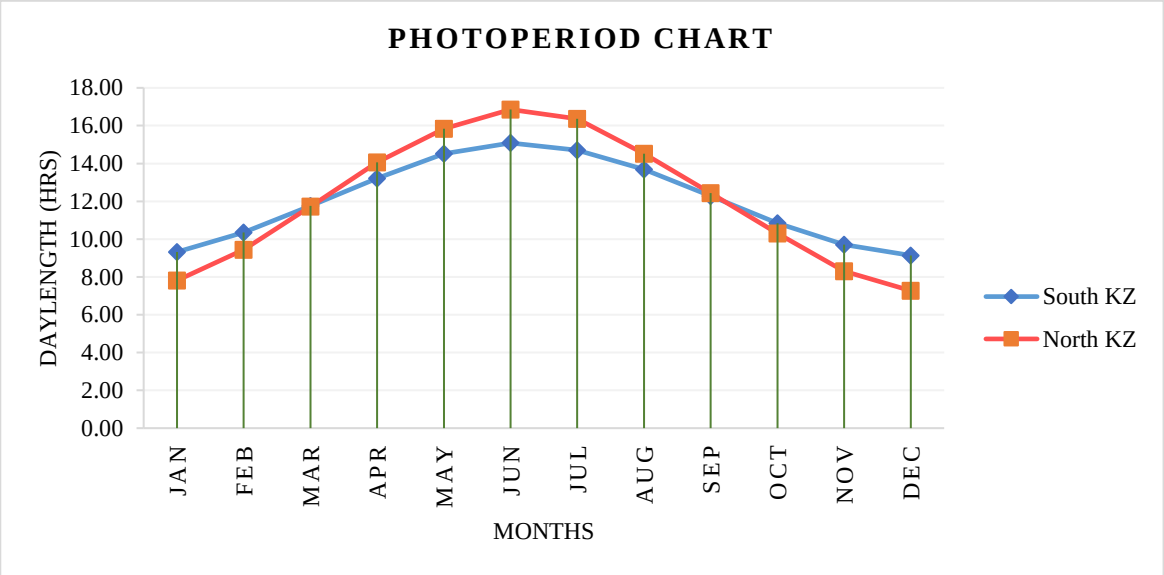


Figure 1.8 Daylength difference in the North and South Kazakhstan

Although days of the North Kazakhstan – the main wheat producing region - are shorter in winter compared to the South, it becomes longer during the spring and summer crop growing months. Daylength are given in hours.

1.3 Limiting factors of agriculture in Kazakhstan

1.3.1 A brief introduction to the environmental factors influencing wheat adaptation in the mega – environments 6 and 12

Twelfth wheat growing mega - environments (ME) are developed by CIMMYT to address the challenges encountered in the wheat breeding are given (Figure 1.9) (Braun et al., 1996). As Kazakhstan is represented by diverse mini environments, it simultaneously belongs to ME12 (mega - environment 12) and ME6 (mega - environment 6) to grow winter and spring wheats respectively. The ME12 covers Southern parts of Kazakhstan and suffers from insufficient precipitation, various rust diseases and low winter survival rate. In contrast, ME6 is subdivided into two “A” and “B” which are high rainfall (typical of Harbin, Heilongjiang, China) and semi-arid (typical of Nursultan, Kazakhstan) areas respectively (http://wheatatlas.org/). Geographic coordinates of the country require breeding materials that have a degree of photoperiod sensitivity, which is different to other spring MEs. Harbin is represented by pre-anthesis drought followed by rainfall during flowering and grain filling. Resistance to Fusarium spp., tan spot, yellow rust, leaf rust, stripe rust and tolerance to sprouting are breeding objectives in this environment. However, central and northern Kazakhstan (12 million ha) and the southern Siberian wheat belts (8 million ha) are very dry representatives of the ME6. The major diseases are leaf and stripe rusts. In addition, drought tolerance specific to the rainfall pattern in the region is needed. Taller wheats generally do better under these conditions.

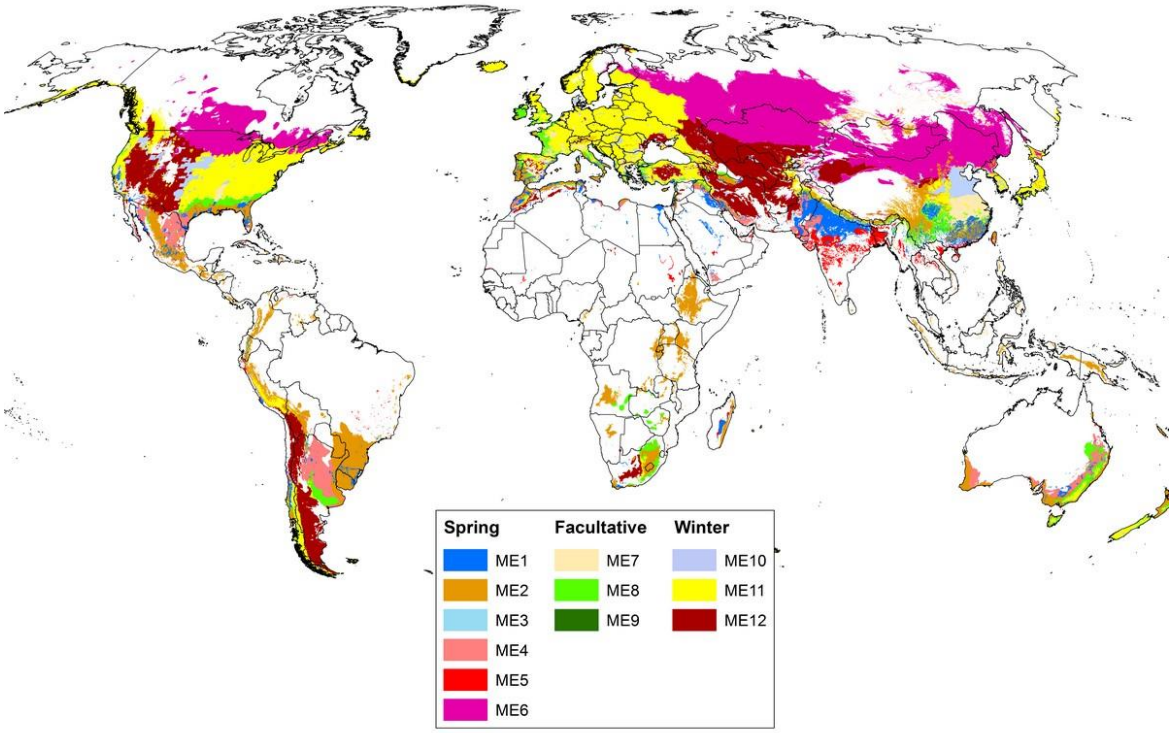


Figure 1.9 Wheat Growing Mega Environments

Kazakhstan encompasses two mega – environments, ME6 and ME12, for winter and spring wheat breeding respectively. Data source: https://hdl.handle.net/11529/10625, CIMMYT Research Data & Software Repository Network, V4 (Sonder, 2016).

1.3.2 Drought and soil erosion, and their management

The main constraining factors of sustainable agricultural production in Kazakhstan are water deficiency (Funakawa et al., 2004) and erosion of soil (Figure 1.10). The total 30 m hectare arable land is entirely vulnerable to low rainfall and soil erosion (Gossen, 2000). The successful management of water content in the soil is key to maximising yield (Kaskarbayev and Kenzhebekov, 2004; Astafev, 2015). Therefore, keeping moisture in the soil during the cropping season is crucial as it allows the seeds to germinate quickly. Failure to do so is likely to delay the wheat harvest time. This in turn, enters the harvest to the season with heavy rain often followed by snow. If this happens, on the one hand farmers will not be able to harvest their crop, on the other, crops remaining under snow throughout winter may cause fungal diseases with lethal consequences in animals as well as in humans in case of its application (Gagkaeva et al., 2011). A common method of soil moisture management is maximum utilisation of winter precipitation which in turn creates favourable condition for all agronomic crops cultivated in the country. The main approaches used for this propose are snow retention, snow accumulation, meltwater retention or/and combined application of the three. Each has its own specific technique to be implemented (Shirvanov and Rybalko, 2001). Although these methods are agriculturally beneficial, they seem to damage the topsoil surface decreasing the humus and fertility as heavy vehicles are employed

during the process (Iorgansky, 2000). Despite this fact, the approach has been very useful in increasing grain yields of cereal crops and is still widely used in northern part of Kazakhstan and locally published sources indicated that yield increased approximately by 0.38 t/ha and 0,56 t/ha in Kazakhstan as well as southern part of Russia (Shirvanov and Rybalko, 2001; Viurkov, 2005).

When it comes to soil management ways for crop cultivation, several new soil management techniques were tested one of which is the conservation tillage method. The method can be classified into five methodologies called zero, minimum (reduced), mulch, ridge and contour tillage and was tested in different environments. The research study concluded that successful deployment of each tillage methodology is dependent on soil type. Of the five, zero tillage (No-till) was offered to be used in low humus content soils with light texture (Busari et al., 2015). In spite of effectiveness of No-tillage in the different ecological zones of northern Kazakhstan with heavy soil texture (Barayev et al., 1978; Astafev, 2017) and of Central Kazakhstan (Yutshenko et al., 2005), it increased number of weeds in south of Kazakhstan. Therefore, the traditional disk cultivation planting method with appropriate level of herbicides (such as Triallat and Illocsan) have been suggested which in turn was useful to combat weeds as well as increase yield by 0.48 t/ha and 0.57 t/ha (Suleimenova, 2000). The disk cultivation was described as the most economical way of incorporating the herbicides into the soil (McWhorter, 1981). Therefore, farmers in Kazakhstan prefer traditional tillage means compared to no-till.

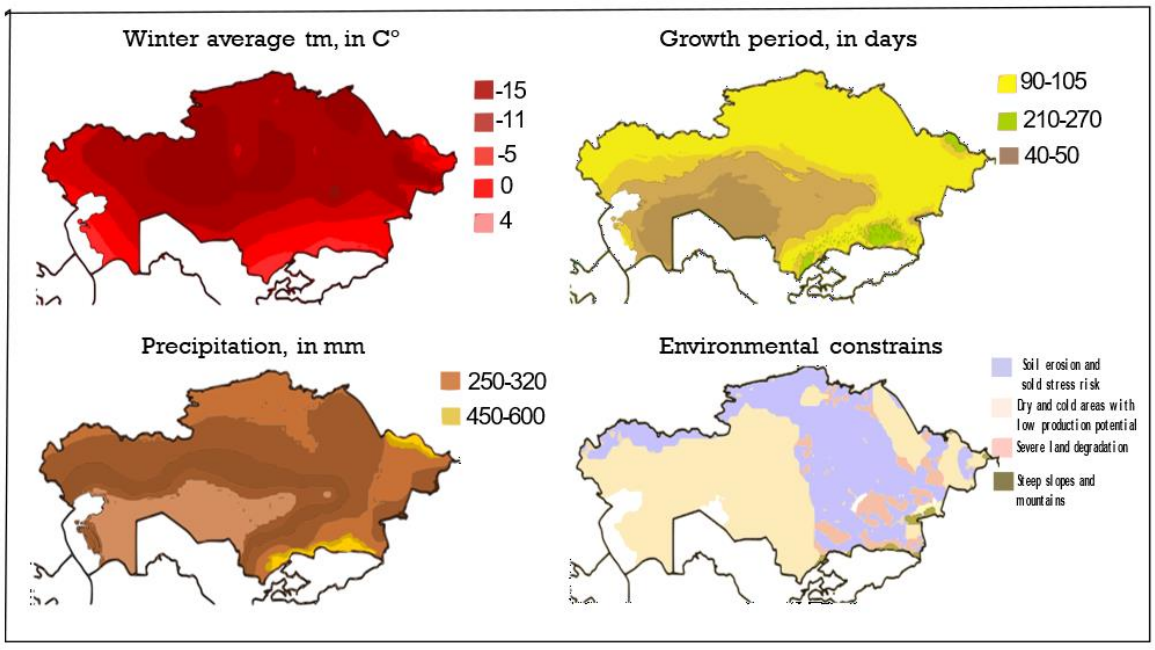


Figure 1.10 The main physical/abiotic constraints in Kazakhstan

1.3.3 Wheat diseases

Wheat diseases, caused by insects, pests, bacteria, fungus and viruses, are another headache of world wheat breeders, which may potentially generate significant economic consequence because of substantial yield and quality loss. Control and prevention of distribution of the new emerging plant diseases or their races is less manageable and this in turn leads world plant breeders to release new genotypes with resistant genes. Because of dry weather in Kazakhstan, wheat breeders also encounter the same challenge mostly every year. Spring wheat, grown at the higher latitudes of the country, often suffers from Septoria-gelmintosporioznye, tan spot, leaf rust and gelmintosporiozno-Fusarium root rot while diseases such as yellow rust, powdery mildew, septoriosis, hard and dwarf smut could infect winter wheat areas which are mainly located in the southern mountainous parts of Kazakhstan (cet al., 2002). The thorough description for cereal crops and their distribution throughout the country, and their infectious diseases is provided by Geshtovt (1986). Among wheat infections three rust diseases, stem, stripe/yellow and leaf /brown, are serious. The last is the most common one. High resistance to these rust diseases could be achieved by evaluating the absence or presence, and composition or pyramiding the associated Lr and Yr genes in wheat cultivars (Singh et al., 2005) and the authors recommended simple ways for developing and identifying tolerant genotypes using controlled environments and real field trials in Central Asia. Due to high costs for phenotyping platforms, including automated or even partly automated greenhouses, this method is not available in most of the research institutions even now. However, a quick pace in the technology development resulted in the emergence of a new phenotyping tools such as 'PhenoBox' which is now available for plant scientists with a small budget (Czedik-Eysenberg et al., 2018). GWAS (Genome Wide Association Study) on hexaploid wheat identified the locations of potential candidate genes for leaf rust resistance on chromosomes 1DS, 2AS, 2BL, 3B, 4AL, 6AS and 6AL, of stripe rust on chromosomes 2AS, 2DL, 3B, and 7DS and for tan spot resistance on chromosomes 1AS, 2AL, 2BL, 3AS, 3AL, 3B, 6AS and 6AL (Juliana et al., 2018). Moreover, the various use of recent advanced genomic data and tools enabled to characterise the resistance genes from the different angles in a greater resolution (Babu et al., 2020; Zhang et al., 2020a; Hafeez et al., 2021). However, many of the identified genes or/and genomic sections are yet to be functionally validated in wider populations for their effective use in pre-breeding and breeding programs. In Kazakhstan and Southern Siberia, despite the dry climate, both the rust diseases - stem and leaf - have a significant negative impact on wheat production in certain seasons, usually every fourth year, with yield losses of up to 30% for leaf rust (Morgounov et al., 2007b; Genievskaya et al., 2022). The recent studies carried out involving hexaploid and tetraploid wheat varieties and hybrid lines from most of the Central Asian states and the rest of the world described stripe rust as very significant disease in the region, identified several important QTLs and also provided tolerant varieties for future crosses (Ziyaev et al., 2011; Kokhmetova et al., 2018; Genievskaya et al., 2022). All of these genetic strategies are important considering the recent stem rust outbreaks in Kostanai, North Kazakhstan and Omsk, Russia (Shamanin et al., 2016; Rsaliyev et al., 2020). Genotyping of

commercial wheat varieties from Kazakhstan and Russian with allele specific DNA marker for the races of wheat tan spot showed that elite wheat cultivars are in different levels of susceptibility for these races (Kokhmetova et al., 2017). In addition, it is known that existing hundreds of isolates of tan spot infect wheats differently (Lamari et al., 2005). One of the main agronomic traits in the region, TGW (Thousand Grain Weight), was reported as being the most affected character by leaf rust (Morgounov et al., 2015). Compared to leaf rust, stripe rust affects crop production every 3-4 years (Shamanin et al., 2016) and can potentially decrease grain yield up to 35 – 45 % (Ziyaev et al., 2011). As only a few studies were conducted using modern molecular genetics teachings on understanding of tolerance of local wheat cultivars, there is still plenty of work to be done to make new varieties with durable resistance genes to particular microenvironments deploying proper MAS (marker assisted selection).

1.3.4 Climate change effects on current and future crop adaptation in Kazakhstan

Observed climate changes across the globe, particularly gradual temperature increases of the earth's atmosphere may have a positive or/and negative effect on adaptation performance of the salient agricultural crops one of which is wheat. In the face of expected global warming there are two likely positive scenarios which exist to alter wheat breeding practices in Kazakhstan: i) extended vegetative period, ii) turning the spring wheat dominated northern parts into winter wheat planted area. Both cases may create favourable growing conditions to enhance grain yield in the country. In case of adverse effects, however, it might cause instability in the yields of not only wheat, yet all cereal crops (Iizumi et al., 2014) and thus might have severe consequences for the global food security. For instance, the study evaluating aridity tendency throughout almost fifteen years period since 1966 till 2015 using CRU (Climate Research Unit) database of the Central Asian countries including Kazakhstan, Uzbekistan Kyrgyzstan, Tajikistan, Turkmenistan, and the northwestern region of China revealed that southeast Central Asia might be affected by ephemeral drought mostly during summertime. In contrast, rare but long-lasting drought with less harmful consequences was observed in the northeast.(Guo et al., 2018). The same article highlighted that north Kazakhstan is getting wetter while southwest part of the country, especially the vicinity of Aral Sea (Kyzylorda Region) became drier. It was reported that an annual average temperature in the Kyzylorda Region have increased from 1.75˚C to 2.25˚C with greater rises in summer comparing to winter (Ragab and Prudhomme, 2002). However, studies show that climate change positively affected wheat productivity in northern and southern parts of Kazakhstan with substantial rises of 0.11 tonnes per hectare in Astana and Kostanay, of 32 t/ha in Petropavl and Kyzylorda (Sommer et al., 2013). This is perhaps due to increased rainfall rate. For example, trends of rainfall in summertime increased by a maximum 10% and 5%, and in winter up to 25% and 10% in the North and the South respectively (Ragab and Prudhomme, 2002). Nevertheless, the recent research based on simulations has predicted potential decrease in wheat production of the country because weather is likely to become warmer and more severe with regular drought periods (Fehér

et al., 2017). If this is a case, then the country should start to launch successful breeding programmes aiming to identify negative climate impacts on wheat adaptation using advanced breeding tools and pick up the best wheat growing conditions and environments, considering the amount of precipitation and soil type, as well as genotypes with combination of desired genes in order to improve wheat production and resistance to climate change. Doing so in turn will contribute future global food security, because Kazakhstan is one of the three giant world wheat exporters, currently supplying one of four world's wheat exports, called "bread basket" along with countries such as Ukraine and Russia (Swinnen et al., 2017).

1.4 Roadmap towards enhancing wheat grain yield in Kazakhstan

1.4.1 Challenges in increasing wheat yield in Kazakhstan

Besides the above mentioned biotic and abiotic challenges, there are many technical aspects to be considered in agricultural sector of Kazakhstan. For instance, without a rotation, the yield losses of 15-20%, 20-25% and more than 25% were observed from the second, third and fourth crops respectively when the wheat was sown after wheat (Kaskarbayev, 1998). Planting date is asserted to be another essential factor to manipulate wheat adaptation and yields. Early sown plants can be hit by drought in May or early June when they are in an important developmental phase (personal communication with breeder colleagues). Late sowing of wheats, on the other hand, might get hit by stem rust, which mostly infects crop in late sowing periods (Shamanin et al., 2012). The yield lost from stem rust is notorious and nears 40 – 50 % (Koishybayev et al., 2008), and at worst it could reach to 60–90% (Movchan, 1998). Stress during the booting stage while spikes are differentiated is much more damaging (Andreeva, 2015). Plant height is of importance in optimising wheat adaptation. Due to short coleoptile, the worldwide benefits of historic *Rht-1* genotypes are likely restricted in hot and dry rainfed environments (Rebetzke et al., 2007) like Kazakhstan. During drilling, sowing depth is another paramount factor to take into consideration to finetune final yield. One of the main causes of low wheat yield are the out-of-date agri infrastructure and lack of systematic pre-breeding programs.

1.4.2 Suggested ways of manipulating the final grain yield in the country

The development of an expedient crop rotation system is essential to increase grain yield. Wheat drilling in the north of Kazakhstan is mostly delayed to take advantage of rainfall in June and July (Morgounov et al., 2001) instead of May and June. This is due to a reduction in the amount of May-June precipitation compared to the June-July (Figure 1.11).

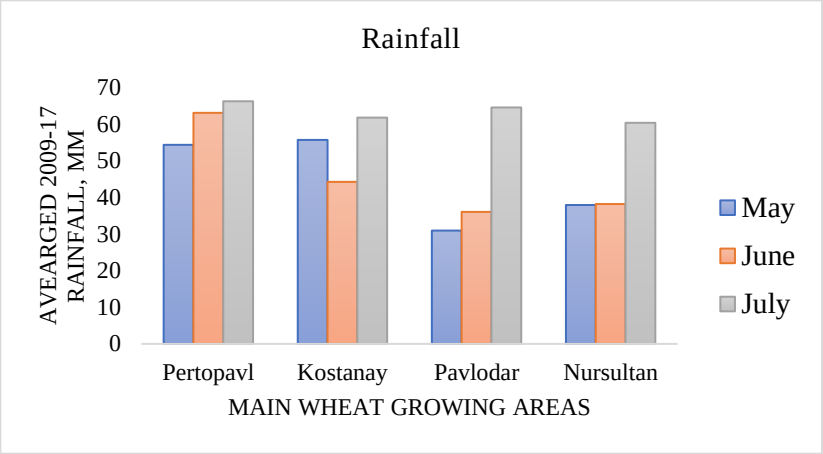


Figure 1.11 May-June rainfall compared to June-July

The most favourable time for spring wheat planting in northern part of Kazakhstan is around May 25th, but wheat can be sown in a time window starting around May 20th (personal communication with breeder colleagues). To the negative impact of late-spring and early-summer drought extreme conditions wheat is better tolerated during growth stage around tillering (Andreeva, 2015). Therefore, it was said earlier that wheat varieties with an extended period of tillering are more adaptable to local conditions of the steppe zone of Northern Kazakhstan (Kuzmin, 1970). These requirements are mostly answered by late-maturing varieties. However, delayed/staygreen genotypes are not preferable considering the short vegetation period and suggested delayed drilling. Therefore, there is a need to release early/standard maturing wheat genotypes with an extended tillering stage. As reduced height genes are less beneficial, it is thought taller wheats perform better. However, plant height as a trait of importance has not been standardised/fixed. Drilling technologies and equipment which are suitable for planting the seeds into the right depth under the soil, in case of plenty of moisture - shallower and in the contrary situation - deeper, also play an essential role in achieving more production and to manage total expenditure (Astafev, 2015).

1.4.3 Genes for improving wheat adaptation and yield

1.4.3.1 Genes of Vernalisation

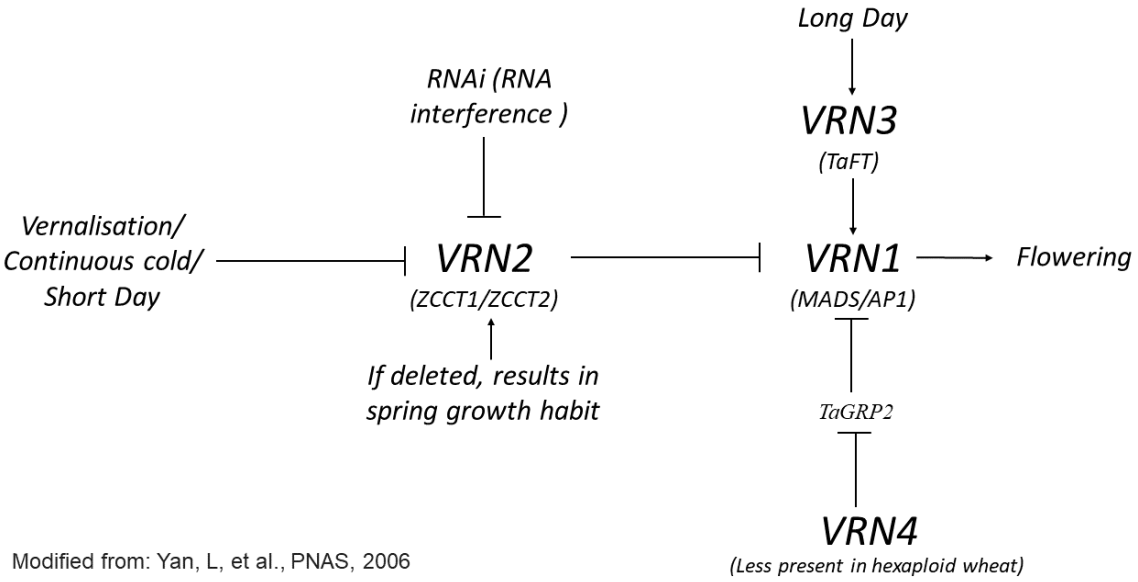


Figure 1.12 Model for VRN interaction

Vernalisation is the requirement of the plant to be subjected to continuous cold (not freezing) to transition from vegetative to reproductive growth. This phenomenon defines the difference between spring and winter wheats. The duration of the cold period inhibits the repressor of the flowering locus and allows the plant to flower. If vernalisation requirement is not fully satisfied the plant cannot complete vegetative growth and pass to reproductive development and so the development of grain. Unvernalised wheat plants stay in early growth stages producing only tillers. Therefore, it is necessary for winter wheat to be vernalised. This is not the case for spring wheat. There are four major *VRN* (*VRN1, VRN2, VRN3* and *VRN4*) genes in wheat (Figure 1.12). Among *VRN* genes, *VRN1* is a major gene explaining most of the natural variation in growth habit in allohexaploid wheat and it is comparatively well-studied (Yan et al., 2004; Fu et al., 2005; Distelfeld et al., 2009; Zhang et al., 2012; Milec et al., 2013; Kippes et al., 2016). *VRN2* is zinc finger and a CCT domain carrying protein that is encoded by *ZCCT1* and *ZCCT2* genes (Yan et al., 2004). *VRN2* is common in diploid wheat and barley and known as an inhibitor of flowering in winter wheat, prior to vernalisation, and thus, nonfunctional *VRN2* determines spring growth habit, but its effect is masked in hexaploid wheat because of gene redundancy (Kippes et al., 2016). Wheat *VRN-3*, located on 7BS chromosome, is orthologous to Arabidopsis *FLOWERING LOCUS T (FT)*, so it was designated as *TaFT* and a positive regulator of *VRN-1* (MADS/AP1). Homozygous dominant alleles (*Vrn-3*) flower first compared to homozygous recessive alleles (*vrn-3*) due to higher level of transcripts (Yan et al., 2006). *VRN-4* known as *Vrn-D4*, residing on chromosome 5DS, is a paralog of *VRN1* as it is estimated to originate as a result of the translocation of a ~290-kb region from chromosome arm 5AL into the proximal region of chromosome arm 5DS (Kippes et al., 2014, 2015). *Vrn-D4* is less present in bread wheat (Goncharov, 2003) and mostly identified in ancient

subspecies *Triticum aestivum ssp. sphaerococcum* (Kippes et al., 2015; Trevaskis, 2015). Molecular mechanisms showed that SNPs (Single Nucleotide Polymorphism) in first intron of the *Vrn-D4* negatively regulate the *TaGRP2* (*GLYCINE-RICH RNA-BINDING PROTEIN 2*) which is known as a repressor of *VRN1* (Kippes et al., 2015) (Figure 1.12). It is not the aim of this thesis to provide a detailed explanation of each important adaptation genes, therefore, we will discuss only *VRN1* as it is main vernalisation gene.

1.4.3.1.1 VRN1

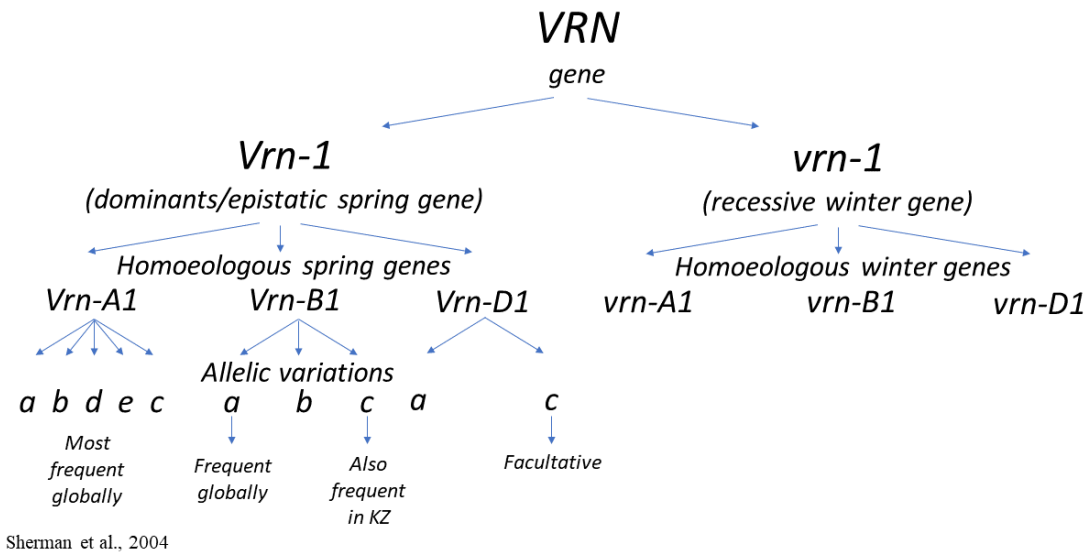


Figure 1.13 Symbolling scheme of the VRN gene

A subtle complication about these genes, whether it be vernalisation and photoperiod response genes, is symbolling them. For instance, spring types, which are epistatic as well as dominant over winter types, of each *VRN* genes were symbolled as *Vrn1*, *Vrn2, Vrn3* and *Vrn4* and the winter types as *vrn1*, *vrn2, vrn3* and *vrn4*. Further, some of the *VRN* genes defining spring growth habit (*Vrn1*, *Vrn2*, *Vrn3* and *Vrn4*) comprise three homoeologous genes on wheat A, B and D genomes. For instance, *Vrn-A1*, *Vrn-B1* and *Vrn-D1* genes exist for a *Vrn1*. Each of these *Vrn-A1*, *Vrn-B1* and *Vrn-D1* genes possess such allelic variations as *Vrn-A1a*/*Vrn-A1b*/*Vrn-A1c*, *Vrn-B1a/Vrn-B1b*/*Vrn-B1c* and *Vrn-D1a/Vrn-D1b* respectively (Figure 1.13). The functions of these variations or mutations at the vernalization loci promote quantitative variation in flowering time (Trevaskis et al., 2007). All three homoeologous *Vrn-1* genes (*Vrn-A1*, *Vrn-B1* and *Vrn-D1*) are located in the close vicinity of a distal end of the long arms on group of 5A, 5B and 5D homoeologous chromosomes of wheat ((Appels et al., 2018)). Among them, *Vrn-A1* is most frequently found as a spring allele and varieties carrying the spring allele of this gene enter heading time earlier than others (Sherman et al., 2004). Spring *Vrn-A1* gene possesses five major alleles – *Vrn-A1a*, *Vrn-A1b*, *Vrn-A1d*, *Vrn-A1e* and *Vrn-A1c* – all of which confer spring growth habit and have allele

specific mutation/s in the promotor or/and 5'UTR or/and intron 1. Compared to *Vrn-A1*, *Vrn-B1* consists of three alleles – *Vrn-B1a* (*former Vrn-B1*), *Vrn-B1b* and *Vrn-B1c* (Milec et al., 2012). Among them, *Vrn-B1a* is widely distributed globally, however, it is likely that *Vrn-B1c* is as important as *Vrn-B1a* to confirm spring growth habit in Kazakhstan. (Milec et al., 2013). In *Vrn-D1a* and *Vrn-D1b* genes, the identical intron 1 deletion makes them different from the wild type *vrn-D1*. However, a SNP which changes cytidylic acid into adenylic acid, at the promotor CArG site of the *Vrn-D1b* is related to facultative growth habit (Zhang et al., 2012) (Figure 1.14).

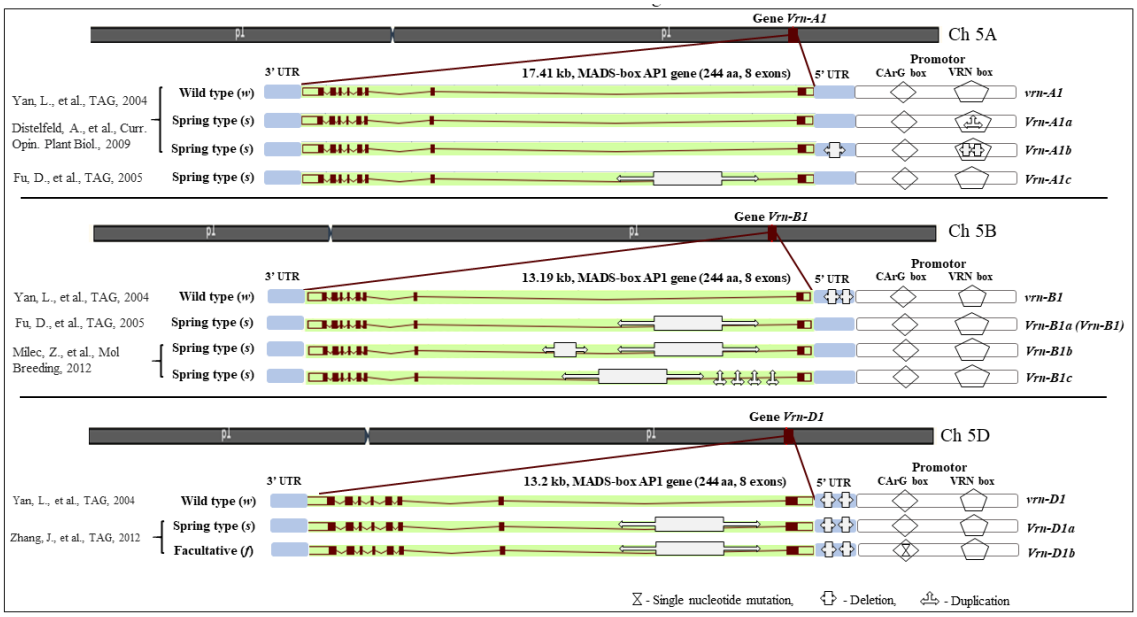


Figure 1.14 Allelic variation at the Vrn-1 genes

Genes are in 3' to 5' orientation

1.4.3.2 Photoperiod genes

Wheat is a long day plant by nature, that means sensitive to photoperiod. Therefore, it flowers earlier under long days (more than 14h). However, genotypes carrying photoperiod insensitive (PI or day neutral) alleles are better adapted to environments with shorter daylength (less than 12h) as they enter ear emergence earlier than those photoperiod sensitive counterparts. Three main homoeologous genes - *Ppd-A1*, *Ppd-B1* and *Ppd-D1* (former *Ppd3 Ppd2* and *Ppd1* respectively in old nomenclature) – localised on the short arms of the wheat group of chromosome 2 are responsible for the genetic control of sensitivity to photoperiod (Turner et al., 2005; Wilhelm et al., 2009) (Figure 1.15). Suffixes "*a*" and "*b*" of these genes indicate insensitivity and sensitivity to the light respectively (McIntosh and Yamazaki, 2013).

Yield advantages of day neutral alleles, *Ppd-A1a*, *Ppd-B1a* and *Ppd-D1a*, combined with semi-dwarf stature, were positively noted during the "Green Revolution" (Borlaug, 1983). Effects of these homoeoalleles on quantitative earliness were different in PI mutants (Shaw et al., 2012). *Ppd-D1a* was described as having a sturdy effect and frequent in wheat gene pool, followed by

Ppd-B1a and *Ppd-A1a* (R Scarth and Law, 1983; Bentley et al., 2011). However, the effect of *Ppd-B1a* on photoperiod insensitivity was stronger than *Ppd-D1a* and the combined effect of *Ppd-B1a+Ppd-D1a* was stronger than *Ppd-B1a* (Tanio and Kato, 2007). Several *Ppd-A1a and Ppd-B1a* alleles and their corresponding haplotypes and haplogroups were identified using different sets of molecular diagnostic markers in different wheat germplasm panels (Nishida et al., 2013; Muterko et al., 2015). Guo et al., (2010) identified six *Ppd-D1* haplotypes and five polymorphic sites two of which are in the promoter region, one in the close vicinity of the 5′ transcription start site of the first intron and other two were found in exons seven and eight (close to the 3′ end of the transcription region) (Guo et al., 2010).

Initial comparative mapping (Beales et al., 2007) found that wheat *Ppd-D1a* is colinear with the barley *Ppd-H1* and had a 2089 bp deletion upstream of the protein coding region which triggered insensitivity where genotypes carrying this deletion flowered early in short and long days. Before wheat, *Ppd-H* in barley was identified and described as a gene of *Pseudo-Response Regulator* (*PRR*) gene family (Turner et al., 2005). *PRR*, especially *PRR7*, is a part of the circadian clock in Arabidopsis and down-regulates *CDF1* , (*CYCLING DOF FACTOR 1*) a repressor of *CO (CONSTANS)*, a key component of the photoperiod pathway (Imaizumi et al., 2005; Nakamichi et al., 2007). However, *PRR* does not affect any of the circadian clock genes *CCA1 (CIRCADIAN CLOCK-ASSOCIATED 1)*, *TOC1* (*TIMING OF CAB EXPRESSION 1*) and *GI (GIGANTEA)* in wheat, but affected *TaCO1* and *TaFT1*, with a reduction in *TaCO1* expression as *TaFT1* expression increased and thus, was said to regulate flowering directly as there was a strong correlation between insensitive alleles, *FT1* and flowering where each insensitive allele increases basal transcription levels (Beales et al., 2007; Wilhelm et al., 2009; Shaw et al., 2012). These changes in gene expression are caused by deletions or transposon insertion within the promoter region as well as copy number variations (CNVs) (Díaz et al., 2012). Seemingly, the large promoter region deletions have a greater effect on decreasing photoperiod sensitivity than CNVs in hexaploid and six tetraploid wheat species (Muterko et al., 2015). However, a recent study showed that CNVs are important source of heading time variation in European durum wheat (Würschum et al., 2019).

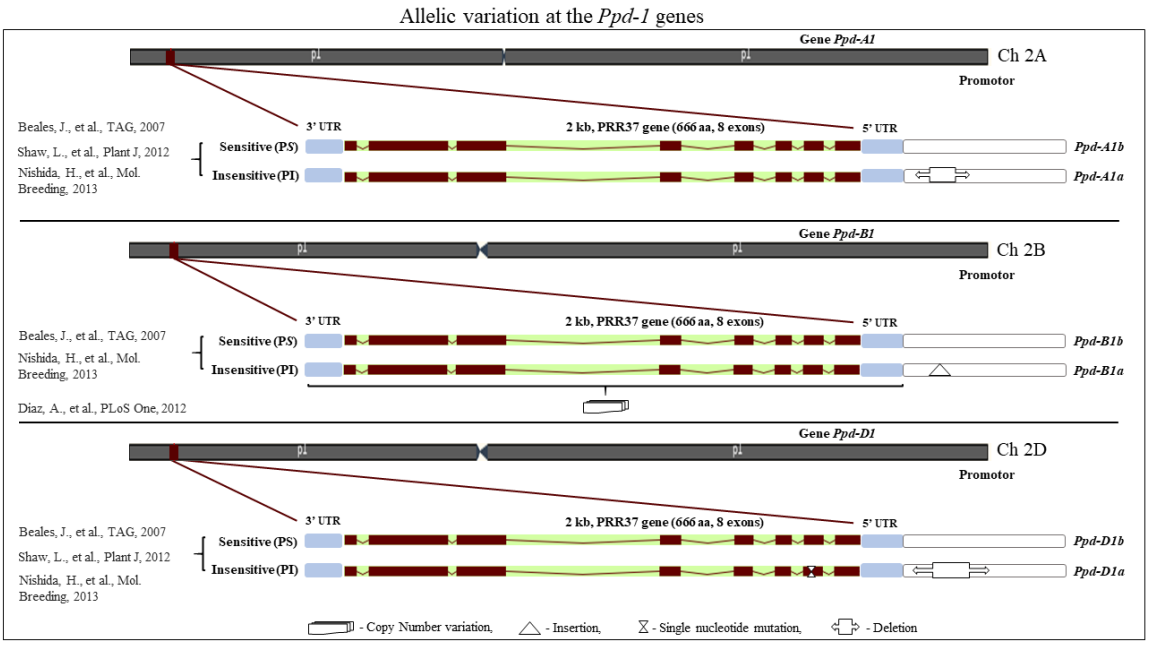


Figure 1.15 Allelic variation at the Ppd-1 genes

Genes are in 3' to 5' orientation

1.4.3.3 Genes of Earliness Per Se

Wheat heading date is mainly controlled by the genes of Vernalization and Photoperiod and earliness per se (*EPS*) pathways. However, the last is comparatively less understood compared to that of vernalization and photoperiod genes in barley and wheat. *Eps* genes are the genes which provide variation in heading date when genotype is completely vernalized and given a maximum duration of photoperiod. The *Eps* regulation of heading date is not dependent on environmental factors such as cold and daylength as in case of *Vrn* and *Ppd* respectively. Even though the effect of *Eps* genes on controlling the time of ear emergence in wheat has been studied in 1980s (Hoogendoorn, 1985), it was neglected due to instability across different agronomic environments (Griffiths et al., 2009). So far, several genes (QTLs) have been found related to earliness *per se* pathways one of which is *Eps-Am1* found in diploid wheat *Triticum monococcum* on chromosome 1Am (Bullrich et al., 2002). Other major two QTLs associated with earliness designated as *Eps-5BL1* and *Eps-5BL2* have been found on wheat 5BL chromosome (Tóth et al., 2003). The chromosome 3A of wheat is known to carry earliness *per se* genes (Miura et al., 1999). *Eps-D1* on 1DL wheat chromosome was fine mapped (Zikhali et al., 2015).

1.4.3.4 Genes of Reduced Height

1.4.3.4.1 General introduction to plant height and genes controlling plant height in wheat

Plant height (PH) is a complex trait in wheat, composed of internode lengths and the length of an ear. In plants, plasticity of PH is controlled by genetic loci and such environmental factors as light intensity, diurnal temperature, and humidity throughout the developmental stages from seed germination to ripening (Mu et al., 2021; Xue et al., 2021). Studying PH is of critical importance in

understanding and improving the adaptation and grain yield of wheat (Griffiths et al., 2012; Würschum et al., 2017; Yu et al., 2020b). For instance, to increase wheat grain yield, one of the main directions of "Green Revolution" in 1960s' was the introduction of semi-dwarf (height reducing) genes (Hedden, 2003) as tall wheats are prone to lodging by wind and rain (Peng et al., 1999) or when high levels of fertilisers are applied (Griffiths et al., 2012). Lodging causes substantial yield penalties and quality drops significantly (Peng et al., 1999). Currently, about 25 height reducing (*Rht*) genes named as *Rht1-Rht25* are known in wheat (McIntosh et al., 2017; Mo et al., 2018). Amongst these, *Rht1 (Rht-B1) and Rht2 (Rht-D1)* (Peng et al., 1999), *Rht3 (Rht-B1c) and Rht10 (Rht-D1c)* (Pearce et al., 2011), *Rht12* (Sun et al., 2019; Buss et al., 2020), *Rht14* (Duan et al., 2022), *Rht18* (Ford et al., 2018), *Rht23* (Zhao et al., 2018) were cloned. *Rht24* and *Rht25* were squeezed to a small interval of 1.85 and 1.85 cM respectively (Tian et al., 2017; Mo et al., 2018). Many *Rht* genes have been found on all three genomes of hexaploid wheat, but not all of them have successfully been used in wheat breeding due to adverse pleiotropic effects.

1.4.3.4.2 The major height-reducing genes and their genetic control

The major ones are historic gibberellic acid (GA) insensitive *Rht-B1b* (*Rht1*) and *Rht-D1b* (*Rht2*) semi-dwarf alleles, located on homoeologous 4BS and DS chromosomes respectively, encoding *DELLA* proteins which repress the function of growth inducer hormone GA (Peng et al., 1999). Thus, semi-dwarf genotypes increased the harvest index significantly in many environments via allocating more nutrients to developing spikes, rather than the straw, resulting in a higher rate of floret survival and increased grain number per spike (Youssefian et al., 1992; Flintham et al., 1997). However, the benefits of these *DELLA* mutants are limited or even reversed in hot and dry rainfed conditions where seeds need to be sown deeper and early establishment, vigour and longer coleoptile/root length are of importance (Rebetzke et al., 2007) to fine-tune final grain yield and adaptation. Especially, their negative effects on significant agronomic traits such as early vigour (Rebetzke et al., 2001), grain yield (Rebetzke et al., 2007), grain size and grain weight (Börner et al., 1993), root (Bai et al., 2013) and coleoptile length (Botwright et al., 2001) have been observed. Therefore, a search for the hot and dry environment specific alternative *Rht* genes without negative pleiotropic effects on mentioned adaptation and yield related traits has always been on the breeder's agenda.

1.4.3.4.2.1 Genetic bases of "Green Revolution" *Rht-1* genes

DELLA proteins are part of plant-specific GRAS family of proteins which act as transcription factors/regulators (Pysh et al., 1999; Locascio et al., 2013; Van De Velde et al., 2017).
The name is originated from a short stretch of amino acids (D-E-L-L-A) in their N-terminus and is well conserved in all plants (Locascio et al., 2013). The N-terminal region of *DELLA* proteins also present the VHYNP domain beside the DELLA itself. By contrast, the C-terminal region contains four domains including two leucine heptad repeats (LHRI and LHRII), a nuclear localization signal

(NLS), and SH2-like phosphotyrosine-binding domain. The N-terminal region is the GA-signalling domain and two highly conserved motifs (DELLA and VHYNP) in this domain are required for GA-induced degradation in rice (Itoh et al., 2002). Moreover, mutations in the N-terminal section confer dwarfism in Arabidopsis, maize and wheat. Particularly, a T-for-C substitution converts the Q64 (CGA) codon to a translational stop codon to form the *Rht-B1b* allele, in contrast, a T-for-G substitution converts E61 (GGA) codon to a translational stop codon in *Rht-D1b* (Peng et al., 1999) (Figure 1.16).

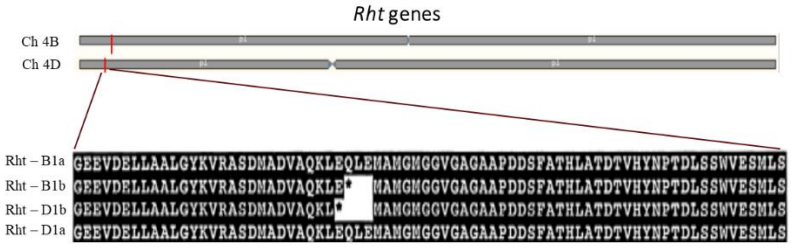


Figure 1.16 Wheat Rht-1a/b alleles

Source: Ensembl Plants and Peng et al., 1999. Translational stop codons are represented by an asterisk.

1.4.3.4.3 Revisiting *Rht* genes without compromising early establishment and growth in wheat

Alternative height reducing genes *Rht8*, *Rht9* and *Rht14* could potentially be used to develop high yielding wheats with longer coleoptile, suitable for semi-arid environments where deep sowing is required, as there was a small correlation between PH and coleoptile length in varieties carrying these semi-dwarf genes (Rebetzke and Richards, 2000; Rebetzke et al., 2007; Amram et al., 2015; Vikhe et al., 2019). However, yield penalties of *Rht8* were observed in the US (Lanning et al., 2012) and UK (Kowalski et al., 2016). The list of wheat alternative *Rht* genes to replace 'Green Revolution' genes also includes recently identified and characterised *Rht12* (Buss et al., 2020), *Rht18* (Ford et al., 2018; Tang et al., 2021), *Rht24* (Tian et al., 2017; Würschum et al., 2017) and *Rht25* (Mo et al., 2018). Although *Rht18* reduced biomass and yield, it increased TGW and harvest index (Yang et al., 2015). Other semi-dwarfism genes, such as *Rht5*, *Rht6*, *Rht8*, *Rht13* enhanced seedling emergence and *Rht4*, *Rht19*, and *Rht12* showed highest for photosynthetic traits while *Rht9*, *Rht16*, and *Rht15* performed best for early seedling growth parameters (Mohan et al., 2021). Among these genes, *Rht15* negatively affects yield component traits, but improves grain quality (protein content (8.7 %), wet gluten content (9.2 %) and starch content (2.3 %) (Zhao et al., 2021) which is an important trading feature of the Kazakh wheat. Moreover, *Rht14* and *Rht16* could potentially be used in breeding for improved establishment (Haque et al., 2011). Here, *Rht14* does not affect root length nor it affects coleoptile length and most of the yield components in macaroni wheat (Vikhe et al., 2019; Duan et al., 2020). Therefore, the benefits of the *Rht14* are yet to be confirmed in bread wheat. A combination of *Rht4* + *Rht8* was reported to reduce PH by decreasing internode length but without affecting coleoptile length and improved yield as well as related

components such as fertile tiller number, TGW and plant biomass of rainfed grown wheat (Du et al., 2018). Recessive *Rht22* shortens PH via reducing cell numbers without affecting internode length (Peng et al., 2011).

1.4.3.4.4 Other spare *Rht* genes

This thesis does not aim to characterise all existing *Rht* genes in detail. However, we do provide readers with general information about all *Rht* genes. For example, Tom Thumb *Rht3* gene, localised on the same chromosome 4BS such as *Rht B1*, severely reduces height and is partially dominant, and more GA (Gibberellic Acid) insensitive than Norin 10 (Gale and Marshall, 1976). The Ai-bian1 (Chinese variety) dwarfing gene on 4DL chromosome *Rht10* is GA insensitive as well, but more severe than *Rht3*. By contrast, *Rht5* is partially dominant, reduces plant height by 50% and associated with a reduction in yield (Gale et al., 1985). *Rht7* is recessive, reduces height by 24% and has negative effect on grain yield (Worland et al., 1980). Hongwei and Zhonghua identified *Rht21* in Chinese wheat "XN0004" (Hongwei and Zhonghua, 1993), however, the existence of this gene is debatable (Börner and Worland, 2002). *Rht11* and *Rht17* were found to reduce leaf elongation rate and coleoptile length (Ellis et al., 2004). *Rht23* confers dwarfness and compact spike phenotype, but decreased grain number per spike and TGW (Zhao et al., 2018). Perhaps, these height reducing genes are not suitable to low precipitation wheat growing environments.

1.4.3.4.5 A brief history and origins of *Rht* genes

The two old Japanese landraces, "Akakomugi" and "Daruma", are natural sources of some wheat dwarfing genes and several others originated from induced mutations such as ethyl methyl sulphate (EMS), MNH, DES, fN, thN and x-ray or Gamma rays. Akakomugi is the originator of the *Rht8* (on chromosome 2D) and *Rht9* (on 7B) genes both of which co-segregate with *Ppd1* insensitive genes (Gale et al., 1985). Thus, the photoperiod-insensitive gene, *Ppd1*, also reduces plant height by shortening the life cycle (Worland et al., 1988). Seeds of Akakomugi were brought to Europe, particularly to Italy, by Nazareno Strampelli at Rieti in 1911. Strampelli's crosses of local varieties with Akakomugi resulted in two famous shortest varieties Ardito and Villa Glori the former of which was used in the parentage of the important Russian variety, Bezostaya 1 (Gale et al., 1985) which in turn was employed in the parentage of our RIL/NIL parent, Pamyati Azieva. Bezostaya 1 carries *Rht-B1a*/*D1a*, *Rht8c* (McIntosh et al., 2000; Zheleva et al., 2006; Chebotar, 2008), *Rht8* (Worland et al., 1998b; Kurkiev et al., 2008), *Rht9* (Bespalova, 2003) and *Rht11*(*Rht-B1e*) (Li et al., 2012). The second variety Daruma has two types; Shirodaruma (white) and Akadaruma (Red) the former of which is the parentage of famous and vitally important wheat semi-dwarf Japanese variety Norin 10. Especially it was obtained from the cross of Shirodaruma (white) with American varieties. As a result, two gibberellin acid insensitive genes *Rht-B1* and *Rht-D1* (initially was

known as *Rht1* and *Rht2*) originated from Norin 10 (Hedden, 2003). *Rht4* was induced by x-ray. By contrast, *Rht5*, *Rht7, Rht15* and *Rht23* are all induced by EMS.

1.4.3.4.6 QTL for plant height in wheat

Plant height is an important characteristic, particularly in common wheat, defining its agronomic performance (Griffiths et al., 2012). Genetic assessment, identification and characterisation of plant height related genomic loci quantitatively or conducting co-called quantitative trait loci (QTL) analyses are the main initial steps of manipulating and optimising plant height. Such analyses have been accelerated with the emergence of molecular markers and other genomic technologies. Since then, numerous height-associated QTL in wheat have been reported (Kato et al., 1999; Peng et al., 1999; Cui et al., 2011; Griffiths et al., 2012; Zhang et al., 2017a; ZHOU et al., 2020). Meta-QTL analysis revealed that all wheat chromosomes, except for 3D, 4A, 5D and three chromosome 7 homeologs, possess plant height genes (Griffiths et al., 2012). Tian and co-workers listed more than 50 QTL localised on all wheat chromosomes (Tian et al., 2017). However, not all QTL are stable. For example, global quantitative trait loci analysis found 33 height associated QTL, almost half of which were stable and these loci were located on chromosomes 1B, 2D, 3A, 4B, 4D, 5A, 6A, 6D, 7A, and 7B (Guan et al., 2018).

1.4.3.4.7 The list of reduced height genes on 6A and 5A wheat chromosomes

Of total 25, five and one height-reducing genes are located on 6A and 5A chromosomes respectively (Figure 1.17). These numbers highly correspond with the number of coding genes each chromosome possesses based on reference sequence of Chinese Spring (Appels et al., 2018). Out of five, two (*Rht16 and Rht25*) and three (*Rht14, Rht18 and Rht24*) height decreasing genes reside on short and long arms of the 6A chromosome respectively. Conversely, no QTL nor gene was reported for 5A short arm, except a single height related gene, independent of *Vrn1*, which was reported a long time ago (Snape et al., 1985). However, the literature search did not show further details of that gene. The underlying genetic mechanism and exact position of *Rht12* are known as the gene was cloned recently (Buss et al., 2020). Among 6A genes, *Rht18* and *Rht14* were fine-mapped (Ford et al., 2018; Duan et al., 2022). All these three genes, *Rht12, Rht18 and Rht14*, have similar height-reducing mechanism and shorten PH via increasing the expression of GA 2-oxidase genes (Buss et al., 2020). Interestingly, it was reported that semi-dwarfism genes on chromosome 6A may potentially be used in wheat breeding for improved establishment (Haque et al., 2011).

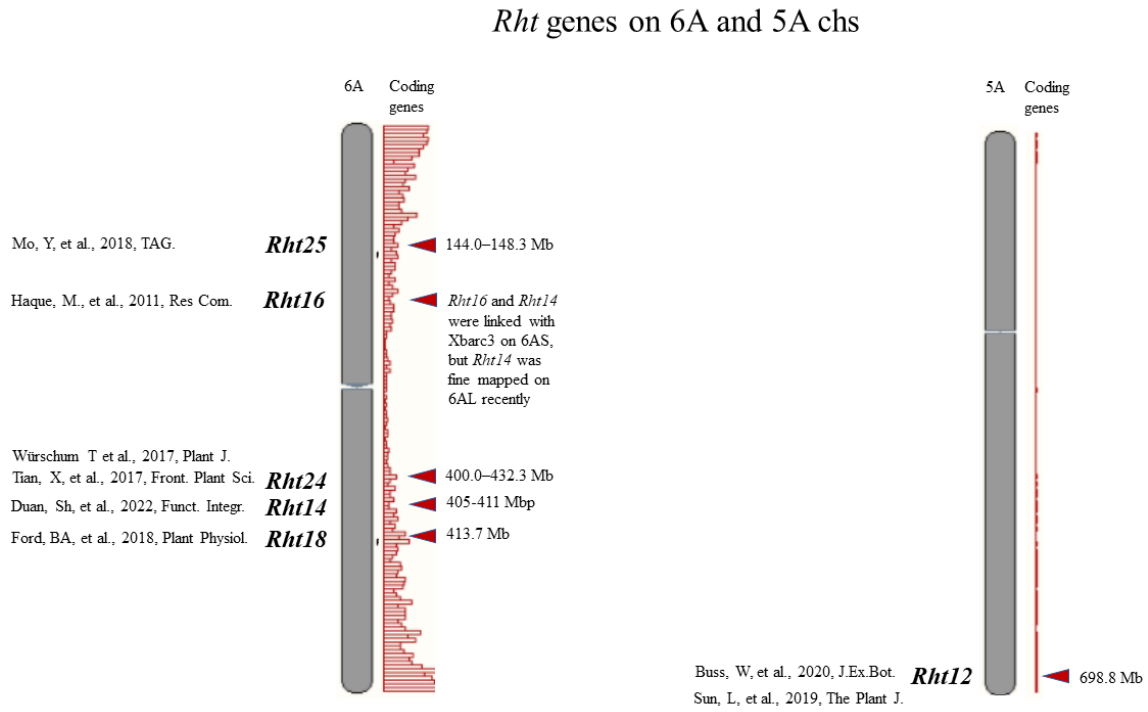


Figure 1.17 Rht genes on wheat 6A and 5A chromosomes

1.4.4 Understanding the broad adaptation of Kazakh bread wheat gene pool through molecular characterisation of *Vrn, Ppd, Rht* and *Eps* genes

In Europe PI wheat varieties can be grown alongside with photoperiod sensitive genotypes (Worland et al., 1998b) and mostly winter wheat is cultivated. Comparatively, most of the wheat cultivars from Kazakhstan are highly photoperiod sensitive (Trethowan et al., 2006) and more than 80% are spring type (Turuspekov et al., 2017a). To assess the allelic state of Kazakhstani wheat for the most common alleles of known adaptation genes, we have genotyped 96 spring wheat accessions using *Vrn* (*Vrn-A1, Vrn-B1, Vrn-D1), Ppd (Ppd-A1, Ppd-B1, Ppd-D), Rht (Rht-B1, Rht-D1), 1RS:1BL* rye-wheat translation and *Eps (TaELF3-D1, TaFT3-A1, TaFT3-B1)* molecular markers using KASP (Kompetative Allele Specific PCR) genotyping platform (Turuspekov et al., 2017b). Results suggested that almost all tested accessions are photoperiod sensitive with no variation at the *Rht-B1* and *Rht-D1* (Table 1.5). In spite of the fact that wheat genotypes from Kazakhstan are spring, tested accessions showed that the wheat gene pool was represented by different haplotypes of *Vrn* genes. In addition, genotyping results based on *Eps* genes showed that most of wheat varieties carry “late” allele. Although the variation at the *Eps* genes is minor (~3-5 days) (Zikhali et al., 2014), sometimes small earliness plays a significant role in such continental environments as Kazakhstan. The differentiation of samples by *1RS:1BL* markers suggested that genotypes carrying this translocation have significantly higher yield in Northern Kazakhstan ($P<0.01$) (results are not presented in the table). At the end it was concluded that long-term breeding efforts of spring wheat in the country led to the selection of photoperiod sensitive genotypes with none of the known semi dwarf mutations at *Rht-B1* and *Rht-D1*. Since wheat is

growing in harsh climate environments with very short vegetation period associated with drought and heat, new genotypes with earlier flowering time are required to be tested. KASP assays for key genes of wheat adaptation are very efficient tools for application in improvement of local breeding projects.

alleles	*Vrn-A1*	*Vrn-B1*	*Vrn-D1*	*Ppd-A1*	*Ppd-B1*	*Ppd-D1*	*Rht-B1*	*Rht-D1*	*TaELF3-D1*	*TaFT3-A1*	*TaFT3-B1*
a	Spring 71	Spring 63	Spring 88	Insensitive (1)	Insensitive (6)	Insensitive (1)	Tall (96)	Tall (96)	Late (77)	CC (96)	GG (7)
b	Spring 25	Spring 33	Spring 8	Sensitive (95)	Sensitive (90)	Sensitive (95)	Short (0)	Short (0)	Early (23)	AA (0)	AA (20)
c											Del (73)
Nei's index	0.385	0.449	0.147	0.019	0.113	0.019	0	0	0.365	0	

Table 1.5 Differentiation of wheat genotypes from Kazakhstan

Commercial wheat varieties from Kazakhstan differentiated by *Vrn-A1*, *Vrn-B1*, *Vrn-D1*, *Ppd-A1*, *Ppd-B1*, *TaELF3-D1*, *TaFT3-A1*, *TaFT3-B1*, *TaFT3-D1*, *Rht-B1* and *Rht-D1* KASP assays. Results of *Vrn* genes revealed that Kazakh spring wheat gene pool exhibit different haplotypes. Varieties with a deletion at *TaFT3-B1* loci are the latest genotypes in short days comparing to other specimens possessing two alleles at the locus.

1.4.5 Quantification of genetic components in wheat/plants

Breeders use quantitative genetics to help identify which genotypes exhibit best performance for target traits with continuous variation and genotype x environment variability (Bernardo, 2020). Thus, the main aim of genetic association studies is to find the underlying genetic loci/locus which cause/s or contribute/s to the change in the architecture of a given trait and then use those genomic locations to improve target traits of agronomic importance (Véronique et al., 2008). Basically, plant geneticists use three approaches: i) development of mapping population/s segregating for the trait of interest, ii) phenotypic assessment of plant performance and iii) genomic evaluation for the trait quantification and linking it to a certain chromosomal part. These and other important techniques will be discussed in the coming sections.

1.4.5.1 Mapping populations

A segregating or mapping population is a population comprised of individuals with different genotypes (Xu, 2013). Most of the traits with high importance for breeding in these populations are quantitative. Genes controlling those traits were mapped mostly using either DH (doubled haploids) or RILs (recombinant inbred lines) segregating for the trait of interest. Both have been used successfully to improve yield and diseases resistance in wheat (Fleury et al., 2010), but each

has its own advantages and disadvantages (Table 1.6). For example, DH breeding technique allows rapid achievement of uniformity which significantly shortens the breeding cycle in not only wheat, but in all crops (Santra et al., 2017). However, this breeding method is associated with a low rate of embryogenesis, regeneration and recombination, high frequency of albinism, segregation distortion, and the low frequency of chromosome doubling to obtain DH (Bernardo, 2009; Dunwell, 2010). Conversely, RIL generation has fewer difficulties, but takes a longer time to reach homozygosity and is associated with a higher rate of recombination which increases the response to selection. Thus, RILs are useful for linkage mapping and quantitative trait loci analysis (Mansur et al., 1996).

Population	Advantages	Limitations
DH	Rapid achievement of uniformity	Low frequency of chromosome doubling, recombination, embryogenesis and regeneration. High frequency of albinism and segregation distortion
RILs	High recombination rate, easy to propagate	Longer time required to reach homozygosity, loss of phenotypic variation in outcrossing plant species in later stages of generation

Table 1.6 Pros and cons of DH and RILs

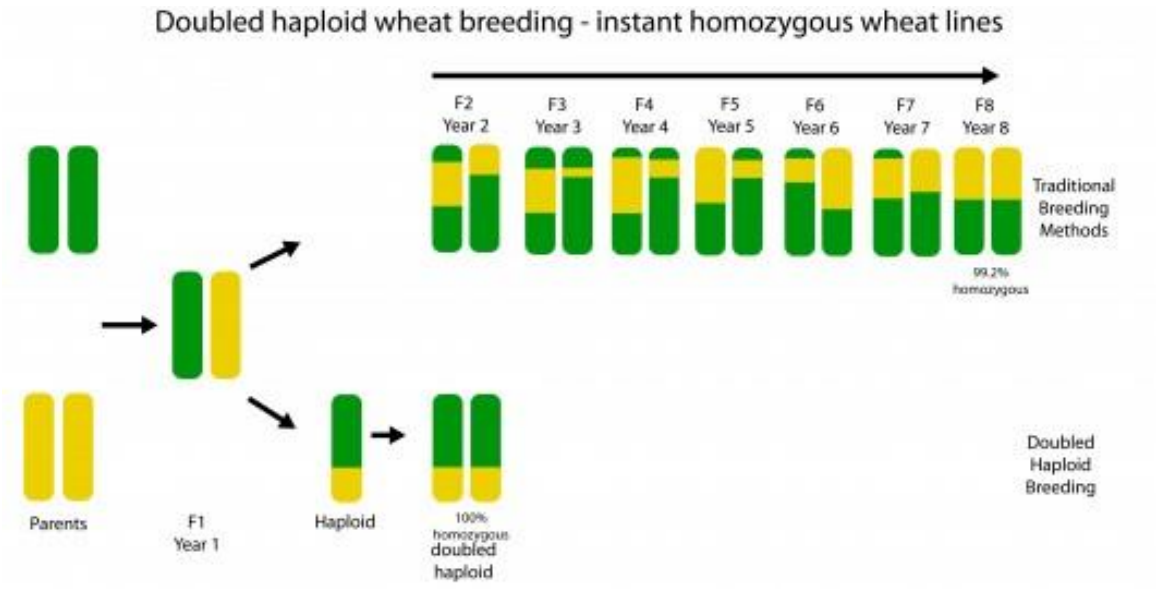


Figure 1.18 DH vs RILs

Source: https://coloradowheat.org/

1.4.5.2 Molecular markers

Molecular markers are the sections of DNA used as genomic landmarks and are a powerful tool to study structure, function, genetic variation and similarity among different genotypes within a species as well as between different species. Molecular markers are often outside of genes which are pieces of the genome which have certain biological functions in contributing to phenotypic variation. Genes are also convertible into molecular markers used to study polymorphism within a population. Thus, a molecular marker can be used to identify the location of a gene in the genome and its function. There are several types of DNA markers. Such markers as RFLP (restriction fragment length polymorphisms), AFLP (amplified fragment length polymorphism) (Becker et al., 1995), RAPID and SSR (simple sequence repeats) (Nagaoka and Ogihara, 1997) are largely outdated now. Instead, SNPs (single nucleotide polymorphisms) are widely deployed in current genetic studies (Collard and Mackill, 2008). The importance of molecular markers in genetic studies is that they are inheritable. This essential phenomenon enables one to track inheritance of a particular gene/molecular marker/trait from generation to generation. This in turn makes it possible to understand deeply the behaviour of the phenotype of interest. Therefore, the construction of genetic maps based on DNA markers is a vital prerequisite for the design of successful molecular breeding programs.

1.4.5.3 Genetic linkage maps

The first concept of a genetic map was introduced by Alfred H. Sturtevant (Sturtevant, 1913).
In general, the distance between two markers in a genetic map is quantified based on recombination frequency/fraction or number of recombination events during gamete formation of sexual cells during meiosis. Recombination frequency/fraction is the ratio of the number of recombined gametes to the total number of gametes produced (Xue et al., 2013). In other words, it is a percentage of chances of genes getting separated. Recombination units or map units between molecular or genetic (two have the same meaning and are used interchangeably) markers are counted in centi-Morgans, named after famous geneticist Thomas Hunt Morgan. If genes or molecular markers are located close to each other, alleles of those genes or molecular markers will usually pass on to the next generation together. Those genes and molecular markers are said to be linked.

A genetic linkage map is a representation of the linear arrangement of genes and molecular markers on chromosomes of an organism whether be it plant species, an animal or human. Genetic maps are an essential instrument in fundamental and applied genetic studies (Stam, 1993) and can be represented in the form of graph and a table. In order to construct genetic maps or to perform linkage analysis i) a segregating population needs to be developed, ii) this population must be genotyped with molecular markers. Nowadays, many genotyping planforms which are based on the application of large numbers of molecular markers are available. As these platforms generate large

amounts of genomic information, the data handling using classic genetic mapping is almost impossible. Therefore, genetic map construction is fully computerised and the statistic behind the screen is complex. Such software packages have been available since the 1980s, particularly LINKAGE-1 (Suiter et al., 1983) and MAPMAKER (Lander et al., 1987), and were widely used. In the 1990s more advanced and user-friendly mapping programmes such as Join Map (Stam, 1993), Gmendel (Holloway and Knapp, 1993), CarthaGene (Schiex and Gaspin, 1997), and in the 2000s, MapManager (Manly et al., 2001), RECORD (Van Os et al., 2005), MSTmap (Wu et al., 2008), MapDisto (Lorieux, 2012) and ASMap (Taylor and Butler, 2017) have been developed mostly as open source software. The large number of individuals used in an experiment for the construction of genetic maps is preferred to deal with obstacles such as genotyping errors and segregation distortion, even though map construction requires fewer samples than QTL mapping (Broman, 2010). Other issues reported in the same article are variations in marker polymorphism, marker density and recombination rate. In addition, large sample size (large size of an experimental population) is important to estimate the comparatively true value of recombination fraction (meaning that the estimation is unbiased) between two loci (Xue et al., 2013) and also is essential to perform statistical analyses such as Chi-Square Test, in order to determine an appropriate p-value (numerical probability) to reject null hypothesis to conclude whether two genes are linked or not when pairwise linkage analysis ends up with distorted results. That means even if two markers do segregate independently, they might appear to be linked as a result of estimation of the recombination frequency showing low recombination frequency (low recombination frequency means markers are linked) with high LOD score (Broman, 2010).

1.4.5.4 Genetic mapping of traits

GWAS and QTL analysis (Quantitative Trai Loci Analysis) are the most important methods allowing plant scientists to identify genetic loci underpinning key agronomic traits of crops and enabling the development of accurate diagnostic molecular markers for systematic breeding strategies (Nantawan et al., 2019). These two mapping strategies both aim to detect QTL through linking the trait with genomic marker. The difference between GWAS and QTL analysis is fundamental and directly related to resolution and power of mapping (Xu et al., 2017). For instance, GWAS involves studying genome-wide set of genetic variants across many genomes (natural population), which are ancestrally similar but differ phenotypically, to identify whether any of those variants are statistically associated with a specific trait or disease resistance (Uffelmann et al., 2021). GWAS takes advantage of historic recombination events and thus the number of recombinations are always higher than that of family – based segregating populations, particularly DH population (Alqudah et al., 2020). Therefore, GWAS provides higher resolution and greater allele numbers (Xu et al., 2017). Unlike GWAS, QTL analysis is done to identify trait-marker association in biparental populations. Thus, it depends on recently created recombination events which in turn provides lower rate of recombination and allele frequency than that of

association or linkage disequilibrium mapping (Xu et al., 2017) but biparental QTL mapping is powerful because two alleles are compared at each locus at equal frequency.

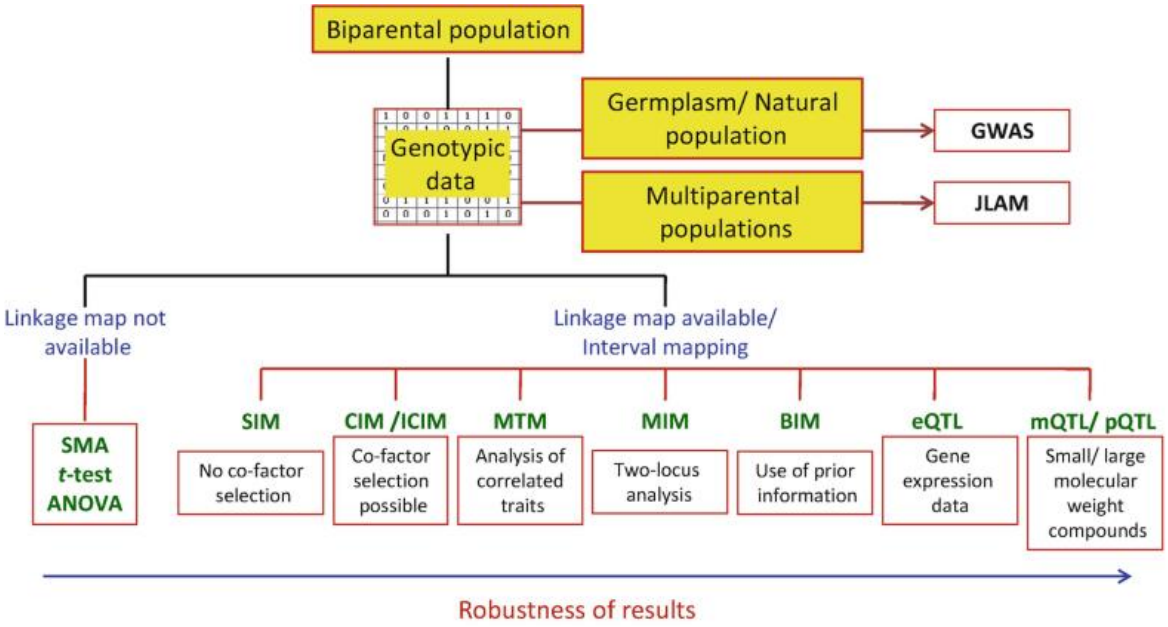


Figure 1.19 Graphical representation of QTL detection methods

The figure demonstrates the graphical representation of various methods of QTL detection. When genetic linkage map is available, the QTL can be identified through simple interval mapping (SIM), composite/inclusive composite interval mapping (CIM/ICIM), multiple trait mapping (MTM), multiple interval mapping (MIM), Bayesian interval mapping (BIM), expression QTL (eQTL) or metabolite or protein QTL (mQTL/pQTL). The criteria used in each of these interval mapping approaches are given in the box below the method. An arrow below shows the relative robustness of results of these methods over one another. When the genetic data is available for the germplasm/natural population, genome wide association study (GWAS) can be performed to map the QTL. Compared to natural population, multiparental populations enable joint linkage and association mapping (JLAM)

Source: (Kulwal, 2018).

1.4.5.5 Trait mapping in the era of next-generation sequencing

Advanced genomic technologies have opened the door to the new era of genetic mapping of traits (Varshney et al., 2020). Trait mapping using WGS and targeted sequencing means such as exome and promotor sequencing are becoming increasingly popular in genetic research and believed to have a clear-cut advantage over pedigree breeding and MAS to enhance genetic gains for complex traits (Crossa et al., 2017). However, to use NGS data properly for capturing genome-wide genetic variation, an urgent need for proper population design, statistical methods and precision phenotyping has arisen (Xu et al., 2017). Recently published wheat genome (Appels et al., 2018) and pan-genomes (Walkowiak et al., 2020) are the initial step in understanding the physical structure of the wheat genome and these tools have already advanced map-based cloning (Jia et al., 2018). For example, sequencing and de-novo assembly of multiple, independent, allelic mutants for 6A chromosome of wheat allowed to functionally characterise *Rht-18* in wheat (Ford et al., 2018). As a result of "targeted chromosome-based cloning via long-range assembly" (TACCA), wheat

leaf rust resistance gene *Lr22a* was cloned within the short period of time (Thind et al., 2017). We have also fully sequenced the whole genomes and exons of NIL parents. It allowed us to target gene coding as well as regulatory areas of the region of interest during fine mapping.

1.4.6 Overview to the next generation plant breeding and the salience of our research

The burst of new generation breeding tools has greatly alleviated the labour-intensive long-lasting solo phenotype-based traditional plant breeding. It also allowed scientists to understand better the functions of genes of agronomic importance. Thus, modern molecular genetics, genomics and bioinformatics methods combined with conventional breeding methodologies is a powerful way to see the desired gene/s in the breeder's favourite variety within a short time. Recently emerged new phenomics and genomics approaches such as speed breeding allow up to six generations of many crops in a year (Watson et al., 2018; Cha et al., 2021) and accelerated gene cloning approaches such as map-based cloning (Munoz-Amatriain et al., 2011), RNA-seq approach (Habib et al., 2018; Chen et al., 2021), next-generation sequencing (Zhong et al., 2018; Bhat and Yu, 2021), GWAS (Juliana et al., 2018), MutChromSeq (Sanchez-Martin et al., 2016), MutRenSeq (Steuernagel et al., 2016), sequence assembly based AgRenSeq (Arora et al., 2019) and bulked segregant exome capture sequencing, BSE-Seq (Dong et al., 2020) have simplified old laborious methods and reduced the cost and time required. However, all these methods require the knowledge of databases understanding and sequencing data management. Moreover, these gene cloning and validation approaches have to be employed with respect to the aim of a study and the type of research material being used. A direct genome editing tool, CRISPR-Cas9 (Cao et al., 2016; Zhang et al., 2020b) and high-throughput phenotyping planforms empowered by GPS (Crain et al., 2018; Song et al., 2021) and temperature, humidity, light sensors, able to monitor the plant throughout its whole life (Gewin, 2017; Zhou et al., 2017) are other methodologies which promise benefits in future plant breeding. Moreover, international cooperation plays a significant role in understanding plant adaptation better as it provides breeders and researchers with an opportunity to share their knowledges and experiences, and wheat with various agro-ecological niches to be tested.

Here, we developed a segregating population, derived from the cross of the wheat varieties Paragon and Pamyati Azieva, specifically Recombinant Inbred Lines (RILs) followed by Near Isogenic Lines (NILs), in collaboration between John Innes Centre (JIC), UK and Institute of Plant Biology and Biotechnology (IPBB), Kazakhstan. The beauty of it is that this is the first mapping population involving wheat cultivar from Kazakhstan. As a result, several QTLs associated with significant agronomic traits have been identified. In Kazakhstan, plant and animal breeding mainly rests on using conventional methods. Although several crop species have been genotyped with various sets of molecular markers, the obtained wealth of genetic information has not been properly captured to intensify breeding programmes yet. Thus, we pioneered efforts to quantify genetic components of Kazakh bread wheat which might be associated with essential agronomic characteristics.

1.4.7 Thesis aims and objectives, and hypothesis

We think that the most important adaptation genes are likely to be fixed or present at high frequency in Kazakh wheat gene pool, but genetically uncharacterised. Additionally, what exact alleles or/and allelic combinations, of those salient adaptation linked genes, enhancing the grain yield is still unknown. Thus, the aim of this PhD thesis was to shed light on how wheat adaptation and yield is controlled genetically in Kazakh environment.

2. Chapter 2: Delineation of adaptation and yield associated loci for Kazakh bread wheat

2.1 Introduction

In order to identify trait associated loci, genetic components of the trait of interest should be quantified by creating new recombination events and then relating each trait with a locus. This can be achieved through the development of segregating populations, creating a genetic map and obtaining phenotypic values of plant characteristics or traits. For this purpose, we developed RILs as they are a powerful tool for the mapping of genes (Broman, 2005) and have been successfully used for the identification and validation of QTLs (quantitative trait loci) underpinning key agronomic traits in staples such as wheat (Griffiths et al., 2012), barley (Alqudah et al., 2020), rice (Shinada et al., 2014), chickpea (Sivasakthi et al., 2018), soybean (Concibido et al., 2004) and maize (Szalma et al., 2007). We also created the genetic map from these RILs and phenotyped them in controlled and non-controlled environments. As a result of the integration of genotypic and phenotypic data, several QTLs were found which are associated with important agronomic traits. Among these QTLs, two plant height (PH) effects residing on chromosomes 5A and 6A were interesting to study further as their additive effects were large, approaching 10cm. We have not seen such an increase in our previous studies with the same UK parent (Paragon) and are really interested to understand why height increasing alleles and relatively tall wheats in general still seem to be important for Central Asia. We hypothesised that PH is an essential contributor of increased adaptation and perhaps of yield as well. Therefore, this chapter provides detailed information about steps of identification of 5A and 6A height increasing QTLs. We will also briefly discuss other QTLs found.

2.2 Materials and methods

2.2.1 RIL development

A segregating or mapping population of bread wheat, particularly RILs consisted of 94 individuals, was developed from a cross between the spring wheat cultivars, Pamyati Azieva (Pam.Az) and Paragon (Par), through Single Seed Descent (SSD) bulk method (Tee and Qualset, 1975) which is widely used among breeders for the last three decades. The first parent - Pamyati Azieva - used in the cross, is well adapted to Kazakhstan. It is originated from Sibirskii NIISKH in Omsk, Russia, and is widely grown in Kazakhstan because it was included into the recommended list of the State Register of the country. Paragon is a UK spring wheat variety. It has been used extensively in the

UK's wheat pre-breeding programs as a parent for the development of Nested Association Mapping (NAM) populations, Near Isogenic Lines (NILs), and mutant populations (Wingen et al., 2017) as it was recognised as the bread making quality benchmark in the UK. Once these two elite varieties were crossed and F1 population was obtained, in each round of the bulk one seed was taken to develop next generation, that is one seed from F2 population was used to obtain F3. Only one seed of F4 population, meanwhile, was employed to form F5 hybrids. RIL development was carried out up by single seed descent to the F6 generation which produced lines with mosaic homozygous genome at most of the loci for the parents. All hybridization experiments and development of RILs were carried out in glass houses and Controlled Environment Rooms (CER) of John Innes Centre (JIC), UK, between 2012-2014.

2.2.2 DNA extraction from Pam x Par RILs and preparation of work DNA for KASP genotyping

In order to extract DNA from Recombinant Inbred Lines, five seeds of each line were sown into Petri dishes and put for 2 days at 4˚C for stratification. After that, they were transferred to an incubator with temperature range of 27-30˚C for 5-6 days for germination. DNA was isolated from 5-6 days old wheat seedlings. Extracted genomic DNA has been subjected to further purification using Qiagene Mini spin columns. Quality and quantity (A260/A280) of extracted DNA were evaluated in Nanodrop spectrophotometer with further visualisation on 1% agarose gels to assess integrity. DNA of each RIL was standardised to a concentration required for genotyping.

2.2.3 Genotyping and genetic map construction

The Illumina iSelect 15K platform was used to genotype RILs. Of 15 thousand SNP markers, 4595 showed polymorphisms in RILs. The initial genetic linkage map of a segregating papulation was constructed by Traitgenetics. Then, the quality control of linkage groups was re-assessed by an opensource statistical environment R, particularly via "ASMap" and "qtl" packages. This was done to identify several factors by which the quality of genetic linkage maps is negatively affected. Such factors are segregation distortion, genotyping error, existence of duplicated markers and the number of missing values/genotypes for markers as well as samples. Therefore, it is imperative to validate the genetic map quality with and without markers with aforesaid issues. If they really drop the genetic map quality, it is better to omit them prior to QTL identification. As almost all software tools based on graphical user-friendly interface, they do not provide users with such as much flexibility to work with data as R does. Two R packages, qtl and ASMap, are user frequently for linkage map construction, the latter of which has been developed recently to overcome the computationally cumbersome combinatoric methods of the former, making the whole process much quicker and easier (Broman et al., 2003; Taylor and Butler, 2017). The advantage of ASMap is that one can still implement many functions of the "qtl" in a way that they are used in the "qtl" package itself.

2.2.4 Phenotypic assessment of RILs

In the winter of 2015 RILs (F6) were sent to Kazakhstan to be grown in various environments. Due to limited number of seeds being sent to Kazakhstan, RILs were only multiplied in the field trials of Institute of Biology and Biotechnology (IPBB) in the first year. In the following 2016, seeds from previous year allowed us to grow RILs in Almalybak (Alm), Almaty region, south-eastern Kazakhstan, in two replicates with two rows each line. The length of rows in field trials was ~1 m and the distance between the seeds sown was approximately 5 cm from each other. At the same time, the segregating population was evaluated in the field and controlled environments of JIC, UK. Field experimental design in the UK included 1m block of plots. In the greenhouse, RILs were grown in 1L pots with cereals mix type of compost.

2.2.5 QTL linkage mapping

Data on phenotyping and genotyping have been combined for QTL analysis. In order to map the markers on wheat chromosomes and detect QTLs associated with traits, two types of mapping software, WinQTLCart 2.5 and R-QTL, have been deployed so as to minimize the statistical bias and increase confidence in the QTLs found. In both, composite interval mapping (CIM), with LOD score of not less than three, for all phenotyped plant traits was conducted to increase the power and precision of QTL detection.

2.3 Results

2.3.1 Genetic Map Statistics and Quality Control

As marker physical and genetic chromosomal locations are known, genetic map construction was not toilsome. However, at the recommended p-value of 1e-3 for the population with up to 100 individuals, two linkage groups were obtained for 2B, 3D, 5D and 7D. The issue was overcome by a slight increase of the p-value, 0.01, however, at the same time lessening the minimum threshold for distinct clusters of markers to appear to be linked. Nevertheless, the maximum spacing of 55 cM was seen for 3D chromosome only, while others, 2B, 5D and 7D, remained between ~41-47 cM once two linkage groups were merged (Table 2.1).

Unmerged Linkage Groups					Merged Linkage Groups				
LG	**n.mar**	**length**	**ave.spacing**	**max.spacing**	**LG**	**n.mar**	**length**	**ave.spacing**	**max.spacing**
1A	269	95.9	0.4	9.3	1A	269	95.9	0.4	9.3
1B	188	139.5	0.7	19.4	1B	188	139.5	0.7	19.4
1D	153	89.9	0.6	11.5	1D	153	90.3	0.6	11.5
2A	182	142.6	0.8	26.4	2A	182	142.6	0.8	26.4
2B.1	50	25.1	0.5	10.4	2B	563	197.7	0.4	41.7
2B.2	513	121.3	0.2	14.7	-	-	-	-	-
2D	222	140.2	0.6	35.0	2D	222	140.2	0.6	35.0
3A	211	166.2	0.8	18.1	3A	211	166.2	0.8	18.1
3B	359	208.8	0.6	15.1	3B	359	209.5	0.6	15.1
3D.1	4	15.0	5.0	10.9	3D	24	119.5	5.2	55.1
3D.2	20	44.9	2.4	33.3	-	-	-	-	-
4A	186	118.6	0.6	18.2	4A	186	118.6	0.6	18.2
4B	95	114.8	1.2	23.7	4B	95	114.7	1.2	23.7
4D	20	16.3	0.9	5.3	4D	20	16.3	0.9	5.3
5A	370	244.6	0.7	32.8	5A	370	244.6	0.7	32.8
5B	391	227.3	0.6	20.5	5B	391	227.3	0.6	20.5
5D.1	3	13.7	6.8	12.6	5D	42	103.4	2.5	42.2
5D.2	39	47.6	1.3	20.5	-	-	-	-	-
6A	403	157.7	0.4	11.5	6A	403	157.7	0.4	11.5
6B	232	155.9	0.7	24.2	6B	232	155.9	0.7	24.2
6D	56	90.8	1.7	36.3	6D	56	90.8	1.7	36.3
7A	318	159.8	0.5	14.1	7A	318	159.8	0.5	14.1
7B	271	142.7	0.5	18.6	7B	271	142.7	0.5	18.6
7D.1	13	4.5	0.4	2.4	7D	40	75.8	1.9	47.7
7D.2	27	23.6	0.9	7.9	-	-	-	-	-
overall	**4595**	**2707.1**	**0.6**	**36.3**	**overall**	**4595**	**2908.8**	**0.6**	**55.1**

Table 2.1 Genetic map summary

LG – Linkage Groups, n.mar – number of markers per LG, length – LG length in cM, ave.spacing and max.spacing – average and maximum spacing between marker per LG.

Although some, but not a large quantity of markers showed a slight segregation distortion towards either AA or BB alleles, p-values from chi-square tests of Mendelian segregation, remained significant (not shown). Moreover, almost all markers performed well, except a few showing more than 20% missing values (not shown). Pairwise comparisons of genotypes allowed the identification of four groups of individuals with a proportion of matching alleles above 80% at the markers. Most of the genotypes shared around 45% alleles at the markers, thereby proving none of them are clones (Figure 2.1). Thus, any recombinant inbred line has not been excluded from the analysis.

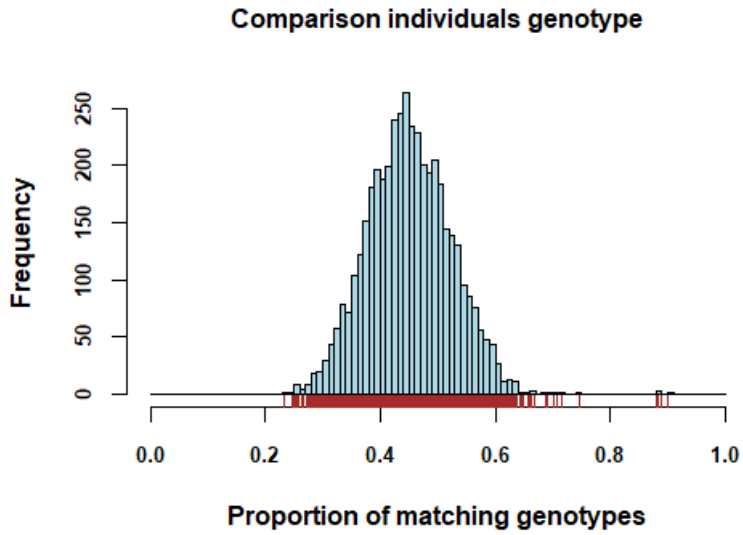


Figure 2.1 Pairwise comparisons of genotypes

The assessment of allelic proportions for RILs provided that 46.7% derived from Pamyati Azieva (red) and 47.5% from Paragon (blue) with remaining 5.8% of being heterozygotes (green) (Figure 2.2). A great deal of missing data was not seen which represented with tiny white dots.

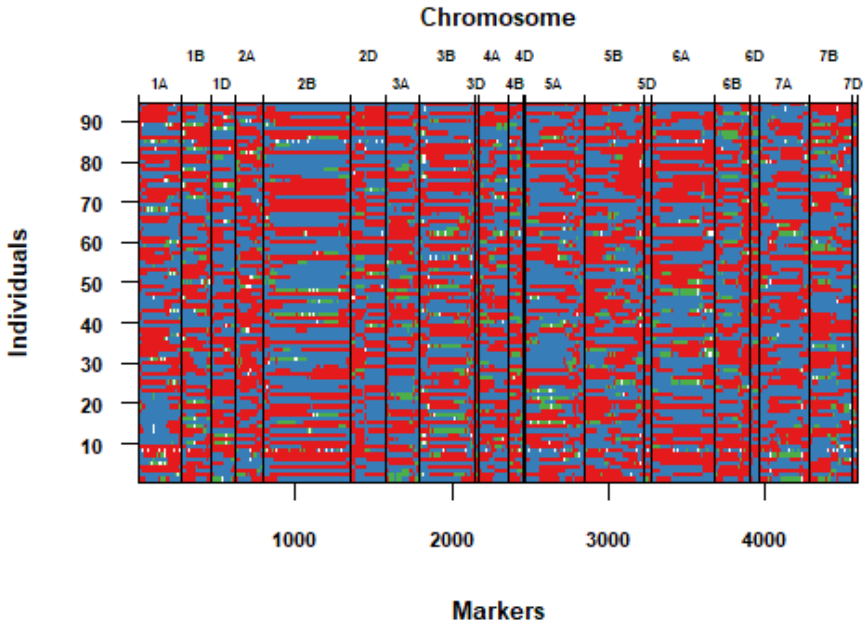


Figure 2.2 Haplotype Map of RILs

Recombination rate or crossovers (COs) of each RIL was assessed by using profileGen function. Based on results, recombination rate of RIL genomes (without distinction of A, B and D genomes) varied between 52-654. However, the recombination rate of eighty-four RILs out of ninety-four was between 52-100 which is consistent with the previous published works on recombination frequency in NAM population (Jordan et al., 2018). The rate of recombination in remaining nine

RILs was 103-244, with only one excessive rate of recombination, 654, which was seen in RIL-8 (pxp8) possibly due to higher level of missing points than the rest (Figure 2.3).

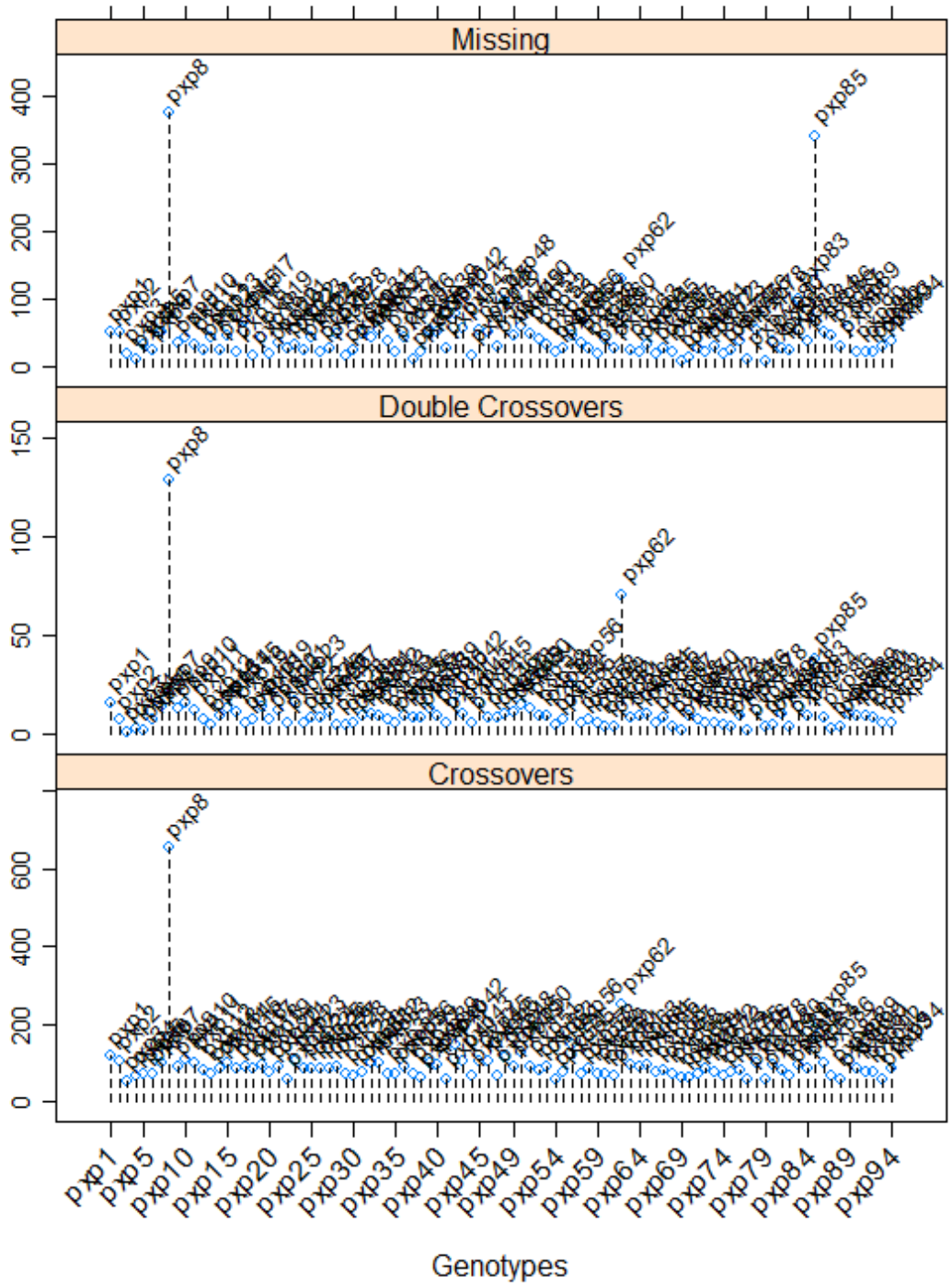


Figure 2.3 Recombination rate in RILs

Genotypes on the x-axis show inbred lines where pxp1 is the RIL-1, pxp5 – RIL5 and so on.

2.3.2 Trait characteristics of RILs

Data on such agronomic traits as Plant Height (PH) and Heading date (HD) were collected in the UK (JIC, Norwich) in controlled and non-controlled environments in 2015 and Plant Height (PH), Peduncle length (PL), Number of spikes per plant (NSP), Main spike length (MSL), Awn length

(AL), Number of spikelets per main spike (NSMS), Number of Kernels per spike (NKS), Weight of grain in the main spike (WGMS), Grain yield per plant (GYP) and Thousand Grain Weight (TGW) of RILs in Kazakhstan (IPBB, Almalybak) in 2016. Of 25 plants in each row in two reps, random 5 plants were involved in collecting the phenotypic data. That is (5 x 2) x 2 = 20 plants of each RIL were observed in total.

2.3.3 The results of analysis of quantitative characters

All measured plant traits were involved in QTL analysis. QTL identification using WinQTLCart 2.5, and R-QTL revealed 27 QTLs in total, residing on all wheat chromosomes, but not in all genomes, 7 of which verified by both statistical software packages. They are 3 QTLs of PH on chromosomes 2B, 5A and 6A, 1 of MSL on 4A and of NKS on 2D, and 2 for WGMS on 2B and 3A respectively. Remaining 20 QTLs linked with PH, PL, MSL, NKS, WGMS and GYP have been detected by both approaches, however, they were located on different wheat chromosomes. QTLs related to AL and NSMS on chromosomes 5A, 5D and 1A, 7D respectively have been identified only by WinQTLCart 2.5. By contrast, TGW QTL on 7A have been detected only by R-QTL. Thus, there were no QTLs found for NSP by both statistical programmes (Table 2.2).

	CHROMOSOMES																				
	1			2			3			4			5			6			7		
GENOMES	A	B	D	A	B	D	A	B	D	A	B	D	A	B	D	A	B	D	A	B	D
TRAITS																					
PH					Q^{WR}								Q^{WR}			Q^{WR}				Q^{W}	
PL				Q^{W}	Q^{W}											Q^{R}	Q^{W}			Q^{R}	
MSL						Q^{W}	Q^{W}			Q^{WR}											
NKS						Q^{WR}		Q^{W}													
WGMS		Q^{W}			Q^{WR}	Q^{W}	Q^{WR}				Q^{R}										
GYP					Q^{W}		Q^{R}					Q^{W}									
TGW																			Q^{R}		
NSP																					
AL													Q^{W}		Q^{W}						
NSMS	Q^{W}																			Q^{W}	

Table 2.2 QTLs detected using two QTL statistical programmes

Most of found QTLs were located on the chromosome 2, while the least on chromosome 1 possessing only two QTLs. Associations of the same genotypic and phenotypic data loaded into two readable statistical software packages resulted in slightly different outcomes. However, PH increasing loci were detected by both statistical programmes. In addition, they were stable in controlled and non-controlled environments. QW and QR - QTLs found by WinQTLCart 2.5 and R-QTL respectively and QWR – by both.

LOD score was not less than 3 for all trait - marker associations. The most important fact is that of 27 QTL, 3 QTL for Plant Height (PH) located on 2B, 5A and 6A wheat chromosomes possessed significant additive effects and were stable in controlled environment (JIC UK) as well as in both

field trials (IPBB KZ, JIC UK). Biometrical analysis showed that two PH increasing alleles on chromosomes 5A and 6A were donated by Pamyati Azieva and PH reducing allele on 2B was inherited from Paragon (Figure 2.4).

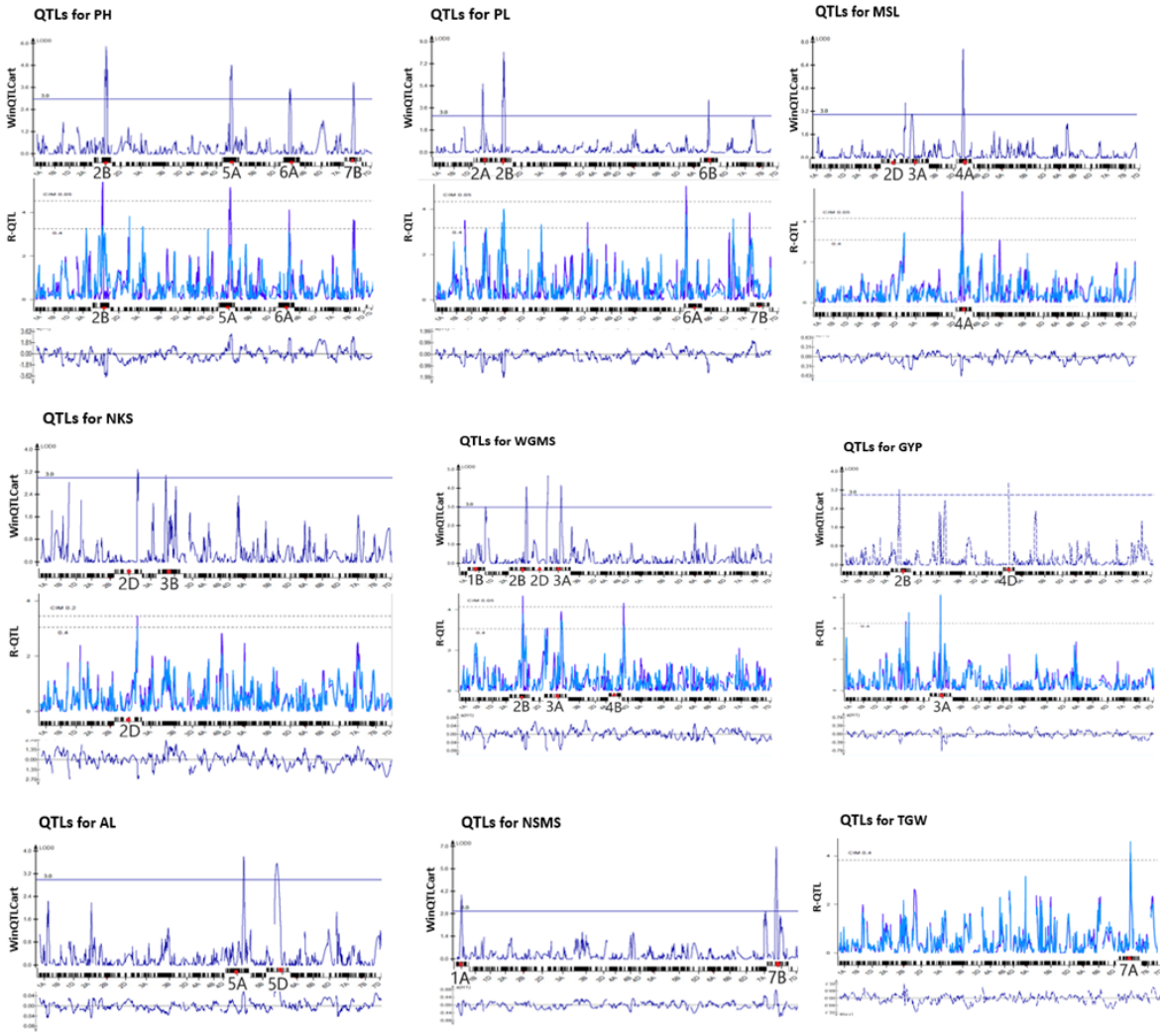


Figure 2.4 The list of QTLs found in RILs grown in Almalybak.

Despite the fact that QTL number varied between two statistical programmes, the graphs showing phenotype and genotype association in both programmes are very similar. For example, even though PH QTL on 7B, QTLs of PL on 2A and 2B, QTL on 2D for MSL, WGMS QTL on 2D and QTL for GYP on 2B have been detected as a QTL by R-QTL statistical programme, genomic overview clearly showed that the peaks with more than 3 LOD score exist in those chromosomes. In contrast, there were quite high signals in WinQTLCart for QTLs for PL on 7B and for GYP on 3A although LOD was slightly smaller than 3. QTLs for AL and NSMS were detected in WinQTLCart and TGW in R only.

Therefore, only QTLs on chromosomes 5A and 6A have been subjected to further genetic studies. More importantly, the responsible sections of the genome for 5A and 6A PH QTLs did not vary anyhow in both controlled and non-controlled environments. Moreover, “closest” and “start and end” DNA markers to QTL peak and flanking regions respectively were almost the same. Therefore, these SNP DNA markers have been converted to KASP DNA markers to facilitate the genotyping process during NIL preparation. Most of the observed quantitative traits were normally

distributed with the only exception being AL which is classified as absent or present. The calculation of the coefficient of determination (R2) between the variables of two reps for PH showed that only 0.26 % of total variation in rep one is not described by the variation in rep two. Therefore, the R2 value of 0.74 explains the fitness of the most variables to the linear model which in turn shows the higher correlation of variables of two replicates (Figure 2.5).

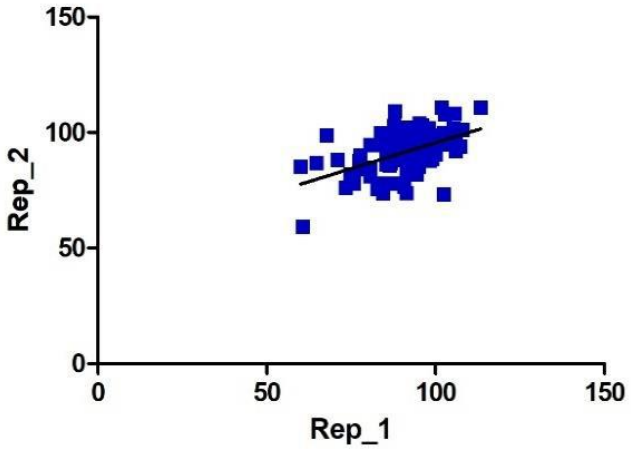


Figure 2.5 The coefficient of determination and p-value for PH

Both were calculated using GraphPad Prism. R square 0.74 and P value < 0.0001

2.3.4 Discussion

Plant height is an important complex plant characteristic which defines plant architecture and could potentially be associated with increased adaptation and thus grain yield (Griffiths et al., 2012). Additionally, like other polygenic traits, PH is controlled by many genes and numerous QTL were reported for PH in wheat in recent years (Cui et al., 2011; Griffiths et al., 2012; Guan et al., 2018). However, no plant height QTL have been reported for Central Asian wheat so far, especially using segregating populations. So, it seems that wheat breeders in Kazakhstan have been selecting for favourable alleles of PH unconsciously and they have never been characterised genetically, thus what specific alleles of PH defines better wheat adaptation in Kazakhstan is still largely unknown. Moreover, we think, PH as a critical trait, influencing plant stature and grain yield, has not been fixed for wheat growing minienvironments of Kazakhstan. This is clearly evident in urgency of studying and searching for new alternative alleles of genes regulating wheat height in Kazakh bread wheat as Kazakhstan is one of the main wheat exporters of the globe. Here, we found two plant height related QTL with enormous additive effects in RILs and validated stand-alone and joined-up effects of both through generating single and double NILs (Near Isogenic Lines, sNILs and dNILs respectively) which will be discussed in chapter 3.

3. Chapter 3: Reshaping the height of the UK spring wheat Paragon for testing its performance in Kazakhstan

3.1 Introduction

The evaluation of relatively pure effect of QTL is manly achieved through the development of Near Isogenic Lines which in turn are generated through backcrossing of the QTL of interest to one of the recurrent parents depending on experimental design. As a result, several rounds of backcrosses allow scientists to have two contrasting alleles of genetic markers flanking that quantitative locus in the common genetic background (Zhou et al., 2005). Thus, eliminating much of the heterogeneous background mutations makes it possible to verify genomic region contributing to the trait of interest, screen for the functional annotation of its gene content, conduct middle resolution mapping of the QTL and tag molecular markers co-segregating with the loci in many plant species including of wheat (Habib et al., 2018; Halder et al., 2021). For these reasons, NILs have widely been used as an essential biological tool to validate QTL contribution for the trait of interest and its pleiotropic effects on other complex traits (Fletcher et al., 2013). For example, NIL sets were successfully used to study QTL controlling fruit quality in melon (Eduardo et al., 2005), eating quality of rice (Cho et al., 2013), salt tolerance mechanisms in barley (Zhu et al., 2020) and the pre-harvest sprouting resistance in bread wheat (Wang et al., 2018). Such examples are plentiful. Importantly, many *Rht* genes and their pleiotropic effects were also studied based on NILs (Appleford et al., 1991; Wojciechowski et al., 2009; Alghabari et al., 2014). Accordingly, this chapter of the thesis aims to provide full information about the development of near isogenic lines for two plant height increasing QTL detected in the Pam.Az (Pamyati Azieva) x Par (Paragon) segregating population. Both PH QTL were donated by Pamyati Azieva and thus NILs were generated in the Paragon background. The effect of each PH QTL allele on height and many yield related traits was tested in a range of environments in the UK and Kazakhstan across 3 years.

3.2 Materials and methods

3.2.1 The delivery of Pamyati Azieva height increasing 5A and 6A QTL alleles to the UK wheat Paragon for NIL development

Once RILs were multiplied, phenotyped, genotyped and QTLs associated with valuable plant traits identified, NILs have been developed so as to trace two PH increasing QTLs of Pamyati Azieva in the Paragon background (Figure 3.1). To implement so, two recombinant lines, PamxPar-5 and PamxPar-16 carrying 6A and 5A PH QTLs respectively, have been chosen as they were heterozygous for the QTL region. However, PamxPar-5 inbred line for 6A QTL was fixed to Paragon at the 5A QTL region and PamxPar-16 for 5A QTL does the same to 6A PH QTL. Two rounds of backcrossing, self-fertilising and collecting of heterozygotes for three QTL flanking DNA markers, allowed us to produce BC2F2 seeds finally for both QTL. During each round of backcrossing, 12 triple heterozygous plants for QTL flanking markers have been selected to produce next generation.

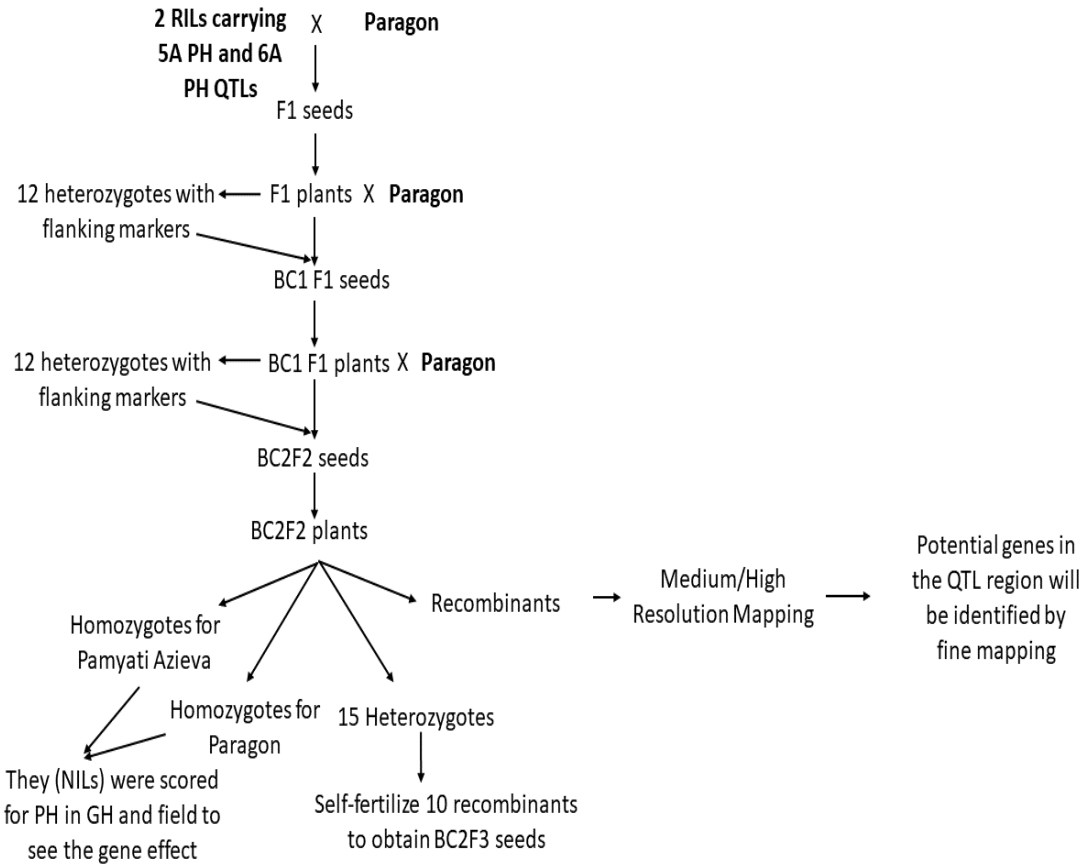


Figure 3.1 NIL&Recombinant development for 5A and 6A PH QTL

Fifteen heterozygous plants from both lines have been chosen in order to self-fertilize to have more hybrids with different recombination combination. Recombinants will be used to fine map the genes of interest.

Produced 94 BC2F2 seeds and two original parents were then genotyped with flanking markers and germinated. Prior to germination, the seeds first were sown on Petri dishes and placed in CER at 6˚C for two days to break the seed dormancy, then moved to CER with 20˚C for further 2-3 days of germination. After that, seedlings transplanted to 96 well tray with "Peat and Sand" soil type and left at CER at 20˚C for two weeks. During this period, DNA was extracted from all 192 plants (94 lines + 2 parents x 2) and DNA samples were genotyped in order to identify heterozygotes, recombinants and homozygotes for both parents. Homozygous lines, that are NILs, carrying Pamyati Azieva and Paragon alleles for both QTLs, on wheat 5A and 6A chromosomes, were grown in 1L pots filled with Cereal Mix soil in a glasshouse (GH) of the JIC. Each NIL was bagged in order to eliminate the possible outcrossing and heights of NILs have been measured (Figure 3.2). Two Sample t-test in R was used to evaluate the significance between Paragon NILs carrying and lacking height increasing allele of Pamyati Azieva.

3.2.2 Field experiments in the UK

3.2.2.1 Seed preparation

Initially, the seeds of NILs were multiplied in the greenhouse at the JIC, UK. About 180 seeds were sown from each original NILs in 2018/19. These seeds were used in 2019/20 and the seeds from 2019/20 were used in the 2020/21 field experiments.

3.2.2.2 Plot design

Each original NILs was sown separately in October 2018 in the field trials of JIC (Norwich) as a $1m^2$ plot experimental design to multiply the seed numbers as well as to measure plant height.

In the 2019/20 growing season, NILs were sown also during winter drilling. The plot dimensions were 1.3m and 5.5m in width and length respectively.

In 2020/21 drilling season, each original NIL was drilled in $1m^2$ plots in randomised manner. Thus, NIL5A(+), NIL5A(-), NIL6A(+), NIL6A(-), dNILs (double NILs or 5A(+)-6A(+)) and nNILs (nonNILs or 5A(-)-6A(-)) were replicated 12, 19, 6, 25, 20 and 10 times. Beside the NIL5A(+) and 6A(+), we have also stacked these two height increasing alleles in Paragon and developed double NILs. More information about the development of double and nonNILs is provided in the coming "**3.3.1.8 Stacking of two height increasing alleles in UK wheat**" section.

3.2.2.3 Data collection

In 2018/19 only the plant height and heading date were scored. For that, plant heights of six random plants within $1m^2$ plots were scored during the early maturation when stem elongation stopped growth.

As in the previous growing season, only PH was scored in 2019/20. For that, we took overall plot height. Moreover, plants from the 50cm distance within the plot (representing the plot), were pulled out carefully and their heights were measured.

In 2020/21, the experiment purpose was to score PH and YSM (Yield per Squared Meter). For PH, height of each plot was measured carefully as each NIL sets were multiplied several times. The combine used to harvest and get YSM was a ZURN 150. It is fitted with an onboard weighing system which is especially for the small $1m^2$ plots. Therefore, we only record the weight of the grain from the plot and not the moisture as the volume of seed coming from the small plots is not sufficient to gain an accurate moisture reading.

3.2.3 Field experiments in KZ

3.2.3.1 The experimental design of the year 2019

3.2.3.1.1 Seed preparation

Compared to the field experiments in the UK, we conducted a large scale field testing in Kazakhstan over the years as the main aim of the thesis was to assess and evaluate possible positive and negative pleiotropic effects of height increasing/decreasing alleles of QTL on adaptation and yield components. For that, the seeds of NILs were multiplied in the greenhouse at the JIC, UK. Multiplied seeds of several original lines to develop NIL were pooled and sent to Kazakhstan in 2018.

3.2.3.1.2 Plot design

The pooled seeds of NIL sets (NIL5A(-/+) and NIL6A (-/+)) and Paragon were sown in Almalybak, Almaty region, south-eastern Kazakhstan, in field trials of the IPBB (Institute of Plant Biology and Biotechnology) during spring drilling of 2019. The experimental design included 6m randomised plots with 2 replicates per line except NIL6A(-) which had limited number of seeds. Fifty seeds of NILs were sown in each 1m row within 6m long plots and the rows were spaced 20cm apart. Thus, 1500 seeds (50 seeds x 30 rows = 1500 seeds) were drilled per each 6m plot.

3.2.3.1.3 Data collection

Random three plants per plot were scored for effective tiller number (ETN), number of kernels per spike (NKS), grain weight per main spike (GWMS), and number of kernels per plant (NKP). This allowed us to have six data points (3 plants x 2 replications) per NIL and Paragon. To assess the diversity in plant height (PH), lengths of internodes – from top 1st to down 5th (Int1 -5), spike length (SL) and number of spikelets per main spike (NSMS) between the NILs and to compare it with Paragon, a wide range of sample size (45 plants per replication, thus 90 in total) was considered. However, traits such as thousand kernel weight (TGW), number of spikes per row (NSR), number of plants per row (NPR), yield per squared meter (YSM), YP (yield per plot), total above-ground biomass (TAGB), number of plants per unit area (NP_1m2) and Normalized Difference Vegetation Index (NDVI) were scored per replication/plot. Thus, each NIL had two data points per trait. YP was calculated through multiplying the YSM by ten.

3.2.3.2 The experimental design of the year 2020/21

3.2.3.2.1 Seed preparation

The seeds were multiplied in the experiment conducted in field trials of JIC (see 3.2.3.1). This allowed us to send seeds from single NIL (not pooled from several NILs) to KZ.

Thus, NIL sets were not sown only in the southeast, also in the central (Arkalyk) and northern (Petropavl) regions of Kazakhstan, the main wheat growing parts of the country, in larger plots in 2020 and 2021 drilling seasons.

3.2.3.2.2 Plot design

The 2020 growing season

In 2020 drilling season, NILs, NIL parents and local check variety (Kazakhstanskaya-10 or Kaz10) were sown in $10m^2$ randomized plots with two replications in Almalybak. In Petropavl, the field experiment with same setting was conducted, but without local standard variety. Experimental design in the Central Kazakhstan was different from the rest. Particularly, NILs were sown in $5m^2$ plots with four replications, two with irrigated and the other two non-irrigated treatments. Unfortunately, the experiment in Central Kazakhstan was not completed due to COVID pandemic restrictions.

The 2021 growing season

In 2021 drilling season plots were larger than that of 2020. Although experimental design was the same as in the 2020, we replicated each NIL four times in $10m^2$ randomised plots. However, the Central Kazakhstan was excluded from the experiment. Therefore, NILs were grown in Almalybak and Petropavl.

3.2.3.2.3 Data collection

Data from the 2020-21 experiments were collected using the standardized phenotyping protocol which was prepared for partners in Kazakhstan to measure height and yield components, namely plant height (PH), lengths of internodes – from top 1st to down 5th (Int1 - 5), spike length (SL), effective tiller number (ETN), number of spikelets per main spike (NSMS), number of kernels per spike (NKS), number of kernels per plant (NKP), main spike weight (MSW), grain weight per main spike (GWMS), grain weight per plant (GWP), thousand kernel weight (TGW), yield per squared meter (YSM), yield per plot ($10m^2$) (YP), number of plants per unit area (NP_$1m^2$), number of spikes per unit area (NS_$1m^2$), seedling number per unit area (SN_$1m^2$) and main spike weight (MSW). The seed number per squared meter (SSM) came from the mean of NKP for each sample multiplied by NP_$1m^2$. Seed Width (SWI) and seed length (SLE) of ~300 grains of each line from two replications were measured using MARViN seed analyser. These data on biomass were used to calculate total above-ground biomass (TAGB) at anthesis and maturity. Harvest index (HI) was calculated from biomass at maturity .

Notably, NRP, YR and NDVI were extra plant characteristics scored in Almalybak 2019 only. However, in that year YP, NS_1m2, SN_1m2, SSM, HI, SWI, SL, GWMS and GWP were not measured in Almalybak. In the first year of field experiment in the North that is Petropavl 2020,

GWMS, GWP and MSW were not scored. Data on SWI and SL exist for Almalybak 2020 and Petropavl 2020 only (Table 3.1).

In the field experiments conducted in 2020-21, wheat developmental stages – germination, tillering, stem elongation, flag leaf emergence, booting, heading, flowering and maturity, and lodging were scored in both environments. However, stem elongation and flag leaf emergence were not scored in 2020 in Almalybak. Information on resistance of NILs and parents to diseases such as *Septoria*, leaf and stripe/yellow rusts is available for Almalybak only in 2020, but present for both in 2021. Seeds were drilled in the North at the depth of 6-8cm from the topsoil.

List of traits	Location	Almalybak			Petropavl	
	ABBR\Years	2019	2020	2021	2020	2021
Plant height	PH	+	+	+	+	+
Internode1	int1	+	+	+	+	+
Internode2	int2	+	+	+	+	+
Internode3	int3	+	+	+	+	+
Internode4	int4	+	+	+	+	+
Internode5	int5	+	+	+	+	+
Spike Length	SL	+	+	+	+	+
Effective Tiller Number	ETN	+	+	+	+	+
Number of spikelets per Main Spike	NSMS	+	+	+	+	+
Number of Kernels per spike	NKS	+	+	+	+	+
Number of Kernels per Plant	NKP	+	+	+	+	+
Main Spike Weight	MSW	+	+	+	-	+
Grain Weight per Main Spike	GWMS	-	+	+	-	+
Grain Weight per plant	GWP	-	+	+	-	+
Thousand Grain Weight	TGW	+	+	+	+	+
Yield per Square Meter	YSM	+	+	+	+	+
Yield per Plot (gramm)	YP	-	+	+	+	+
Number of Plants per $1m^2$	NP_$1m^2$	+	+	+	+	+
Number of Spikes per $1m^2$	NS_$1m^2$	-	+	+	+	+
Seedling number per $1m^2$	SN_$1m^2$	-	+	+	+	+
Seed per Square Meter	SSM	-	+	+	+	+
Harvest Index	HI	-	+	+	+	+
Seed width	SWI	-	+	-	+	-
Seed Length	SLE	-	+	-	+	-
Number of Plants per Row	NPR	+	-	-	-	-
Yield per Row	YR	+	-	-	-	-
Total Above Ground Biomass	TAGB	+	+	+	+	+
NDVI	NDVI	+	-	-	-	-

Table 3.1 The list of scored traits on NILs grown in two locations in Kazakhstan for three (Almalybak) and two (Petropavl) years.

ABBR = Abbreviation

3.2.3.2.4 Data treatment and statistical analysis

Data treatments and Statistical analyses were mostly done in the statistical environment R, but some minor arrangements and calculations in Microsoft Excel. Outliers for each measured trait were removed prior to conducting any calculations including ANOVA. Tukey's method which considers the variables with coefficients outside of Q1 and Q3 (Q=quartile) as outliers was the base approach for their detection. Depending on a sample size, either ANOVA (for sub-sample) or one-sample t-test (for whole-sample) were conducted to assess a significance level of measured traits.

3.2.4 Statistical analysis

ANOVA or LMM or T-Test was used depending on data structure to model the data in order to obtain overall significance for each measured trait. Then, the same model was involved in further statistical calculations to estimate the scored trait parameters such as mean, estimated differences between the means (EST) of group levels, standard error of the mean (SEM), variance, standard deviations (SD), upper and lower confidence intervals (UCI and LCI respectively) and significance values (p-value) for NILs and NIL comparisons.

Homogeneity of variances of group levels was evaluated using Bartlett's. This was particularly done to choose appropriate method, among a variety of approaches such as Tukey's honestly significant difference (HSD), pairwise t - tests with pooled SD and Games-Howell's HSD, to conduct post-hoc statistical test between group levels. Among these, Games-Howell was chosen as it was reported to provide reasonable outcome for pairwise comparisons of treatments or groups with unequal variances (Shingala, 2015). Games-Howell's HSD is similar to Tukey's test, albeit it does not assume normality, equal variances or sample sizes (Ruxton and Beauchamp, 2008). However, multiple comparisons of traits possessing two data points were analysed based on Tukey's HSD as games_howell_test fuction in R to conduct Games-Howell's HSD does not handle smaller sample size less than three data points. Also, Tukey's HSD is more powerful than Bonferroni when testing large numbers of means (Field, 2009).

Association between scored plant characteristics was assessed based on Pearson's correlation coefficients. Plots were produced using available tools such as ggplot2 in R.

3.3 Results

3.3.1 Field experiment results in the UK and KZ

3.3.1.1 The location of the experiment in Kazakhstan and its geographical and meteorological characteristics

3.3.1.1.1 Almalybak 2020

The field experimentation on NILs was carried out on the stationary field trial of the "Kazakh Research Institute of Agriculture and Plant Growing" LLP (Limited Liability Partnership). The

centre is located in the town of Almalybak, near Almaty, in the Southeast part of the country (48.0 N, 77.0 E, 740m above the sea level). From now onwards, Almalybak will be referred as Alm (short for Almalybak with years of field trials, e.g., Alm 2019, Alm 2020 and Alm 2021 for Almalybak 2019, Almalybak 2020 and Almalybak 2021 respectively) throughout the thesis. The number of days with temperatures below 0^0C ranges from 125-130. The average Long Term Annual Average (LTAA) precipitation is 360-400 mm. The soil cover of the experimental plot is light chestnut soil.

In 2020, the amount of precipitation was 3 and 8 times higher in April and June compared to LTAA (Table 3.2) with 3 times an increase in rainfall for the season. However, monthly average temperature was higher comparing to LTAA for seasonal months, except June where the large rainfall increase was observed.

Months	Precipitation, mm			Temperature, ^{o}C		
	Monthly average	LTAA	DLTAA	Monthly average	LTAA	DLTAA
March	52.7	48.8	3.9	6.4	0.7	5.7
April	146.7	56.5	90.2	14.2	10.4	3.8
May	73.5	61.6	11.9	18.7	16.4	2.3
June	426.3	53.9	372.4	16.5	21.2	-4.7
July	38.1	26.6	11.5	24.4	24.1	0.3
Season	737.3	247.4	489.9	80.2	72.8	7.4

Table 3.2 The main meteorological indicators of the growing season 2020, Almalybak

LTAA = Long-Time Annual Average, Deviation from the LTAA = DLTAA.

3.3.1.1.2 Almalybak 2021

The growing season 2021 was unfavourable for agricultural crops. The monthly average temperature was higher comparing to LTAA for all seasonal months. Although the average amount of precipitation was higher than the LTAA in March and May and was the same in April, June experienced severe drought in which the month had only 40% of precipitation comparing to LTAA. July had uneven rainfall throughout the month where no rain fell in last ten days (LTD) the month. Moreover, this drought period was accompanied by extreme heat wherein daily maximum and minimum temperature reached to almost 40^0C and 30^0C respectively. The monthly precipitation accounted for 22.8 mm which was 3.8 mm less than the norm. In general, drought and heat in June and July reduced the yield dramatically as plants are in an important developmental stage such as grain filling in these two months.

Months	Precipitation, mm			Temperature, ^{o}C		
	Monthly average	LTAA	DLTAA	Monthly average	LTAA	DLTAA
March	117.9	48.8	69.1	4.1	0.7	3.4
April	56.3	56.5	-0.2	12.4	10.4	2.0
May	81.6	61.6	20.0	19.4	16.4	3.0
June	20.9	53.9	-33.0	23.1	21.2	1.9
July	22.8	26.6	-3.8	26.9	24.1	2.8
Season	299.5	247.4	52.1	85.9	72.8	13.1

Table 3.3 The main meteorological indicators of the growing season 2021, Almalybak

LTAA = Long-Time Annual Average, Deviation from the LTAA = DLTAA.

3.3.1.1.3 Petropavl 2020

The experiment was conducted in 2021 at the "North-Kazakhstan Agricultural Experimental Station" LLP. The breeding station is located in the steppe zone of the North Kazakhstan region. From now onwards, Petropavl will be referred as Pet (short for Petropavl with years of field trials, e.g., Pet 2020 and Pet 2021 for Petropavl 2020 and Petropavl 2021 respectively) throughout the thesis. The climate of the zone is arid, with moderate heat. The amount of precipitation is 240-330 mm. The vegetation period varies in the range of 136-137 days, with HTC (hydrothermal coefficient) of 0.8-0.7. The landscape is characterized by the absence of forests and flat with large number of shallow hollows occupied by lakes. The soil of the experimental plot is an ordinary carbonate heavy loamy chernozem with a neutral and slightly alkaline reaction, the pH of the water extract is 7.8-8.1. The content of humus is 4.5 - 5.0%, nitrate nitrogen determination by the disulfophenol method according to Grandval-Lage in the soil layer 0-40 cm 16.6 mg / kg of soil, and mobile phosphorus according to the method of Machigin B.P. in a layer of 0-20 cm 10.0 mg/kg of soil and potassium according to the method of Machigin B.P. 630 mg/kg soil.

According to the data of the North-Kazakhstan Breeding Station (Shagalaly meteorological station), May 2020 in the north of the region turned out to be abnormally hot and windy. The maximum temperature reached 33.5-35.6^{0}C, the GDD (Growing Degree-Days) at the end of the month exceeded the LTAA by 276^{0}C. The hottest was the second ten days (STD) of the month, where the average daily temperature was +20^{0}C (norm +13^{0}C), exceeding the norm by + 7^{0}C. At the same time, the amount of precipitation recorded was 28.1 mm, which is at the level of the norm. Despite the severe drought experienced in this growing season, normal germination rate was detected. The duration of the sowing – germinating period was 7 days.

The first summer month, June, of the year 2020 is marked by very contrasting meteorological conditions: first ten days (FTD) and STD were extremely dry - 1.1 mm and 1.8 mm (Table 3.4) and respectively, these are only 8% and 16% of the norm. However, the situation with severe drought

conditions was significantly improved by precipitation in LTD of July (33 mm). In general, 35.9 mm or 82% of the climatic norm fell during the month. The accumulation of heat in the current year is ahead of the norm - at the end of June, the GDD was 1280^{0}C, with an average LTAA of 1069^{0}C (excess heat amounted to 211^{0}C).

Months	Precipitation, mm						Average monthly temperature, ^{0}C					
	TD of the month			Monthly	LTAA	DLTAA	TD of the month			Monthly	LTAA	DLTAA
	FTD	STD	LTD				FTD	STD	LTD			
May	-	-	-	28.1	28	0.1	-	-	-	20	12.7	7.3
June	1.1	1.8	33	35.9	44	-8.1	17.6	17	14.7	16.4	18.5	-2.1
July	66.6	0.2	8.8	75.6	71	4.6	20.7	24.6	19.1	21.4	20	1.4
August	2.6	16.7	2.3	21.6	50	-28.4	25	16.2	18.3	19.8	17.2	2.6
Season				161.2	193	-31.8				19.4	17.1	2.3

Table 3.4 The main meteorological indicators of the growing season 2020, Petropavl

TD = Ten Days (of the month), FTD = First Ten Days, STD = Second Ten Days, LTD = Last Ten Days, LTAA = Long-Time Annual Average, Deviation from the LTAA = DLTAA

July 2020 experienced extremely uneven rainfall throughout the month. In general, the amount was 75.6 mm, or 106% of the LTAA. In fact, almost the entire amount of moisture (66.6 mm) fell out in 2 days at the beginning of FTD. In the following STD and LTD, the amount of precipitation was extremely low - 0.2 mm and 8.8 mm, or 1% and 34% of the norm respectively. The average daily temperature during the month was 1.4^{0}C above the norm, the maximum reached 35.5^{0}C.

August of the year was also hot and dry. The average daily temperature was 19.8^{0}C, which is 2.6^{0}C above the norm. In addition to the faced elevated temperature, the amount of precipitation was almost 2 times lower than the climatic norm - 21.6 mm, while the norm was 50.0 mm, which amounted to 43%. The prevailing weather conditions significantly accelerated the onset of wax ripeness of wheat. The beginning of the phase was fixed at the beginning of the first decade of August, i.e., the growing season of wheat in comparison with long-term observations was reduced by ~10 days.

In total, 133.1 mm of precipitation fell during the summer of 2020, which, with an average annual rate of 165.0 mm, amounted to 81% of the norm. The average temperature for the summer months was within 19.2^{0}C, which is 0.6^{0}C warmer than the norm.

Thus, the agro-meteorological conditions of 2020 for the growth and development of agricultural crops are characterized as arid, with early summer and August drought and a pronounced July maximum of precipitation.

3.3.1.1.4 Petropavl 2021

According to the data of the North-Kazakhstan Breeding Station (Shagalaly meteorological station), May 2021 in the north of the region turned out to be abnormally hot. The average temperature for the month was 18.1^{0}C, which is 5.4^{0}C higher than the norm (Table 3.5). If FTD, STD and LTD are considered, then the excess of temperature identified was at +2.5^{0}C, +5.2^{0}C, +8.2^{0}C, respectively. The maximum temperature reached 34.5-37.4^{0}C. At the end of the month, GDD (growing degree days or the sum of positive temperatures) amounted to 638^{0}C, which exceeds the LTAA by 126^{0}C. During the month, frequent winds were observed, drying up the topsoil. The amount of precipitation for the month was only 10.1 mm, or 36% of the norm. Despite the severe drought experienced in this growing season, normal germination rate was detected. The duration of the period of sowing – germinating was 9-10 days.

Months	Precipitation, mm						Average monthly temperature, ^{0}C					
	TD of the month			Monthly	LTAA	DLTAA	TD of the month			Monthly	LTAA	DLTAA
	FTD	STD	LTD				FTD	STD	LTD			
April	1.7	0.2	21.1	23	24	-1	-0.1	4.2	7.5	3.9	4.3	-0.4
May	6.1	1.6	2.4	10.1	28	-17.9	13	18.2	22.8	18.1	12.7	5.4
June	8.8	10	3.2	22	44	-22	14.5	18.3	19	17.2	18.5	-1.3
July	0.4	64.6	4.8	69.8	71	-1.2	23.7	17.8	20.8	20.8	20	0.8
August	17.2	0.9	11	29.1	50	-20.9	21.3	19.3	20.6	20.4	17.2	3.2
Season				120.9	165	-44.1				19.5	18.6	0.9

Table 3.5 The main meteorological indicators of the growing season 2021, Petropavl

TD = Ten Days (of the month), FTD = First Ten Days, STD = Second Ten Days, LTD = Last Ten Days, LTAA = Long-Time Annual Average, Deviation from the LTAA = DLTAA

In June, there was also a significant deficit of precipitation. In total, 22.0 mm fell during the month (in FTD, STD and LTD of the month 8.8 mm, 10.0 mm, 3.2 mm, respectively). This accounts for only 50% of the norm (the norm 44.0 mm). At the same time, the average daily temperature was 17.2^{0}C which is below the norm by 1.3^{0}C. The GDD at the end of the month reached 1154^{0}C, while the norm was 1069^{0}C, exceeding the LTAA by 85^{0}C. High-quality and timely seedlings of wheat were obtained mainly due to good autumn-winter moisture. The beginning of the growing season of the year 2021 is characterised by harsh dry conditions, i.e., unfavourable for the cultivation of major crops, since the lack of precipitation in critical phases of agricultural development. crops and high temperatures for a long time had a negative impact on the initial period of growth and development of plants.

July 2021 was marked by extremely uneven rainfall throughout the month. The precipitation was 69.8 mm, or 98% of the LTAA. In fact, almost the entire amount of moisture (64.6 mm) fell in the second decade of the month. Thus, the expected full amount of July precipitation manifested,

which had a positive effect on the development of agricultural crops. However, in the first and third decades, the amount of precipitation was extremely low - 0.4 mm and 4.8 mm, or 2% and 18% of the norm. cultures. The average daily temperature during the month was 20.8^{0}C, which was normal.

August 2021 was hot and dry. The average daily temperature was 20.4^{0}C, which is 3.2^{0}C above the norm. At the same time with the elevated temperatures, the amount of precipitation was significantly below the climatic norm - 29.1 mm, at a norm of 50.0 mm, which amounted to 58%. The prevailing weather conditions accelerated the ripening of wheat. The beginning of the phase was recorded in the second decade of August. The GDD for the period reached 2429^{0}C, exceeding the norm by 245^{0}C.

In total, 120.9 mm of precipitation fell during the summer, which, at a rate of 165.0 mm, amounted to 73%. The average temperature for the summer months was within 19.5^{0}C, which is 0.9^{0}C warmer than the norm.

Thus, the agro-meteorological conditions of 2021 are characterised as unfavourable for the growth and development of agricultural crops, extremely dry, with an early summer and August drought and a pronounced maximum of precipitation in July.

More accurately, the moisture conditions are reflected by HTC, because in addition to precipitation, the temperature regime needs to be taken into account (the ratio of the amount of precipitation and their evaporation to the sum of positive temperatures, reduced by 10 times). In May and June of the current year, a catastrophically severe drought of the HTC was observed - 0.18 and 0.43, respectively (Table 3.6). In July, the HTC is 1.08 at the LTAA, i.e., precipitation availability was satisfactory. In August, the HTC - 0.46, i.e., conditions correspond to a catastrophically severe drought.

Months	Growing season 2021		LTAA	
	HTC	Levels of drought	HTC	Levels of drought
May	0.18	catastrophically severe drought	0.71	mild drought
June	0.43	catastrophically severe drought	0.79	mild drought
July	1.08	satisfactory	1.14	satisfactory
August	0.46	catastrophically severe drought	0.94	satisfactory

Table 3.6 HTC (hydrothermal coefficient) for the growing season 2021, Petropavl

Agronomy of the trials and Growing conditions

Spring work consisted in closing/saving the moisture with BIG-3 harrows, intermediate cultivation with the simultaneous application of Ammophos (nitrogen-phosphorous fertilizer) (50 kg/ha), pre-sowing cultivation SZS-2.1 to a depth of 6-8 cm, followed by rolling with ringed rollers. The drilling of genotypes was carried out at the optimal time for the crop - May 15, with a small-sized selective drilling machine SSN-7 in pairs, in four repetitions, the plot area was 10 m^2. In the tillering phase, herbicide treatment of crops was carried out with an AVAGRO sprayer with a tank mixture Galantny 12 g/ha, Efir extra 0.4 l/ha, Centurion 0.2 l/ha. During the growing season, inter-rows were cleaned from the weeds and plots were looked after.

Phenological observations were carried out according to the phases of wheat development: germination, tillering, stem elongation, flag leaf emergence, booting, heading, flowering and ripening (based on the protocol provided).

Genotypes were screened for resistance to leaf (*Puccinia triticina*) and stem (*Puccinia graminis*) rusts and *Septoria*. The assessment for the resistance to brown rust and *Septoria* was carried out twice: at the end of the grain filling and beginning of the wax deposition. For stem rust, the last screen was carried out before the harvesting. The resistance screen was conducted based on Duveiller et al. (2018).

3.3.1.2 The plant heights (PH) of NILs for 5A and 6A QTL

3.3.1.2.1 Pamyati Azieva's height increasing 5A and 6A QTL alleles made the UK wheat Paragon tall in glasshouse

Genotyping results of the 94 BC2F2 seeds with QTL flanking markers controlling PH allowed us to identify 9 (16 initially) lines homozygous for Pamyati Azieva alleles, 12 (16 initially) lines homozygous for Paragon alleles and, 15 heterozygous lines (these lines are for obtaining more heterozygous recombinants in case of running out homozygous recombinants to zoom in the QTL region) out of initial 94 samples for PH related QTL on wheat 5A chromosome. Likewise, 16 (17 initially) lines homozygous for Pamyati Azieva, 7 lines homozygous for Paragon and, 15 heterozygotes lines were identified out of initial 94 samples for PH related QTL on wheat chromosome 6A.

Conducted Two Sample t-test using R showed that the height means of NILs, designated as NIL5A(+), NIL5A(-) and NIL6A(+), NIL6A(-), have been significantly different (Figure 3.2 and Figure 3.3). Here, NIL5A(+) and NIL5A(-) are near isogenic lines, with and without the PH increasing allele of Pamyati Azieva respectively, in the genetic background of Paragon. Likewise, NIL6A(+) and NIL6A(-) are near isogenic lines with and without the PH increasing Pamyati Azieva's allele respectively in the genetic background of Paragon.

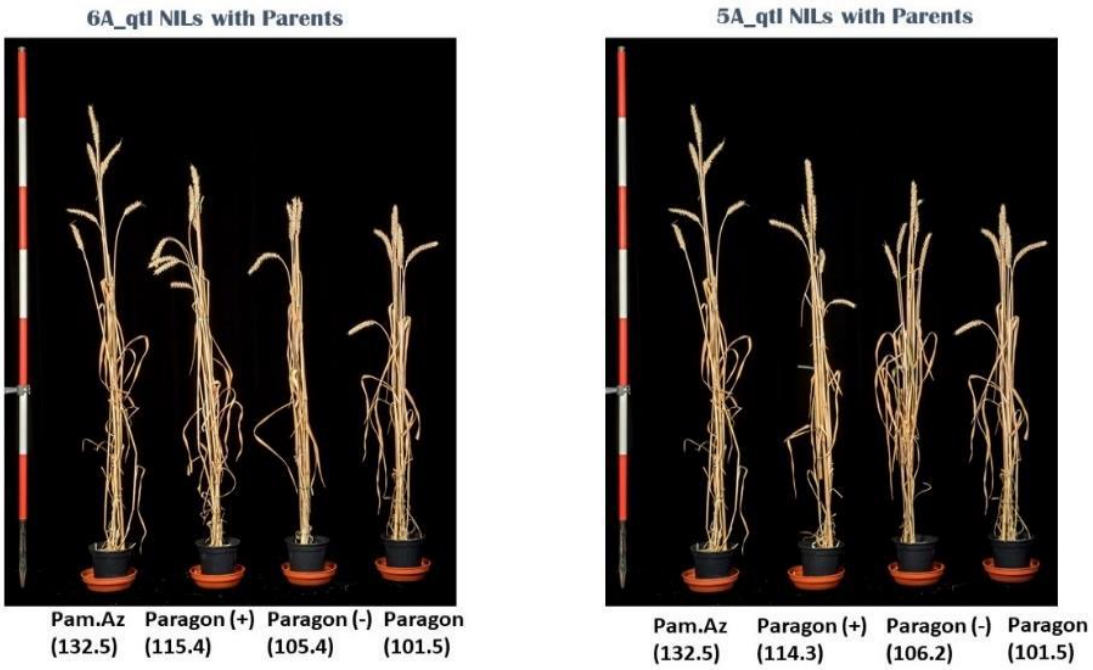
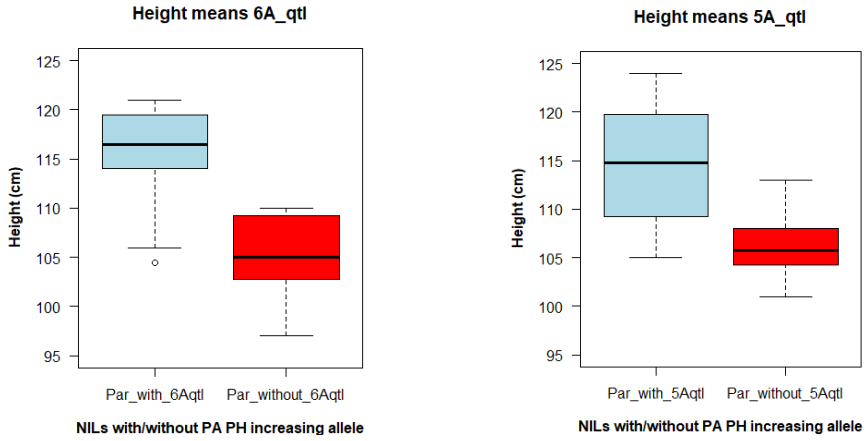


Figure 3.2 6A (left) and 5A (right) Near Isogenic Lines (in the middle) with parents Pam.Az = Pamyati Azieva and Paragon on the right and left of each figure respectively.

Plant height means from glasshouse have shown statistical significance. A QTL additive effects were 10cm and 9cm for 6A and 5A PH QTL respectively.



a) b)

Figure 3.3 Paragon NILs with and without Pam.Az PH increasing alleles on 6A (a) and 5A (b) chromosomes.

Figure 3.3 a) represents plant height of 6A NILs with parents in centimetres and Y-axis shows NILs with/without plant height increasing allele. P value of 8.038e-05 indicates that true difference in means is not equal to 0. During a two-sample t-Test the group variances were assumed to be not equal. PA = Pamyati Azieva, PH = Plant Height. The means of NILs with and without PH increasing allele are 114.2 cm and 106.1 cm respectively. QTL additive effect was ~10 cm.

Figure 3.3 b) represents plant height of 5A NILs with parents in centimetres and Y-axis shows NILs with/without plant height increasing allele. P value of 0.0001257 indicates that true difference in means is not equal to 0. During a two-sample t-Test the group variances were assumed to be not equal. PA = Pamyati Azieva, PH = Plant Height. The means of NILs with and without PH increasing allele are 115.4 cm and 105.1 cm respectively. QTL additive effect was ~10 cm.

3.3.1.2.2 Testing the tall and short NILs of the 5A and 6A QTL in the UK

3.3.1.2.2.1 Norwich 2018/19

Fifteen original NIL5A(+) and fourteen NIL5A(-), and seventeen NIL6A(+) and seven NIL6A(-) were grown in 1m² plots in the fields of John Innes Centre. The main purpose of an experiment was to multiply the seeds and measure the plant height (PH) of NILs. However, we also scored heading data to evaluate the possible pleiotropic effect of PH QTL on flowering date.

As a result of ANOVA, both traits – PH and DTEM (Days To Ear Emergence or Heading Date) – showed statistical significance. However, multiple comparisons test did not reveal any significant difference between the means of DTEM of NILs and the NIL parents (data not shown), but did for plant height (Table 3.7).

NILs	Mean	SE	DF	LCI	UCI
NIL5A(-)	82.8	0.39	330	82	83.5
NIL5A(+)	90	0.36	330	89.3	90.7
NIL6A(-)	85.1	0.55	330	84	86.2
NIL6A(+)	91	0.35	330	90.3	91.7
Pam	97	1.46	330	94.1	99.9
Par	86	1.46	330	83.1	88.9

a)

Trait	Contrasts	EST	LCI	UCI	SED	DF	p.adj
PH	NIL5A(-) - NIL5A(+)	7.21	5.49	8.92	0.421	149	****
	NIL5A(+) - Pam	7.03	3.17	10.9	0.699	6.37	**
	NIL5A(+) - Par	-3.97	-10.1	2.17	1.06	5.54	ns
	NIL6A(-) - NIL6A(+)	5.96	4.41	7.51	0.376	91.6	****
	NIL6A(+) - Pam	5.97	2.11	9.83	0.697	6.3	**
	NIL6A(+) - Par	-5.03	-11.2	1.11	1.06	5.51	ns
	Pam - Par	-11	-17.2	-4.76	1.22	8.49	**

b)

Table 3.7 NIL performance in the UK, 2018/19

a) provides PH means in cm, SE (Standard Error), UCI and LCI (lower and upper confidence intervals at the 95% significance level)

b) Pairwise comparisons of NILs on PH in JIC 2018/19. Contrasts = NIL sets and Parents. Pam = Pamyati Azieva and Par = Paragon. EST = estimated difference on trait means between contrasts. SED = standard error of a difference between the means of two contrasts. p.adj = p-values at the 95% significance level adjusted for Tukey's method. Significance codes: 0 '***' 0.001, '**' 0.01, '*' 0.05 or ns = nonsignificant. LCI and UCI are lower and upper confidence intervals respectively.

3.3.1.2.2.2 Norwich 2019/20

3.3.1.2.2.2.1 Data from plot height

Height data from the plots revealed a significant difference between the means of NIL5A(-) and NIL5A(+) and NIL6A(-) and 6A(+). Moreover, heights of NIL5A(+) and 6A(+) were significantly taller than that of Paragon (Figure 3.4).

3.3.1.2.2.2.2 Data from 50cm distance

Multiple comparisons of the height data from the plants collected from 50cm also found the height of the NIL5A(+) to be significantly different from that of NIL5A(-) (Figure 3.5). Although, the difference in the heights of NIL6A(+) and NIL6A(-) was almost 4cm, this difference between the means was not statistically significant. Nevertheless, NIL5A(+) and NIL6A(+) were significantly taller than the recurrent parent Paragon.

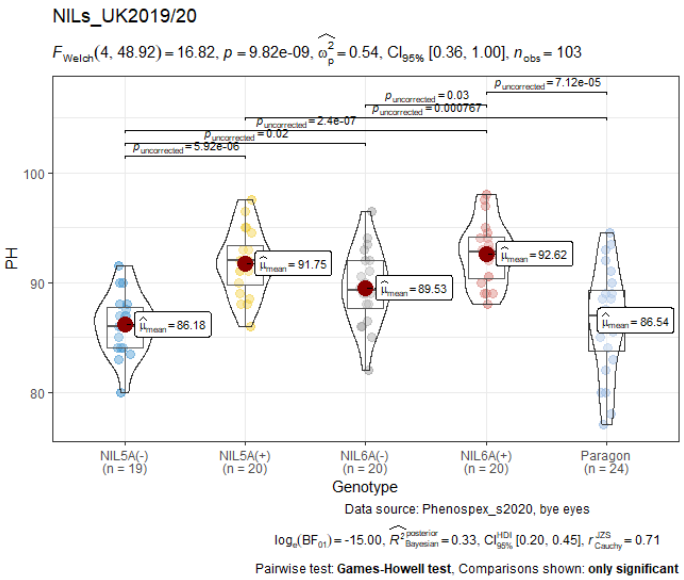


Figure 3.4 Plant height of NIL sets and Paragon, UK 2019/20

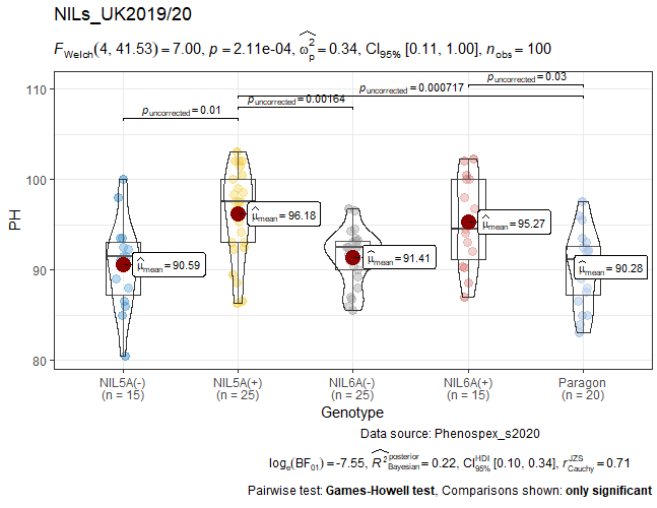


Figure 3.5 Plant height of NIL sets and Paragon, UK 2019/20

3.3.1.2.2.3 The year 2020/21

Conducted experiment was important as this was the first time when double height increased Paragon that is dNILs were tested in the field conditions. The top scored traits were PH and YSM. Both revealed overall statistical significance (***). Additionally, when the means of both NILs possessing 5A and 6A PH increasing alleles were compared with Paragon and short NILs, they were significantly taller (Figure 3.6, Table 3.8 and Figure 3.7). Importantly, expression of both height genes made Paragon almost as tall as Pamayti Azieva in the natural environment in the UK. Moreover, what was essential is that dNILs did not lodge, although Pamyati Azieva did partially (Figure 3.8), because of the excessive rainfall during the mid and late grain filling.

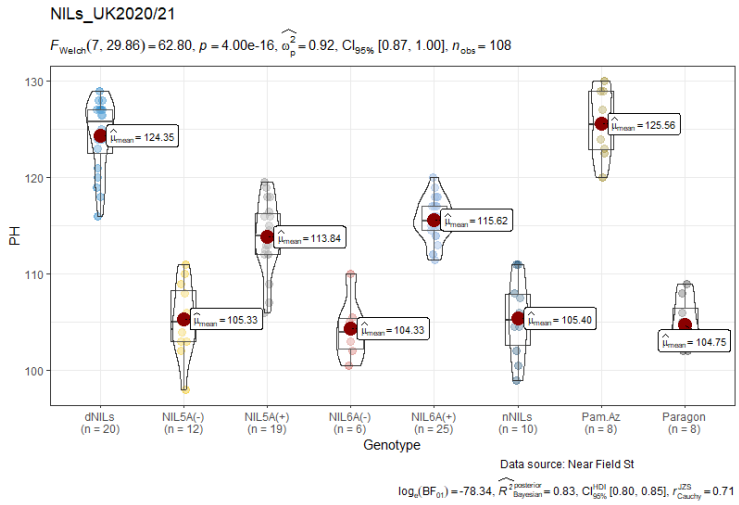


Figure 3.6 Plant height of NIL sets, dNILs and Paragon, UK 2020/21

Contrasts	Trait	EST	LCI	UCI	SED	DF	p.adj
dNILs - NIL5A(+)	PH	-10.5	-14.4	-6.65	0.85	36.8	****
dNILs - NIL6A(+)		-8.73	-11.8	-5.66	0.67	29.1	****
dNILs - nNILs		-18.9	-24.3	-13.6	1.09	16.6	****
dNILs - Pam.Az		1.21	-4.29	6.72	1.09	13.1	ns
dNILs - Paragon		-19.6	-23.9	-15.3	0.90	17.8	****
NIL5A(-) - NIL5A(+)		8.51	3.93	13.1	0.98	23.8	****
NIL6A(-) - NIL6A(+)		11.3	5.11	17.5	1.01	6.06	**
nNILs - Pam.Az		20.2	13.8	26.5	1.30	15.8	****
nNILs - Paragon		-0.65	-6.27	4.97	1.14	15.5	ns
Pam.Az - Paragon		-20.8	-26.6	-15	1.14	12.9	****
dNILs - NIL5A(+)	YSM	0.00389	-0.146	0.154	0.03	28.9	ns
dNILs - NIL6A(+)		0.0681	-0.0467	0.183	0.03	39.4	ns
dNILs - nNILs		0.0822	-0.0328	0.197	0.02	22.1	ns
dNILs - Pam.Az		-0.466	-0.619	-0.313	0.03	14.2	****
dNILs - Paragon		0.109	0.00268	0.214	0.02	21	*
NIL5A(-) - NIL5A(+)		-0.0788	-0.278	0.121	0.04	22.5	ns
NIL6A(-) - NIL6A(+)		-0.0376	-0.195	0.12	0.03	12.9	ns
nNILs - Pam.Az		-0.548	-0.702	-0.395	0.03	12.4	****
nNILs - Paragon		0.0263	-0.0813	0.134	0.02	13.9	ns
Pam.Az - Paragon		0.575	0.425	0.724	0.03	10.6	****

Table 3.8 NIL performance in the UK, 2020/21

Contrasts = NIL sets and Parents. EST = estimated difference on trait means between contrasts. SED = standard error of a difference between the means of two contrasts. p.adj = p-values at the 95% significance level adjusted for Tukey's method. Significance codes: 0 '***' 0.001, '**' 0.01, '*' 0.05 or ns = nonsignificant. LCI and UCI are lower and upper confidence intervals respectively.

Figure 3.7 dNILs compared to Paragon, UK 2020/21

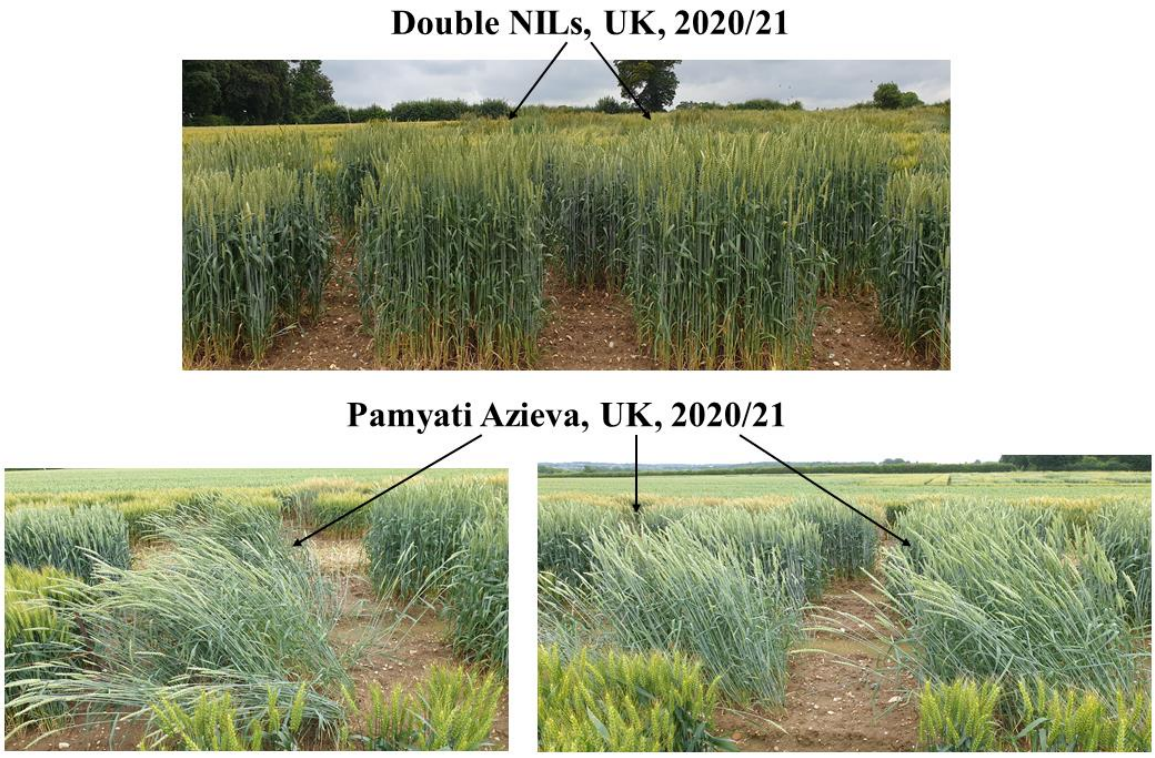
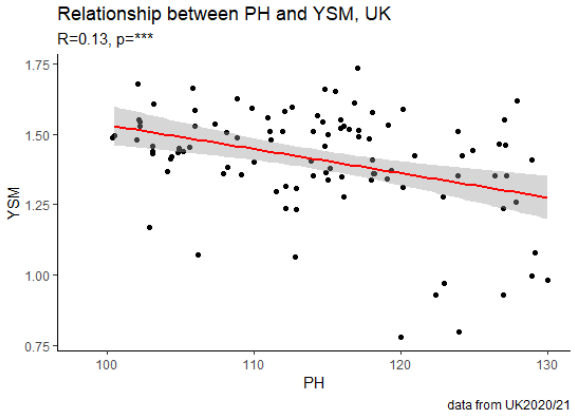


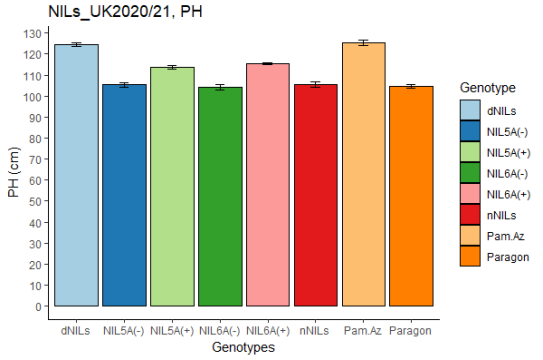
Figure 3.8 Lodging, UK 2020/21

The data showed that increased height has an adverse effect on grain yield in the UK. Accordingly, when PH was correlated with YSM, a significant negative association was detected (Figure 3.9). The NIL5A(-), NIL6A(-), nNILs and Paragon yielded approximately the same with no significant difference between them. Conversely, the taller NILs did not yield significantly less than shorter

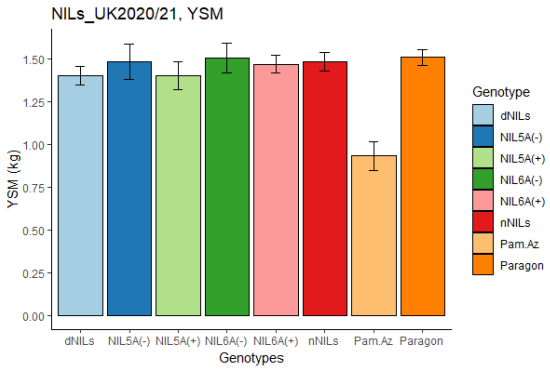
NILs and Paragon. Moreover, the reduced grain yield of dNILs compared to NIL5A(+), NIL6A(+) and nNILs did not show significance although the trend goes down with the increased height. Pamyati Azieva had the lowest yield compared to NIL5A(+), NIL6A(+), nNILs and Paragon all of which were statistically significant (see Table 3.8 for p values). Although dNILs were as tall as Pamyati Azieva with small but not significant difference of 1.2 cm between the means, its grain yield was significantly higher than that of Pamyati Azieva.



a)



b)



c)

Figure 3.9 Correlation of PH with YSM, UK 2020/21 growing season

a) PH was significantly negatively correlated with YSM in the UK. b) and c) show PH and yield, respectively, of both NILs sets, dNILs, nNILs and parents in absolute values.

3.3.1.2.3 Testing the tall and short NILs of the 5A and 6A QTL in Kazakhstan

Overall plant height varied considerably across the sites as well as years (Figure 3.10) with Alm 2019 and Alm 2020 the environments in which the greatest heights were observed with grand heights of 91.8 cm and 96.1 cm respectively. Pet 2020 and Pet 2021 produced a shorter crop, as expected with the difference in rainfall during the growing season for these sites (161.2mm and 120.9mm, see Table 3.4 and Table 3.5). Temperature seems to have less of an effect (Table 3.2, Table 3.3, Table 3.4 and Table 3.5). Alm 2021 was as short as the wheat grown in Pet. This was due to the low rainfall across the nation resulting in 57% height reduction compared to the previous year. The low rainfall of 2021 also caused height reduction in Pet compared to 2020 of 75%.

Interestingly, although the 5A effect when identified as a QTL in the PamxPar mapping population and initially discovered in the JIC glasshouse and then in the field in the UK and Kazakhstan, when studied as 5A+/- NILs the height effect was not observed in two out of three growing years in Almalybak (Figure 3.10). In fact, NIL5A(+) was shorter in these two, 2019 and 2021, growing seasons. In 2021, its height reduction, compared to NIL5A(-), was almost significant at the 5% level (0.06). However, in 2020 the NILs behave more like the originally identified QTL with NIL5A(+) being significantly taller than NIL5A(-). The phenomenon of reversing allelic height effect explains the fact that the QTL was not identified in all environments in the original QTL studies. Compared to Alm, no height difference was observed between 5A NIL sets in any of the experiments conducted in Pet.

As opposite to 5A QTL, the height increasing effect of 6A QTL was stable across the sites and growing seasons. The NIL6A(+) was always significantly taller than its counterpart NIL6A(-) (Figure 3.10).

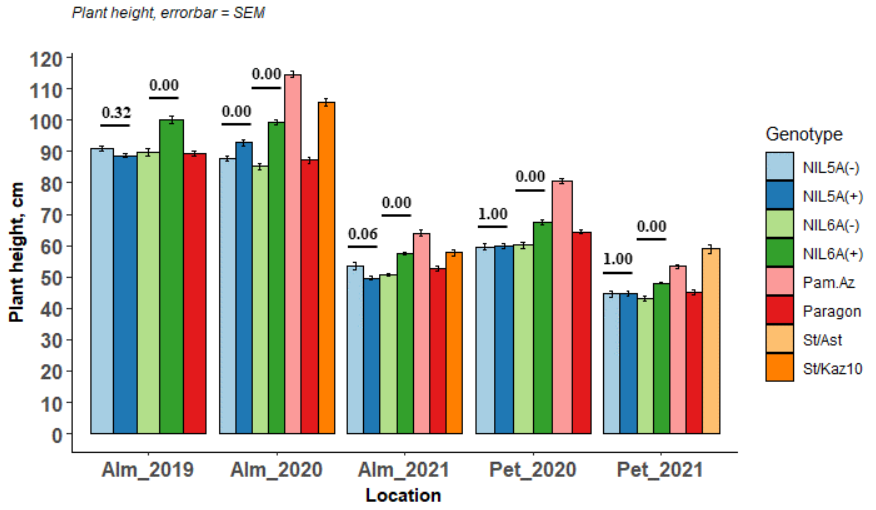


Figure 3.10 The heights of NILs, parents and local check varieties across the years and sites

St = Standard or Local check variety. Ast = Astana is the local check variety in Petropavl. Kaz10 = Kazakhstanskaya 10 is the local check variety in Almalybak.

3.3.1.2.3.1.1 Genotyping NIL5A to evaluate the possible genetic effect of the 6A QTL on PH of 5A (-) and 5A (+) NILs grown in Almalybak 2019

When the 5A height effect silenced in the first field trial in Kazakhstan, we had two hypotheses: i) it could be sampling error and/or ii) actually 5A NILs could be segregating at some part of the 6A loci although the recombinant lines used for NIL development were mostly fixed at the loci. Therefore, the seeds of all NIL pairs, including Paragon, from the same experiment was sent back to the UK to conduct genetic assessment. From that bag of seeds, random 24 grains of each NIL were sown to extract DNA. Once DNA was isolated, to evaluate the genetic state of 5A and 6A NILs at the 5A and 6A QTL chromosomal regions, we genotyped both NILs with 5A and 6A QTL flanking markers (QTL start, peak and end markers). Thus, NIL5A(-) and NIL5A(+) were genotyped with 5A as well as 6A QTL flanking markers and NIL6A(-) and NIL6A(+) with 6A and 5A QTL flanking markers. As a result, regardless of being (-) or (+), 5A NILs were segregating at the 6A QTL start marker, but for small chromosomal region (Figure 3.11).

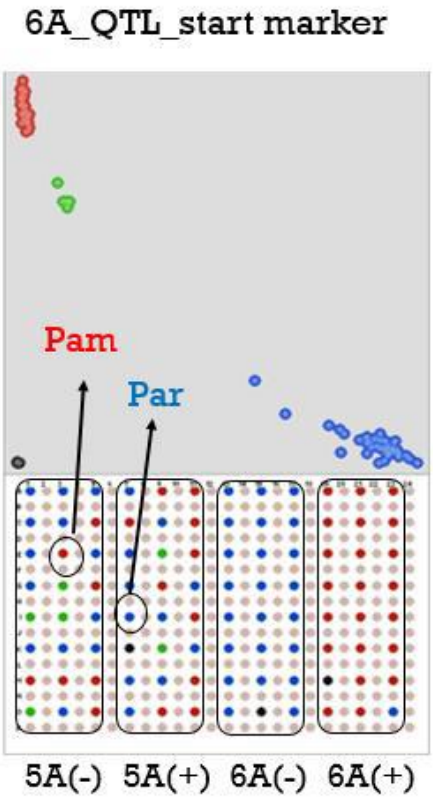


Figure 3.11 Genetic assessment of the 5A NIL set at the 6A PH QTL and 6A NIL set at the 5A PH QTL

Therefore, to determine whether the genetic state of 5A (-) and 5A (+) NILs at the 6A QTL start marker has an overall effect on plant height, plants which are fixed for Pamyat and Paragon at the 6A QTL start marker have been grown in the glasshouse in randomized complete blocks. During phenotyping, the tallest (canopy) spike of each plant was measured for plant height.

As a result of plant height data analysis, the significant difference in plant height between 5A(-) (6A="Pam" (fixed for Pamyati Azieva at the 6A QTL start marker)) and 5A(-) (6A="Par" (fixed for Paragon at the 6A QTL start marker)), and between 5A(+) (6A="Pam") and 5A(+) (6A="Par") NILs was not detected. Therefore, it was assumed that the Pamyati Azieva's plant height increasing allele at the 6A_start_marker (please note that "Pamyati Azieva" state is only for 6A_start_marker not for 6A_peak and 6A_end_markers) has no genetic effect on plant heights of NIL5A(-) and NIL5A(+) (Figure 3.12).

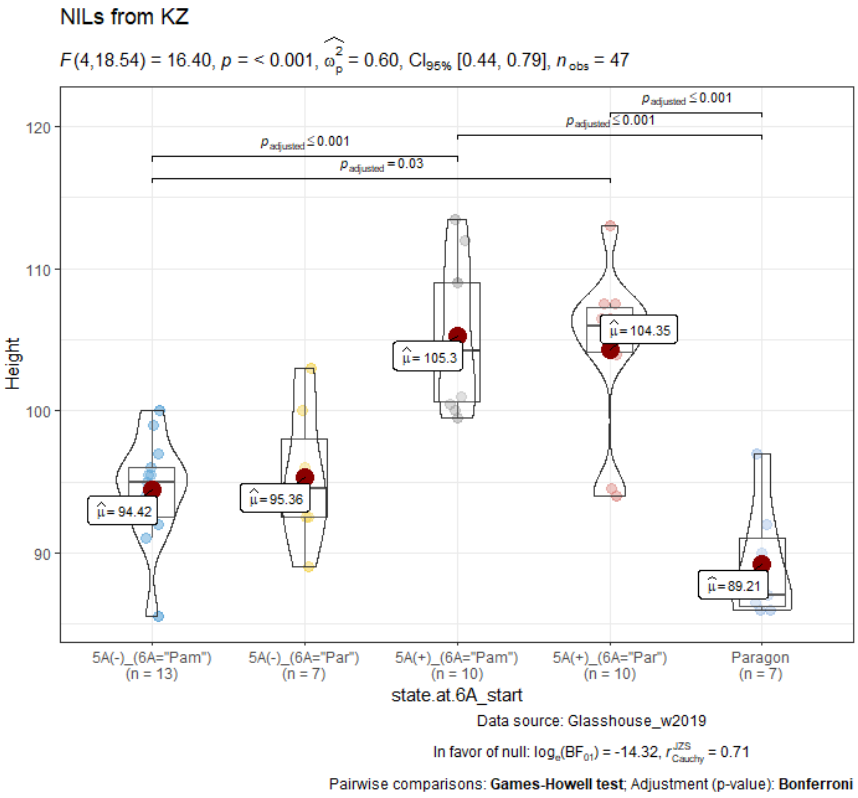


Figure 3.12 PH of the 5A NILs grown in Almalybak 2019 based on genetic state at the 6A_start_marker

5A(-)_(6A=Pam) = 5A(-) NILs fixed for Pamyat at the 6A_start_marker. 5A(-) (6A=Par) = 5A(-) NILs fixed for Paragon at the 6A_start_marker. 5A(+)_(6A=Pam) = 5A(-) NILs fixed for Pamyat at the 6A_start_marker. 5A(+)_(6A=Par) = 5A(-) NILs fixed for Paragon at the 6A_start_marker

Then, plant height data of 5A (-) and 5A (+) NILs was combined (not taking the 6A_start_marker state into account) and analysed. As an outcome, 5A (-) NILs were short regardless of carrying positive plant height increasing allele of Pamyat at the 6A_start_marker with less than 1cm difference in the means between them. Likewise, despite the 6A_start_marker state, 5A (+) NILs were taller than their 5A (-) NILs counterparts. An additive effect was ~10cm as expected (Figure 3.13).

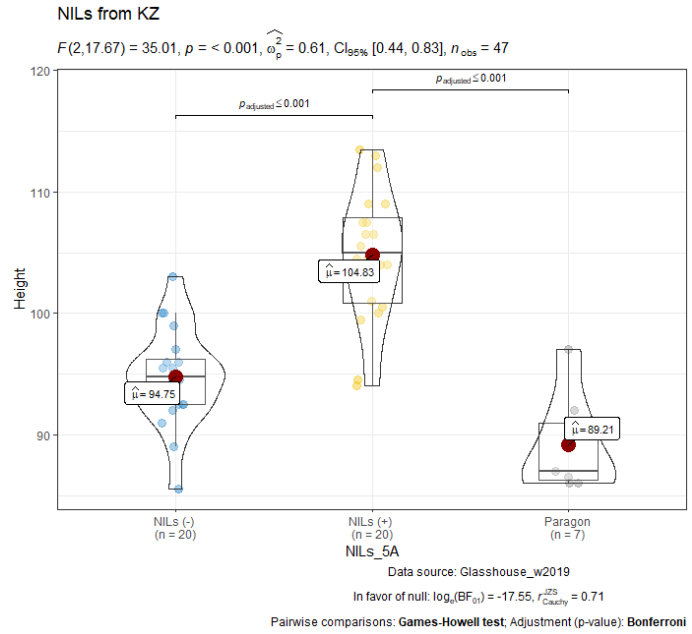


Figure 3.13 PH of the 5A NILs grown in Almalybak 2019 regardless of the genetic state at the 6A_start_marker

Figure shows that heights of 5A (-) NILs were significantly different from those of 5A (+) NILs with the QTL additive effect of ten. Moreover, statistical significance was found between 5A (+) NILs and Paragon likewise. Despite some difference between the plant height means of 5A (-) NILs and Paragon, it was not statistically different from each other.

3.3.1.3 Comparing the developmental characteristics in NIL sets

3.3.1.3.1 Almalybak

There was no obvious difference observed throughout growth stages between tall and short NILs for both QTL in Alm 2019, 2020 and 2021. However, compared to Pamyati Azieva, NIL sets for both QTL started lagging behind for 1-2 day at the tillering and for 3-4 days at the elongation, flag leaf emergence and booting developmental stages. This difference widened for 6 and 7 days for 5A and 6A QTL respectively at the heading and flowering stages. Despite the gap, NILs managed to ripen at the same time as Pamyati Azieva.

3.3.1.3.2 Petropavl

In Pet 2020 and 2021, NIL sets regardless of QTL germinated at the same time between themselves and compared to parents and Astana (local check, tested in Pet 2021 only). However, Pamyati Azieva and Astana matured ~10 days earlier than NIL pairs in both testing seasons. Interestingly, Paragon ripened exactly 3 days after Pamyati Azieva in both testing years.

There was no difference between NIL5A(+) and NIL5A(-) until stem elongation when 3 day earliness was observed in NIL5A(+) relative to NIL5A(-) in Pet 2020, but no difference was observed in Pet 2021. However, the 3 day difference in Pet 2020 was shortened to one day at flowering, but again widened to 5.5 days at physiological maturity. Although tall and short 5A NIL did differ throughout the developmental stages, NIL5A(+) matured 3 days earlier than NIL5A(-).

In 6A NIL pairs, no difference was seen until maturity when 6A(+) ripened 2 and 1.5 days earlier than 6A(-) in Pet 2020 and 2021 respectively.

3.3.1.4 Yield benefits of the 5A and 6A QTL in Alm and Pet

Overall grain yield was also highly variable between seasons and sites, as was the plant height (Figure 3.14). Crop production and yield was highly dependent on the amount of precipitation. For example, a significant increase in rainfall gave rise to the highest grain yield in Alm 2020 compared to two other growing seasons, Alm 2019 and 2021, with the fold increase of 2.2 and 7.5 respectively when the grand means of YP were compared (Figure 3.14). Likewise, a large improvement in yield was observed in Pet 2020, double compared to Pet 2021. However, as opposed to Alm, no increase in the rainfall was observed in Pet 2020 as it was in Pet 2021. Instead, in both growing years the seasonal precipitation rate dropped by 16.5 and 27.8 percent compared to LTAA. Nevertheless, the amount of rainfall in the last ten days (LTD) of June and the first ten days (FTD) of July seem to play a crucial role in determining the final grain yield of wheat and might explain the observed large yield difference between Pet 2020 and Pet 2021. Although there was 25% reduction in the amount of seasonal rainfall in Pet 2021 compared to Pet 2020, it is likely that the rainfall in LTD of June and FTD of July is of more importance as the precipitation rate in this time window differed significantly between 2020 and 2021 in Pet. For instance, LTD of June and FTD of July experienced an excessive rate of rainfall, 36 and 67 mm respectively, in 2020 as opposed to LTD of June and FTD of July, 3.2 and 0.4 mm respectively, in Pet 2021. However, although no increase in the rainfall was detected in June, July 2021 had the same large amount of rainfall, 65 mm, as in FTD of July 2020, but it fell in the second ten days (STD) of the month which was followed by no increase in grain yield. Despite the fact that the rainfall in STD of July in Pet 2020 was even smaller than that of the FTD of July in Pet 2021, it did not bring about the reduction in the final yield. This conclusion is highly consistent with suggestions highlighting the importance of the late June and early July rainfall in improving the crop yield in Pet (Morgounov et al., 2001).

Mean yield gain of the NIL5A(+) was observed in Pet compared over both seasons although this was not statistically significant in either case. In Alm there was also no statistically significant difference observed for yield in Alm 2019 (negative with $p = 0.39$) and Alm 2020 (positive with $p = 0.73$) although there was actually a significant reduction in mean yield of 5A+ NILs in Alm 2021 ($p=0.04$). Combined YP data of two growing seasons in Pet did not provide extra statistical power ($p<0.59$, EST = 0.23 t/ha) although the direction of the effect was the same in both seasons (positive) (Figure 3.15).

The height increasing allele for 6A also showed no significant effect (at 5% level) on yield in any of the environments tested. However, the mean yield of lines carrying the height increasing allele was lower than their short counterparts in Alm 2019 (no pairwise comparisons conducted due to limited seed number of 6A (-)) and Pet 2021 (p = 0.16). There were cases when 6A height increasing locus increased grain yield, but the p-values were higher than that of yield reduced seasons. However, the longer long term yield benefits of the 6A PH increasing QTL are likely restricted based on combined yield data from two seasons. Thus, we hypothesise that breeding against height increasing allele should not influence yield negatively, yet it could be the way to fine-tune the final grain yield in the region. However, extra field trials on 6A QTL NIL pairs are needed to support the provided hypothesis. In case of 5A height increasing QTL, we suggest keeping the allele in Pet as the relative yield increase was observed in this main wheat growing area. However, due to an observed significant reduction in grain yield in the southeastern regions (Alm) of the country in 2019, it is recommended to breed against the Pamyati Azieva's allele at the 5A QTL locus. Further studies to confirm the benefits of the 5A QTL should aim at introgressing allele into other foreign cultivars adapted to other regions as the local cultivars are mostly fixed at the locus. However, there could be other unexplained molecular systems contributing to grain yield in NIL5A(+). To show whether the yield gain in NIL5A(+) is truly due to 5A PH QTL or not, the reverse selection to height increasing allele is the option to take for wheat breeders in Kazakhstan.

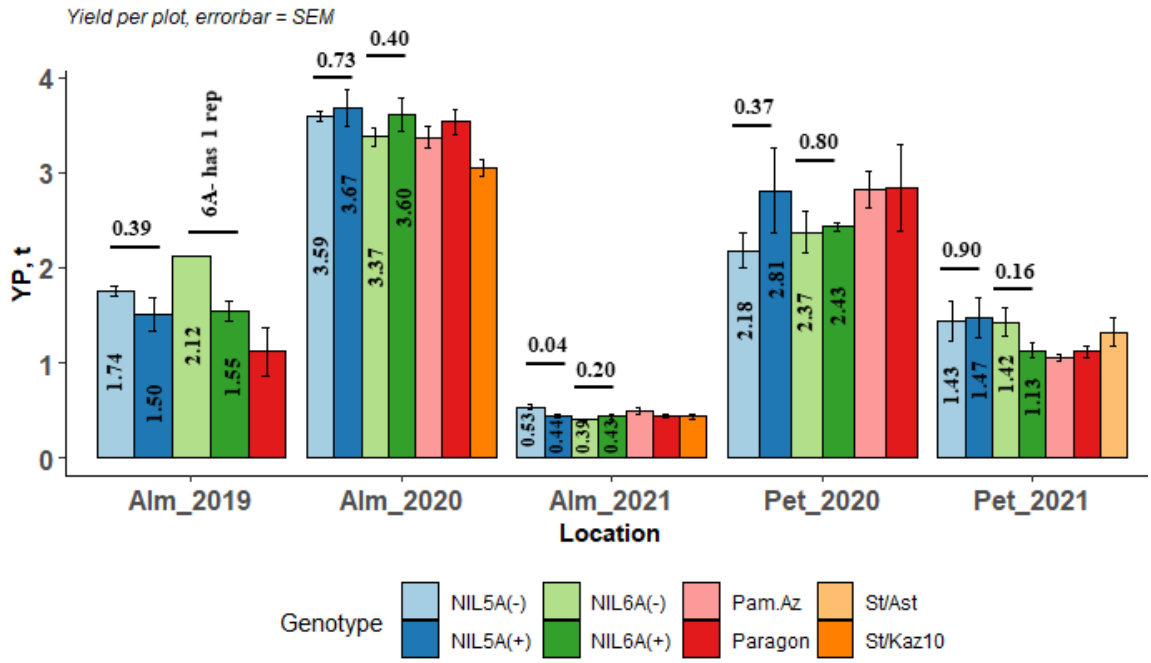


Figure 3.14 The yields of NILs, parents and local check varieties across the years and sites

In general, any yield advantage was not observed from both Pamayari Azieva tall alleles over the Paragon short alleles in Near Isogenic Lines in Kazakhstan. Instead, grain yield of 5A+ was significantly lower than that of 5A- in Almalybak 2021 (p=0.04).

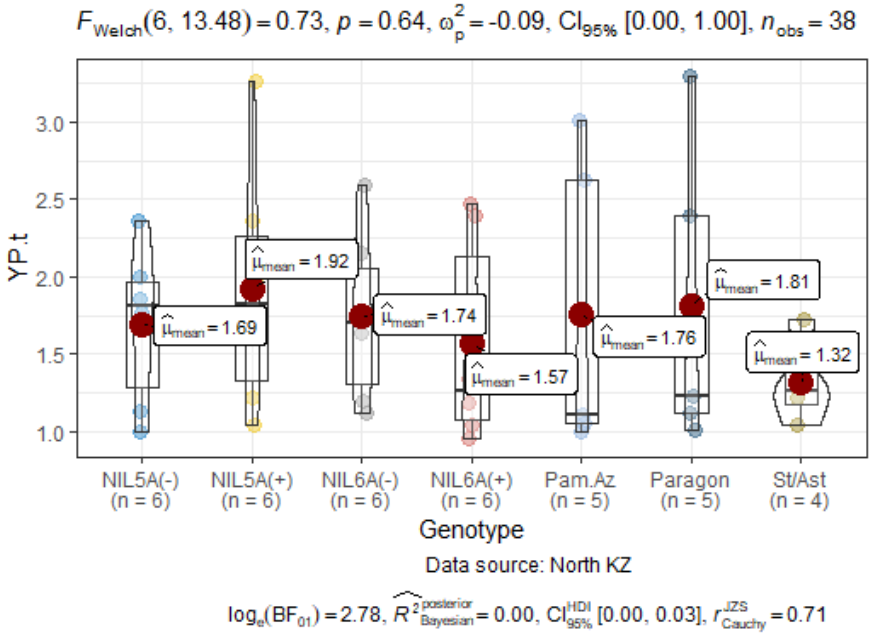


Figure 3.15 Combined yield data across the growing years for Pet

3.3.1.5 Dissecting the benefits of grain yield

Obtained data on yield components of NIL pairs were not consistent between growing seasons. We saw insignificant decrease in TGW of NILs possessing the height increasing allele of the 5A QTL donated by Pamyati Azieva's. By contrast, TGW of NIL6A(+) was higher than that of NIL6A(-) in four environments out of five (Figure 3.16). However, none of the comparisons displayed statistical importance except Pet 2020 which was relatively close to the significance level at the 5% (p=0.09).

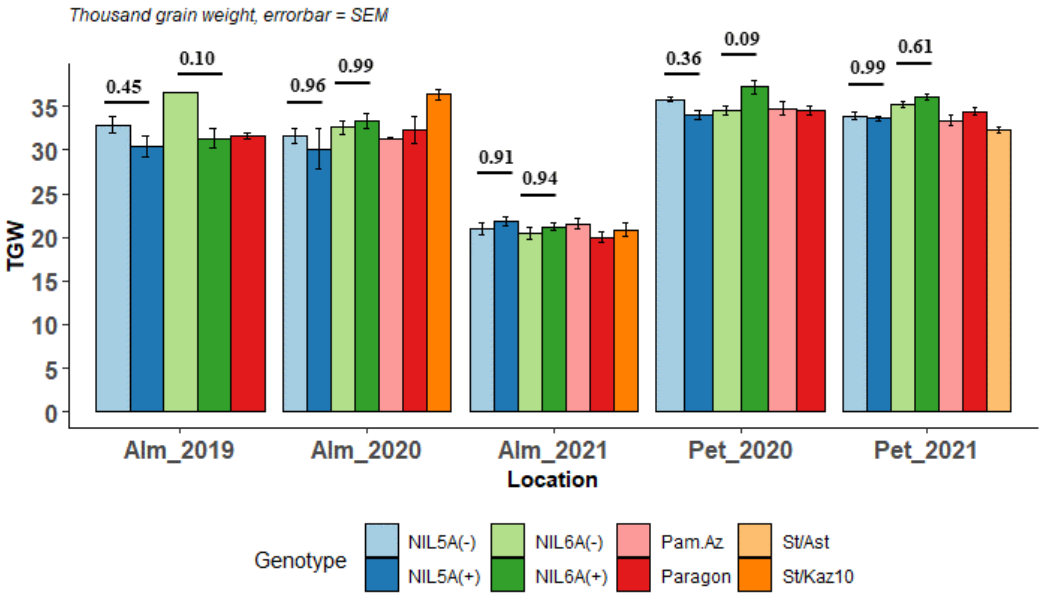


Figure 3.16 TGW of the 5A and 6A QTL NIL pairs

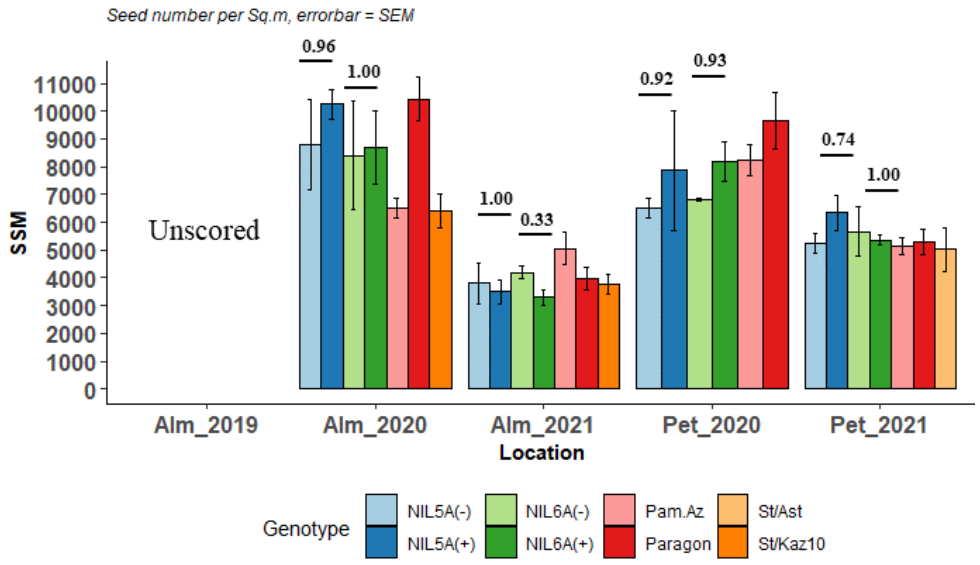


Figure 3.17 SSM of the 5A and 6A QTL NIL pairs

The SSM trend of 5A NIL pairs was directly opposite to what was detected in TGW (Figure 3.16 and Figure 3.17), but similar to that of YP (Figure 3.14). However, neither single environment nor combined data on SSM showed statistical significance. When it comes to the 6A NIL pairs, a partial contrast was observed between TGW and SSM, but again none of them were significant. A little, but not significant mean increase except Pet 2020 (p=0.00), towards NIL5A(+) was observed between 5A NIL pairs in both yield determinants such as plant and spike number per squared-meter (Figure 3.18 and Figure 3.19). Although, most of the comparisons were statistically insignificant, the directions in the means between YP, SSM, NP_1m^2 and NS_1m^2 were similar in 5A QTL lines. This was also partially true for 6A NIL pairs, but as in 5A NILs, none of the estimated differences between the means showed statistical importance.

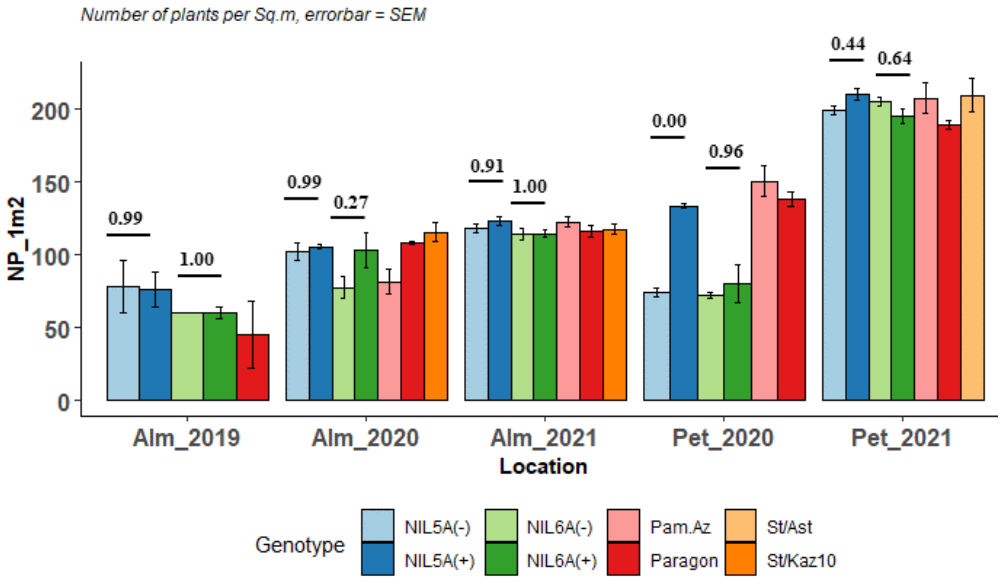


Figure 3.18 NP_1m^2 of the 5A and 6A QTL NIL pairs

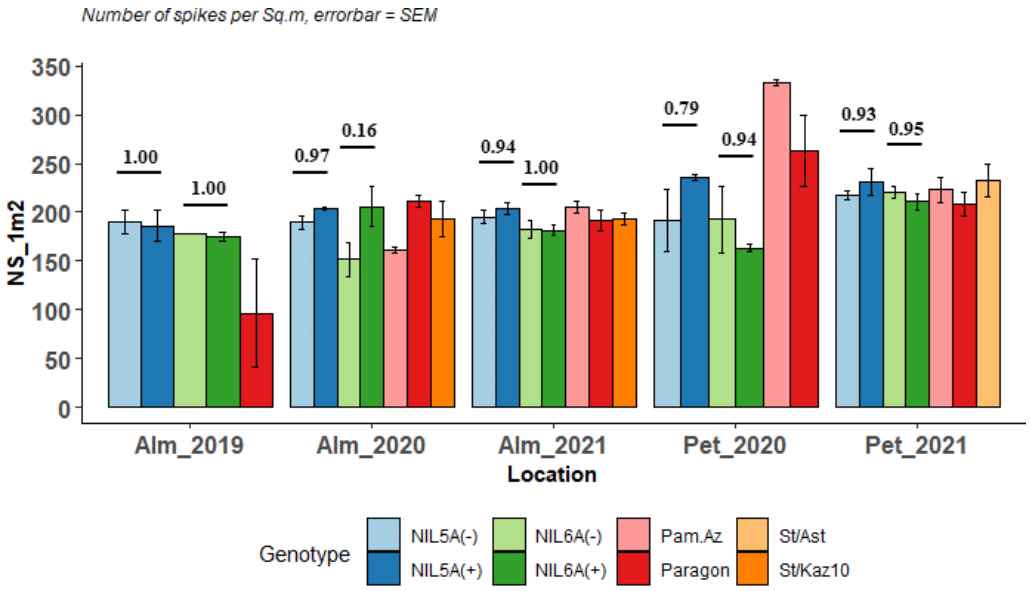


Figure 3.19 NS_1m^2 of the 5A and 6A QTL NIL pairs

Looking at TAGB and HI and comparing them between environments and NIL pairs was useful to understand some of physiological bases of grain yield. Like PH, the TAGB in Alm was in general significantly higher than that of Pet in 2020 (p=0.00 with EST of - 58.9 (EST always belongs to second variable of the contrasts throughout the thesis)). However, the substantial reduction in grain yield in Alm 2021 gave rise to directly opposite scenario in TAGB (p=0.00, EST= 14.6) (Figure 3.14 and Figure 3.20) despite the fact that plants were still significantly taller in Alm compared to Pet (Figure 3.10). Please note that in these figures the p-values stated above are not shown as the environments are being compared and not the NIL sets.

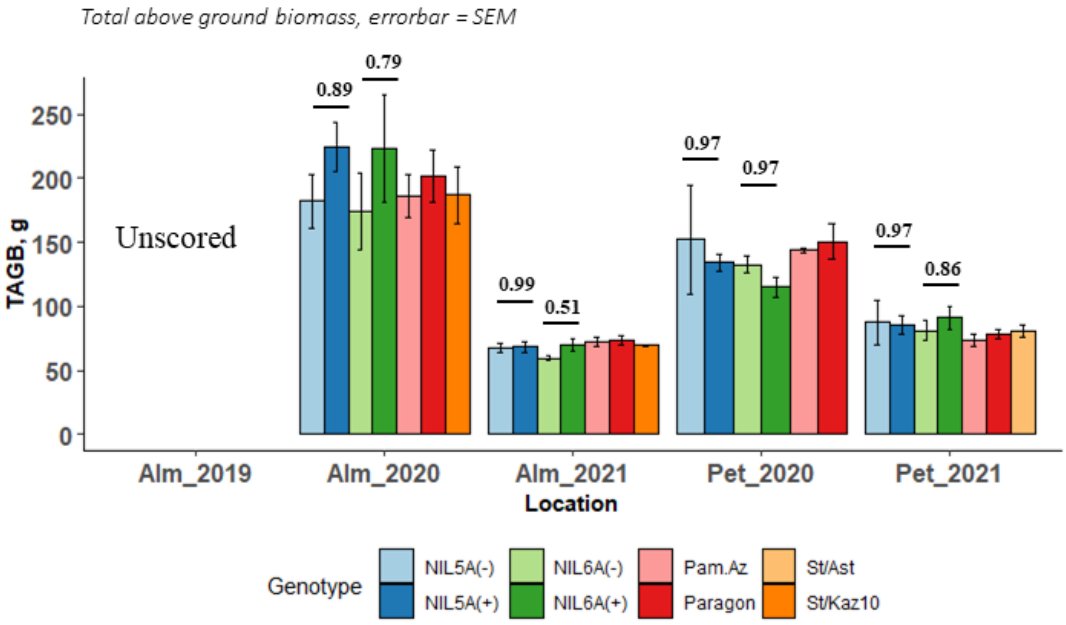


Figure 3.20 TAGB of the 5A and 6A QTL NIL pairs

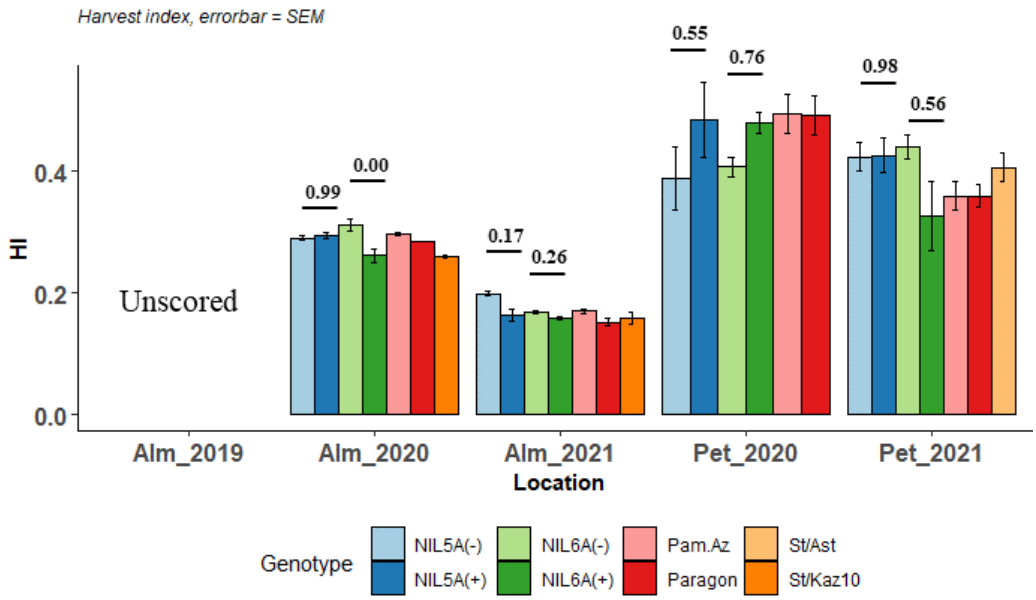


Figure 3.21 HI of the 5A and 6A QTL NIL pairs

Despite the large differences in PH between Alm and Pet which were higher in Alm, the resource allocation to grains (RAG) was significantly higher in Pet in both 2020 and 2021 growing seasons (p=****) with EST of 0.17 and 0.22 respectively (Figure 3.21). EST was even higher (0.20) when two-year data on HI for each environment were combined and compared showing that RAG in Pet is twice as much as in Alm (p=****) (Figure 3.22).

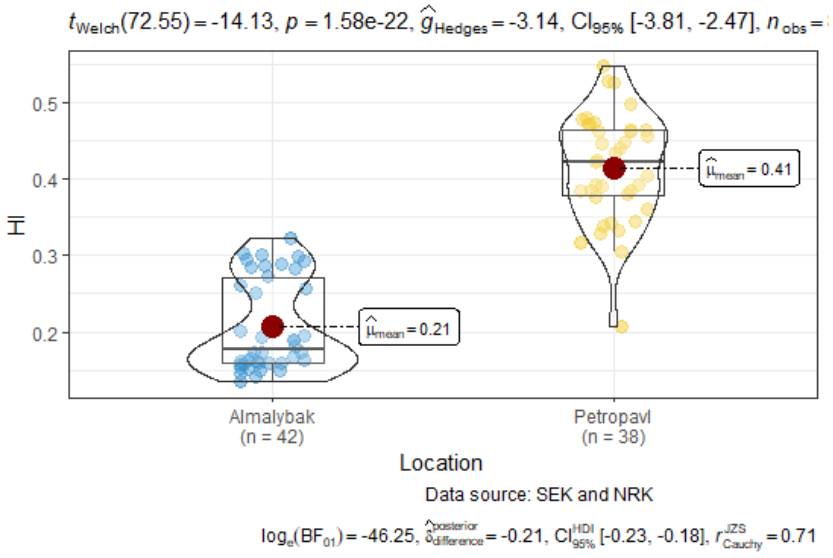


Figure 3.22 Difference in HI (2020+21) between Alm and Pet

The marginal increase in the absolute mean of HI or RAG was seen, as it was in YP, SSM, NP_1m^2 and SN_1m^2, across the growing seasons for NIL5A(+) relative to NIL5A(-), but in fact it did not add any value to the statistical analysis. As opposite to NIL5A(+), the NIL6A(+) had reduced

means of HI in three out of four environments. Among, one of two growing seasons in Alm, Alm 2020 revealed significance at the 5% (p=0.00) and second (Alm 2021) did not, but also had lower p-value (p=0.26) compared to Pet 2021 (p=0.56). However, these respective significant and insignificant drops in HI or RAG in Alm 2020 and 2021 resulted in significant reduction in number of kernels per main spike and thus per plant of tall NIL of 6A QTL compared to the short (Figure 3.23 and Figure 3.24). Accordingly, grain weight per main spike and per plant also reduced substantially in Alm (Figure 3.25 and Figure 3.26).

Compared to NIL5A(-), NKS and GWMW of the NIL5A(+) also reduced almost significantly at the 5% level (p=0.08) in Alm2021.

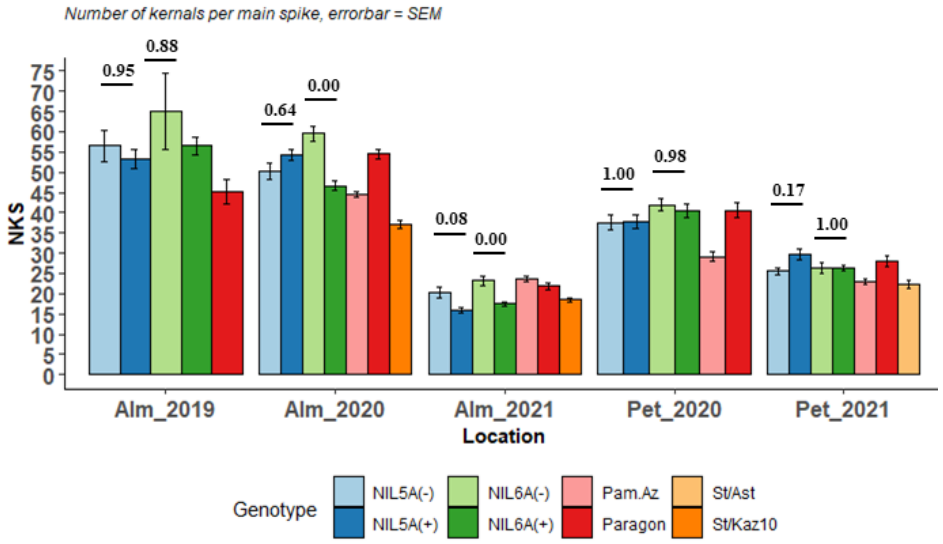


Figure 3.23 NKS of the 5A and 6A QTL NIL pairs

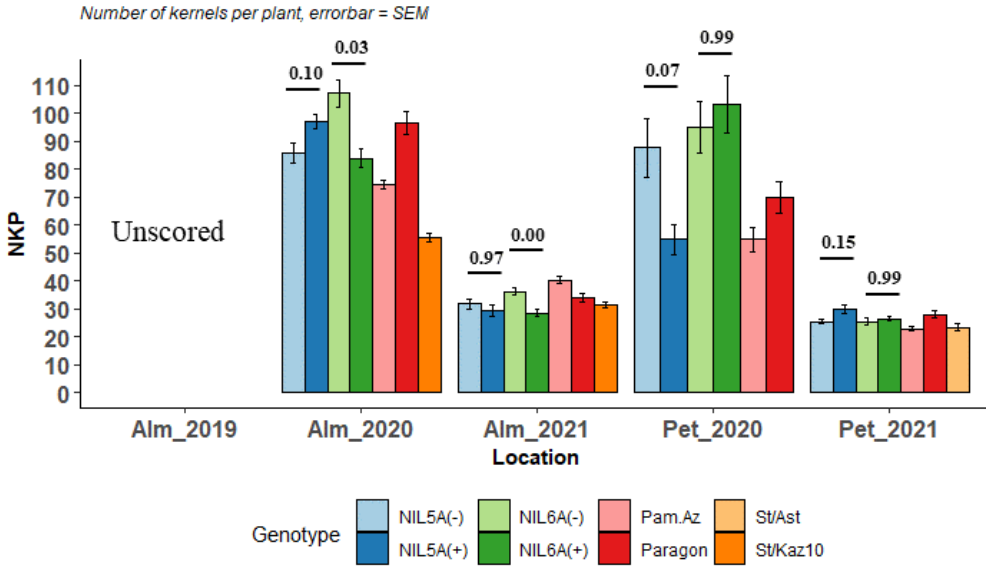


Figure 3.24 NKP of the 5A and 6A QTL NIL pairs

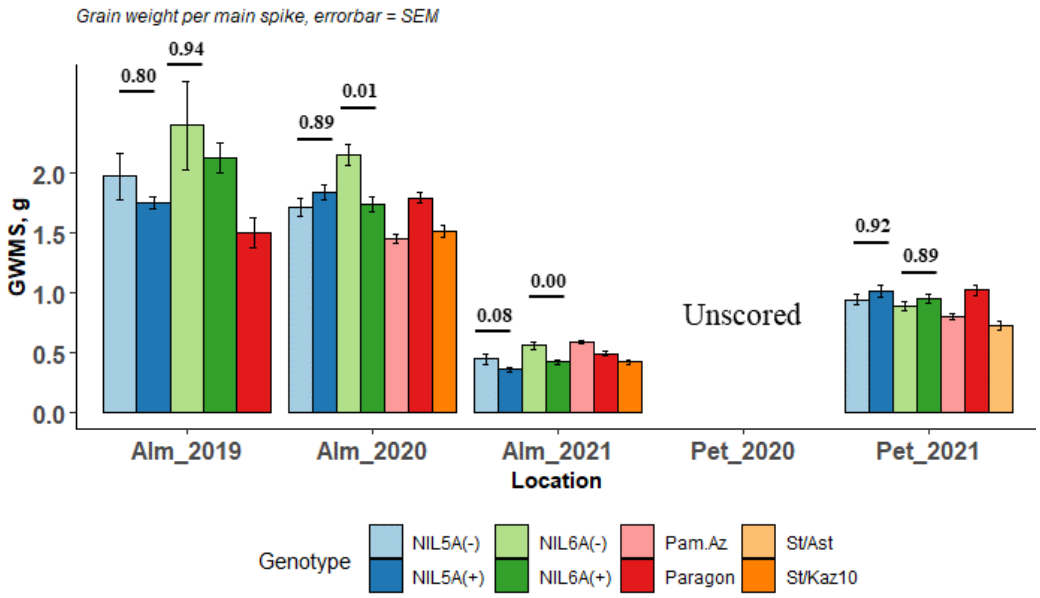


Figure 3.25 GWMS of the 5A and 6A QTL NIL pairs

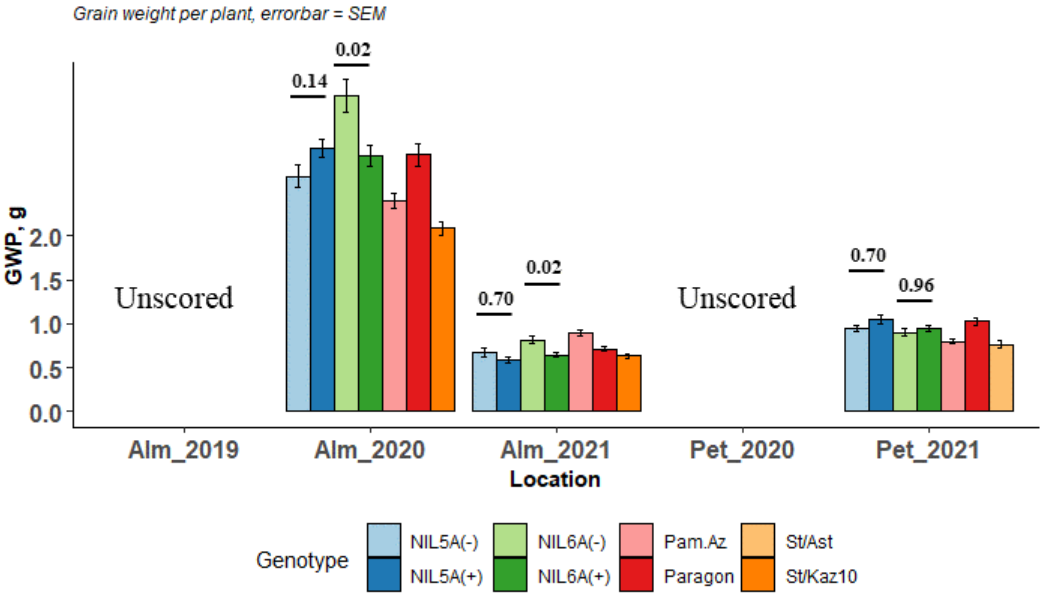


Figure 3.26 GWP of the 5A and 6A QTL NIL pairs

The number of effective tillers did not differ significantly between both NIL pairs. However, countrywide, plants grown in Alm set 0.61 spikes more than that of grown in Pet. This difference between the means provided a strong statistical significance (p=0.00, plot is not shown).

3.3.1.6 Dissecting the plant height components

To help understand the action of the 5A and 6A PH QTL, we measured the length of plant height components (PHCs) including the main spike and internodes (from the top 1st which is the peduncle to the bottom 5th) (Figure 3.27 and Figure 3.29). Surprisingly, the main spike length (SL) of tall NILs for 6A QTL significantly reduced in Alm 2020 (p = 0.05) and 2021 (p = 0.00),

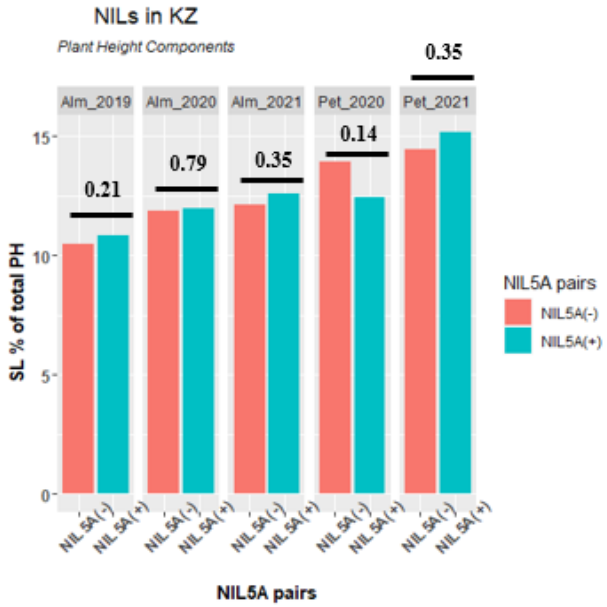
although it was substantially longer by 0.7cm in Alm 2019 (p = 0.01) compared to short siblings. By contrast, no difference was observed in SL in any of the field experiments in Pet. The 5A allele which increased height in the UK had no effect on SL (although this was not measured in the UK when the height difference was expressed) with no mean difference in Alm and contrasting but insignificant small mean differences in Pet.

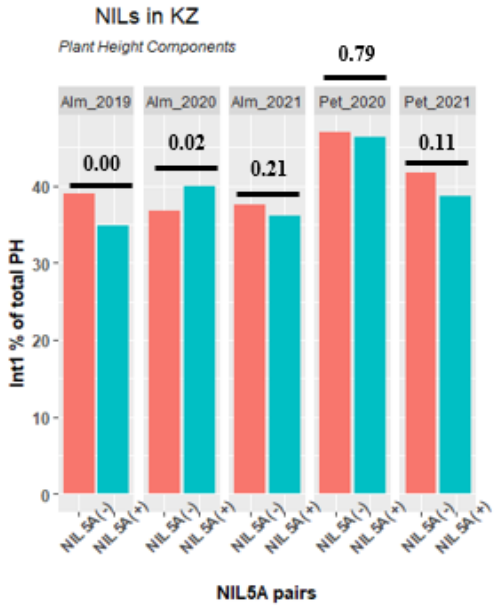
The increased height in NIL6A(+) resulted in the stable and significant increase in the first internode across the environments and growing seasons excluding only Pet 2020 the p-value of which was slightly higher (0.20) although the EST of 2.8cm observed between NIL pairs supported the trend. The positive and significant effect of the plant height increasing allele of 6A QTL on other PHCs such as internode-2 and -3 was also expressed in all three growing seasons at Alm 2019, Alm 2020 and Pet 2021 (only Int3), but not in Pet 2020 and Alm 2021. Interestingly, although PH difference was not observed, the tall allele of 5A QTL had almost the same positive effect on internode-2 and -3 in Alm, except for Alm 2019 when no difference was observed. Both height increasing alleles seem to have no detectable effect on internode-4 and 5. The only case when 6A(+) made the 4th internode statistically longer than that of 6A(-) was Alm 2020. In Alm 2019 and 2021 the mean length of internode -4 of NIL6A(+) did not show statistical difference despite being longer. Only a few plants had 5th internode in Pet regardless of allelic state at the 5A and 6A QTL loci. In general, like PH, there was a significant difference in all PHCs between Alm and Pet.

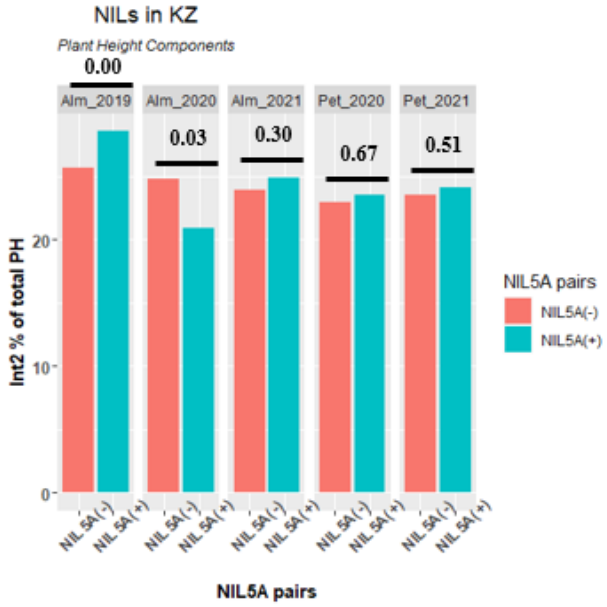
Besides the absolute values of PHCs, we calculated the relative proportion of each component to the total plant height in both 5A and 6A QTL NIL pairs across the environments to assess if more growth has contributed into one particular PHC per line and compare the proportion of each component between environments. For example, although the spikes were significantly longer in Alm relative to Pet (p = 0.00, EST = -1.7), its relative proportion accounted for only 11.2% compared to 13.4% in Pet (Figure 3.28 and Figure 3.30). The same was true for internode -1, which is the peduncle length, with significant mean difference of ~6%. However, the stem height was mainly compensated by contributing towards more lower internodes such as internode-4 and -5 in Alm. In fact, it was not these lower parts of the stem, yet upper, especially the internode -1 which contributed the highest percentage to total stem length, because the peduncle length, and thickness of the node are both considered important in wheat breeding with local breeders widely believing these to be an indicator of drought tolerance (personal communication). The same concept is also largely accepted within the international community (Bogale et al., 2011; Farshadfar et al., 2013; Mursalova et al., 2015; Naoura et al., 2019). The relative proportion of PHCs of NILs for both QTL displayed approximately the same outcome. If the environmental and allelic effects are ignored from the global data on NIL pairs, we can deduce that the mean SL, Int1, Int2, Int3, Int4 and Int5 proportion accounts for 11.9%, 30.6%, 23.5%, 14.6%, 8.72% and 1.7% of the total height

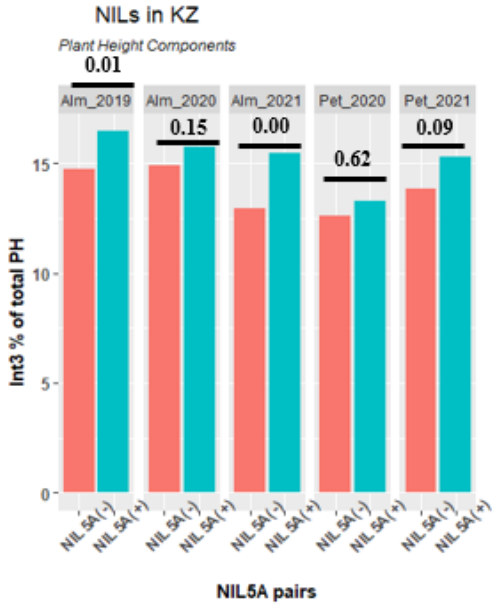
respectively. When the data on SL of each NIL sets were considered, the lower contribution of 6A(+) allele was observed compared to 6A (-) allele which was significant in Alm 2020 and Pet 2021 with p-values of 0.00 and 0.01 respectively. Conversely, no significant difference was observed in SL proportion in 5A NIL pairs. As seen in real values of internode-3, the tall 5A QTL allele seems to have a consistent increasing effect across the five growing seasons the two (Alm 2019/21) and one (Pet 2021) of which were significant and close to significant respectively.

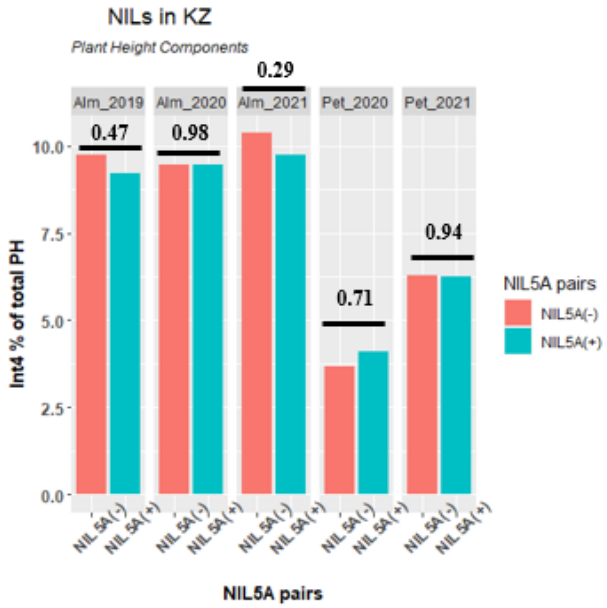
Figure 3.27 The absolute values of PHCs in NILS for 5A QTL

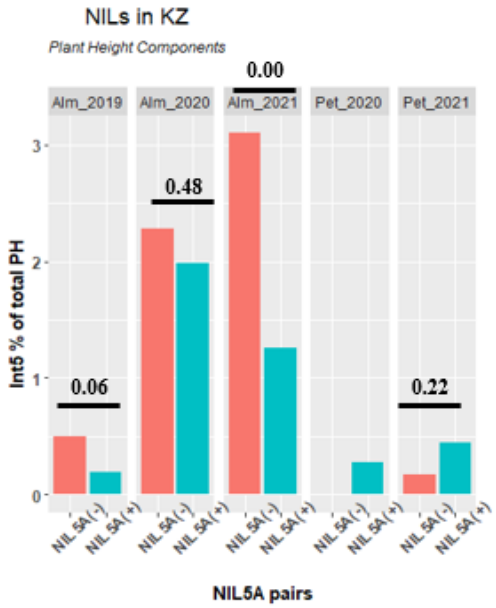

NILs in KZ
Plant Height Components
Alm_2019
Alm_2020
Alm_2021
Pet_2020
Pet_2021
0.21
0.79
0.35
0.14
0.35
SL % of total PH
NIL5A pairs
NIL5A(-)
NIL5A(+)
NIL5A pairs


NILs in KZ
Plant Height Components
Alm_2019
Alm_2020
Alm_2021
Pet_2020
Pet_2021
0.00
0.02
0.21
0.79
0.11
Int1 % of total PH
NIL5A pairs


NILs in KZ
Plant Height Components
Alm_2019
Alm_2020
Alm_2021
Pet_2020
Pet_2021
0.00
0.03
0.30
0.67
0.51
Int2 % of total PH
NIL5A pairs
NIL5A(-)
NIL5A(+)
NIL5A pairs


NILs in KZ
Plant Height Components
Alm_2019
Alm_2020
Alm_2021
Pet_2020
Pet_2021
0.01
0.15
0.00
0.62
0.09
Int3 % of total PH
NIL5A pairs


NILs in KZ
Plant Height Components
Alm_2019
Alm_2020
Alm_2021
Pet_2020
Pet_2021
0.47
0.98
0.29
0.71
0.94
Int4 % of total PH
NIL5A pairs
NIL5A(-)
NIL5A(+)
NIL5A pairs


NILs in KZ
Plant Height Components
Alm_2019
Alm_2020
Alm_2021
Pet_2020
Pet_2021
0.06
0.48
0.00
0.22
Int5 % of total PH
NIL5A pairs

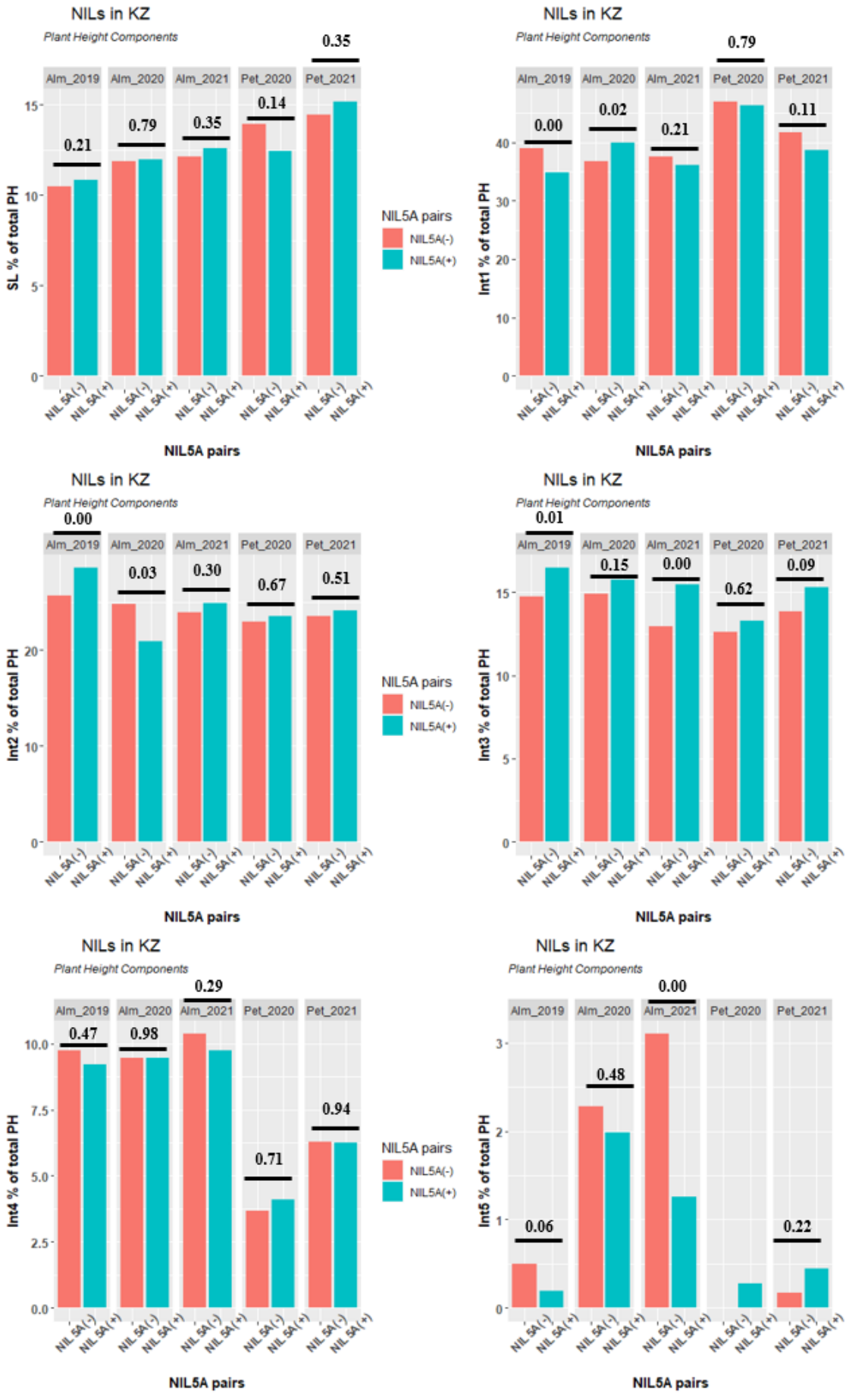


Figure 3.28 The relative percentage of PHCs to the total plant height in NILS for 5A QTL

Figure 3.29 The absolute values of PHCs in NILS for 6A QTL

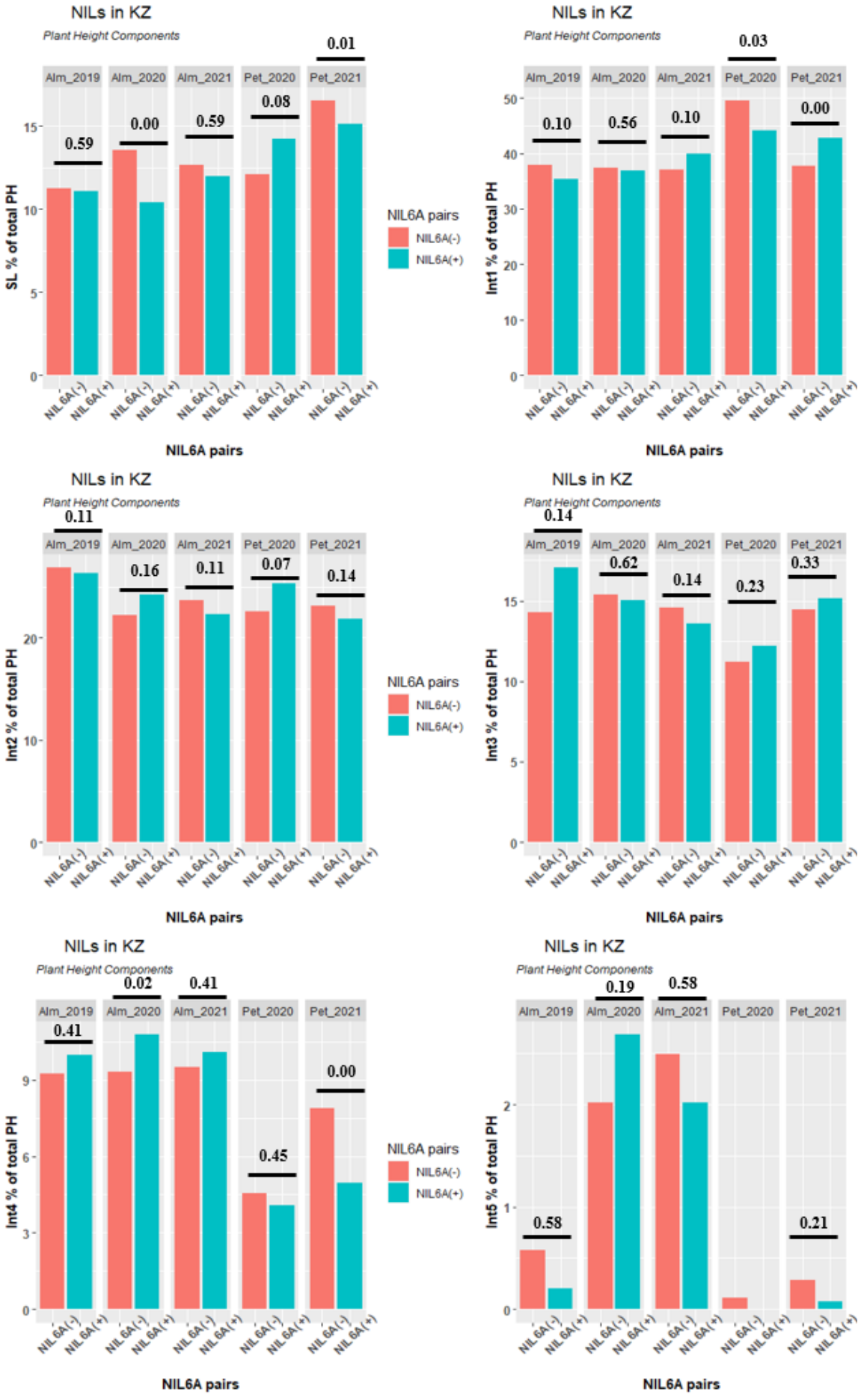


Figure 3.30 The relative percentage of PHCs to the total plant height in NILS for 6A QTL

3.3.1.7 Coleoptile length

3.3.1.7.1 Introduction

We have already mentioned in the section 1.4.3.4.2 that the benefits of *DELLA* mutants are limited or even reversed in hot and dry rainfed conditions where seeds need to be sown deeper and early establishment, vigour and longer coleoptile/root length are of importance (Rebetzke et al., 2007; Amram et al., 2015). Perhaps, that was the main reason why the use of high yielding semi-dwarf *Rht-B1b (Rht1) and Rht-D1b (Rht2)* genes are likely restricted in Kazakhstan where the wheat is spring type and rainfed? This question remains unanswered as the sophisticated molecular characterisation of effects of the alleles of these important adaptation genes on early establishment and vigour are still less studied or unstudied in Kazakh wheat germplasm. No environment specific plant height gene has been identified and characterised for this region. Thus, we have no clear idea how plant height genes which the breeders have been unconsciously breeding for, affect yield and adaptation, including the early establishment although they certainly do. If longer coleoptile length is an important determinant of early establishment and vigour in hot and dry rainfed environments like Kazakhstan, we hypothesised that height related genes in wheat from Kazakhstan should not have a negative effect on the length of coleoptile. We also hypothesised that the identified height increasing alleles of the 5A and 6A QTL might increase coleoptile length simultaneously as the two can be correlated positively (Gulnaz et al., 2011). Therefore, the aim of the experiment was to assess effect of PH increasing alleles of 5A and 6A QTL on the coleoptile length.

3.3.1.7.2 Materials and methods

We followed the method "cigar roll" to screen the length of coleoptile (Bai et al., 2013). In total, the coleoptile length of 127 uniform-sized seeds was assessed. Out of 127, 20 and 12 assessions belonged to NIL5A(+) and NIL5A(-) respectively. NIL6A(+) and NIL(-) were represented by 26 and 17 samples respectively. Twenty seeds of double NILs (dNILs) which are NIL5A(+)_6A(+) were included to the experiment. Pamyat and Paragon were replicated 10 times each. *Rht-B1b and Rht-D1b*, on the background of Saitama and Alchemy respectively, were both represented 4 times. *Rht B1xD1* on the background of breeding line H117 was also replicated 4 times.

Overall, 12 cigar-rolls were prepared, 7 and 6 of which had 11 and 10 seeds respectively and the seeds in each cigar-roll were randomised. The seeds were placed 2cm apart horizontally on moist "blue-roll" towel-paper covered by A4 paper (Figure 3.31 and Figure 3.32). After that, cigar-rolls were wrapped and tightened in the middle with the rubber band. Once, the wrapping of all 12 cigar-rolls was complete, they were placed vertically in 1000mL glass beaker 1/3 of which was filled with water. Then, glass beaker was placed in a darkened black box with slightly opened lid to allow air flow and left at the room temperature for 6-7 days. After 7 days, the coleoptile length of each seedling was measured, in millimeters, from the base of the seed to the coleoptile tip (Figure 3.33).

Figure 3.31 Cigar-rolls for coleoptile length screening

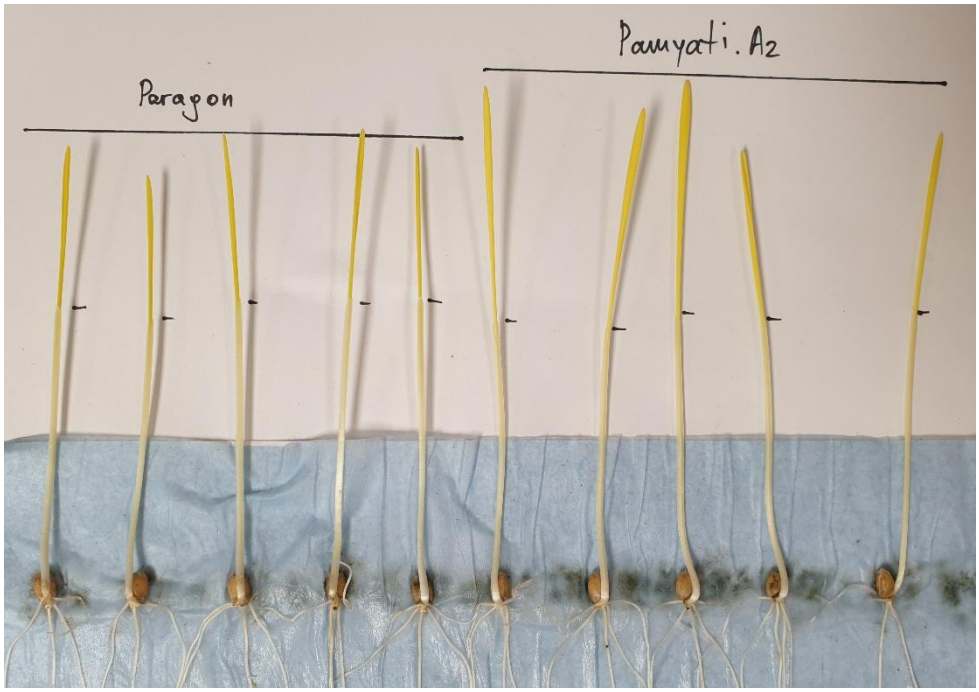


Figure 3.32 The length of coleoptiles of Paragon and Pamyati Azieva

Dash "-" = tip of coleoptiles

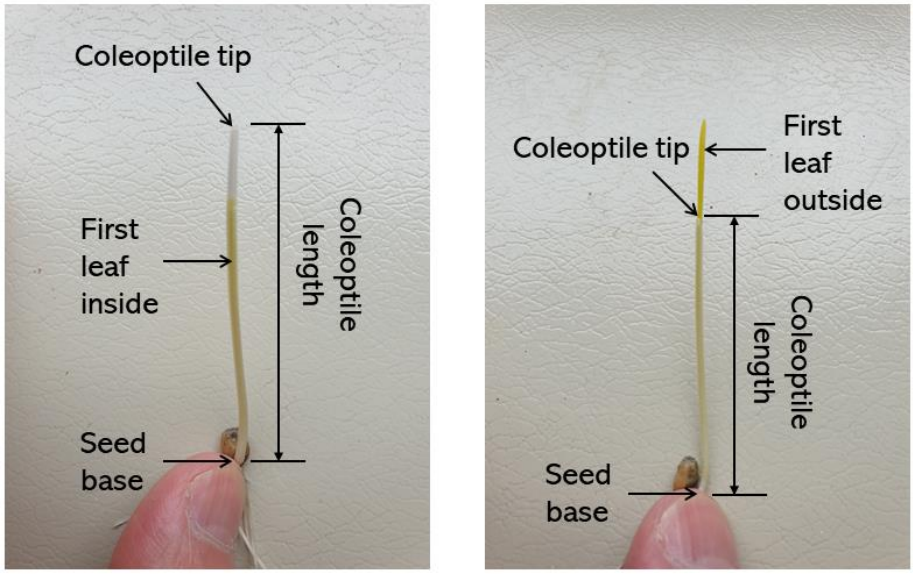


Figure 3.33 Coleoptile length measurement

On the left: First leaf inside the coleoptile. **On the right**: First leaf emerged from the coleoptile.

3.3.1.7.3 Experiment results

Overall statistical significance was detected as a result of the analysis of variance (p = ***, with F value of 12.6). To evaluate how different the factor levels are, we further conducted multiple comparisons analysis. Consequently, no difference in the coleoptile length between the parents or both NIL sets including the dNILs was observed. However, NIL sets, dNILs and parents had substantially longer coleoptiles than that of *Rht-B1b, Rht-D1b* and *Rht-B1b+Rht-D1b* controls (Figure 3.34). Interestingly, although coleoptiles did not differ between Pamyati Azieva and Paragon, NIL5A(+) and NIL5A(-) and NIL6A(+) and NIL6A(-), apparently the length of the first leaf differentiated clearly (Figure 3.32). Images of NIL sets are not given as they had almost the same height profile.

Rht-B1b+Rht-D1b severely reduced the coleoptile length by 71.1% relative to 5A and 6A QTL regardless of the allelic state. The sole negative effect of *Rht-B1b* and *Rht-D1b* on the length of coleoptile accounted for 29.2% and 44.6% respectively.

Based on our data, it is likely that UK wheat breeders also breed for height genes without compromising coleoptile length in spring wheat varieties such as Paragon.

Regarding the results of our experiment on coleoptile screening, tall and short alleles of the 5A and 6A PH QTL neither increase nor reduce the length of coleoptile.

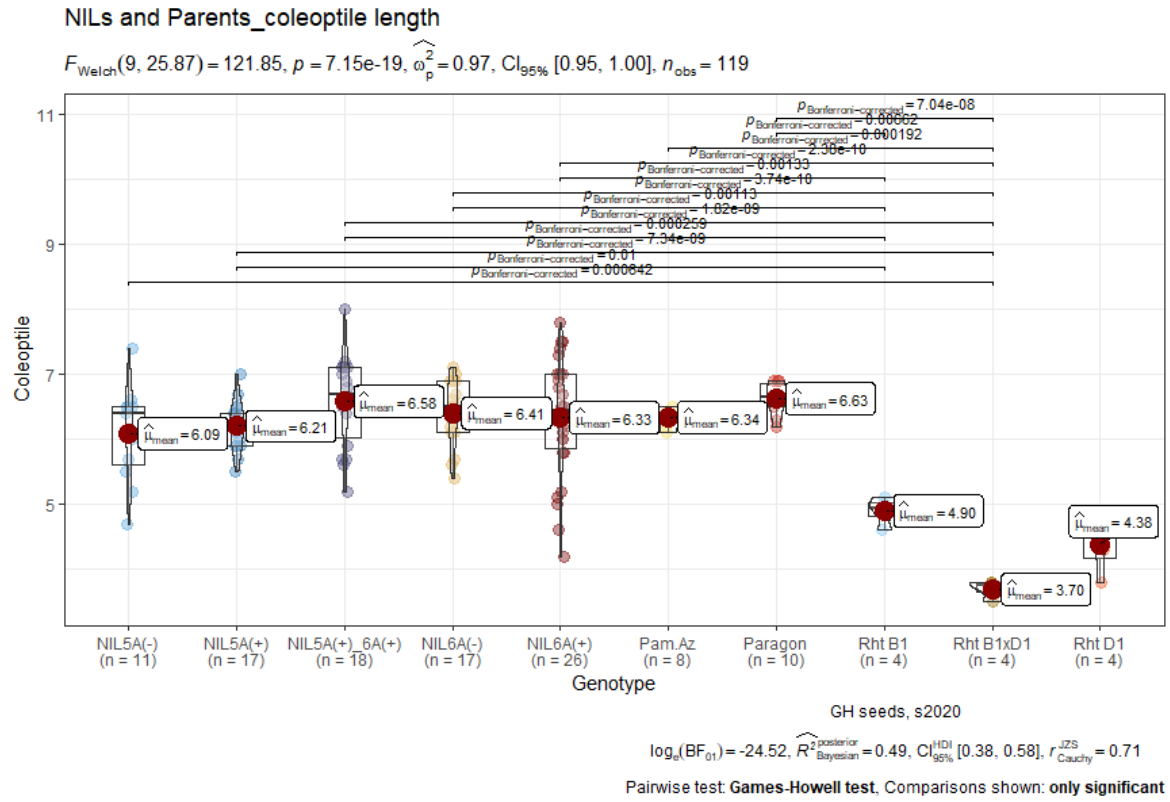


Figure 3.34 Coleoptile length of NILs compared to parents and widespread Rht genes

3.3.1.8 Stacking of two height increasing alleles in UK wheat

3.3.1.8.1 Introduction

The data so far has shown consistent height effects of the 6A QTL in the UK and KZ. While for 5A height effect was clearer in the UK but much less stable in KZ. In neither case was any height effect significantly associated with yield. It is possible that both alleles do improve performance on Kazakh conditions, but they need to be present together. QTL pyramiding has long been a successful way to stack the genes of agronomic importance into a common genetic background in crops. There are plenty of published studies, however to illustrate some, for example QTL pyramiding of eight QTL improved the grain quality and resistance of popular elite wheat cultivar PBW343 to rust diseases (Tyagi et al., 2014). The similar marker-assisted QTL pyramiding resulted in the development of a new genotype tolerant to blast and bacterial blight and vigorous cold at the fertilization stage in rice (Shinada et al., 2014; Jiang et al., 2015) and scald and mildew in barley (Hautsalo et al., 2021). Likewise, individual and combined effects of *Rht* genes on yield were evaluated (Flintham et al., 1997). Therefore, this section of the thesis aimed at developing NILs with doubled height effect via combining two height increasing genes on wheat 5A and 6A chromosomes in the background of UK wheat Paragon and assessing the performance of double NILs (dNILs) relative to single (sNILs), non (nNILs) and parents. The main trait we stressed was PH which can provide valuable insights into how the genes underlying the QTL act on crop development. In the long term (not possible within the 4 years of PhD) these materials will be

valuable resources to further test the main hypothesis of this work which is that adaptation to Kazakh conditions requires height increasing QTL.

3.3.1.8.2 Materials and methods

We have conducted ten reciprocal hybridisations (5AQTL x 6AQTL and 6AQTL x 5AQTL) of two PH QTL on 5A and 6A wheat chromosomes. Nine of ten crosses were successful. In order to self-fertilize these obtained nine hybrid lines, initially, they have been grown in CER (Controlled Environment Room) in 96 well tray with peat and sand soil type to germinate, then plants were transferred into 1L pots with cereal mix and grown alongside with recombinants in a greenhouse during the winter of 2018. Then, 96 seeds from each 9 plants were grown for DNA extraction and then genotyped with flanking markers of 5A and 6A QTL.

The double NILs for 5A and 6A PH increasing QTL with their counterparts, single and non-NILs (non-NILs lack the 5A and 6A PH increasing QTL regions of Pamyati Azieva and are almost Paragon, but still have some Pamyati Azieva background which might affect slight increase in plant height than real Paragon) were grown in glasshouse in three complete randomised balanced blocks. All plants were harvested and threshed. Plant height and height related traits, as well as yield and yield related traits were scored.

In the field experiment in Alm 2021, double, single and non-NILs with parents and check variety, Kaz10, were grown in two replications in $1m^2$ randomised plots. A random three plants were used to assess the plant height.

3.3.1.8.3 Results of the experiment

3.3.1.8.3.1 Self-fertilisation to recover homozygotes possessing two PH QTL

Although all ten crosses were successful, one cross provided only one seed and it did not survive while self-fertilising. Thus, nine hybrid lines have been harvested and threshed, and provided enough seeds for the further experiments.

As we adjusted all QTL flanking markers (start, peak and end) to Mendelian segregation factor, the expected segregation ratio for two genes was one out of sixteen (15:1) (Figure 3.35). Thus, 96 seeds from each 9-hybrid line (96 x 9 = 864 plants were grown in total) were subjected further for DNA isolation and then genotyping to find lines carrying both 5A and 6A increased plant height QTL. Because of limited seed number for some hybrid lines and germination issues, overall number of plants genotyped was 768, not 864.

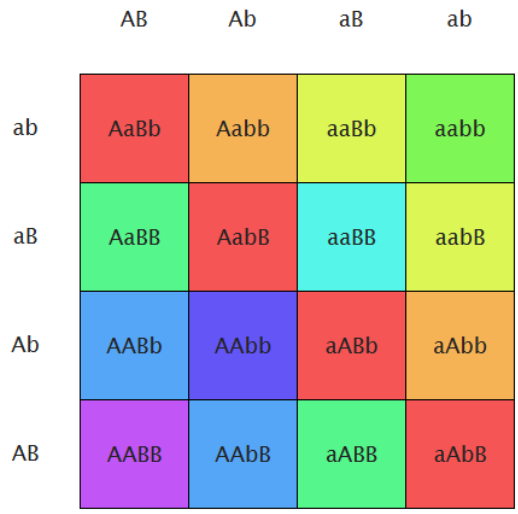


Figure 3.35 Segregation ratio of two PH genes

3.3.1.8.3.2 Genotyping to recover homozygotes possessing two PH QTL

For identification of plants possessing both PH increasing alleles of Pamyati Azieva 768 plants were genotyped by six molecular markers (3 flanking marker of 5A QTL and 3 of 6A QTL). Out of 768, 20 were double NILs (dNILs or NIL5A(+)_6(+)), 11 non-NILs (nNILs or NIL5A(-)_6A(-)) lacking both PH increasing alleles, but still have some Pamyati.Azieva genetic background), 10 plants for each single NILs (sNILs or either sNIL5A(+)_6A(-) or sNIL5A(-)_6A(+)) (Figure 3.36). All these recovered lines were grown in a glasshouse to multiply the seeds for further field experiments on them and to score the plant heights and other traits.

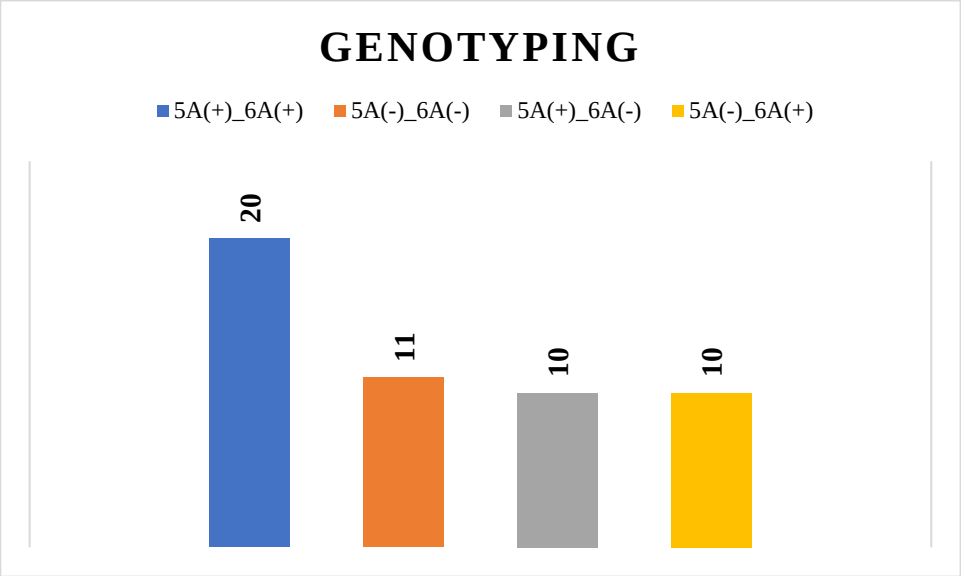


Figure 3.36 Genotyping results for dNIL recovery

3.3.1.8.3.3 Glasshouse experiment on double NILs

During phenotyping, the tallest (canopy) stem of each plant was measured for plant height. In addition, the same stem (the tallest) was used to measure internodes. For spike length (SL), mainly the main spike length was scored. We counted number of spikelets of a main spike for the spikelet number per spike (NSMS). Also, thousand grain weight (TGW), seed number per plant (SNP), seed width (SWI) and length (SLE) were measured. However, only plant height data will be shown as at this exact stage we are less interested in comparative yield performance.

Unequal spread of light

The fact to note is that, due to the unequal spread of light effect on the top canopy of plants within and between blocks, the data on PH was analysed separately for plants under the direct and indirect light cover. Statistics, including such sample parameters as population mean, the 95% lower and upper confidence intervals, standard deviation, minimum/maximum values, ranges and standard error of the mean were calculated. One-sample t-test was conducted to find out the 95% lower and upper confidence intervals for the means of each light effect, No (direct) and Yes (indirect), separately (Table 3.9).

Light effect	Sample size	Mean	LCI	UCI	SD	Min	Max	Range	SE
No (direct)	28	97.73	93.73	101.72	10.3	80	118	38	1.95
Yes (indirect)	30	103.8	100.26	107.33	9.46	86	119.5	33.5	1.73

Table 3.9 Light effect on plant height

Estimated difference between the means of plant height of the plants under and outside of the direct light effect was 6 cm. Despite the observed 6cm difference for minimum plant height, the maximum values were close with only 1.5cm difference. Sample parameters as population mean (Mean), standard deviation (SD), minimum (Min) and maximum (Max) values, Ranges, standard error (SE) and standard deviation (SD) of the mean were calculated with the "describe" function of "phych" library in R. One-Sample t-test was used to determine the 95% lower and upper confidence intervals for the means.

To determine whether the difference in the means between two groups (No and Yes) was significant or not, two-sample t-test was used. Prior to conduct hypothesis testing (two-sample t-test), first we checked whether the data distributed normally or not and second, equality/homogeneity of variances of two groups (yes, no) was tested. As a result, plant height data with/without a direct effect of the light showed the relative normal distribution and the p-value for pooled variances of two groups was not significant, p-value = 0.65, which states the homogeneity of variances. Two-sample t-test revealed that there is statistically significant evidence of a difference in the mean of plant height from glasshouse with/without light effect, with p-value of 0.022* (Figure 3.37).

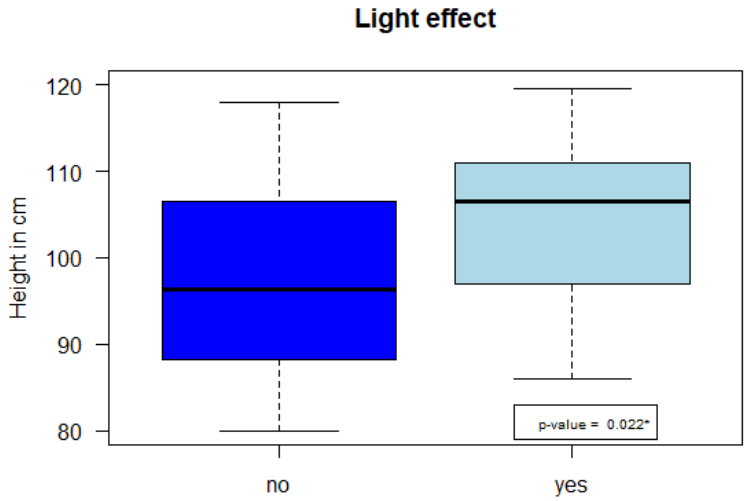


Figure 3.37 Two-sample t-test of light effect on plant height

Figure shows the difference in plant height of plants (double, single and non-NILs) with/without (yes/no) direct light effect. The difference between the means of with and without light effect was 6 and standard error of difference between the means with and without light effect was 2.5. Confidence intervals at the 95% was between 0.9 and 11.3 (if the difference in the means is 6, CI mean that samples were shorter by 5.9 cm (6-0.1) and taller by 5.3cm (11.3-6) respectively).

Despite the significant difference between the plants under direct and indirect lights, the double NILs were the tallest, and single and non-NILs were medium and shortest in both treatments respectively. The recurrent parent, Paragon, was even shorter than non-NILs in both treatments. Thus, although the light affected only overall height, the relative heights of genotypes were unaffected. ANOVA was conducted to evaluate the proportion of variance in the data explained by genotypes, light effect and replication (blocking) and to retrieve p-values (Table 3.10).

	DF	Sum Sq	Mean Sq	F-value	P
Genotype	4	3171.8	792.95	24.6287	3.258e-11 ***
Light	1	867.7	867.66	26.9491	4.002e-06 ***
Rep	2	150.7	75.35	2.3403	0.107
Residuals	49	1577.6	32.2		
Total	56	5767.8	1768.16		

Table 3.10 The proportion of variance in the plant height data explained by factors

Table 3.10 represents ANOVA summary. The model was fitted to assess the effects of genotype, light and replication effects on plant height. Genotype and light showed the significance, but not the replication effect. The proportion of total variation in plant height phenotype explained by genotype and light were 56% and 15% respectively. A bit more than 2% variation was explained by replications. The rest ~27% remained unexplained.

Linear Mixed Models

Once the proportional variation in plant height was calculated, Linear Mixed Models (LMM) was used to estimate such trait parameters as plant height (PH), the lengths of internodes, ETN, SL and NSMS (SNPS). During data analysis (LMM), the replication and plant number in each replication

was set as a random effect to analyse data for PH and all internode lengths as the replication effect was not significant. Therefore, genotypes (double, single and non-NILs) and the light effect were treated as fixed effects in the mixed model, because the means of both were statistically significant (Table 3). However, the replication effect was significant when SL was analysed. Thus, the genotypes, light effect as well as replication were treated as fixed effect in LMM.

Before conducting pairwise comparisons, Bartlett's test was used to determine homogeneity of variances for the genotype and light. The variances of both were not significant (p-value = 0.5063 at the degrees of freedom = 4 for the genotype effect and p-value = 0.4833 at the degrees of freedom =1 for the light effect).

Starting with PH as it is the main plant characteristic under the current thesis, its overall p-value was statistically significant. Moreover, to estimate the differences between the means of genotypes, pairwise analysis was performed. The results of pairwise comparisons between genotypes/contrasts revealed that double NILs were significantly different from 5A and 6A NILs, non-NILs and Paragon. Plant height means of NIL5A(+) and NIL6A(+) were not statistically different between each other although they were different from nNILs and Paragon. There was no significant difference between nNILs and Paragon. Additive effect of one QTL was ~10cm, of both was at around 20cm (Figure 3.38).

Figure 3.38 The height of Paragon double-, single- and non-NILs and wild type Paragon

P values are adjustment by Bonferroni's method.

3.3.1.8.3.4 The first field testing of the dNILs

Stacking the two height increasing loci of Pamyati Azieva made Paragon taller than either single NILs (sNILs). However, no significant difference between double and the 6A single NILs was observed although almost 2cm mean height increase was identified between the two (Figure 3.39). Interestingly, in Alm 2021 experiment, the local check variety Kaz10, which was significantly taller than any of the single NILs in previous experiments in Alm, did not show a significant height increase in relation to sNIL6A. We would like to conduct more experiments in the future to validate whether the observed height differences are stable across the seasons. However, it was impossible within the scope of this PhD work due to time limitations. In future experiments we will also study the association between the double QTL which increases height and other traits of agronomic importance.

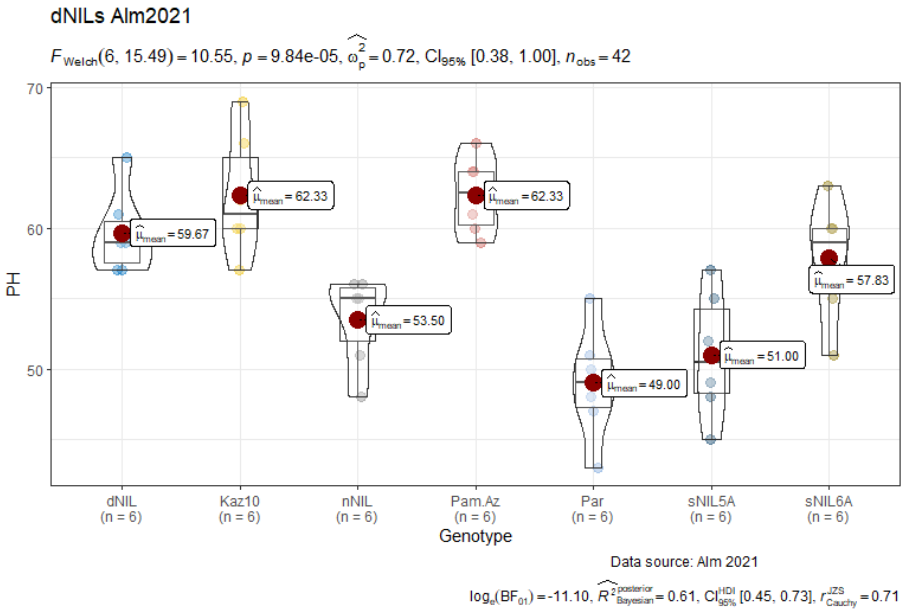


Figure 3.39 Plant height of single-, double-, non-NILs, parents and Kaz10 from the first experiment on dNILs in Alm2022

3.3.1.9 Quality

The grain quality is an extremely important characteristic for Kazakh wheat breeders. Therefore, we have evaluated quality related traits such as the content of moisture, protein, fat, crude fiber, ash and starch. Among these, protein content is paramount as this is the foremost characteristic which attracts Kazakh wheat customers. Thus, we will focus on protein content outcome. It is important to note that grain quality parameters were assessed by the "Laboratory of biotechnology, physiology, plant biochemistry and product quality assessment" of the LLP "Kazakh Research Institute of Agriculture and Plant growing" based on the Kjeldahl method according to GOST 10846-91 (GOST (Russian: ГОСТ) refers to a set of international technical standards maintained by the Euro-Asian Council for Standardization, Metrology and Certification (EASC), a regional standards organization operating under the auspices of the Commonwealth of Independent States (CIS)).

We used one-way t-test in R to calculate p-values. Prior to significance test, we calculated the mean and all its corresponding properties. The protein content of NILs in Almalybak was as high as that of parents. Interestingly, local standard cultivar Kaz10 had less protein relative to NILs, Pamyati Azieva and Paragon (Table 3.11 and Table 3.12).

In Pet, overall protein content was not as high as in the Alm. NILs of 5A PH QTL had almost the same protein. However, it is likely that 6A PH QTL positively affected quality improving the protein content by 0.2 and 0.7% in Alm and Pet respectively. In both environments NIL6A(+) had the highest values for protein exceeding even Pamyati Azieva. It is important that NILs could preserve a high content of protein. Paragon's ability to make high protein in the Alm was not detected in Pet.

Genotype	Protein (%) mean	SD	SE	CI	DF	P
NIL5A(-)	14	0.64	0.45	5.72	11	***
NIL5A(+)	13.9	0.71	0.5	6.35		
NIL6A(-)	14.1	0.21	0.15	1.91		
NIL6A(+)	14.3	0.07	0.05	0.64		
Pam.Az	14.1	0.42	0.3	3.81		
Paragon	14.1	0.92	0.65	8.26		
Kaz10	13.4	0.14	0.1	1.27		

Table 3.11 Protein content, Alm 2020

SD = Standard Deviation, SE = Standard Error, CI = Confidence Interval, DF = Degrees of Freedom, P = overall significance level from one-way t-test.

Genotype	Protein (%) mean	SD	SE	CI	DF	P
NIL5A(-)	12.8	0.28	0.2	2.54	13	***
NIL5A(+)	12.7	0.07	0.05	0.64		
NIL6A(-)	12.2	0.28	0.2	2.54		
NIL6A(+)	12.9	0.14	0.1	1.27		
Pam.Az	12.8	0.14	0.1	1.27		
Paragon	12	0.21	0.15	1.91		

Table 3.12 Protein content, Pet 2020

SD = Standard Deviation, SE = Standard Error, CI = Confidence Interval, DF = Degrees of Freedom, P = overall significance level from one-way t-test.

3.4 Discussion

3.4.1 Initial hypothesis

The power of a reductionist approach was used to test the main hypothesis of this work by taking phenotypic data from homogeneous lines, that, as much as possible differed only for the trait of interest. This is useful in dissecting the specific trait effects and pleiotropy for other potentially important adaptation and yield related traits. Therefore, when the 5A and 6A height increasing QTL with large additive height effects (~10cm) were identified, both were considered as possibly important in improving wheat adaptation in Central Asia (CA). It is interesting that CA still produces tall wheat varieties when most parts of the world have seen selection towards shorter varieties. This is even true for Canada with environmental conditions quite similar to northern Kazakhstan (Kaut et al., 2009; Chen et al., 2016). It is generally accepted that this trend of height reduction results in increased grain yield as a result of increased dry matter allocation to grains and it certainly reduces the risk of lodging (Rebetzke and Richards, 2000; Gulnaz et al., 2011). These points made it interesting to ask how the single and combined height increasing QTL perform and interact respectively in an isogenic background.

To address these questions and identify the importance of height increasing alleles, located on wheat 5A and 6A chromosomes, in relation to other plant characteristics near isogenic lines were developed carrying these alleles in the genetic background of the UK spring wheat Paragon. Subsequent large scale field experiments in the UK and Kazakhstan shed light on the performance of these height related alleles across these two nations.

3.4.2 The 6A QTL stably promotes PH in KZ but the 5A QTL exhibits strong GxE

In these experiments the plant height was prioritised as it is the main trait of interest being studied in this thesis. It is important to see that QTL studies are validated and that the NILs are fit for the purposes proposed. The association of plant height with other significant adaptation and yield components was also assessed. Based on the data obtained, the height controlling mode of action of these two QTL seems different. The effect of the 6A locus on height is stable being observed in all experiments in the UK and Kazakhstan including the controlled environment. However, the height effect of 5A QTL was silenced in four out of five experiments in Kazakhstan, only expressing its height effect in Alm 2020. By contrast, in Alm 2019 and 2021 the (+) allele, compared to (-), was shorter in the latter season with a p -value of 0.06, so representing GxE crossover event. In the two Pet experiments no height difference was ever identified between NIL5A(+) and NIL5A(-) (Figure 3.10).

3.4.3 Grain yield and its components compared between NIL sets

Looking at yield components of the 5A QTL isogenic lines, we saw no significant effect on grain yield in Alm 2019 and 2020, instead a significant reduction in YP associated with the “tall” allele

was observed in Alm 2021 (p = 0.04). In this season yield components such as NKP, GWMS and HI also reduced in NIL5A(+) expressing relatively lower p-values of 0.17, 0.08 and 0.08 at the 5% significance level. So, in the Alm experiments the lack of a height effect means that no comments can be made on the idea that increasing height might confer some benefit in Kazakhstan. However, there is some evidence that the 5A Pamyati Azieva allele, which certainly does increase height in the UK, has some positive effect on yield components in this northern region which is the main wheat growing area in Kazakhstan and the target environment of varieties like Pamyati Azieva.

In the northern site (Pet) there was a non-significant mean increase in both years (Pet 2020 (p = 0.37) and 2021 (p = 0.90)) in favour of 5A(+). It must be stressed that this is not supported by statistical significance but there was certainly no reduction in yield as a result of this introgression. Looking at potential causes of any yield increase that leads to the observation of increased number of plants (NP_1m^2, p = 0,01 and p = 0.44 in Pet 2020 and 2021 respectively). In addition, the NIL5A(+) produced more spikes (SN_1m^2) and seeds (SSM) compared to NIL5A(-), but not significantly.

Looking for the hypothesised beneficial effects of the tall allele of the 6A QTL it was shown that a significant increase in yield components associated with the tall Pamyati Azieva allele was never observed. Although these comparisons were not statistically important, the p-values were lowest in the seasons when a mean yield reduction was observed. Moreover, the NIL6A(+) in yield components such as NKS, NKP, GWMS, GWP and HI significantly underperformed when compared to NIL6A(-) in Alm 2020 and 2021 with only HI being non-significant (p = 0.26).

3.4.4 Plant height components of 5A and 6A height effects

While the height effect of the 5A(+) allele was not observed in Kazakhstan, it certainly made the internode-3 significantly longer in tall NILs compared to the short in Alm (only Alm 2021 was not significant). It was also obvious when its relative proportion was calculated. In addition, the internode-2 of NIL5A(+) increases significantly relative to NIL5A(-) when its positive height effect is retained in Alm, but not in Pet. However, there is an insignificant mean increase in absolute values, as in proportional percentage, in Pet. Considering that some significant increase in yield components was observed it is interesting to speculate on how this change in the partitioning of internode growth might have influenced these yield components. One theory is that by reducing biomass commitments to internodes more resources can be allocated to the growing spike (Rivera-Amado et al., 2019). It could be that the 5A+ allele increases the strength of source traits at the time that internode 3 is extending. This results in a longer internode but possibly the developing spike also benefits from a “surge” in resources allowing more grain to be set.

The tall allele of the 6A QTL has a significant negative effect on the length of the main spike in Alm, but not in Pet. However, it stably increased the length of internode-1 (peduncle length) across environments and growing seasons. When it comes to internode-2 and -3, they were significantly longer in NIL6A(+) relative to NIL6A(-), again in Alm only (only Alm 2021 was not significant). This shows that tall alleles at the 5A and 6A loci possess an overlapping positive influence on internode-2 and -3 in Alm.

3.4.5 The 5A and 6A PH effects do not increase coleoptile length

It was shown that 5A and 6A height increasing QTL do not correlate positively with coleoptile length. Our hypothesis "increased height mostly increases coleoptile length and thus does early establishment and growth" (Gulnaz et al., 2011) was rejected as the experiment showed that they neither increase nor reduce the coleoptile length (Figure 3.32 and Figure 3.34). However, the increasing effects of both QTL on the first leaf length were clearly visible.

3.4.6 The 5A and 6A PH QTL independently increased the Paragon height

Stacking the two height increasing alleles stably increased the height of Paragon by about 20cm in the glasshouse and field in the UK. This is the additive effect of two height increasing alleles of 5A and 6A QTL when combined. Bearing in mind the hypothesis that a height increase could be beneficial for yield in Kazakhstan it would be interesting to grow double NILs (tall) in environments such as Pet. This might result in the sub significant yield increases of 5A being expressed more fully.

3.4.7 The quality remains intact

Quality is a complex trait incorporating many grain characteristics but protein content is a major factor. The dynamics in protein content is mainly dependent on environmental changes and the level of nitrogen applied although the genetic impact also should not be neglected (Hussain, 2002; Iqtidar et al., 2006; Chen et al., 2012). As an example, our data showed that approximately 1% less of protein was synthesised in Pet compared to Alm, but it was within the range of the country's quality demand which meets the "superior" and "medium" quality grades (Table 3.11 and Table 3.12). NIL5A pairs shared almost the same quality with minor 0.1% difference in both regions. The tall 6A NIL possessed the highest content of protein in both regions (perhaps this is the main reason of breeding for 6A tall allele), which is an important criteria in Kazakhstan, although previous reports show a negative linear relationship between PH and grain quality in crops (Amirthadevarathinam, 1983; Ya-wen et al., 2005; Islam et al., 2014; Barley, 2021). Therefore, it is likely that the grain quality in Kazakh bread wheat is independent of PH. Whereas Paragon's quality was as good as Pamyati Aziva in the Alm, its protein dropped to the lowest in Pet. Surprisingly, the local check variety, Kaz10, in Alm had also the lowest protein content. Having

said all of these, we have to admit that the quality check results are available for growing season 2020 only.

3.5 Summary

The 5A and 6A Paragon NILs validated the height QTL identified. They provided some evidence that 5A height increase could be of benefit in terms of yield (at least some of yield components), but this is less likely for 6A. The effects of the QTL are additive in the UK but the 5A effect is almost always lost in KZ experiments.

4. Moving towards the identification of causal genes

4.1 Development of recombinants for 5A and 6A QTL regions

4.1.1 Introduction

Once the trait of interest is identified and validated in a mapping population and isogenic background respectively, most geneticists try to get as much accurate markers as possible co-segregating with the trait and to narrow down and/or identify a causal gene/s underpinning the trait of interest. Traditionally, it is achieved through map-based cloning methods. Recombinant development is one of these techniques and thus has largely been used in fine-mapping procedures. Recombinants are generated directly from genotyping large number of segregants at any stage of the hybrid/filial development or/and through several backcrosses of the hybrid line of interest within the population to the recurrent parent and identifying any crossovers (Xue et al., 2013; Wang et al., 2021). The latter technique takes longer but possess greater accuracy as diminishes the possible background genetic noise. There are large number of examples when QTL regions were squeezed, mapped physically and fine-mapped in wheat based on recombinants. For instance, Liu with co-workers used 382 recombinants derived from F_7 heterozygous plants to fine-map resistance to Fusarium head blight in wheat (Liu et al., 2006). Wheat flag leaf width gene, *TaFLW1*, which increases photosynthetic capacity and yield potential was mapped using one hundred and ten recombinants with homogenised genetic background which belonged to ten recombinant classes (RC) (Xue et al., 2013). Sixty-one recombinants identified in F_2 population were helpful to map the gene, *TIN4*, contributing to ideal plant architecture in common wheat (Wang et al., 2021). Similarly, we also developed 43 and 52 homozygous recombinants, which belonged to 8 and 12 RC, for 5A and 6A PH QTL respectively.

4.1.2 Materials and methods

4.1.2.1 Identification of heterozygous recombinants for 5A and 6A QTL chromosomal regions from BC2F2 seeds

Besides NILs, genotyping results of the 94 BC2F2 seeds with QTL flanking markers allowed us also to identify 17 and 29 heterozygous recombinants, for 5A and 6A PH QTL respectively (out of initial 94 samples). These heterozygous recombinants for 5A and 6A PH QTL region have been

classified into 7 and 9 recombinant classes respectively according to their recombination combination (data not shown). However, the number of recombinant classes changed while saturating both QTL regions with DNA markers from exome capture and whole genome sequencing of NIL parents. Ultimately, we ended up having 8 and 12 recombinant classes for 5A and 6A PH QTL respectively.

4.1.2.2 Recovery of homozygous recombinants for 5A and 6A QTL chromosomal regions from heterozygotes

Twelve seeds of each of these heterozygous recombinant lines, for both 5A (17) and 6A (29) QTLs, were sown in 96 well tray with peat and sand soil type and initially grown in CER. Thus, DNA was isolated from self-fertilized 552 heterozygous recombinants to identify and grow only homozygous recombinants in 1L pots with cereal mix soil type in a glasshouse. All 552 heterozygous recombinants, of which 204 plants (17x12=204) for 5A QTL and 348 (29x12=348) plants for 6A QTL, were genotyped with QTL flanking markers. As a result, we identified 43 and 52 homozygous recombinants, which belonged to 8 and 12 RC, for 5A and 6A QTL respectively.

4.1.2.3 The height assessment of recombinants in controlled and non-controlled environments

Plant heights of identified homozygous recombinants for both QTL on 5A and 6A wheat chromosomes were assessed in controlled and field conditions in the years of 2018 and 2019 respectively.

In a glasshouse, each homozygous recombinant was sown and grown in 1L pots with the cereal mix soil and experimental design incorporated completely randomised blocks.

In order to support the initial glasshouse data used for medium resolution mapping of genes for both increased plant height QTL, all recombinant lines of two QTLs were sown in randomised $1m^2$ plots in the field of JIC in three replicates. Each $1m^2$ plot was represented by 6 rows. The experimental design of each $1m^2$ plot consisted of three rows of Paragons along with three randomly selected recombinant lines, regardless of their RC and the total number of $1m^2$ plots in the experiment was 140. Two out of the three Paragons in each $1m^2$ were sown on the margins of each plot to eliminate the edge effect on recombinants. We measured the heights of 5 plants per line in the plot ($1m^2$). Thus, in total 15 plants per line were involved in height measurement in the field. This helped to compensate for the relatively small sample size in the glasshouse.

4.1.2.4 Statistical analysis

Analysis of Variance (ANOVA) was carried out to check the statistical significance of means of recombinant heights for 5A and 6A QTL. Bartlett's test was used to determine whether RC

variances are homogenous or not. Post-hoc statistical test between RC was evaluated through calculating a common SD for all groups and this in turn was used for all multiple comparisons. P-values for multiple comparisons were adjusted based on the Bonferroni's method. Test for association between recombinant classes for both QTL was performed using Pearson's product moment correlation coefficients. All statistical procedures were computed and plots were conducted and produced in R respectively.

4.1.3 Results

4.1.3.1 Homozygous recombinants were identified for 5A and 6A PH QTL region

As a result of genotyping the heterozygous recombinants, out of 204 plants, 56 lines were chosen, 44 of which were homozygous recombinants and 12 new heterozygous recombinants for 5A QTL. By contrast, out of 348 heterozygous recombinants 188 plants have been chosen, 70 of which were homozygous recombinants and 48 new heterozygous recombinants for 6A QTL. Among these, 44 homozygous recombinants for 5A QTL and 70 homozygous recombinants for 6A QTL have been grown in 1L pots with cereal mix soil in the glasshouse in randomised blocks. New heterozygous recombinants were used to generate new homozygous recombinants, but the experiment was not completed because of the COVID-19 pandemic.

4.1.3.2 Height difference between recombinant classes for 5A and 6A QTL in the glasshouse and field

4.1.3.2.1 Height measurements on homozygous recombinants in the glasshouse

Plant heights of all homozygous recombinant plants were measured. The sample size of the glasshouse experiment was smaller than that of the field conditions. To check the normal distribution of both data, all recombinant classes (RC) were plotted as a population rather than each RC independently because of the small sample size in the greenhouse. Q-Q plots were produced using R by plotting empirical quantiles of the data against the theoretical standard normal quantiles. Height data points of the recombinants in glasshouse and field conditions followed the trend line quite closely which indicates that heights of recombinants for 5A and 6A QTL were normally distributed (Figure 4.1).

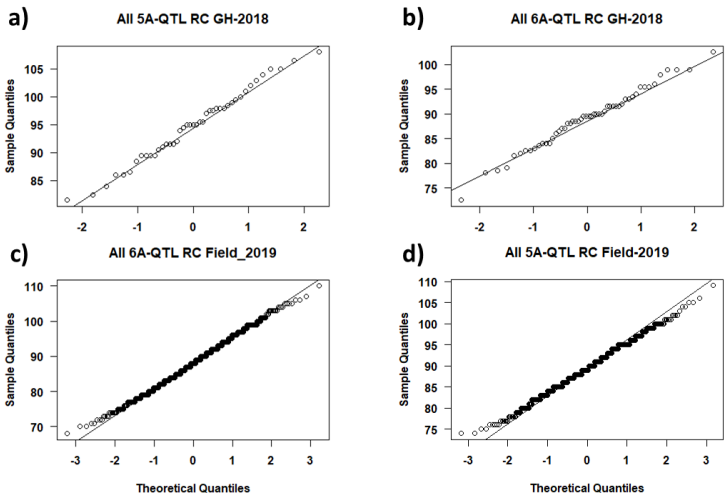


Figure 4.1 Q-Q plots show PH data distribution of 5A and 6A recombinants

PH data of homozygous recombinant plants for 6A (leaf: a) and c)) and 5A (leaf: b) and d)) was normally distributed.

Analysis of Variance revealed significant difference among the means of both QTL with F value of 9.0807 and P value of 3.576e-06 (***) for 5A-QTL and F value of 4.0405 and P value of 0.001733 (**) for 6A-QTL. The assumption of homogeneity of variances between RC for 5A QTL was not validated as result of the Bartlett's test. Thus, RC variances were significantly different from each other ($p < 2.2e-16$ in the field and $p= 0.03$ in the glasshouse) (Table 4.1).

Recombinant classes	mean	SE	df	LCI	UCI
RC-01	85.7	2.8	35	80	91.4
RC-02	91.9	1.62	35	88.7	95.2
RC-03	90.1	1.98	35	86.1	94.1
RC-04	93.9	2.17	35	89.5	98.3
RC-05	96.5	3.43	35	89.5	103.5
RC-06	96.8	1.98	35	92.7	100.8
RC-07	102.9	2.17	35	98.5	107.3
RC-08	99.4	1.83	35	95.6	103.1

a)

Recombinant classes	mean	SE	df	LCI	UCI
RC-01	99	3	40	92.9	105.1
RC-02	94.2	2.45	40	89.2	99.1
RC-03	91.1	1.6	40	87.9	94.4
RC-04	83.5	2.45	40	78.6	88.4
RC-05	89.5	4.24	40	80.9	98.1
RC-06	87.6	1.9	40	83.8	91.4
RC-07	88.7	1.73	40	85.2	92.2
RC-08	85.7	1.41	40	82.9	88.6
RC-09	87.1	2.12	40	82.8	91.4
RC-10	89.8	3	40	83.7	95.8
RC-11	97.7	1.9	40	93.9	101.5
RC-12	82.4	1.9	40	78.6	86.2

b)

Table 4.1 Plant heights of RC for 5A (a) and 6A (b) QTLs in the glasshouse

RC = "Recombinant Classes, mean = "Mean", SE = "Standard Error of the Mean", df = "Degrees of Freedom", LCI (Lower Confidence Interval) and UCI (Upper Confidence Interval).

Therefore, pairwise comparisons between the RCs for both loci were conducted based on the calculation of a common SD for all RC and this in turn was used for all multiple comparisons. Pooled SD (calculating a common SD) can be useful if some groups are small as in our case. We also used Tukey's honestly significant difference (HSD) test (with 95% family-wise confidence level) and Games-Howell's HSD method to assess the significance between RC. Obtained outcomes from these three ways of calculating significance level between the groups were mostly consistent. Thus, the results of t - tests with pooled SD with p – values calculated using Bonferroni's method are provided for all multiple comparisons for both PH QTL (Table 4.2 and Table 4.3)

	RC-01	RC-02	RC-03	RC-04	RC-05	RC-06	RC-07	RC-08
RC-01	-	0.97	1.00	0.47	0.39	0.06	0.00	0.01
RC-02	-	-	1.00	1.00	1.00	0.97	0.01	0.10
RC-03	-	-	-	1.00	1.00	0.44	0.00	0.04
RC-04	-	-	-	-	1.00	1.00	0.12	0.97
RC-05	-	-	-	-	-	1.00	1.00	1.00
RC-06	-	-	-	-	-	-	0.74	1.00
RC-07	-	-	-	-	-	-	-	1.00
RC-08	-	-	-	-	-	-	-	-

Table 4.2 Pairwise comparisons between RCs of 5A QTL using t tests with pooled SD

P – values calculated using Bonferroni's method. PH data were collected from the glasshouse.

	RC-01	RC-02	RC-03	RC-04	RC-05	RC-06	RC-07	RC-08	RC-09	RC-10	RC-11	RC-12
RC-01	-	1.00	1.00	0.01	1.00	0.12	0.22	0.01	0.12	1.00	1.00	0.00
RC-02	-	-	1.00	0.17	1.00	1.00	1.00	0.22	1.00	1.00	1.00	0.02
RC-03	-	-	-	0.56	1.00	1.00	1.00	0.67	1.00	1.00	0.53	0.05
RC-04	-	-	-	-	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
RC-05	-	-	-	-	-	1.00	1.00	1.00	1.00	1.00	0.94	1.00
RC-06	-	-	-	-	-	-	1.00	1.00	1.00	1.00	0.03	1.00
RC-07	-	-	-	-	-	-	-	1.00	1.00	1.00	0.05	0.83
RC-08	-	-	-	-	-	-	-	-	1.00	1.00	0.00	1.00
RC-09	-	-	-	-	-	-	-	-	-	1.00	0.03	1.00
RC-10	-	-	-	-	-	-	-	-	-	-	1.00	1.00
RC-11	-	-	-	-	-	-	-	-	-	-	-	0.00
RC-12	-	-	-	-	-	-	-	-	-	-	-	-

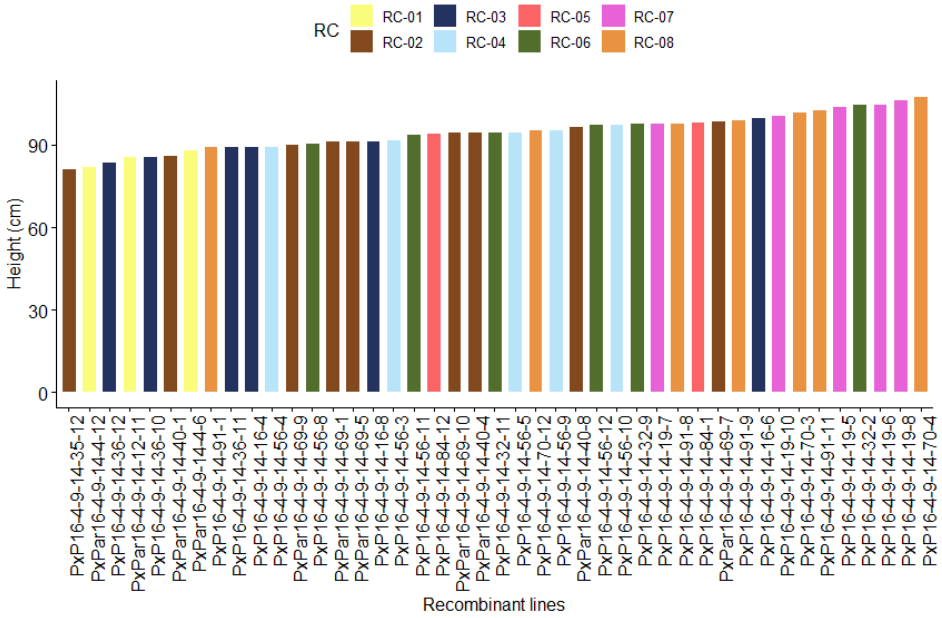
Table 4.3 Pairwise comparisons between RCs of 6A QTL using t tests with pooled SD

P – values calculated using Bonferroni's method. PH data were collected from the glasshouse. Data were derived as for Table 4.2.

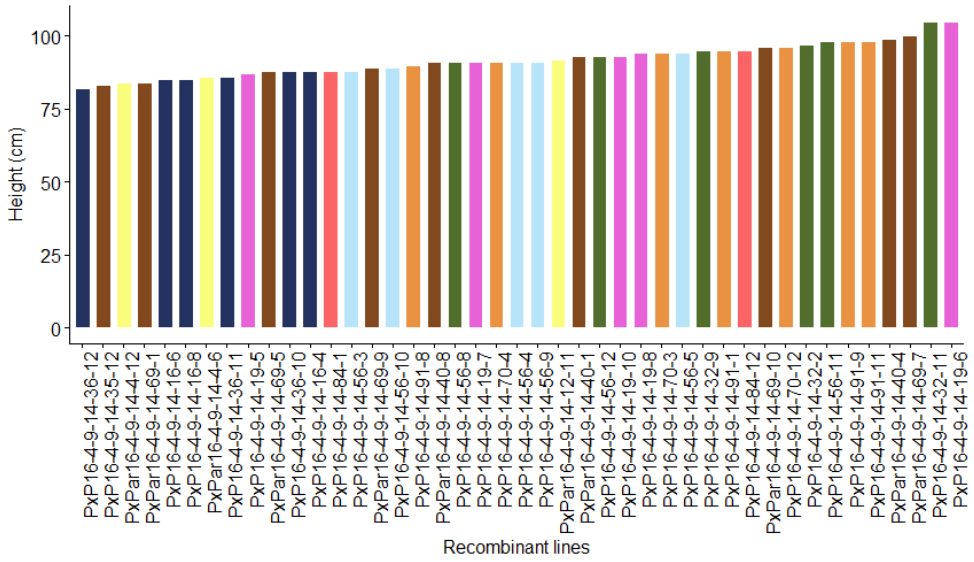
Despite the fact that obtained data from glasshouse was slight variable because of a smaller sample size, glasshouse effects such as unequal watering, spread of sunlight, light shed and heating, it was extremely helpful to conduct initial medium resolution mapping to identify potential coordinates of the genes, underpinning QTL, in the wheat genome.

4.1.3.2.2 Height measurements of homozygous recombinants in the field and comparison with the data from glasshouse

Plant heights of all recombinants were measured in the glasshouse and field and the data were processed using appropriate methods of statistics. The small sample size of the glasshouse height data was compensated by the height data obtained from field trials. Overall, plant heights of 645 and 776 plants for 5A and 6A QTL respectively were measured. Data from glasshouse well matched the data from the field. Thus, short recombinant classes in the glasshouse were relatively short in the field and tall recombinant classes in the glasshouse were relatively tall in the field (Figure 4.2 and Figure 4.3)



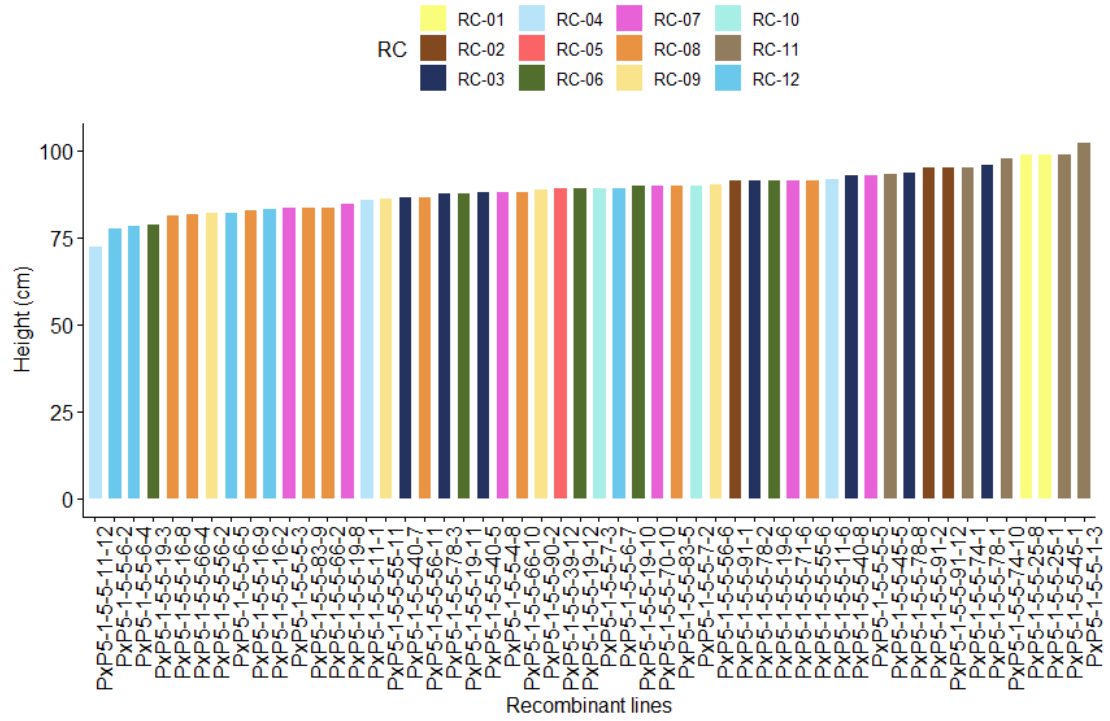
a)



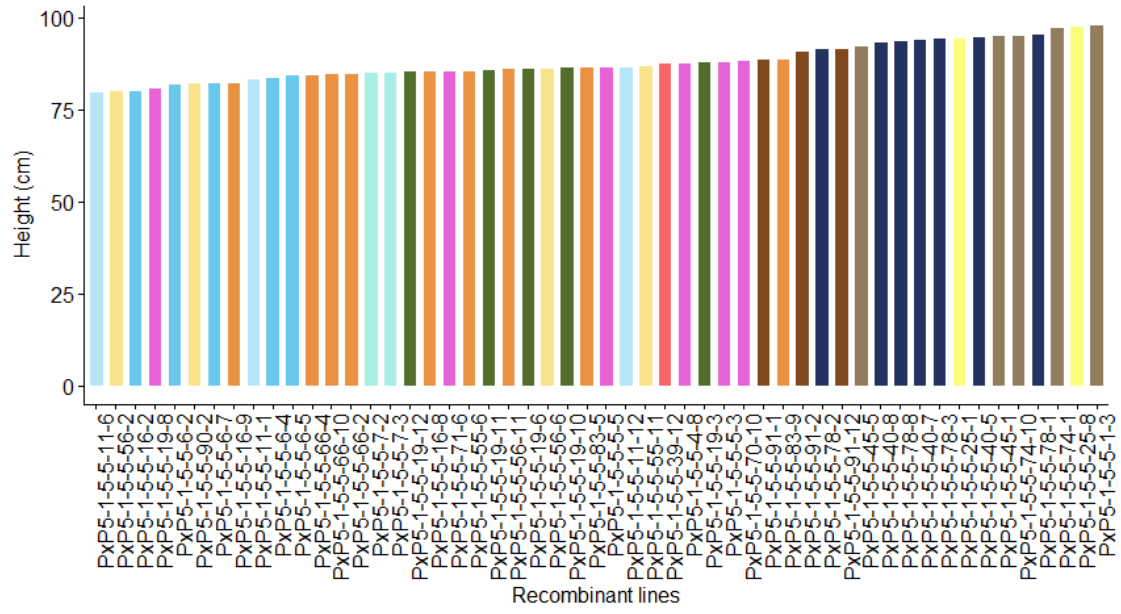
b)

Figure 4.2 Height data on RC for 5A QTL in the greenhouse (a) and field (b)

Note: Barplots of the field data include only random 43 samples out of 645.



a)

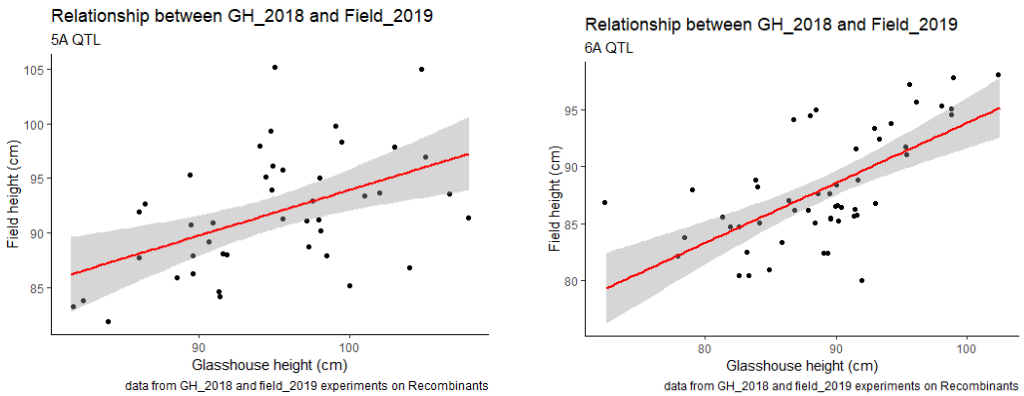


b)

Figure 4.3 Height data on RC for 6A QTL in the greenhouse (a) and field (b)

Note: Barplots of the field data include only random 52 samples out of 776.

Pearson's correlation coefficient test was conducted for the data on plant height of recombinant lines from the glasshouse and field both 5A and 6A QTL. Results showed some sort of linear relationship between plant height data obtained from the glasshouse and field (Figure 4.4). Samples followed independent normal distributions.



a) b)

Figure 4.4 Correlation between glasshouse and field data on plant heights of recombinants

Scatter plots represent the positive correlation between glasshouse and field data on plant height with P value of 8.94e-04 (***), correlation coefficient of 0.49 and Confidence Intervals at 95% was 0.22- 0.69 for 5A QTL (a), and with P value of 4.6e-08 (***), correlation coefficient of 0.59 and Confidence Intervals at 95% was 95%: 0.45-0.78 for 6A QTL (b).

Moreover, plant height distributions of recombinant lines grown in the glasshouse and field were compared. Analyses revealed that despite the fact that variation in plant height of recombinants grown in the glasshouse was wider than that of field for 5A QTL, the means were similar. However, plant height variation of recombinants grown in the glasshouse for 6A QTL was almost the same with very close means (Figure 4.5). In the next section of this chapter, we discuss how we have successfully combined map-based cloning with the exome and whole genome sequencing data for target marker development.

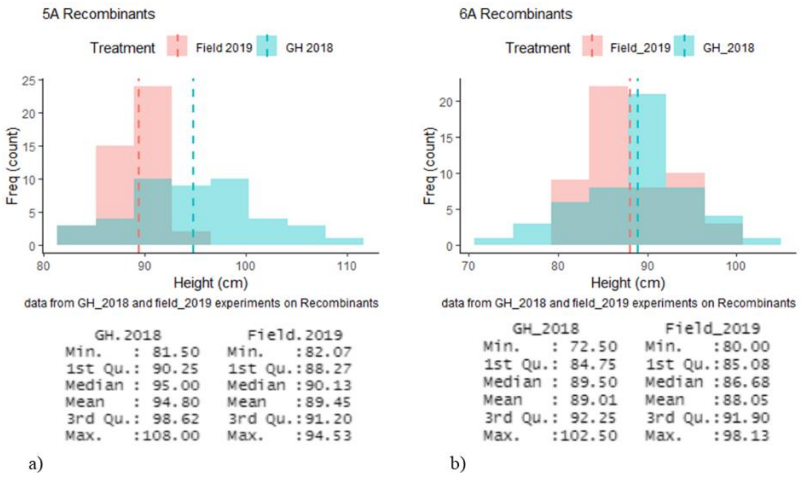


Figure 4.5 The plant height distributions of recombinants for 5A (a) and 6A QTL (b) grown in the glasshouse and field

4.2 Capturing the exons and WGS of Pamyati Azieva and Paragon for targeted marker development and delineation of the region of interest

4.2.1 Introduction

Although map-based cloning is time-consuming, its efficacy in gene identification is rising with the unprecedented improvements in whole genome and targeted sequencing technologies which have improved the marker saturating approaches significantly (Peters et al., 2003). Thus, advances in genomic tools have opened the door to the methods of forward genetics to study larger genomic regions or QTL controlling the trait of interest in a greater resolution, especially when map-based cloning is restricted due to lack of genetic exchange events during meiotic recombination. In this regard, the most accurate genome of hexaploid wheat has been published and about hundred thousand genes have been physically ordered and placed on seven wheat chromosomes representing three A, B and D genomes (Appels et al., 2018). Moreover, DNA sequence assemblies of ten wheat varieties were generated, and the positions and sizes of chromosomal centromeres were determined (Walkowiak et al., 2020). Importantly, all these data were integrated into available genomic databases and resources (Howe et al., 2020; Hassani-Pak et al., 2021). Therefore, the wheat reference genome quickly became a great tool for wheat scientists and breeders to discover any part of the wheat genome which they are interested in. Particularly, it is important to identify all genes under the QTL region and make their functional annotation to predict what possible particular gene is contributing to protein biosynthesis, in elucidating the plant trait the QTL is associated with. Thus, map-based cloning became relatively straightforward in organisms in which whole genomic data are available. To take advantage of these available genomic tools in wheat, we have fully (Whole Genome Sequencing) and partially (protein coding regions co-called “Exome Capture”) sequenced the wheat varieties which we used to develop RILs, NILs and recombinants. Both sequencing data were useful to develop targeted genetic maker design and narrow the QTL regions down substantially. They also were helpful to visualise any genomic areas of interest with the QTL regions in the genome viewer software packages.

4.2.2 Materials and methods

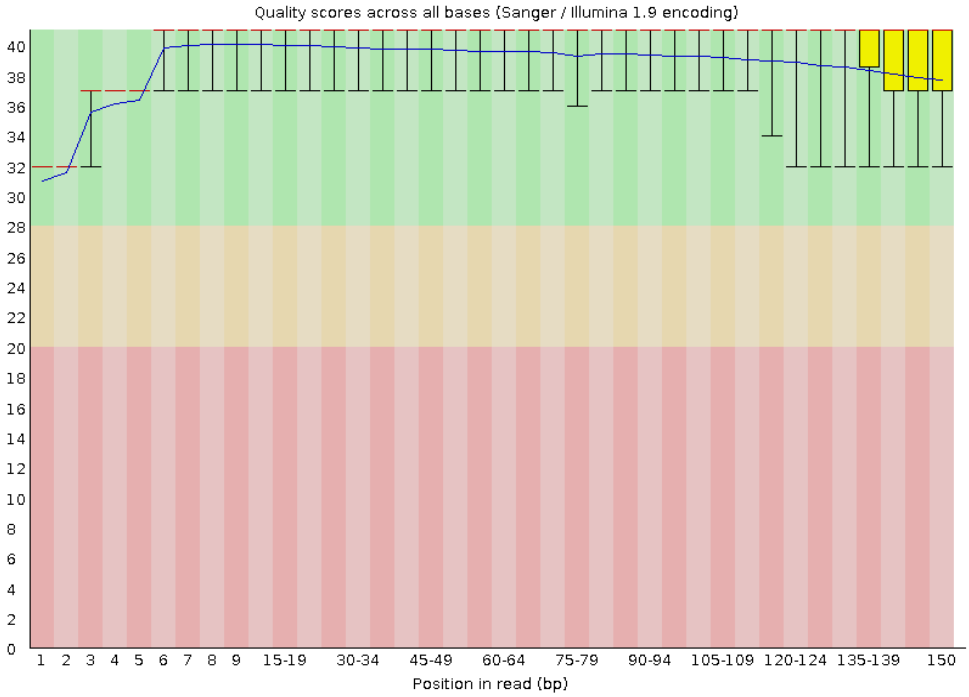
4.2.2.1 Sample preparation

To shrink and fine-map the QTL regions, we had the protein coding regions of parents captured and whole genomes sequenced. High – quality genomic DNAs of parents, Pamyati Azieva and Paragon, were extracted using DNeasy Plant Mini Kit of Qiagen and Oktopure platform to conduct exome capture and WGS respectively. Once genomic DNA was isolated with a required concentration to sequence coding regions, it was sent to the private commercial organisation, “Novogene”, the service of which has been utilised to do the analysis. Entire genomes of parents were sequenced during sequencing of the “Watkins population”, the WatSeq side project going in Dr Simon Griffiths group.

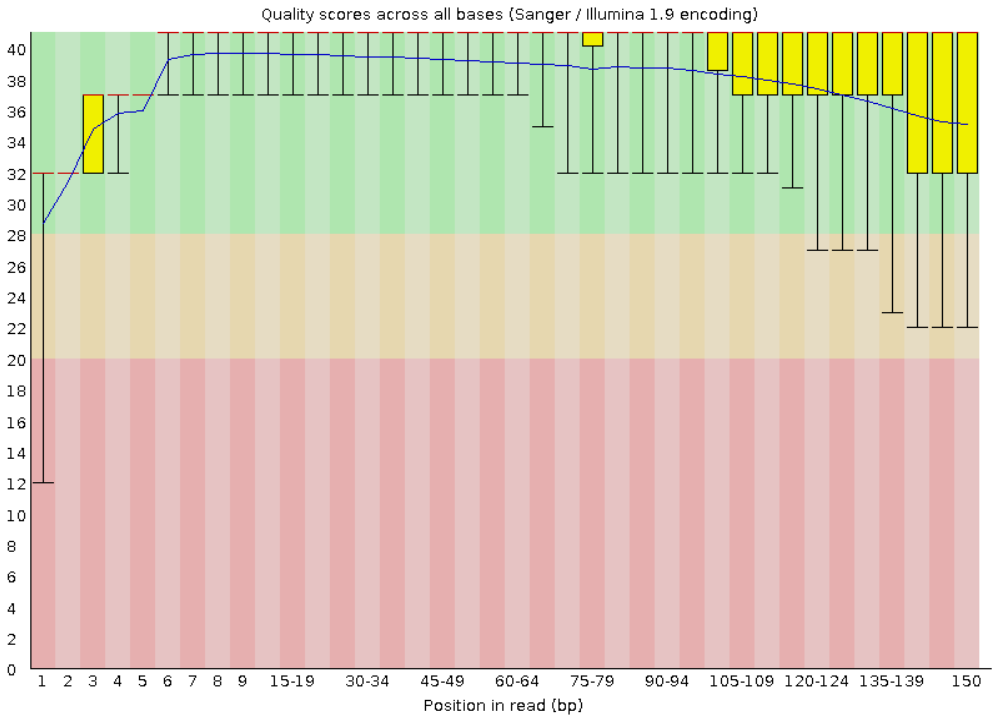
4.2.3 Exome capture results

4.2.3.1 Quality control of exome capture reads

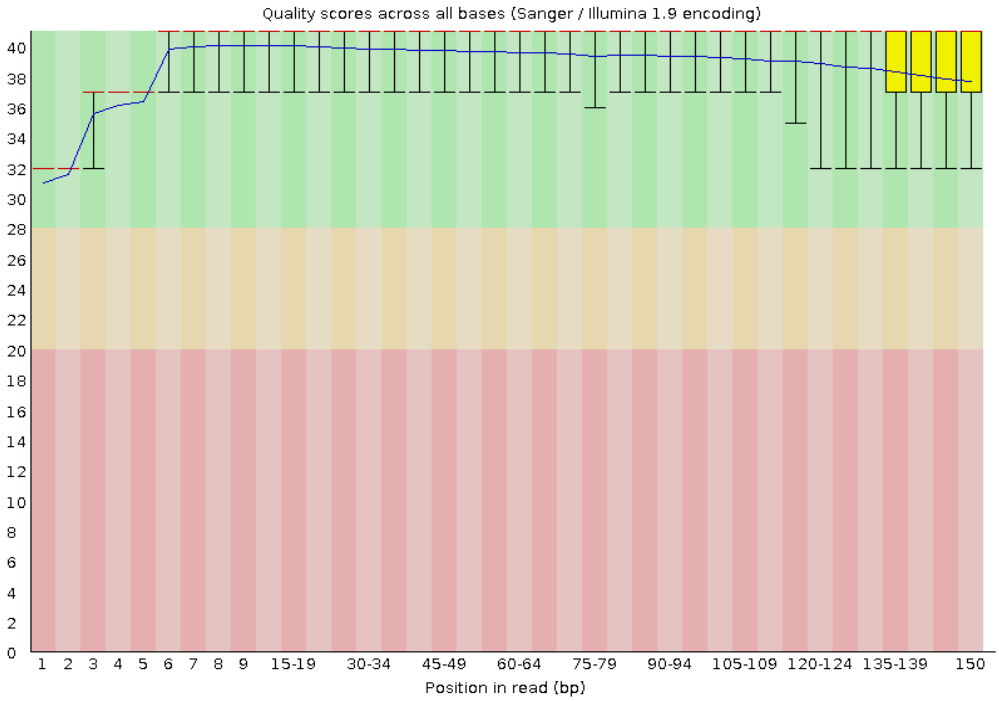
Once the commercial organisation received DNA from us, they prepared gDNA library for exome capture and released raw sequence data in FASTQ file format using Illumina PE150 platform which released Paired-end Reads with ~150 bp length. Guaranteed quality of the reads was Q30≥80%. QC tests showed the high quality of the reads for both parents (Figure 4.6a, b c, d).



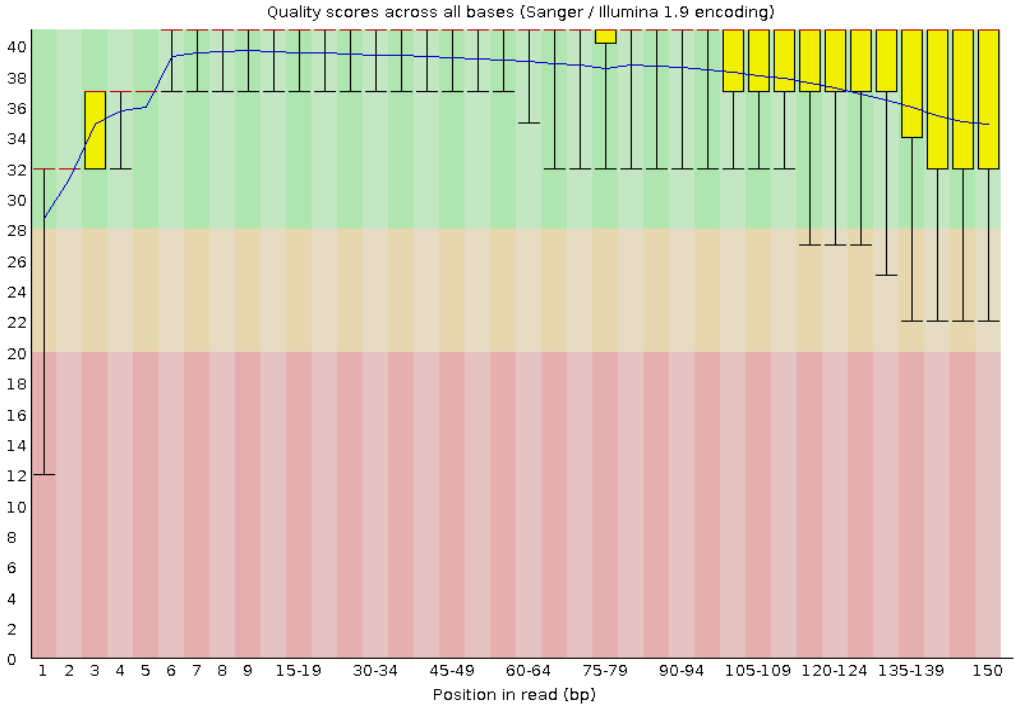
a)



b)



c)



d)

Figure 4.6 Per Base Quality

The figures 4.6a, and b, provide the quality of reads based on position from the beginning to the end of sequence for Pamyat Azieva reads 1 and 2 respectively. Figures 4.6c, and d for Paragon. Each plot shows the quality per base pair. The quality scores are given on y-axes. The position of reads (bp) from left to right are provided on x-axes. The background colour of the graph which divides the y – axes displays the quality per base pair as “good” (green), acceptable (orange) and poor (red).Yellow bars (boxes) represent the inter-quartile range (25-75%) and upper and lower whiskers represent the 10% and 90% points. The mean quality is shown as a blue line.

4.2.3.2 Raw Sequence Alignment to the wheat reference genome and variant calling

The exome capture data of parents were analysed and aligned to Chinese Spring (CS) Refseqv1.0. During the alignment, the latest alignment and variant calling tools were utilized, particularly BWA-MEM (bwa-0.7.5) algorithm of BWA software package for aligning the initial raw exome capture reads to the reference genome. Various options of SAM (Sequence Alignment Map) and BAM (Binary Alignment Map) tools were used for manipulating and analysing the alignment files. Especially, "samtools view" and "samtools sort" functions of the samtools-1.9 version were used to convert initial SAM files to BAM and to sort obtained BAM files for further analyses respectively. As the indexing of sorted BAM files is required by IGV (Integrative Genomics Viewer) or any other genome viewer software packages, "samtools index" was exploited to index our sorted BAM files for both parents.

PCR duplicate marking and removal in sorted BAM files was conducted by Picard-2.21.2 software, particularly MarkDuplicates option was used.

Software package freebayes-1.3.1 with standard filters was employed to call variants, that are SNPs and In-Dels, from the alignment. Finally, source programs such as tabix-0.2.6 and bcftools-1.8 were used to index a VCF (Variant Call Format) file, containing called variants, and extract specific columns (such as chromosome, position, mutations and quality) from it respectively. All bioinformatics analyses were done using HPC (High Performance Computing) cluster available for NBI (Norwich Research Park) scientific community.

During the alignment of exome capture analysis, two parents were aligned to the CS reference genome at the same time, thus, a multi-sample VCF file containing all possible allelic combinations was generated. Therefore, allele value combinations used to obtain variants/mutations are given in Table 4.4.

Paragon	Pamyat.Az
0/0	0/1
0/1	0/0
0/0	1/1
1/1	0/0
1/1	2/2

Table 4.4 Allele value combinations

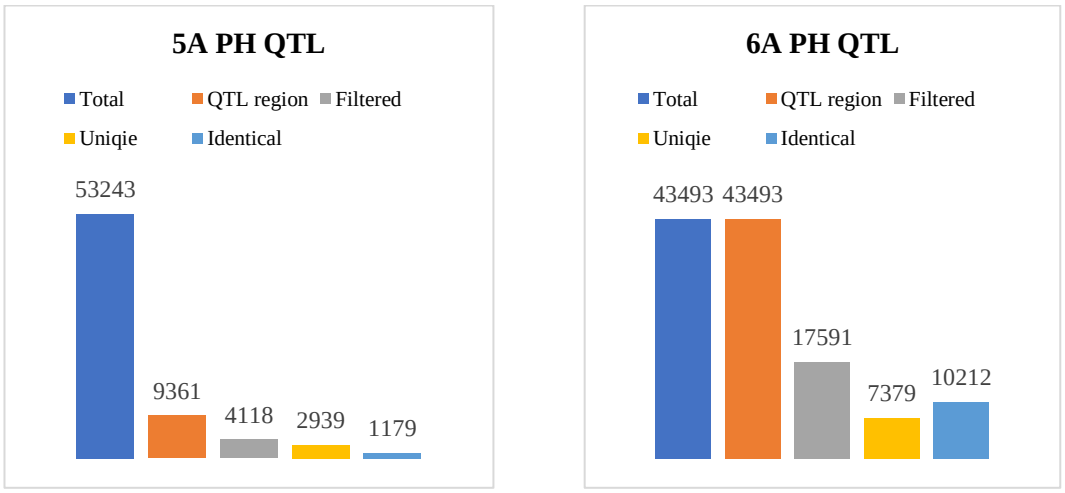
The table illustrates the most common allele value combinations used for designing new SNP KASP DNA markers. Here, 0/0 represents the reference, and 1/1 and 2/2 for parents respectively. Heterozygous mutations are shown as 0/1 and they were used very rarely in marker design.

4.2.3.3 The 5A QTL variants

The total number of mutations (SNPs, InDels) found were 53243. Out of 53243, 9361 mutations were pulled out for the QTL region. After trimming, based on QUAL>=30 and DP=4, only 4118 mutations were left. The number of unique (Pamyat Azieva carries alternative allele comparing to Paragon or vice versa) and identical (Paragon and Pamyati Azieva carry the same allele, but both different from CS) mutations were 2939 and 1179 respectively, out of 4118, covering entire QTL region.

4.2.3.4 The 6A QTL variants

The total number of mutations (SNPs, InDels) found were 43493 (this number is the number of mutations pulled out for QTL region as it covered almost entire chromosome). After trimming, based on QUAL>=30 and DP=4, only 17591 mutations were left from initial 43493. The number of unique (Pamyat Azieva carries alternative allele comparing to Paragon or vice versa) and identical (Paragon and Pamyati Azieva carry the same allele, but both different from CS) mutations were 7379 and 10212 respectively, out of 17591, covering entire QTL region.



a) b)

Figure 4.7 Number of variations found in the coding region for 5A PH QTL (a) and 6A PH QTL (b)

4.2.3.5 Physical mapping of the 5A and 6A PH QTL

4.2.3.5.1 Inner QTL region gene content assessment

The alignment of QTL flanking molecular markers of 5A and 6A PH QTL to wheat reference sequence (RefSeqv1.0) revealed that 5A QTL is located in the long arm of the 5A wheat chromosome, spanning the 37 Mb region under which about 473 genes are localised. In comparison, 6A QTL is covering the vast chromosomal region of 513 Mb, almost entire chromosome, and there are 2563 genes are located (Figure 4.8).

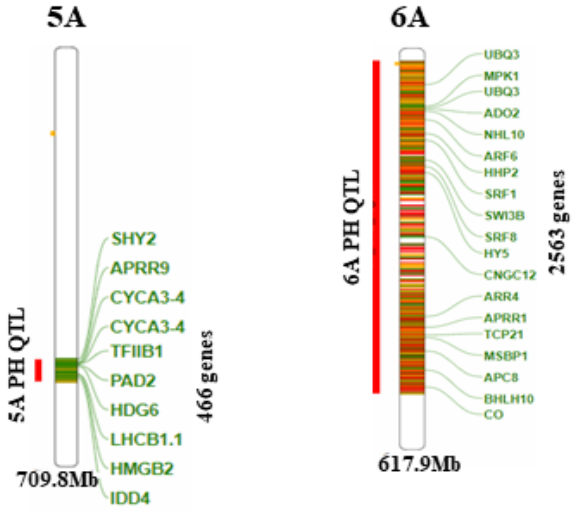


Figure 4.8 Unmapped 5A and 6A QTL regions

The figure demonstrates initial QTL regions for 5A and 6A PH QTL in a mapping population. The positions of some candidate genes are shown.

4.2.3.5.2 Homoeologous specific SNP KASP molecular marker development

While saturating the initial QTL regions with KASP markers, we focussed on SNPs only. KASP genotyping requires three markers (two unique non-labelled forward primers with different SNPs at the 3' end and different tail sequences at the 5' end, and one reverse or common primer) to amplify the mutations (SNPs in our case). Thus, 231 homoeologous specific DNA markers were designed for 77 SNP KASP markers to genotype RC for 6A PH QTL. Likewise, to genotype RC of 5A PH QTL, 123 homoeologous specific DNA markers were designed for 41 SNP KASP markers (Table 4.5 and Table 4.6). Although the results of exome capture were used to design all these markers, some of them did not hit protein coding regions. This is due to an older version of the wheat reference genome which was used for the bait design. Some of the markers did not work or/and were uninformative.

№	Genomic coordinate	Target gene	FAM	VIC	Common	Informative	Function
1	528114084	TraesCS5A02G317100	GGTGGTTCACCATTTGCTCCTTT	GGTGGTTCACCATTTGCTCCTTC	AAGGGACCAAAATAATGCTCATTCGGTA	yes	
2	528261326	TraesCS5A02G317200	CTGTAATTCATTTTTTTAATGAAG	CTGTAATTCATTTTTTTAATGAAA	GCTAACAGAATTGCAGGG	yes	
3	531236012	TraesCS5A02G453200LC	GCGTGTGGGCTTAAATGTTGTAT	GCGTGTGGGCTTAAATGTTGTAG	TGGTTACATTCTCTACTAC	no	shoot system morphology trait
4	533554208	TraesCS5A02G320500	GTGTCCTGTTGCTGGCATTGTTT	GTGTCCTGTTGCTGGCATTGTTG	CATCACGATAGAATGCAAGT	yes	shoot system morphology trait
5	534253659	TraesCS5A02G321500	GCCAGTGAATGGGCTCCTTCCC	GCCAGTGAATGGGCTCCTTCCA	GTGGTTCTGGGGGAGAGTCAA	no	
6	535578125	TraesCS5A02G324200	CCATTCTAGAGTGAGGGACAAGA	CCATTCTAGAGTGAGGGACAAGC	TGGAATTTTCTGGATTAATATCT	yes	stem internode
7	536773464	TraesCS5A02G326300	CTAAGCTCTAGCTATGCAAAAA	CTAAGCTCTAGCTATGCAAAAG	ATATCACGGTCTTCAAGCCAC	yes	plant height
8	537952270	TraesCS5A02G328900	ACCTCGCCTTCTCCTCCTACCT	ACCTCGCCTTCTCCTCCTACCC	CATTGGAAGCTCCCGGAATGG	yes	coleoptile emergence stage
9	540052320	TraesCS5A02G330500	CCCAGGCCGCGTGTTCTTACA	CCCAGGCCGCGTGTTCTTACG	GCAACCGGATAGAAGTAGAGG	yes	shoot system
10	542384407		CCGAGCGCTCCATCCTCGTCA	CCGAGCGCTCCATCCTCGTCG	GAAGTGGAGGCTAGGATGTAG	no	
11	536261749	TraesCS5A02G325600	GCTCACGCTCGCCCCCCCA	GCTCACGCTCGCCCCCCCC	TGTTGCGGCCTTGCCCGA	no	
12	536814443		CTCCTTCCAGCGCCCCCTT	CTCCTTCCAGCGCCCCCTC	GATTTGAATCCTCAAATCA	no	
13	538350743	TraesCS5A02G329500	ACTCGCACTGTGCTCCGGAGCAT	ACTCGCACTGTGCTCCGGAGCAA	CTCCTTGTACACCTTCCCTGGCC	no	root number
14	540369986	TraesCS5A02G331500	CGTACTCACAAAATACTTC	CGTACTCACAAAATACTTT	ACAAATAGTTTCTGTTTTA	no	
15	540822721	TraesCS5A02G332400	GCGCCGGCGTGGCGTGGACCTT	GCGCCGGCGTGGCGTGGACCTG	CTCCTTGCGACGGCGGCCATAG	no	
16	544154759	TraesCS5A02G334600	CACTTCGGGTTAGCTAATTGTTA	CACTTCGGGTTAGCTAATTGTTG	GACAGCAGAGTCGACAAGGTA	yes	shoot apical meristem
17	544613546	TraesCS5A02G335700	TTGGACCCCTAGAATGTTTCC	TTGGACCCCTAGAATGTTTCT	CCTAATTTTAATCAGCGTCAG	no	coleoptile emergence stage
18	545047977	TraesCS5A02G335900	GGTGGCCATGGCAGCTGGTAT	GGTGGCCATGGCAGCTGGTAC	GCGCCATGCCACGGAGTACGC	no	seedling coleoptile
19	546206195	TraesCS5A02G338100	CTCCTTGATCCGTGTGGAC	CTCCTTGATCCGTGTGGAT	AATTATGCATGTGCTTTTAGA	no	
20	546239406	TraesCS5A02G338300	CGCCAGCAGCGCCGCCTTCAA	CGCCAGCAGCGCCGCCTTCAG	GCGTACAGGCGCGTGGAGGAG	no	plant height and coleoptile
21	546704568	TraesCS5A02G339400	TGGAGCGCGCGGTAGCCGCGA	TGGAGCGCGCGGTAGCCGCGG	CCTCTCCGACCCCGGCTTCGC	no	stem internode
22	546972873	TraesCS5A02G340300	CCTTCTTTCACTAGTATTGGC	CCTTCTTTCACTAGTATTGGT	TGTAACTTTCAAGGCACTAAA	yes	hypocotyl endodermis
23	547344101	TraesCS5A02G341000	CAAGAAACTTGATGACGCATTGT	CAAGAAACTTGATGACGCATTGC	TCGAGAGGTAAGTGACATGCT	no	
24	547818913		TGTGTGTTGATGTTTGTTTGCT	TGTGTGTTGATGTTTGTTTGCC	GATATCCTATCCAAGCCTTCCC	no	
25	547613987	TraesCS5A02G341800	AAACAGCAAAGCTGAACAAGT	AAACAGCAAAGCTGAACAAGC	CAAGCAGAGTTTGTATACTTA	no	
26	548234636	TraesCS5A02G343400	GCATGATGTACACCGCAGACG	GCATGATGTACACCGCAGACT	GAGGGACCTCCAACCGGGTTC	yes	stem internode
27	549154332	TraesCS5A02G344600	ACAGTGTAAGATAGTAAGATAAG	ACAGTGTAAGATAGTAAGATAAC	CAATTTTCCTGCTCCTATTGCCT	no	shoot system morphology
28	550645755	TraesCS5A02G347600	CATGGCCAAAATACTTGCTGG	CATGGCCAAAATACTTGCTGA	GAACAGAAGCATCACCTCCGC	yes	
29	552515212	TraesCS5A02G349400	TGCCGAGCACATTCAGAGA	CTTGCCGAGCACATTCAGAGC	AGGGAAGGAGCTGCAGAAGATGAAA	yes	
30	552602170		CGAAGAACCAACGATGGC	CGAAGAACCAACGATGGA	TGTCTTCTCGTCGTTGCA	yes	
31	552777433	TraesCS5A02G489700LC	ATCGTGTTGGAATTATGTGTCC	ATCGTGTTGGAATTATGTGTCT	TTTGAAAATAGTTGCA	no	
32	554188721	TraesCS5A02G351600	AACTTGATCTATTTCTAAATTTG	AACTTGATCTATTTCTAAATTTT	CCACAAGAGTTCTATGAAGTTC	no	grain number
33	555541006	TraesCS5A02G352700	GATTGGCCATTCATTTGAG	GATTGGCCATTCATTTGAT	GCTTCTTTGTATTAGATT	no	
34	555548303	TraesCS5A02G352700	ATCTAAGGAAACCATTGATTGACG	ATCTAAGGAAACCATTGATTGACA	AGCGCAGTAAGTCCAGACT	yes	
35	558684533	TraesCS5A02G356300	GCTGCCCAAAGCTTACCTCAGC	GCTGCCCAAAGCTTACCTCAGG	GATCAGAGCTAACTGATCA	yes	
36	561662630	TraesCS5A02G359900	AGAGAGTAGGGCAGACAG	AGAGAGTAGGGCAGACAA	AACGACATCACCGCCCTG	yes	
37	559038834	TraesCS5A02G356800	TGCTCCGGCAGAAGGACACA	TGCTCCGGCAGAAGGACACG	GGTGAGCGAGGAGGAGCT	no	
38	559039564	TraesCS5A02G356800	TACATGGCGTCGGCTGTC	TACATGGCGTCGGCTGTT	AATACGAGGGTTTGGTGG	no	
39	561662630	TraesCS5A02G359900	CATGAACAACAAGAAAAAAAAAC	CATGAACAACAAGAAAAAAAAAA	GTGATACCTGCTTTGCCT	yes	
40	562991229	TraesCS5A02G361800	TCGCACGGAGCGAGTCATCA	TCGCACGGAGCGAGTCATCG	TGACCCGGACTTCTCCAA	yes	
41	565540546	TraesCS5A02G365500	CATCTTAGATAGCTGAAGGC	CATCTTAGATAGCTGAAGGT	TCCTGACCAAATCAGTCAC	yes	

Table 4.5 KASP SNP markers for 5A QTL

Functional annotation is available for some genes only.

№	Genomic coordinate	Target gene	FAM	VIC	Common	Informative	Function
1	18705935	TraesCS6A02G037800	ATTCTAAAGTAAACCTTACACAGATGTTGT	CTAAAGTAAACCTTACACAGATGTTGC	CTAGAGTTAAGCATGTAAGGTTACACAGTT	yes	
2	21519942	TraesCS6A02G041000	GAAGAAAAGTGATAGTGATA	GAAGAAAAGTGATAGTGATG	CCCATCGTCAACGTAACT	yes	
3	22954779	TraesCS6A02G043900	ACCGTTGCCATGGGTGGATCGT	ACCGTTGCCATGGGTGGATCGC	CTCCGTTGCAATTCTCAG	yes	
4	22954779	TraesCS6A02G044700	ATGCCGAACCCGCCGCACCCC	ATGCCGAACCCGCCGCACCCG	TGAAGGCGACGGGGGACTT	yes	
5	23602641	TraesCS6A02G045800	CCTTTCGCCACTGTTGGCA	CCTTTCGCCACTGTTGGCG	CGACAAAATGGAACCA	no	
6	23602809	TraesCS6A02G045800	ATTGCCATGGCTATCATAC	ATTGCCATGGCTATCATAG	ACAATGGACACAGTGGTG	no	
7	24458056		ATAGTGTCTCTGATACATT	ATAGTGTCTCTGATACATC	TCAACCGCCTGCACACAC	yes	
8	24921964	TraesCS6A02G048900	ATCCACAACCGCAAACACAA	ATCCACAACCGCAAACACAG	CAATTTCGCAGCAACAAGTA	no	
9	25697521	TraesCS6A02G050300	GCTGTCAGCGAATAAAAGCGC	GCTGTCAGCGAATAAAAGCGG	TGAGAACCCATGTTATTGCTG	yes	
10	25775827	TraesCS6A02G060200LC	ACCATCAAATTTGAATCT	ACCATCAAATTTGAATCG	TTAACGCTTTACCAGA	no	
11	25776587	TraesCS6A02G060200LC	GGAAGCATGTCTATTACAA	GGAAGCATGTCTATTACAC	ACGATAGATTCTTCTTGA	no	
12	26530636	TraesCS6A02G050800	GAATTGCTGACGCAAAAGT	GAATTGCTGACGCAAAAGC	TACAAGATCGAGAGCACA	yes	
13	26565185	TraesCS6A02G050900	GAATTGCAAGAGGGCCAAATC	GAATTGCAAGAGGGCCAAATT	GAACTTACATGTTCGACT	no	
14	26735596	TraesCS6A02G051300	CGAAGCAACGACGATGAGGAT	GAAGCAACGACGATGAGGAG	GCGTACGTGTGGCTCTGTC	yes	
15	32865132		CGCGAAGCACCGCCCCA	CGCGAAGCACCGCCCCT	GCGGAGGGTGATATCAG	yes	
16	44094117	TraesCS6A02G074800	AACTGTCCTGCAATAAGAT	AACTGTCCTGCAATAAGAA	GGCAGCTGCAGTGCACCAG	yes	
17	53260463	TraesCS6A02G085200	TAAATACGGTCAAACTTTG	TAAATACGGTCAAACTTTC	GCACGAACATGTAGAGGGA	yes	stem internode
18	65066575	TraesCS6A02G097800	GCACCGGACGCACCGGCA	CACCGGACGCACCGGCG	CATCAGAAGACGCCTAGATGTCCAT	yes	
19	63783428	TraesCS6A02G096600	TGCTGTCTGCTGTTTCTCCG	TGCTGTCTGCTGTTTCTCCA	GCTCCCTGAAGCCTGCAG	no	
20	68128977	TraesCS6A02G100500	GGTGAAGGTAAGCACATATTTTT	GGTGAAGGTAAGCACATATTTTC	CCTGACAATAGAAGGCCG	yes	
21	73672142	TraesCS6A02G105000	CATGGAAACTGGTAGGAGAT	CATGGAAACTGGTAGGAGAC	CCAGTCATTAGGTATATGTG	yes	
22	73566796	TraesCS6A02G104800	CTCTACTGTGACTACTCTAGCG	CTCTACTGTGACTACTCTAGCT	TAAAGAAGGGGATGGCACA	no	
23	79958221	TraesCS6A02G110800	TGACGGCCTCATCCAGGTG	TGACGGCCTCATCCAGGTT	TTTTCTACCGTAACTACAT	yes	
24	104413593	TraesCS6A02G132400	CGCGACGAAGAGCGCGATGCC	CGCGACGAAGAGCGCGATGCT	CCCCGACGTGAGCGCCAAC	yes	
25	107211012	TraesCS6A02G135600	TGCTCCGTCTTCCTGCATGCG	TGCTCCGTCTTCCTGCATGCA	CATCTAGTGTACCTTTGAGGT	yes	shoot system
27	109987873	TraesCS6A02G138600	AGCTACCATGAACTTCGTAGT	AGCTACCATGAACTTCGTAGC	GCCAAATAGGCGCCCTACCTC	yes	plant height
28	112593834		CGGCTACCCTGTGCCCCGACG	CGGCTACCCTGTGCCCCGACC	TCAACTCCTCTGACATCGGCT	no	
29	115412680	TraesCS6A02G140800	CTAGCGGGTGAGAAGCCGGAC	CTAGCGGGTGAGAAGCCGGAG	GCGACCGCGGCAACCCGGTAC	yes	hypocotyl length
30	116706958	TraesCS6A02G141800	GCTGAATATGTCCTGGATATGGAT	GCTGAATATGTCCTGGATATGGAG	CTCCATCTTCAAGCTACCTGG	no	stem internode
31	116949974	TraesCS6A02G193500LC	GCTCCATCATACATTTGGTCCC	GCTCCATCATACATTTGGTCCT	ACCGAAAAATGGATAGATGACCC	yes	
32	117713598		GCAAGCCCCGGATGCGACGAG	GCAAGCCCCGGATGCGACGAC	GCGATCCAGCAGAGCATGTCG	yes	
33	118181303		GCGTCTCTGACTTGGGCACCG	GCGTCTCTGACTTGGGCACCT	TACAGTTGTGGTGGACCTTGG	yes	
34	119186730		GCGAAATGAAAGCTTGAAGCA	GCGAAATGAAAGCTTGAAGCG	GCTTGCCTAGCATCTGCCCGA	no	
35	120275935		AATATGAATTCATAGATCAAC	AATATGAATTCATAGATCAAT	ATTTCAATGCAACCCATGTCA	yes	
36	117412598	TraesCS6A02G142600	TATGAGCGCGTTGAAGTAAGT	TATGAGCGCGTTGAAGTAAGC	CAGCTTGCAGTTGCGTCCGCA	no	
37	117412598	TraesCS6A02G142600	TATGAGCGCGTTGAAGTAAGT	TATGAGCGCGTTGAAGTAAGC	CCTCCCTCTCTATTGTCAGCG	no	
38	121616502	TraesCS6A02G143900	CCCCACCCCCTCCTCCGACCG	CCCCACCCCCTCCTCCGACCC	TTCGCCATTGCCAACGACCAG	no	
39	121885552	TraesCS6A02G144700	GTGCATAATTATTGATCTGCAATGA	GTGCATAATTATTGATCTGCAATGG	GTTTTAATTTATCTCTTGCGAACCG	yes	
40	124715996	TraesCS6A02G146200	CGTGTGCAATGTGTTGTTATGG	CGTGTGCAATGTGTTGTTATGC	GCCTGCCCAGTGGTTGTG	yes	
41	125496419	TraesCS6A02G146500	CGACAGCGGCAACCTCAAT	CGACAGCGGCAACCTCAAC	GTACGCGTCCGGTATCTCC	yes	

continued on next page

continued from previous page

№	Genomic coordinate	Target gene	FAM	VIC	Common	Informative	Function
42	127779586		CGGTGAGGCAGGTTAGCA	CGGTGAGGCAGGTTAGCG	GCCCAGATCCCGTTACCG	yes	
43	129340989	TraesCS6A02G148200	TGCGCAGACGCCGCCGCCGCT	TGCGCAGACGCCGCCGCCGCC	GGTCGTGGCACTCCTGGCCTC	no	shoot system
44	131444641	TraesCS6A02G149400	GAGAAGAGGAGCAGAGCTGAG	GAGAAGAGGAGCAGAGCTGAA	AGGTGATCTTGAGCAGAGTTG	yes	stem internode
45	132411901		ACGAGGAGCTCAGCCGCATCG	ACGAGGAGCTCAGCCGCATCA	TGGGCTCCACGGCGGCGTCCA	no	
46	132412042		TACCATTGTGATTCACATGTC	TACCATTGTGATTCACATGTA	AATACACTGTAGAGTAATTTG	no	
47	132412346		TATAGTGCTAAAACTCGTACA	TATAGTGCTAAAACTCGTACG	AGTAATCAAATTTATGCATCA	yes	
48	134683037	TraesCS6A02G150700	CGACGACAACCGCGTCGGCCC	CGACGACAACCGCGTCGGCCA	TTTGTTTTGGTACGTATAATC	no	
49	134683626		TCGCCGTCCCGCCCTTCCCTC	TCGCCGTCCCGCCCTTCCCTT	ACACGGCGGTGGGGAGAGGTA	yes	
50	135067509	TraesCS6A02G151000	AGGTAACACCAGCACACTGCA	AGGTAACACCAGCACACTGCG	TGGACTTGGACACTTCGACGA	yes	
51	137977057		CAAGCTTCAAACGACTCCGA	CAAGCTTCAAACGACTCCGG	GAAGCTTGCAATGGAGCGAC	yes	
52	142784074	TraesCS6A02G154100	CCTGTTTCTGATCTCAGTCAATATG	CCTGTTTCTGATCTCAGTCAATATA	GTGGATATTGCAAAGGTCAGC	yes	
53	144524574	TraesCS6A02G155700	GGTCGTTTCAACAAACAGCTG	GGTCGTTTCAACAAACAGCTC	GTGCTCGTCCCTTCTATCCTA	yes	
54	146445114		TCCCTCTGTTCCCAAATACGA	TCCCTCTGTTCCCAAATACGG	CCTCTACGCCCACTTCACTC	no	
55	146744128		CGGTTCAGGGTACAGTGTAGAG	CGGTTCAGGGTACAGTGTAGAC	CAAATTTTCAGATCCACCGCT	no	
56	146743603	TraesCS6A02G222100LC	TCCCTTGGCATTGCATACAA	TCCCTTGGCATTGCATACAG	GACATTCAGGATCTCCCCTTG	yes	
57	148248467	TraesCS6A02G158100	CAGCTCGTGAAGACGATTAAC	CAGCTCGTGAAGACGATTAAT	CCGTCTGCTGCAGCAACAACA	yes	
58	148530677	TraesCS6A02G158200	ACCTTTGGAAAGTTTGTCACTACT	ACCTTTGGAAAGTTTGTCACTACC	CCAGTCAACCTGTTGCAGTTG	yes	
59	150395091		CGTGGCAGGAGGAAGAATAAAT	CGTGGCAGGAGGAAGAATAAAC	GGATGCCTACGCCACTGC	yes	
60	153351294	TraesCS6A02G160100	CTCTTTACATTTGGCAGTGAC	CTCTTTACATTTGGCAGTGAT	AATGCTATGTTCTGTCAATTC	yes	shoot system
61	176574357	TraesCS6A02G169100	CGTGCGGTCCCGCCGCTACGC	CGTGCGGTCCCGCCGCTACGT	ACGCCGTCGTCTGCCAGGAGC	yes	shoot system
62	201107376	TraesCS6A02G179000	TCCCTCTTCACCAGAGGCGTT	TCCCTCTTCACCAGAGGCGTC	GGTACTCACAGGCCGCGGGC	yes	Auxins promote stem elongation
63	201107376	TraesCS6A02G179000	TCCCTCTTCACCAGAGGCGTC	TCCCTCTTCACCAGAGGCGTG	GGTACTCACAGGCCGCGGGC	no	Auxins promote stem elongation
64	215528325	TraesCS6A02G184500	GCATCCTCTACCACCTCCTCC	GCATCCTCTACCACCTCCTCT	CGCTCCGGTGGAAGCCGAGGC	no	shoot system
65	224281364	TraesCS6A02G187700	TAATTAAGCTTCACCTTGATG	TAATTAAGCTTCACCTTGATA	TCCTCTGTAGGTACGTGATCA	yes	shoot apical meristem
66	250202610	TraesCS6A02G190100	CCTACCTCCCTGCAACCGAGC	CCTACCTCCCTGCAACCGAGT	TTTGGCCTCTCTAAGCCCTAA	yes	
67	258373792	TraesCS6A02G191400	GCATCCTCTACCACCTCCTCC	GCATCCTCTACCACCTCCTCT	CGCTCCGGTGGAAGCCGAGGC	no	shoot system morphology trait
68	277274586	TraesCS6A02G195100	CGAAGGCTGGTAATCTTTGAG	CGAAGGCTGGTAATCTTTGAA	TAAACGAGGTGTGGTGTAGTC	yes	shoot system morphology trait
69	299335643	TraesCS6A02G197800	CAACAAATTTGCTGTCAGCGC	CAACAAATTTGCTGTCAGCGT	GTTATCAGGAGAAGCTTTATA	yes	stem internode
70	402952170	TraesCS6A02G218100	CTCAACACGGAAAATAAG	CTCAACACGGAAAATAAA	TTGCGAACACACTTGAAAA	yes	coleoptile emergence stage
71	410244708	TraesCS6A02G383400LC	CGCCGTGGTGGGAAGCCAG	CGCCGTGGTGGGAAGCCAT	ACGACGACAGTGATATCGT	yes	plant height
72	420412399	TraesCS6A02G223700	GTGGGCACTTTTTGATCAG	GTGGGCACTTTTTGATCAC	CTATCTCCAGGTGCTCCTCACT	yes	hypocotyl length
73	426315033	TraesCS6A02G226500	TGAATGGCCATGGCAGAGGAT	TGAATGGCCATGGCAGAGGAC	GTGCGTACCTTGAAGTTGCCC	yes	stem internode
74	429227551		GTTCGAATTACCAAAAATAT	GTTCGAATTACCAAAAATAA	TCATGGAACAGATCCTGCG	yes	
75	441431948	TraesCS6A02G233300	TGCTGTTGGAAGGTTCAGAT	TGCTGTTGGAAGGTTCAGAG	ACCATATATAGACAGCTTAT	no	shoot apical meristem
76	495254850	TraesCS6A02G268600	TTCCTTGCTCTTTTCCTTT	TTCCTTGCTCTTTTCCTTG	TACAGATTATAAAGAGTTT	no	
77	531522332	TraesCS6A02G298200	TCGAGAACGCATCATGCACAC	CTTCGAGAACGCATCATGCACAT	CGCGATGATCCAGGAGAGGAATTTT	yes	

Table 4.6 KASP SNP markers for 6A QTL

Functional annotation is available for some genes only.

4.2.3.5.3 Genetic mapping of the 5A and 6A PH QTL

In the section 4.1 of chapter 4 we mentioned about the development of homozygous recombinants belonging to several recombinant classes (RCs) for 5A and 6A PH QTL. Moreover, we provided height data on those recombinants in controlled and non-controlled environments with comprehensive experimental design and appropriate statistical analysis used to describe the data. In the previous section of this chapter, we have given details of KASP marker development. Here, we combined the obtained genetic and phenotypic data to map the 5A PH QTL.

Overall, 5A PH QTL recombinants were genotyped using 41 SNP KASP markers some of which were not informative. Figure 4.4 a showed a significant relationship among glasshouse and field height data on recombinants although the variation between the RC was smaller in the field compared to in the glasshouse (Figure 4.5a). Using these data, we mapped the 5A PH QTL to a physical interval of ~18.8Mb, from initial 37Mb, between 528.1 – 546.9Mb based on CS reference genome (Table 4.7 and Figure 4.9).

	RC-01	RC-02	RC-03	RC-04	RC-05	RC-06	RC-07	RC-08
RC-01	-	0.97	1.00	0.47	0.39	0.06	0.00	0.01
RC-02	1.00	-	1.00	1.00	1.00	0.97	0.01	0.10
RC-03	0.47	0.37	-	1.00	1.00	0.44	0.00	0.04
RC-04	1.00	0.12	0.00	-	1.00	1.00	0.12	0.97
RC-05	1.00	1.00	0.06	1.00	-	1.00	1.00	1.00
RC-06	0.24	0.00	0.00	1.00	1.00	-	0.74	1.00
RC-07	0.40	0.01	0.00	1.00	1.00	1.00	-	1.00
RC-08	0.07	0.00	0.00	1.00	1.00	1.00	1.00	-

Table 4.7 P – values from the pairwise comparisons between RCs for 5A QTL in glasshouse (upper triangle) and field (lower triangle) conditions

In comparison, the total number of designed SNP KASP markers for genotyping of the 6A PH QTL recombinants were 77, 25 of which did not show polymorphism. As in 5A recombinants, the data obtained in the glasshouse were strongly correlated with the data collected in the field, however, the variation in plant height was consistent in 6A compared to 5A recombinants (Figure 4.5b). The combination of genetic and phenotypic data on 6A recombinants thus allowed us to conduct medium resolution mapping locating the 6A QTL into ~28Mb physical location from the initial 513Mb region which spanned almost the entire chromosome. The mapped region covered the interval between 120.2 – 148.2Mb on 6A chromosome of the Chinese Spring (Table 4.8 and Figure 4.10).

	RC-01	RC-02	RC-03	RC-04	RC-05	RC-06	RC-07	RC-08	RC-09	RC-10	RC-11	RC-12
RC-01	-	1.00	1.00	0.01	1.00	0.12	0.22	0.01	0.12	1.00	1.00	0.00
RC-02	0.00	-	1.00	0.17	1.00	1.00	1.00	0.22	1.00	1.00	1.00	0.02
RC-03	1.00	0.02	-	0.56	1.00	1.00	1.00	0.67	1.00	1.00	0.53	0.05
RC-04	0.00	0.00	0.00	-	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
RC-05	0.00	1.00	0.00	0.31	-	1.00	1.00	1.00	1.00	1.00	0.94	1.00
RC-06	0.00	0.00	0.00	0.10	1.00	-	1.00	1.00	1.00	1.00	0.03	1.00
RC-07	0.00	0.00	0.00	0.15	1.00	1.00	-	1.00	1.00	1.00	0.05	0.83
RC-08	0.00	0.00	0.00	0.60	1.00	1.00	1.00	-	1.00	1.00	0.00	1.00
RC-09	0.00	0.00	0.00	1.00	0.61	0.33	0.47	1.00	-	1.00	0.03	1.00
RC-10	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	-	1.00	1.00
RC-11	1.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-	0.00
RC-12	0.00	0.00	0.00	1.00	0.04	0.00	0.00	0.01	1.00	0.54	0.00	-

Table 4.8 P – values from the pairwise comparisons between RCs for 6A QTL in glasshouse (upper triangle) and field (lower triangle) conditions

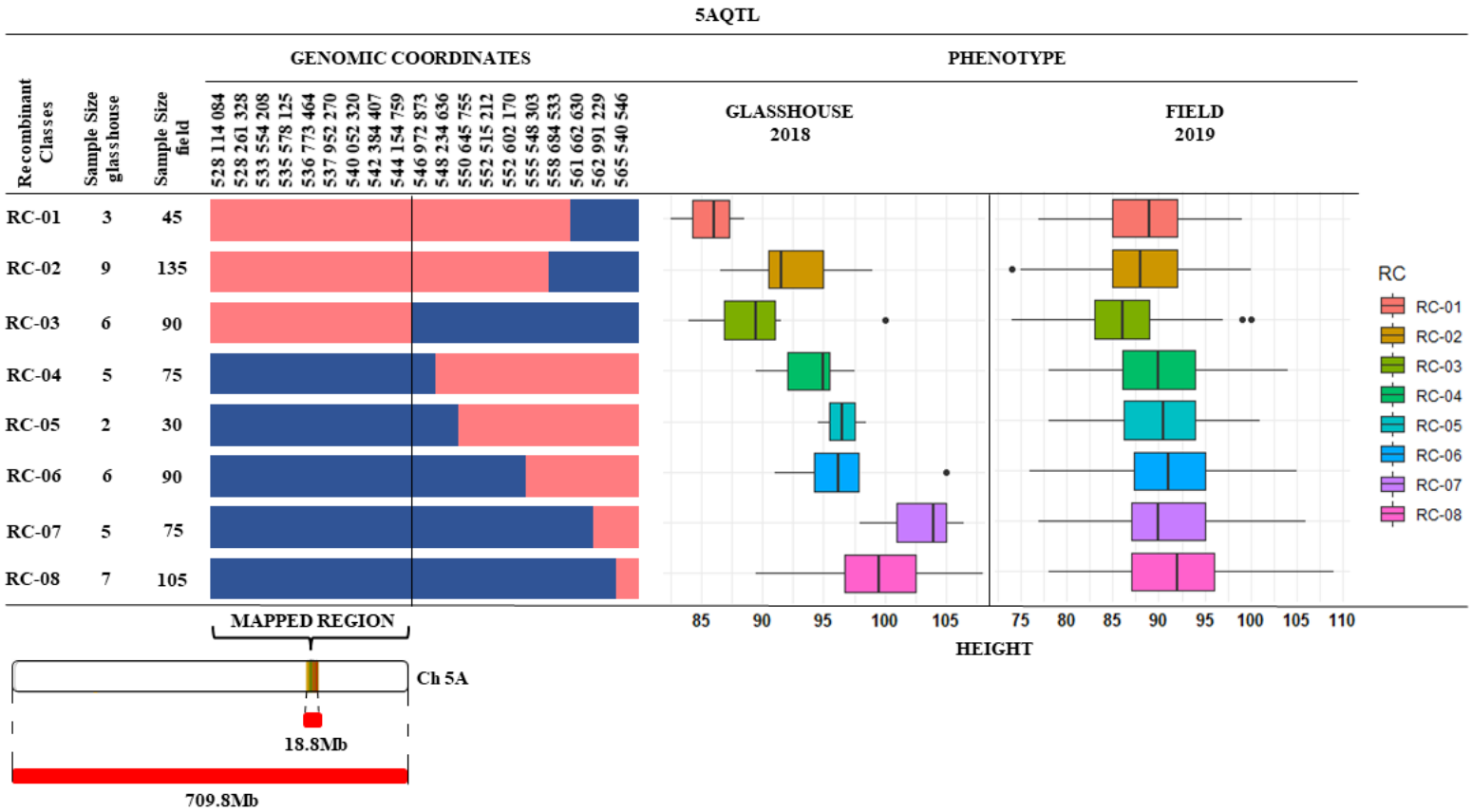


Figure 4.9 5A QTL mapped region

Physical genomic positions are on top of the map. Coral and blue correspond Paragon and Pamyati Azieva respectively. Plant heights of RCs obtained from GH and field are shown on the right as box plots. On the left-hand side of the map are RCs with their corresponding sample sizes (number of plants) in GH and field. The entire chromosome and mapped interval and their corresponding physical distances are given on bottom of the map.

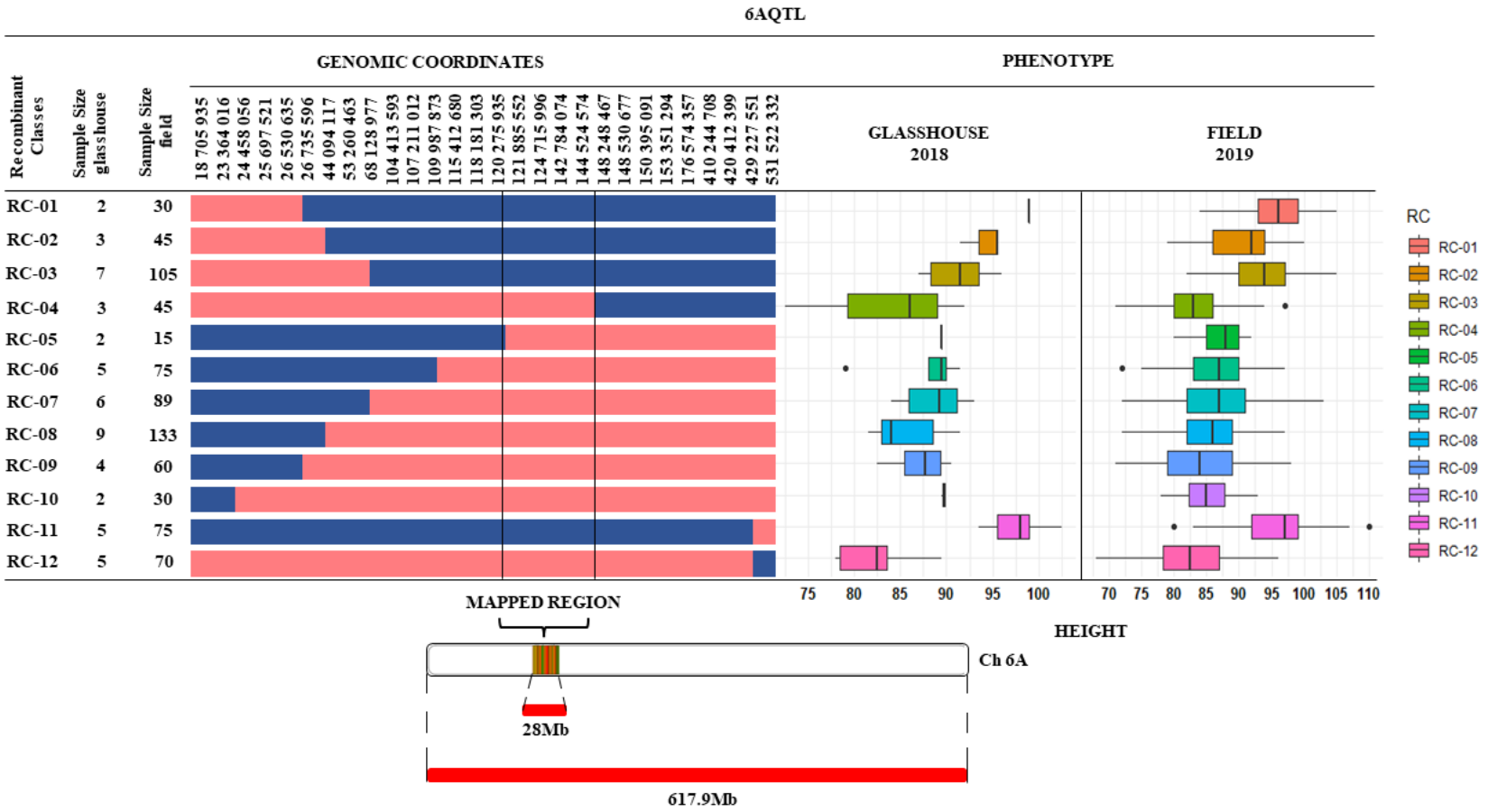


Figure 4.10 6A QTL mapped region

Physical genomic positions are on top of the map. Coral and blue correspond Paragon and Pamyati Azieva respectively. Plant heights of RCs obtained from GH and field are shown on the right as box plots. On the left-hand side of the map are RCs with their corresponding sample sizes (number of plants) in GH and field. The entire chromosome and mapped interval and their corresponding physical distances are given on bottom of the map.

4.2.4 WGS outcomes to discover mapped regions

4.2.4.1 Quality control and variant calling

As the parents were sequenced with modern wheat varieties representing the world genetic diversity and historic Watkins's collection landraces, the QC and SNP calling were carried out by partners. Despite this we have also made the alignment and called the variants, the results of partners were used for further mapping processes as the variant calling in large populations provides an extra power for the exotic variants to be discovered and thus increases the confidence on discovered variants/mutations.

In general, the alignment of parents to Chinese Spring reference genome revealed that Pamyati Azieva possess more than twice as many differential expressed mutations as Paragon based on the size of VCF files obtained as a result of the alignment.

4.2.4.2 Further discovery of the 5A and 6A QTL mapped interval

To squeeze the mapped 5A and 6A QTL intervals we took advantage of data from WGS. Overall, 61503 SNPs were found, for the region of interest (528.1-546.9Mb) within the 5A QTL region, between Pamyati Azieva and Paragon based on reference sequence of Chinese Spring. In contrast to 5A QTL, the 6A QTL mapped interval comprised 7676 SNPs only although the region was wider.

The 61503 and 7676 variants for 5A and 6A respectively covered coding, non-coding and regulatory regions as opposed to only coding regions in exome capture. None of the found variants were used for the marker design due to lack of recombination events within the regions of interest. Thus, the data were used for visualising the mapped genomic intervals and find if there are some interesting insertions and deletions.

4.2.5 The functional annotation of genes for mapped intervals and shrinking the 5A and 6A QTL into smaller interval

The 5A PH QTL was mapped to physical distance of ~18.8Mb on the long arm of wheat 5A chromosome. The region is located between 528.1 – 546.9Mb on the CS genome and in total contains 168 genes among which the number of height related genes is 21. These height-associated genes are localised between 528.2 – 540.5Mb thereby allowing us to squeeze the mapped interval into 12.5Mb region. The list of genes for the 5A QTL mapped interval included: *IAA1* - Indole-3-acetic acid, *FLS2* - Flagellin-sensing 2, *DAW1* - DUO1-activated WD40 1, *AXR1* - NEDD8-activating enzyme E1 regulatory subunit AXR1, *SHI* - Short internodes, *RLCK185* - Receptor-like cytoplasmic kinase 185, *GAUT1 (TaGT)* – Galacturonosyltransferase and the wheat ortholog is *TaGT, SGR* - Staygreen/shoot gravitropism, *APRR5* - Arabidopsis pseudo-response regulators, *ODO1* - Odorant1, *PAP14* - Probable plastid-lipid-associated protein 14 (chloroplastic), *RLCK176*

- Receptor-like cytoplasmic kinase 176, *PAE3* - Pectin acetylesterase 3, *AtGH9B5* -Endoglucanase B class, *ER2* – ERECTA, *SRK6* - Serine protein kinase receptor, *HDG6* - Homeodomain glabrous6, *GLC1* - Glucan endo-1,3-beta-glucosidase, *NORK* - Nodulation receptor kinase, *ASK3 (TaSKP)* - Shaggy-related protein kinase gamma and wheat orthologue is *TaSKP, MSI1* - Multicopy suppressor of ira1 and *TRAB1* - bZIP transcription factor (Table 4.9).

The 6A PH QTL was mapped to physical distance of ~28Mb on the long arm of wheat 6A chromosome. The region is located between 120.2 – 148.2Mb on the CS genome and in total contains 109 genes among which the number of height related genes is 23. The list of genes for the 5A QTL mapped interval included: *SPS1* - Sucrose-phosphate synthase 1, *SRS1* - Small and round seed1, *PAP9* - purple acid phosphatase 9, *AtMYB48* - MYB-Related Transcription Factor, *HOS58* - Homeobox protein knotted-1-like 2 of *Orysa Sativa* (Rice)), *RLCK185* - Receptor-like cytoplasmic kinase 185, *RR24* - response regulator 24, *WRKY18* - WRKY transcription factor 18, *CCR1* - Cinnamoyl CoA reductase, *RUB1* - ubiquitin-like protein related to ubiquitin, *HHP2* - Heptahelical transmembrane protein 2, *TULP4* - tubby-like proteins, *CER3* - eceriferum3, *GHD7* - Grain number plant height and heading date 7, *AP1* - Apetala1, *4CL1* - 4-coumarate--CoA ligase 1, *UVR8* - Ultraviolet-B receptor UVR8, *SRF1* - Strubbelig-receptor family 1, *ERF8* - Ethylene response factor 8, *TULP4* - Tubby-like proteins, *BZIP9* - Basic leucine zipper 9, *PRK1* - Pollen receptor like kinase 1 and *APL* - Altered phloem development (Table 4.10)

Further, height related genes, with corresponding articles, were pulled out based on the information available in database. Looking at each of these height related genes individually for both QTL mapped intervals, we found that most of them are multitrait genes contributing to important adaptation and yield related traits such as plant growth and development (PGD), plant height (PH), quality (QL), disease resistance (DR), stay green (SG), Yield (YD), drought tolerance (DT), abiotic stress tolerance (AST), chloroplast protein degradation (CPD), enhanced biomass yield (EBY), nodulation (NDL), flowering time (FT), self-incompatibility (SI), cold tolerance (CT), slow growth (SLG), plant architecture (PA), regulators of floral fragrance (RFF) and Salinity tolerance (ST) (Table 4.9 and Table 4.10).

Moreover, the databases showed that the genes provided above have wide range of molecular functions such as Protein Binding (PB), DNA Binding (DB), Damaged DNA Binding (DDB), miRNA Binding (mRB), DNA Binding Transcription Factor (DBTF), Protein Binding Transcription Factor (PBTF), miRNA Binding Transcription Factor (mRBTR), Auxin Receptor Activity (ARA), Auxin Binding (AB), Auxin Resistant (AR), Auxin Efflux Transmembrane (AET), ATP Binding (ATPB), Protein kinase activity (PKA), Protein serine kinase activity (PSKA), Protein threonine kinase activity (PTHKA), Protein tyrosine kinase activity (PTYKA), Transcription cis-regulatory region binding (TCRRB), Iron-sulfur Cluster Binding (ISCB),

Accession	Gene ID	Chr	Start	Score	Expressed organ	Mol.Function	Trait	Source
TRAESCS5A02G317200	IAA1	5A	528260775	15.48	RAM, SAM	DBTF, PBTF, mRBTF, ARA	PH, PGD	CS Timpte, et al., 1992, Planta LC Strader, et al., 2002, BMC B Singla, et al., 2006, J Exp Bot
TRAESCS5A02G317700	FLS2	5A	528558253	4.65	AN	PSKA, PTHKA, PTYKA, ATPB, PB	PH, DR	S Sarowar, et al., 2019, Mol Plant Pathol
TRAESCS5A02G318500	AXR1	5A	529088921	21.98	Int2, SP, SPK, GR, SAM, RAM	AR, AET, DB, mRB, PB	PH	C Lincoln, et al., 1990, Plant Cell AS Knöller, et al., 2010, J Exp Bot
TRAESCS5A02G319200	SHI	5A	530301334	14.41	EM, SPK, AN	TCRRB	PH	H Lütken et al., 2010, Plant Biotechnol J
TRAESCS5A02G319400	RLCK185	5A	531082002	2.36	RT	PKA, PHA	PH, DR	D Couto and C Zipfel, 2016, Nat Rev Immunol D Li, et al., 2009, Plant Bio J
TRAESCS5A02G319800	GAUT1 (TaGT)	5A	533020331	16.43	GR, SP, STG, OVR, Int2	GLCA	PH, QL, PGD	TK Pellny, et al., 2012, Plant Physiol Ch Wang, et al., 2019, BMC Genomics
TRAESCS5A02G319900	SGR	5A	533071317	2.71	ANT, FL	ISCB, CB	SG, SLG, PA	AN Farhood, et al., 2020, Sys Rev Pharm
TRAESCS5A02G320300	APRR5	5A	533290702	18.83	Int2, FLB	SSDB, TCRRB	PH, YD, FT	S Fowler, et al., 1999, EMBO J H Gao, et al., 2014, Proc Natl Acad Sci U S A H Sun, et al., 2020, Front Plant Sci
TRAESCS5A02G320500	ODO1	5A	533553547	4.87	PDC	TCRRB, SSDB	RFF, PA	MP Fenske, et al, 2015, Proc Natl Acad Sci U S A
TRAESCS5A02G321200	PAP14	5A	534107499	9.78	RT, FLS, Int2	APA, PPB	DT, PGD, CPD	S Frank, et al., 2019, J Exp Bot Y Jiang, et al., 2020, PeerJ
TRAESCS5A02G322300	RLCK176	5A	534816681	2.36	LM, GL, AN	PHA, PSKA	PH, DR	D Couto and C Zipfel, 2016, Nat Rev Immunol D Li, et al., 2009, Plant Bio J Y Ao, et al., 2014, Plant J
TRAESCS5A02G322400	PAE3	5A	534986400	2.61	STG, OVR	TMRP, PKA	PH, DR, AST	A Souza, et al., 2014, Planta F Philippe, et al., 2017, BMC Genomics
TRAESCS5A02G322700	AtGH9B5	5A	535136781	10.76	PDC	CLA, CHB	PH, EBY	M Glass, et al., 2015, J Integr Plant Biol Y Wang, et al., 2016, Biotechnol Adv
TRAESCS5A02G323800	ER2	5A	535519780	4.65	ANT	IPB, TMRP	PH, YD, DT	J Zheng, et al, 2015, PLoS One M Kulkarni, et al., 2017, Front Chem TA Yasir, et al, 2018, Biol Plant
TRAESCS5A02G323900	SRK6	5A	535528236	7.01	FLS	IPB, TMRP	PH, SI	SJ Hiscock, 2002, Genome Biol D Li, et al., 2009, Plant Bio J
TRAESCS5A02G329600	GLC1	5A	538581665	9.78	ANT	GLC	DR	RN Pudake, et al., 2009, Acta Agron Sin
TRAESCS5A02G330200	HDG6	5A	539504334	21.98	STG, OVR, GR	DBTF, SSDB	PH, FT	W Chew, et al., 2013, Int J Mol Sci HB Patil, et al., 2018, Tree Physiol
TRAESCS5A02G330500	NORK	5A	540050948	2.36	FLS, FLB	PSKA, SB	NDL	G Endre, 2002., 2002, Nature
TRAESCS5A02G331300	ASK3 (TaSKP)	5A	540364015	19.85	STG, OVR, AN	ARA, AB, CFPB	PGD, AST	MJ Hong, et al., 2013, Mol Biol Rep G Ameen, 2019, PhD thesis
TRAESCS5A02G331900	MSI1	5A	540395209	2.35	SAM, SAX	DDB, RB	PGD	J Rodrigues, et al., 2009, Sex Plant Reprod
TRAESCS5A02G332000	TRAB1	5A	540593912	12.93	SAM, SAX	TCRRB, DBTF, SSDB	PH, AST, CT	T Hobo, et al, 1999, Proc Natl Acad Sci U S A W Zong, et al., 2016, Plant Physiol

Table 4.9 The functional annotation of genes for the 5A QTL mapped interval

Accession	Gene ID	Chr	Start	Score	Expressed organ	Mol.Function	Trait	Source
TRAESCS6A02G144800	SPS1 (TaSPS1)	6A	121919225	4.71	SAM, GR, PDC	SSA	PH, PGD, YD	CK Castleden, et al., 2004, Plant Physiol K Ishimaru, et al., 2003, Planta
TRAESCS6A02G145000	SRS1	6A	122679378	4.87	EM, STG, OVR		PH, GRS	F Shang, et al., 2020, Agronomy
TRAESCS6A02G145200	PAP9	6A	123446047	15.15	RAM	PB	PGD	X Zhu, et ak., 2019, J. Plant Biol.
TRAESCS6A02G145400	AtMYB48	6A	123521873	9.78	PDC		DT, ST	H Xiong, et al., 2014, PLoS One
TRAESCS6A02G145500	HOS58	6A	123536071	7.5	LF		PH, PGD	M Sheng, et al., 2002, Plant J.
TRAESCS6A02G145700	RLCK185	6A	124265146	2.36	PDC	PKA, PHA	PH, DR	D Couto and C Zipfel, 2016, Nat Rev Immunol D Li, et al., 2009, Plant Bio J
TRAESCS6A02G146200	RR24	6A	124714684	2.92	RT, SAM	DB, TCRRB	PH	JM Worthen, et al., 2019, Development
TRAESCS6A02G146900	WRKY18	6A	126344720	9.86	LM	NA	DT, ST	I Ahmad, et al., 2019, Biol. Plant.
TRAESCS6A02G147300	CCR1	6A	127189408	2.59	FLB		PH, PGD	R Ren, et al., 2022, pre-print, under-review in BMC T Goujon, et al., 2003, Planta
TRAESCS6A02G147500	RUB1	6A	127604722	21.17	SAM	mRB, DB, UBT	PH, DR	AW Woodward, et al., 2007, Plant Physiol Y Yang, et al., 2017, Plant Cell Rep.
TRAESCS6A02G149300	HHP2	6A	131230859	4.87	SPK	SRA	PGD, CT	HG Lee and PJ Seo, 2015, Plant J.
TRAESCS6A02G149400	TULP4	6A	131443174	4.87	SAM	DBTF	AST	Mwang, et al., 2018, Gene.
TRAESCS6A02G150000	CER3	6A	133755943	4.36	LF, CLP, SPK	OXA	PH, DT	H Kim et al., 2019, The Plant Cell
TRAESCS6A02G150900	GHD7	6A	134947729	16.26	GL, FL	DBTF	PH, YD, FT	W Xue, et al., 2008. Nat. Genet.
TRAESCS6A02G151400	AP1	6A	135433120	4.87	PDC, FLS		PH,FT, PDC	Hyu, et al., 2020, Sci. Hortic
TRAESCS6A02G151700	4CL1	6A	135557573	2.52	PDC		PH, PGD	Jyang, et al., 2011, Plant Physiol. Biochem.
TRAESCS6A02G153100	UVR8	6A	139127657	2.27	SAM	PKA, PB, IPB	PGD	C Shi and H Liu, 2021, Plant Physiol
TRAESCS6A02G153600	SRF1	6A	141358991	9.89	RT, STG, OVR, GR	PKA, PSKA	PH, PGD	D Chevalier, et al., 2005, PNAS
TRAESCS6A02G153700	ERF8	6A	141851251	4.87	LM, GL		PGD, AST, DR	M Dubois, et al., 2018, Trends Plant Sci.
TRAESCS6A02G154100	TULP4	6A	142780132	9.78	ANT	DB	AST	Mwang, et al., 2018, Gene.
TRAESCS6A02G154600	BZIP9	6A	143376818	4.87	PDC, ANT, SAX		PH, AST, CT, DT	P Das, et al., 2019, Rice (N Y) P Agarwal, et al., 2019, Sci Rep H Wang, et al., 2022, Front Plant Sci
TRAESCS6A02G154900	PRK1	6A	143697737	2.36	ANT	PKA, PB	PGD, DR	SH Shiu and AB Bleecker, 2001, Sci STKE
TRAESCS6A02G155400	APL	6A	144247225	20.39	RT	DBTF	PH	M Bonke, et al., 2003, Nature Y Zhang, et al., 2017, PLoS One

Table 4.10 The functional annotation of genes for the 6A QTL mapped interval

Chlorophyll Binding (CB), Sequence-specific DNA Binding (SSDB), Protein Homodimerization Activity (PHA), Transmembrane Receptor Protein (TMRP), Identical Protein Binding (IPB), Glucan endo-1,3-beta-D-glucosidase activity (GLC), Steroid Binding (SB), Cullin family protein binding (CFPB), Ribosome Binding (RB), Cellulase activity (CLA), Carbohydrate Binding (CHB), Acid Phosphatase Activity (APA), Pyridoxal Phosphate Binding (PPB), Glycosyltransferase Activity (GLCA), Ubiquitin-protein Transferase (UBT), Signaling Receptor Activity (SRA), Oxidoreductase Activity (OXA), Protein Kinase Activity (PKA) and Sucrose Synthase Activity (SSA) (Table 4.9 and Table 4.10).

We also were able to pull out the information about the expressed organs of genes. These expected gene expression organs should provide a valuable clue for future studies to filter out the most suitable candidate gene/s to conduct higher resolution mapping as there were plant organs such as Root apical meristem (RAM), Shoot apical meristem (SAM), Shoot Axis (SAX), Awns (AN), - Embryo (EM), Spike - (SP), Spikelets (SPK), Grain (GR), Stigma (STG), Ovary (OVR), Anther (ANT), Flag leaf (FL), Flag leaf blade (FLB), Flag leaf sheath (FLS), Peduncle (PDC), Roots (RT), Lemma (LM), Glumes (GL), Endosperm (ESP), Leaves (LF), Coleoptile (CLP) and Grain shape (GRS) that we could focus and selectively target (Table 4.9 and Table 4.10).

4.3 Discussion

4.3.1 Defining the mapped intervals

Creation and identification of new recombinant events within a genomic region of interest are essential prerequisites for the QTL fine mapping in a map-based cloning. In this respect, we generated homozygous recombinants for 5A and 6A plant height regulating loci. Identified recombinants for both PH QTL were assigned to several recombinant classes and examined in the glasshouse and field. The plant height profile of these RCs obtained from glasshouse followed the field PH data profile closely (Figure 4.9 and Figure 4.10).

However, the variation in PH of 5A RCs was compressed in the field leaving some pairwise comparisons between tall and short RCs being non-significant although relative height difference was observed. Even if we mapped most of the contributing height effect within an interval of 528.1Mb – 546.9Mb, it is likely that other half of the QTL region also adds value to the height quantitatively, based on PH data of RC-05 – RC-08, when the mapped interval possesses tall alleles. Comparatively, when the mapped interval is fixed to Paragon, the other half of the QTL region has less or no influence on PH regarding height profile of RC-01 – RC-03. A clear example is provided by RC-03 and RC-04. The RC-03 was shortest in the glasshouse as well as field. Regardless of the large estimated difference between the means of these two RCs, the statistical significance in the glasshouse was debunked, but it was retrieved in the field due to increased sample size (Table 4.7).

As opposite to RCs of 5A QTL, the 6A QTL RCs were much more informative and stable displaying almost the same height profile across the RCs in the glasshouse and field (Figure 4.10). This was extremely useful to map the QTL more precisely.

4.3.2 Predicting the potential candidate genes

The functional gene analysis conducted retrieved PH related genes for both QTL mapped genomic intervals. The further investigations of each gene aimed at finding the most suitable candidates among these genes. In this respect, we investigated important characteristics of each gene such as molecular function, expressed organ and other traits the gene is associated with, besides PH, using wheat gene expression atlas and evidence-supported gene discovery tools (Winter et al., 2007; Ramírez-González et al., 2018; Hassani-Pak et al., 2021). However, the expression pattern is provided for some genes only.

The functional gene analysis identified 21 candidate genes for the 5A QTL mapped genomic interval. The search for the most likely candidate genes among these lists of genes resulted in the identification of eight genes, out of 21: *IAA1*, *AXR*1, *SHI*, *RLCK185*, *GAUT1*, *APRR5*, *HDG6* and *TRAB1*. Surprisingly these genes also possessed the highest scores of hitting the trait with the exception of *RLCK185* (Table 4.9). Despite this fact, *RLCK* from subfamily VII (PATTERN-TRIGGERED IMMUNITY COMPROMISED RECEPTOR-LIKE CYTOPLASMIC KINASE 1 (*PCRK1*)) mediates BAK1-dependent (BRI1-ASSOCIATED RECEPTOR KINASE 1 also known as SOMATIC EMBRYOGENESIS RECEPTOR KINASE 3 (SERK3)) PTI (PRR-triggered immunity or PAMP-triggered immunity)) responses and *RLCK185* is the member of the rice *RLCK* family VII (Sreekanta et al., 2015; Couto and Zipfel, 2016). In turn, *OsBAK1* gene, which is the closest relative of *AtBAK1*, was used as a molecular tool to improve rice architecture for high yield (Li et al., 2009). Overexpression of *OsBAK1* (intracellular, not extracellular *OsBAK1*) shortened PH (similar to the rice BR-insensitive mutant plants) and altered cell length of the internode-2 and -3. Moreover, *OsBAK1* expression changed important agricultural traits of rice such as grain morphological features, resistance to disease and leaf erectness where the latter is regulated by brassinosteroids (BRs) partially in *Oryza sativa* (Li et al., 2009). The erect leaf phenotype was something we noticed in NIL5A(+) when they were sown in the UK, but the trait was not stable across the years. The development of leaf architecture in *Arabidopsis thaliana* was demonstrated to be controlled by *ERECTA* which encodes a receptor-like kinase and is proposed as a candidate for determining transpiration efficiency of plants (Masle et al., 2005). Interestingly, the mapped interval also contained *ER2* (*ERECTA2*) which might have contributed to the altered leaf angle in NIL5A(+). The wheat orthologs, *TaER1* and *TaER2*, were mainly demonstrated to improve mainly drought tolerance (DT) and yield (Zheng et al., 2015; Kulkarni et al., 2017; Yasir et al., 2018). In

addition, the HDG6 also improved DT, but delayed the flowering which in turn increased PH (Chew et al., 2013; Patil et al., 2017).

Besides, *RLCK185*, three genes, *IAA1*, *AXR1*, and *TRAB1*, are also induced by or involved in the regulations of well-known plant phytohormones such as auxin (AUX) and abscisic acid (ABA) respectively which play a crucial role in controlling many plant growth and developmental aspects including plant height (Hobo et al., 1999; Singla et al., 2006). In addition, the *IAA1* is found to be highly expressed in SAM/RAM and *AXR1* in the internode-2 (Figure 4.11), as well as *APRR5*.

Considering that the yield gain in NIL5A(+) was mainly due to increased germination rate, there is a special gene, *TRAB1*, a bZIP transcription factor, which interacts with *VP1* and mediates ABA-induced transcription. The *VP1* in turn conferred maturation and dormancy in plant seeds by activating genes responsive to the stress hormone abscisic ABA (Hobo et al., 1999). A bZIP transcription factor enhanced DT through the regulation of ABA signalling in rice and it was found that especially *OsbZIP23* TF positively regulates *OsPP2C49* the overexpression of which increased PH significantly relative to wild type (Zong et al., 2016). Here, we can see that bZIPs are key players in ABA signalling pathway and ABA, through the interaction of *AtbZIP39/ABI5*, blocked seed germination and early seedling development (Jakoby et al., 2002).

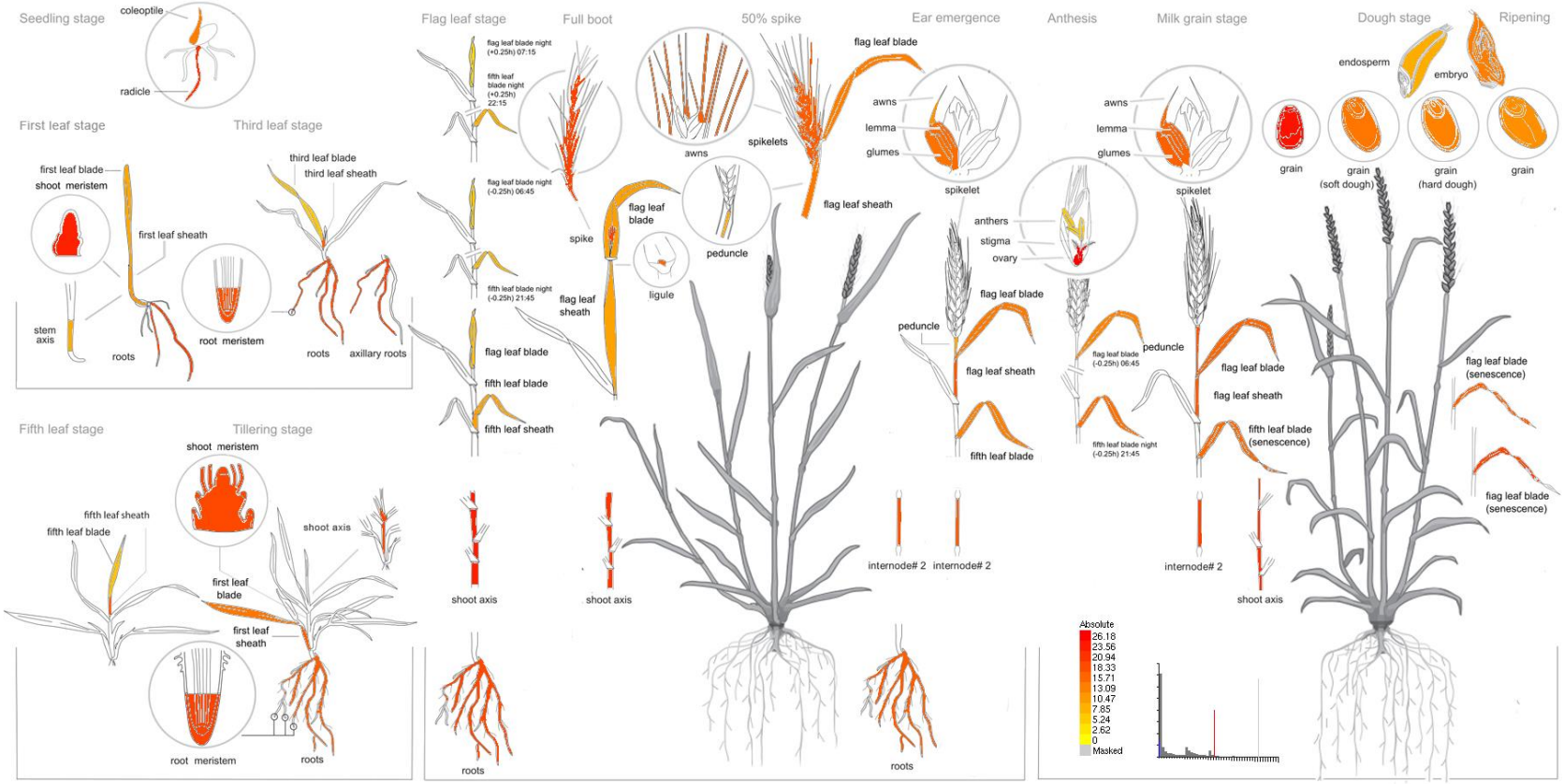


Figure 4.11 The expression pattern of AXR1

The expression pattern is based on RNA-seq data from Azhurnaya spring wheat

Despite that fact that each filtered 23 genes, for the 6A QTL, conferred the alteration in PH, there were some genes which mostly were involved in direct regulation of plant height. Therefore, we selected those nine genes as promising candidates to be researched first in future studies. The list included *RLCK185, RR24, CCR1, RUB1 (TaRUB1), GHD7, AP1, 4CL1, BZIP9* and *APL*. Importantly, as in the 5A QTL mapped interval, two genes, *RLCK185* and *BZIP9* (bZIP transcription factor), were also found to be controlling height in the 6A QTL mapped region (Li et al., 2009; Couto and Zipfel, 2016; Agarwal et al., 2019; Das et al., 2019; Wang et al., 2022). In addition, the accessions related to *RLCK185* and *BZIP9* in wheat were highly expressed in the PDC (Winter et al., 2007; Ramírez-González et al., 2018) (Figure 4.12). This supportive evidence automatically puts these genes on the top of the list to be studied as the effect of 6A PH QTL was conditioned to peduncle length. The next candidates, *AP1* and *4CL1*, are also attractive as they are intensely synthesised in the PDC. Regarding *AP1*, two genes, *MiAP1-1* and *MiAP1-2*, were cloned in mango (*Mangifera indica L*.) the latter of which reduced PH of transgenic *Arabidopsis* (Yu et al., 2020a). The second gene, *4CL1*, was involved in the regulation of stem growth (Yang et al., 2011). The *RR24* (RRs are TF and play significant role in cytokinin signalling in rice) and *RUB1* (also defines DR) both reduced plant height (Woodward et al., 2007; Yang et al., 2017; Worthen et al., 2019). The next promising candidate genes, *GHD7* and *APL* which encode CCT domain protein and MYB TF, respectively, had major influence on PH and both enhance yield potential in rice (Bonke et al., 2003; Xue et al., 2008; Zhang et al., 2017b). The last candidate, *CCR1*, mediated the regulation of plant height in *Brassica napus* (two SNPs in the promoter region of *BnCCR1;* BnaC03g60490D) and *Arabidopsis* (Goujon et al., 2003; Ren et al., 2022).

All of this provided evidence makes all aforementioned genes perfect candidates to consider. The close examination of coding, non-coding and regulatory regions, including of promotors for the genes of interest revealed several SNPs and InDels showing the high genetic diversity between parents for the region of interest. Therefore, further comprehensive genetic studies need to be carried out to pinpoint the underlying causal polymorphism/s.

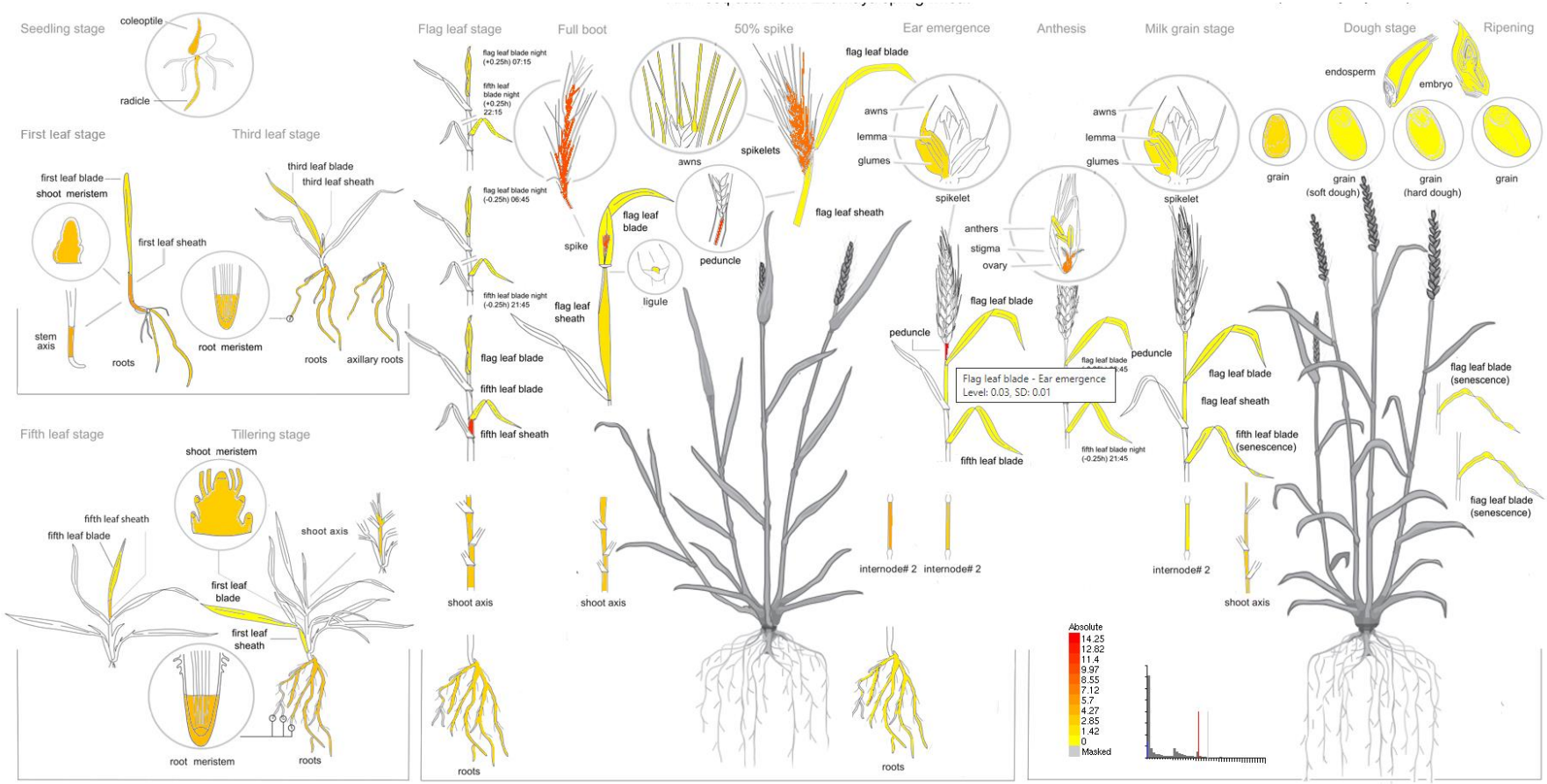


Figure 4.12 The expression pattern of RLCK185

The expression pattern is based on RNA-seq data from Azhurnaya spring wheat

5. Broader analysis of genetic variation in Central Asian wheat

5.1 Introduction

Breeding for only favourable alleles (targeted breeding) led to the limited genetic diversity in contemporary crop cultivars including wheat. Nevertheless, important gene sources are kept as wild relatives, landraces and traditional cultivars in gene banks and/or conserved by farmers who continue to grow old cultivars and even landraces. A recovery of those alleles for genes of agricultural importance has always been an important target for breeders when a new strain of diseases emerges or to improve adaptation to overcome issues caused by abiotic stresses (Duveiller et al., 2018). With the advent of or/and advance in new molecular breeding technologies, tracking and delivering these essential genomic sections, discovered in one particular genotype, to the next became easier and quicker (Varshney et al., 2020). Especially, molecular tools such as DNA markers, co-segregating with the trait of interest, make it possible to transfer the gene/s and alleles between accessions with great precision compared to conventional selection methods. Genetic fingerprinting of wider populations adapted to diverse environments with thousands of genetic markers allows specific inspection of exotic genomic regions under selection pressure in certain environments in a greater resolution than when these populations are only compared phenotypically. In this chapter, the assembly of a new CA wheat panel is described, alongside its genetic fingerprinting and comparison with a European wheat panel and historic Watkins's landraces collection known as GEDIFLUX and Watkins collection respectively (Wingen et al., 2014). The GEDIFLUX is a collection comprised of 473 Western Europe and UK elite varieties and the Watkins collection consisted of more than 7000 accessions (but we used the core collection consisting of 825 lines) collected from local markets in 32 countries in the 1930s and thus nearly captures global wheat diversity (Wingen et al., 2014; Aradottir et al., 2017). Therefore, in this chapter we also discuss how we used these valuable genetic resources to validate the marker effects which were used to fine-map the two height related genes studied in this PhD research and to estimate their allelic frequencies.

5.2 Materials and methods

5.2.1 Plant resource preparation

The germplasm panel consisted of 489 wheat accessions from all over the CA countries, Kazakhstan (276 spring, 64 winter and 3 facultative wheat varieties), Uzbekistan (30 winter), Kyrgyzstan (19 winter and 11 facultative), Afghanistan (7 spring), Tajikistan (9 winter), Russia (25 winter and 1 facultative), Turkey (30 winter), Ukraine (2 winter) and others (12 winter). The panel was assembled as a result of the Central Asian Workshop, funded by BBSRC, which took place in Astana (currently Nur-Sultan) and Almaty in May 2018, Kazakhstan. By the end of the year 2018 the panel was sent to and received by John Innes Centre. The new germplasm panel has been named as **C**entral **A**sian **W**heat **B**reeding **In**itiative (CAWBIN) although there are some wheat accessions from Russia (southern Siberia which is located in Eurasia) which were provided by

Central Asian partners. Thus, we want to stress that the naming does not bear any political and regional rationale.

The CAWBIN was taken through one round of SSD (Single Seed Descent) purification during this PhD work (2018 - 2020) to minimise heterogeneity and heterozygosity. For that the entire population was grown in the JIC glasshouse. Some lines were not viable at all and thus excluded from the final collection. Seventy five lines gave SSD purified seeds, but they were not viable. Therefore, while multiplying, the original seeds were used for these lines. The lines with F1 SSD purified seeds and the seeds from original lines (these lines either did not survive during SSD purification or seeds were not viable) were sown, bagged, grown, harvested as well as threshed in the glasshouse. Then the glasshouse seeds were sent to KZ and grown in Alm 2021.

5.2.2 Genetic fingerprinting

From the SSD purified plants, high-quality genomic DNA was extracted with Oktopure at the JIC using the services of the in-house genotyping platform. Then genomic DNA was sent to and genotyped using 35K Axiom® wheat HD Genotyping Miroarray by Bristol University (Burridge et al., 2018). The panel was genotyped, and data were received. Because of array related issues, 73 out of 475 lines failed, 15 lines did not pass QS and 10 lines did not cluster well. These fails and lines with QC problems were re-sent and re-genotyped.

Watkins and GEDIFLUX were already genotyped with the same microarray and platform.

5.2.3 Assigning the physical positions to markers

The raw genetic data of CAWBIN had 35143 SNP markers in total with genotyping call codes - NoCall (-1), AA (0), AB (1) and BB (2) where AA (0) and BB (2) are for homozygotes, AB (1) and NoCall (-1) are heterozygotes and missing genotypes respectively. We then converted the homozygotes genotyping calls to single nucleotides. Accordingly, heterozygotes and missing genotyping calls were converted to IUPAC nucleotide codes using R. The IWGSC physical coordinates were assigned to each axiom marker, based on 16 genetic maps developed at JIC, with Python.

5.2.4 Quality control of the genetic data

Quality control of the genotype data was performed with PLINK version 1.9. We removed SNPs and individuals having more than half of their values missing. Genotyping data then were masked and imputed by widely used method LD KNNi in TASSEL software v5.2.73 (https://doi.org/10.1534/g3.115.021667). The number of sites and nearest neighbours for imputation were set as 525 and 85 respectively with the default maximum distance between sites to find LD. For the purpose of masking, we used default settings.

5.3 Results

5.3.1 Quality control

The total genotyping rate accounted for almost 97 percent of markers. Fifty-six SNPs of initial 35143, did not pass the QC, and therefore 35087 genomic regions were used to assess the genetic relatedness of a germplasm panel. In contrast, all samples had more than 50% of their genotypes present at the SNPs positions. More than ninety four percent imputation accuracy between original, masked and masked imputed HapMaps was detected.

5.3.2 The genetic fingerprints of wheat breeding in Central Asian wheat compared with GEDIFLUX and Watkins landraces

The Watkins landraces were used as a source to evaluate wheat breeding progress and historical relationships in Central Asia compared to Europe (EU, GEDIFLUX). Particularly, it was interesting to understand how the genetic history of wheat breeding in the world shaped the Central Asian wheat compared to European. In addition, it was used to assess how much the Central Asian wheat has diverged from the ancestral roots compared to European. For that we used "Principal Coordinates Analysis" (Gower, 1966), known as PCA analysis, to calculate the genetic distances between individuals (Figure 5.1). The results revealed that CAWBIN and GEDIFLUX overlap with the Watkins collection which certainly can theoretically be explained. However, CA and EU lines grouped far from each other, showing a minor genetic interaction between the two. Those bridging lines were found to be from Germany. Looking at the origins of these lines we found that that the lines clustered closer to CAWBIN and GEDIFLUX were from GDR (German Democratic Republic) and FRG (Federal Republic of Germany) respectively. This obviously shows the signature of politics in plant breeding and perhaps partially explains the fact of CAWBIN and GEDIFLUX being genetically distant/distinct. Of course, there are other important factor to consider such as plant habitat. For instance, EU mostly sows winter wheats compared to spring dominated CA (mainly KZ and RUS). Within CAWBIN a high genetic relatedness between Kazakh and Russian wheats, compared to other Central Asian states, was identified which is not a "surprise" (Figure 5.1) as the two have been sharing wheat breeding history since Soviet times and still have a close partnership in plant breeding especially of wheat. Moreover, it could also be due to the higher selection pressure in wheat breeding programs in these two countries as Kazakhstan and Russia are the main wheat exporters in the region.

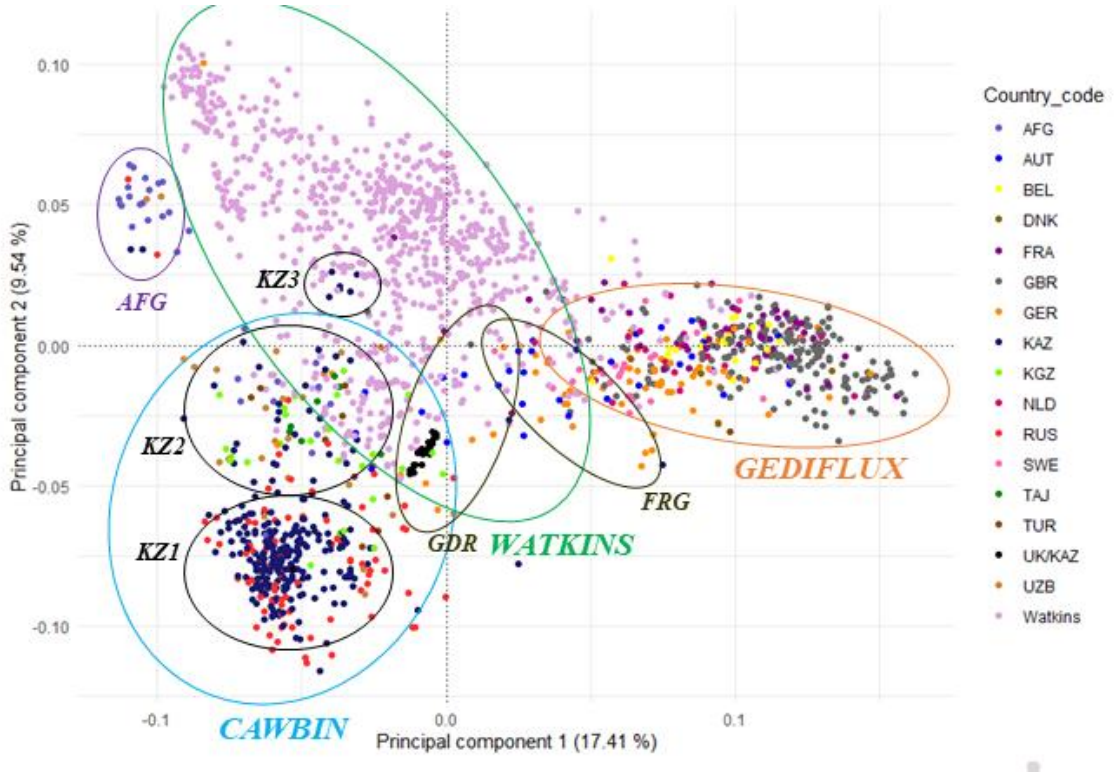


Figure 5.1 PCA of the world accessions and Watkins landraces

The PCA plot includes wheat accession from countries such as Afghanistan (AFG), Austria (AUT), Belgium (BEL), Denmark (DNK), France (FRA), Great Britain (GBR), Germany (GER), Kazakhstan (KAZ), Kyrgyzstan (KGZ), Netherlands (NLD), Russia (RUS), Sweden (SWE), Tajikistan (TAJ), Turkey (TUR), UK/KAZ (NILs), Uzbekistan (UZB) and Watkins (Watkins collection).

PCA analysis divided the varieties from Kazakhstan into three main clusters (KZ1, KZ2 and KZ3). The main cluster – KZ1 - overlapped closely with the Russian varieties and has minimal overlap with Watkins. These lines were mostly spring habit. The second group – KZ2 – is distinct from the main cluster and overlaps with the cluster containing lines mainly belonging to other CA countries (CAC) such as UZB, KGZ and TJK. Surprisingly, half of them were a mixture of spring and winter growth. In the latter case it is understandable as all of the CAC grows winter wheats except KZ and southern part of RUS. However, in the former case where relatively large number of Kazakh spring wheats grouped with other CAC and some of the Watkins collection is something interesting as the vast majority of these varieties were developed in and after 2000's. It is possible that they are the result of the recent wheat breeding programs established after the dissolution of USSR. The third group – KZ3 – contained *T.timopheevii* and *T.millitinae* crossed lines developed by Prof A Abugalieva (may Allah grant her with the highest status in Jannah) at the LLP "Kazakh Research Institute of Agriculture and Plant growing" (KRIAPG) and they cluster within the Watkins collection. When the PCA outcome is looked at from the different angle, PC2 against PC3, this group clearly is distant from the rest (Figure 5.2). Interestingly, there was a fourth group, containing six lines from CAWBIN, which tightly grouped with AFG lines. However, there were

only two samples (C-008_KAZ_W and C-368_KAZ_S) from KZ. The others were two and two from from UZB (C-072_UZB_W and C-053_UZB_W) and RUS (C-387_RUS_S and C-442_RUS_S) respectively. Interestingly, five of them were varieties released between 2000-2012, except C-387_RUS_S which was developed during USSR in 1980's. When we looked at the varieties possessing facultative growth habit in Kazakh gene pool, there were two separate clusters one of which went into KZ1 and the second grouped with KZ2 (Figure 5.4).

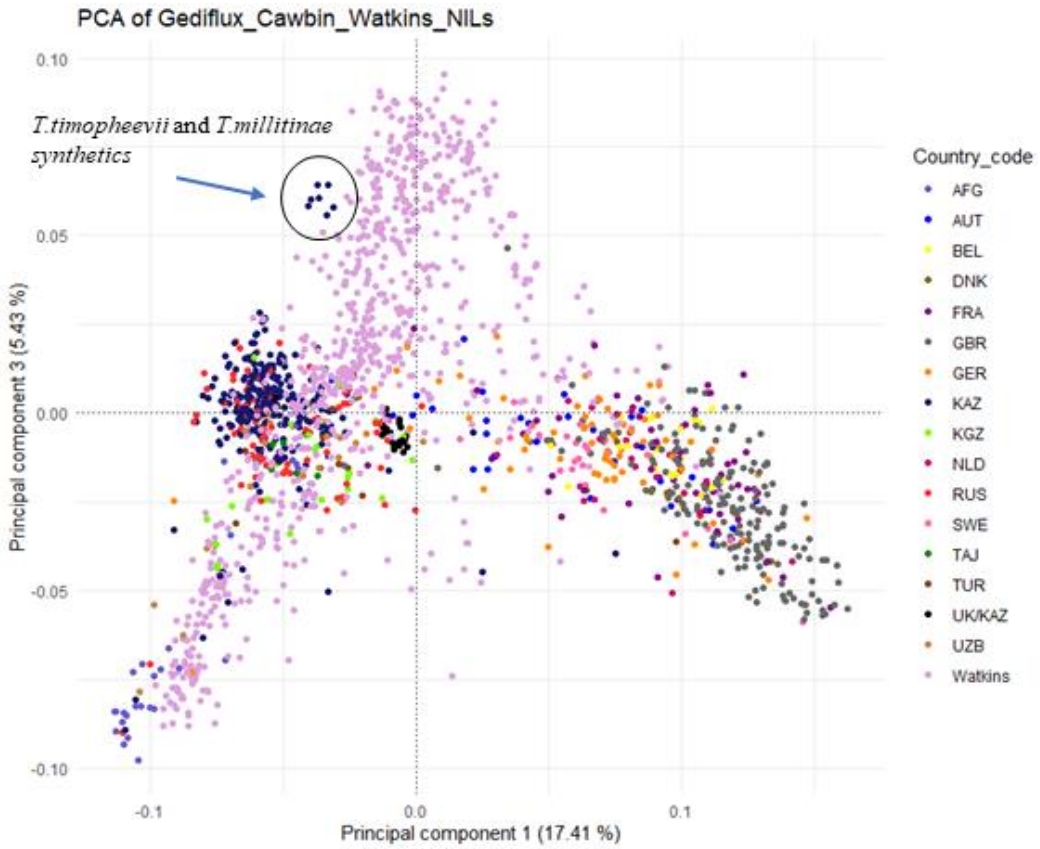
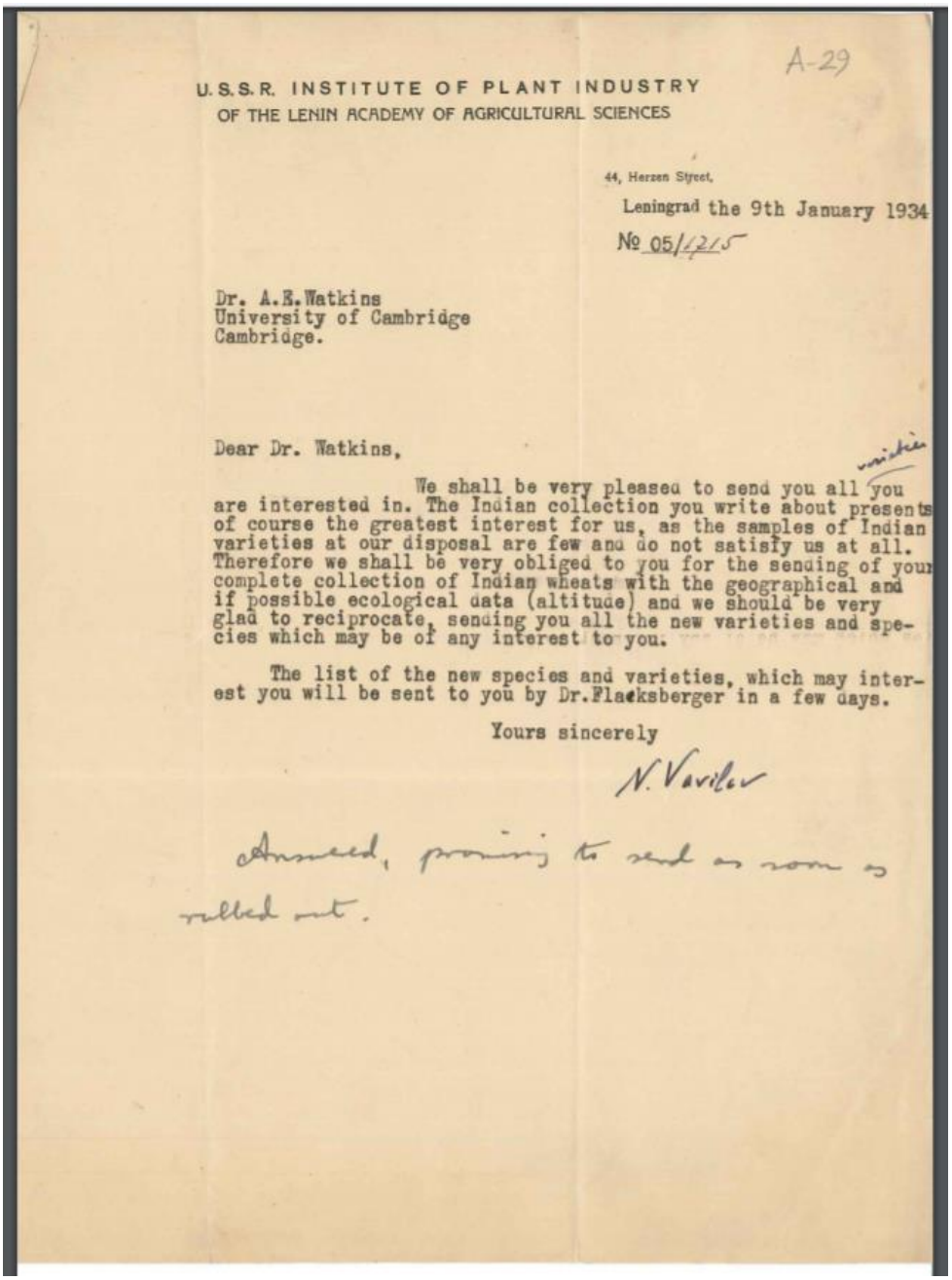


Figure 5.2 PCA of T.timopheevii and T.millitinae lines

In the previous PCA plot (Figure 5.1) *T.timopheevii* and *T.millitinae* lines clustered together with Watkins lines under KZ3 group. For their better visualisation, PCA1 was plotted against PCA3.

The lines in the Watkins collection overlapping with the second KZ and CAC are found to be mainly from USSR (Figure 5.5). Perhaps these lines were sent to Watkins by Vavilov (Figure 5.3) and CA could be the region of origin of these lines. There were other USSR lines in the Watkins collection clustering far from the rest. They were close to Afghan lines which created a distinct cluster from the rest. However, almost half of the AFG lines grouped with KZ2 where mostly CAC dominate.



A-29

U.S.S.R. INSTITUTE OF PLANT INDUSTRY
OF THE LENIN ACADEMY OF AGRICULTURAL SCIENCES

44, Herzen Street,
Leningrad the 9th January 1934
№ 05/1215

Dr. A.E.Watkins
University of Cambridge
Cambridge.

Dear Dr. Watkins,

We shall be very pleased to send you all you are interested in. The Indian collection you write about presents of course the greatest interest for us, as the samples of Indian varieties at our disposal are few and do not satisfy us at all. Therefore we shall be very obliged to you for the sending of your complete collection of Indian wheats with the geographical and if possible ecological data (altitude) and we should be very glad to reciprocate, sending you all the new varieties and species which may be of any interest to you.

The list of the new species and varieties, which may interest you will be sent to you by Dr.Flaeksberger in a few days.

Yours sincerely

N. Vavilov

Answered, promising to send as soon as ruled out.

Figure 5.3 The letter of N Vavilov to A. E. Watkins

The genetic data were split by chromosomes and are being analysed. However, the results are not provided as this was not the main purpose the PhD research aimed at. Rather we intended to use this important genetic resource to validate the markers co-segregating with the height loci.

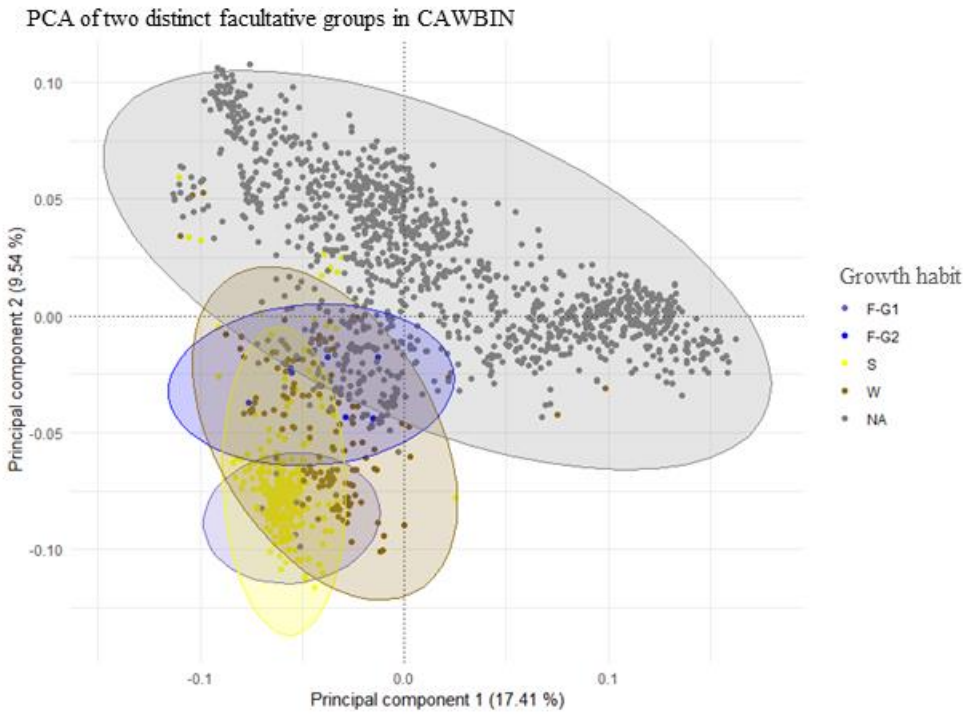


Figure 5.4 The growth habit of wheat from Kazakhstan

F-G1 and F-G2 refer facultative group - 1 and groups – 2 respectively. S and W are groups with spring and winter growth habit. NA is Watkins collection (the growth habit of Watkins lines were not of interest).

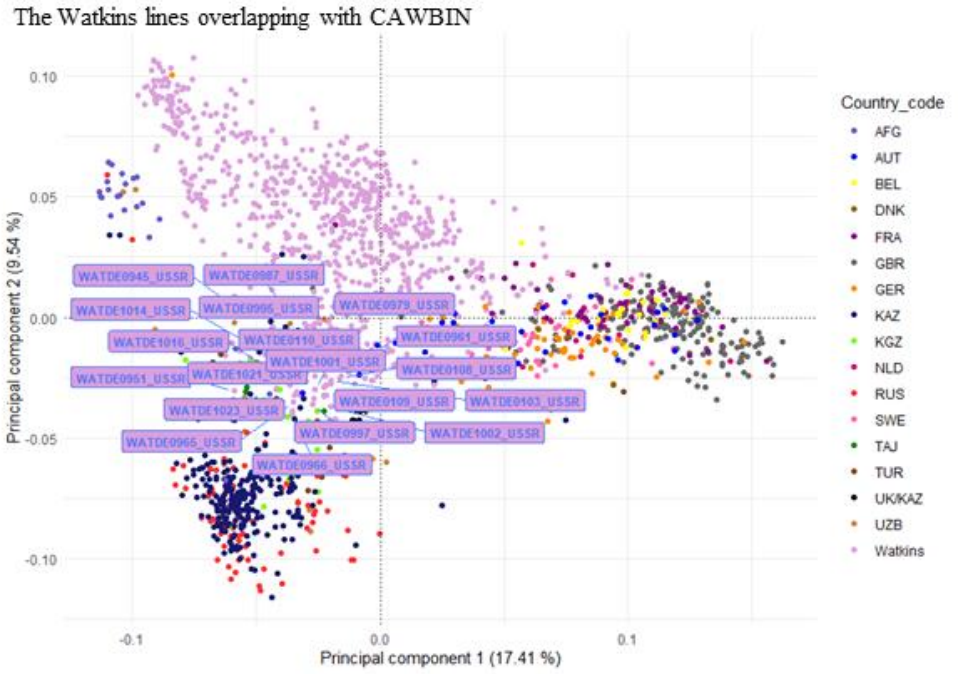


Figure 5.5 USSR lines within the Watkins collection

The Watkins lines overlapping with KZ2 and CAC are labelled. Country code refers to nations such as Afghanistan (AFG), Austria (AUT), Belgium (BEL), Denmark (DNK), France (FRA), Great Britain (GBR), Germany (GER), Kazakhstan (KAZ), Kyrgyzstan (KGZ), Netherlands (NLD), Russia (RUS), Sweden (SWE), Tajikistan (TAJ), Turkey (TUR), UK/KAZ (NILs), Uzbekistan (UZB) and Watkins (Watkins collection)

5.3.3 Harnessing the CAWBIN and GEDIFLUX population for marker validation

Once the two height increasing loci were fine-mapped using recombinants, the genotyping of CAWBIN and GEDIFLUX with the flanking markers allowed the validation of the co- molecular markers in terms of their association with PH in independent and diverse populations. In this respect, CAWBIN was grown in Alm 2020, Kazakhstan. The GEDIFLUX field data however were collected in 2011 and 2016 in the UK.

Genotyping the CAWBIN with three markers (at the beginning, peak and end of the mapped interval located at physical positions 528.1Mb, 537.9Mb and 546.9Mb) controlling the 5A height allele resulted in the identification of four haplotype blocks (HTB) (Figure 5.6). The two "talls" (HTB-01 and 02) were significantly taller that the two "shorts" (HTB-03 and -04) with no significant height difference between themselves. Importantly, the vast majority of lines (68%) in CAWBIN were fixed for Pamyati Azieva allele at the locus and this was the first HTB (Figure 5.7).

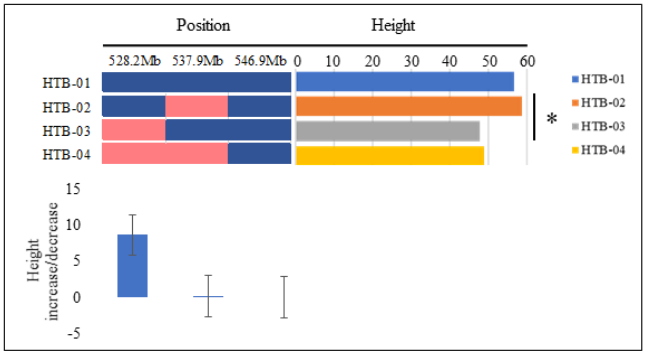


Figure 5.6 Allelic combination of HTB for CAWBIN at the 5A locus

Coral and blue correspond to Paragon and Pamyati Azieva respectively. PH of each HTB is given on the right of the map as horizontal bar plots (cm). Asterisk indicates 95% statistical significance level between two "tall" and two "short" HTBs. Significance codes: 0 '***' 0.001, '**' 0.01, '*' 0.05 or ns = nonsignificant.

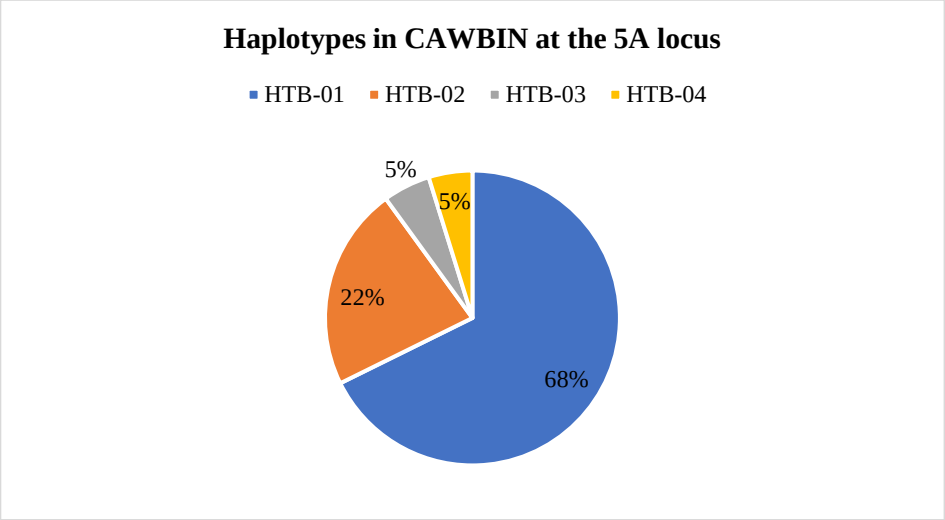


Figure 5.7 The relative percentage of haplotype groups in CAWBIN at the 5A locus

In the second HTB, the lines had a recombination event taking place approximately at the 537.9Mb genomic position. This block comprised 22% of the total. The third and fourth haplotype groups each had 5% of lines. The HTB – 03 was fixed for Paragon at the first marker (528.1Mb). The last block, HTB – 04, was fixed also for Paragon at the peak and end markers for the mapped interval. The height difference was noticeable and significant between these HTB which could be due to background mutation including at the 6A loci. In addition, this variation in plant height does not add value to our understanding of how 5A Pamyati Azieva allele controls height (locus is silenced at the environment where it was constantly bred for and always expresses as height locus in the UK).

Compared to CAWBIN, the western European panel (GEDIFLUX) was highly diverse at the 5A locus. We identified 13 major HTB (there were even more, but these were the main ones with the comparable samples sizes) in the panel. In total 12 markers, located within the mapped interval, were used for genotyping. Although, the number of HTB in GEDIFLUX was high, there were also two groups with clear height difference. The first one includes HTB-01, HTB-02 and HTB-03. These were taller than the second group which included the rest starting from HTB-04 to HTB-13. Pairwise comparisons were conducted and showed that many of these comparisons are not significant. Among "tall" HTBs, only HTB-02 was significantly taller than HTB-08, 09, 10, 11 and 13 (Figure 5.8).

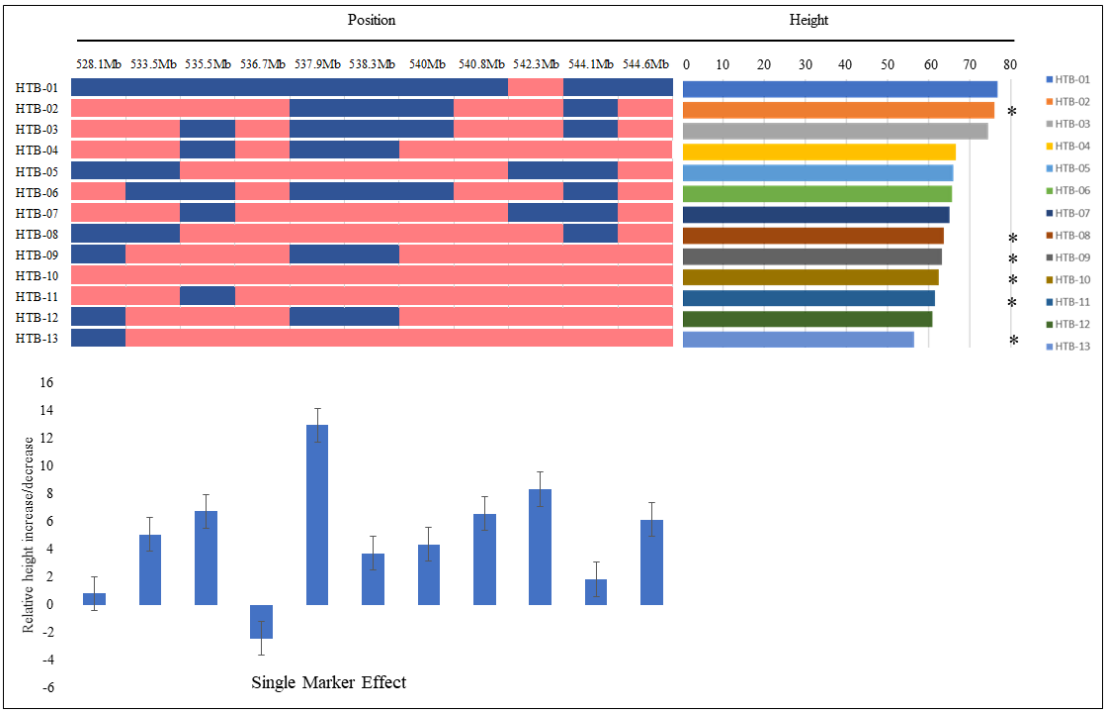


Figure 5.8 Allelic combination of HTB for GEDIFLUX at the 5A locus

Coral and blue correspond to Paragon and Pamyati Azieva respectively. PH of each HTB is given on the right of the map as horizontal bar plots (cm). Single marker effect on the bottom of the map. Asterisk indicates 95% statistical significance level between "tall" HTB-02 and "short" HTBs. Significance codes: 0 '***' 0.001, '**' 0.01, '*' 0.05 or ns = nonsignificant.

The molecular markers we developed for the 6A mapped interval were tested on CAWBIN. This genetic data then was combined with the field data on CAWBIN which was grown and phenotyped by Prof Yerlan Turuspekov's team in Alm 2021. Most of the lines were carrying Pamyati Azieva's allele for the entire mapped interval in CAWBIN as expected. There were only two haplotype groups. The HTB-01 and HTB-02 was entirely fixed for Pamyati Azieva and Paragon respectively (Figure 5.9). Proportionally, HTB-01 dominated significantly over the HTB-02 (Figure 5.10).

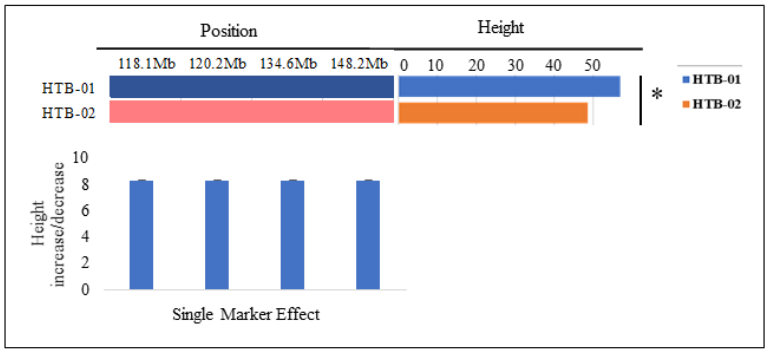


Figure 5.9 Allelic combination of HTB for CAWBIN at the 6A locus

Coral and blue correspond to Paragon and Pamyati Azieva respectively. PH of each HTB is given on the right of the map as horizontal bar plots (cm). Asterisk indicates 95% statistical significance level the between "tall" and "short" HTB. Significance codes: 0 '***' 0.001, '**' 0.01, '*' 0.05 or ns = nonsignificant.

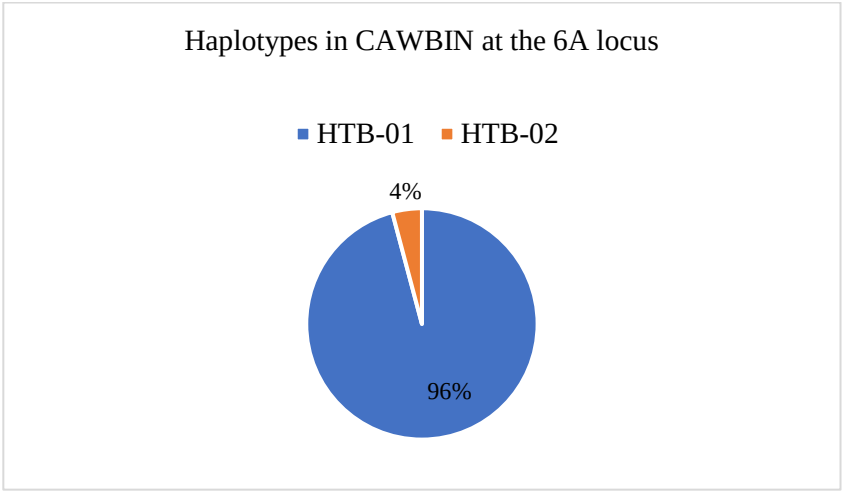


Figure 5.10 The relative percentage of haplotype groups in CAWBIN at the 6A locus

When the 6A markers for the mapped interval were run on GEDIFLUX, again two haplotype blocks were identified the first of which consisted of 405 and the second of 33 lines. So, in total 438 lines were useful out of initial 473 to form the HTBs in GEDIFLUX. Interestingly, the first group (HTB-01) which was the largest in GEDIFLUX was entirely fixed for Pamyati Azieva's allele at the 6A mapped interval as in CAWBIN. The height difference between the two HTBs, which showed the same additivity when the locus was identified in PamxPar, was significant at the

5% level (Figure 5.11). The relative proportion of HTB-01 compared to HTB-02 in GEDIFLUX is provided in Figure 5.12 and showed similar results to what was observed in CAWBIN. The only difference was that the lines carrying the second haplotype (HTB-02) was fixed for Pamyati Azieva at around 134Mb region. Therefore, it would be interesting to look at and compare HTB-02 of CAWBIN with HTB-02 of GEDIFLUX with the consideration of the allelic state of each at the 5A locus as it might also contribute to the overall height.

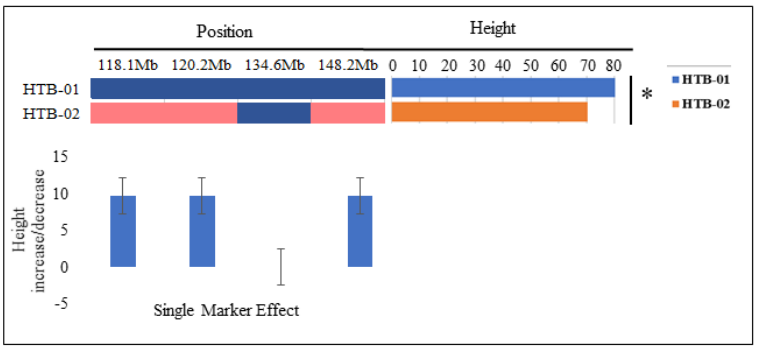


Figure 5.11 Allelic combination of HTB for GEDIFLUX at the 6A locus

Coral and blue correspond to Paragon and Pamyati Azieva respectively. PH of each HTB is given on the right of the map as horizontal bar plots (cm). Asterisk indicates 95% statistical significance level the between "tall" and "short" HTB. Significance codes: 0 '***' 0.001, '**' 0.01, '*' 0.05 or ns = nonsignificant.

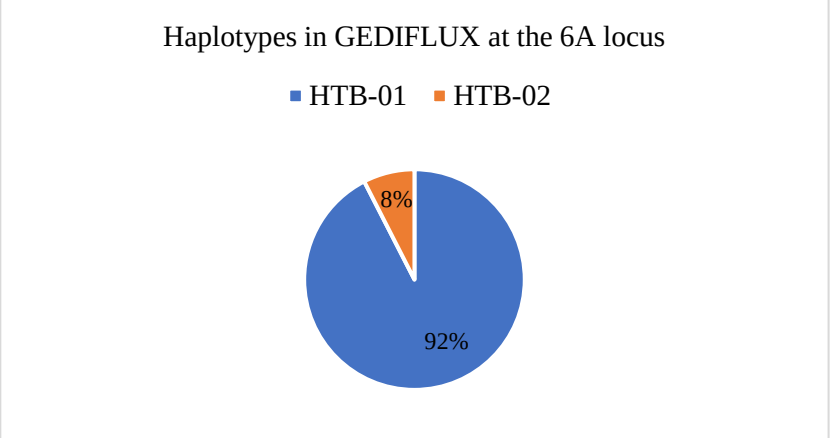


Figure 5.12 The relative percentage of haplotype groups in GEDIFLUX at the 6A locus

So, comparing and contrasting the different HTBs at the 5A and 6A locus in CAWBIN and GEDIFLUX provided some indication that both loci could be increasing or/and to some extent are contributing to the height regulation in these two large populations adapted to two different environmental conditions. However, the observed height difference between HTBs of these populations might be due to the contribution of 6A locus when the difference was detected in HTBs at the 5A or vice-versa. To address this question, we combined the GEDIFLUX data at the 5A and 6A loci. When data were joined, it was extremely difficult to keep track of samples in HTBs

identified at the 5A locus earlier. Nevertheless, 320 lines out of initial 473 formed 12 new HTBs in GEDIFLUX (Figure 5.13). From the total 12 HTBs, six HTBs – HTB-01, HTB-02, HTB-07, HTB-08, HTB-10 and HTB-11 were considered as "tall" at this stage. Among these "tall" HTBs, haplotype group HTB-02 was the tallest and HTB-01 was one of the tall groups compared to other HTBs. The vast majority of loci in these two haplotype blocks were fixed for Pamyati Azieva, HTB-01 being almost entirely fixed for Pamyati Azieva (Figure 5.13).

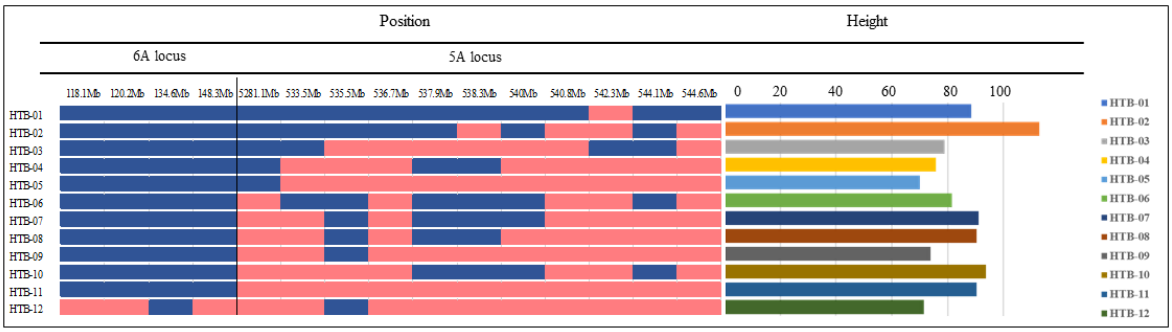


Figure 5.13 Allelic combination of HTB for GEDIFLUX at the 5A and 6A locus

Coral and blue correspond to Paragon and Pamyati Azieva respectively. PH of each HTB is given on the right of the map as horizontal bar plots (cm).

Surprisingly, this haplotype group – HTB-01- composed four lines, Mironowskaya-Jubileinaya (Mironovskaja 50), Mironovskaya-808, Miras and Frista, the first two of which are famous USSR wheat varieties released in 1950-60's. The third variety, Miras, was developed in the 1980's and its pedigree include MIRONOVSKAYA-808/LUTESCENS-2539, UKR(LUTESCENS-4471)//VIGINTA (MIRAS (wheatpedigree.net)). No data is available for "Frista". Looking at the tallest haplotype (HTB-02), it was identified that it is consisted of five old varieties (Dr. Lassers Dickkopf, Loosdorfer Austro Bankut Grannen, Stamm 101, Ritzlhofer Neu and Tschermaks Weisser Begrannter Marchfelder) which were released in 1940 – 50s (Accessions List (wheatpedigree.net)).

The multiple comparisons test which was carried out to assess the statistical significance level between HTBs, obtained from the combined 5A and 6A loci, in GEDIFLUX showed that the plant height of HTB-01 and HTB-07 are not statistically different neither from "tall" nor "short" HTBs although there were considerable estimated differences between the means (Figure 5.14). However, HTB-02 was significantly taller than all including some of the "tall" HTBs such as HTB-08, HTB-10 and HTB-11. Presumably, this extra "tallness" is due to the fact that this specific haplotype block contained the pre-green revolution developed accessions (as mentioned earlier). The plant height of remaining "tall" HTBs that are HTB-08, HTB-10 and HTB-11 were significantly taller than that of some "short" HTBs but not of all (Figure 5.14).

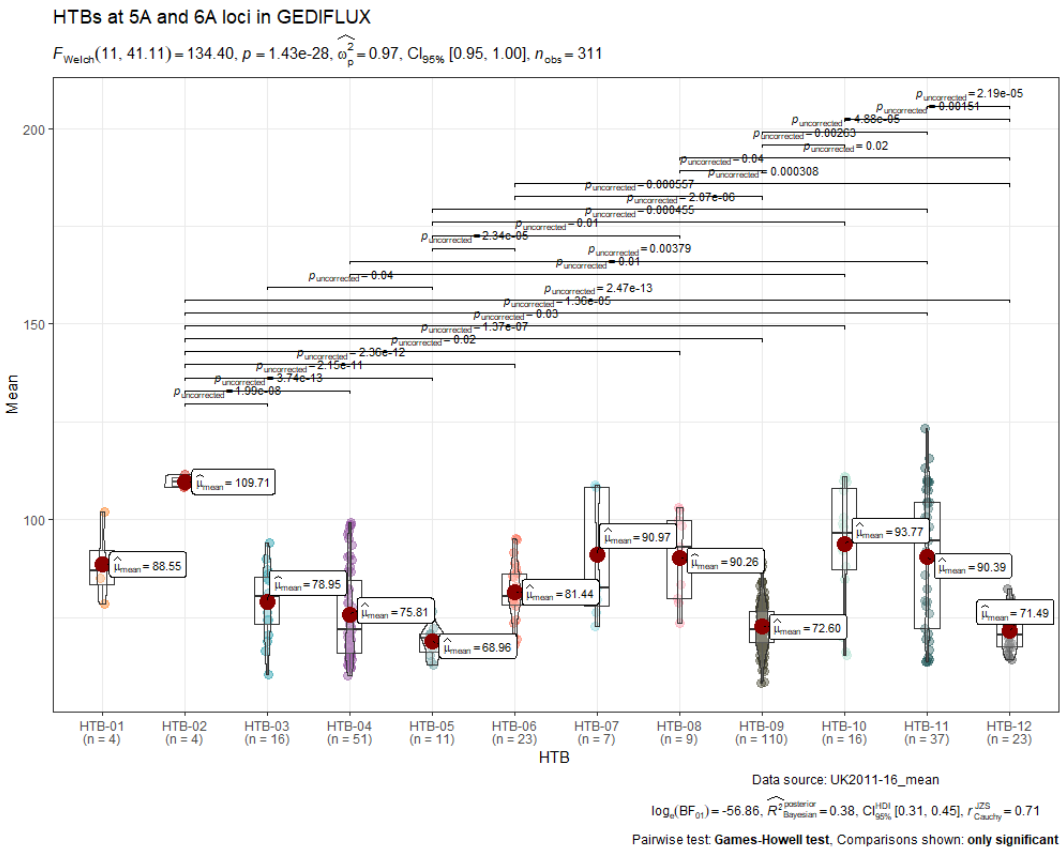


Figure 5.14 Significance test of HTBs at 5A and 6A loci in GEDIFLUX

So, this finding indicates that when 5A and 6A are combined they might influence the plant height at some point, but not always.

The combination of 5A and 6A genotypic and phenotypic data resulted in five main haplotype groups for CA wheat (Figure 5.15). The HTB-01 and HTB-03 were significantly taller than HTB-02, HTB-04 and HTB-05 (Figure 5.16). In addition, these two haplogroups possessed a large sample size compared to others, HTB-01 being the largest (75% or 224 lines out of total 280) (Figure 5.17). Although, the total sample size of CAWBIN was 438 when HTBs were discovered, 36% of them did not have PH data. Therefore, the total number of lines which had PH data was 280 and this data was equally distributed among HTBs.

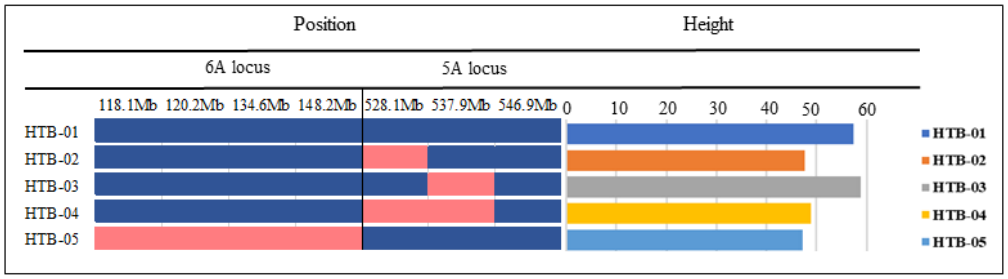


Figure 5.15 Allelic combination of HTB for CAWBIN at the 5A and 6A locus

Coral and blue correspond to Paragon and Pamyati Azieva respectively. PH of each HTB is given on the right of the map as horizontal bar plots (cm).

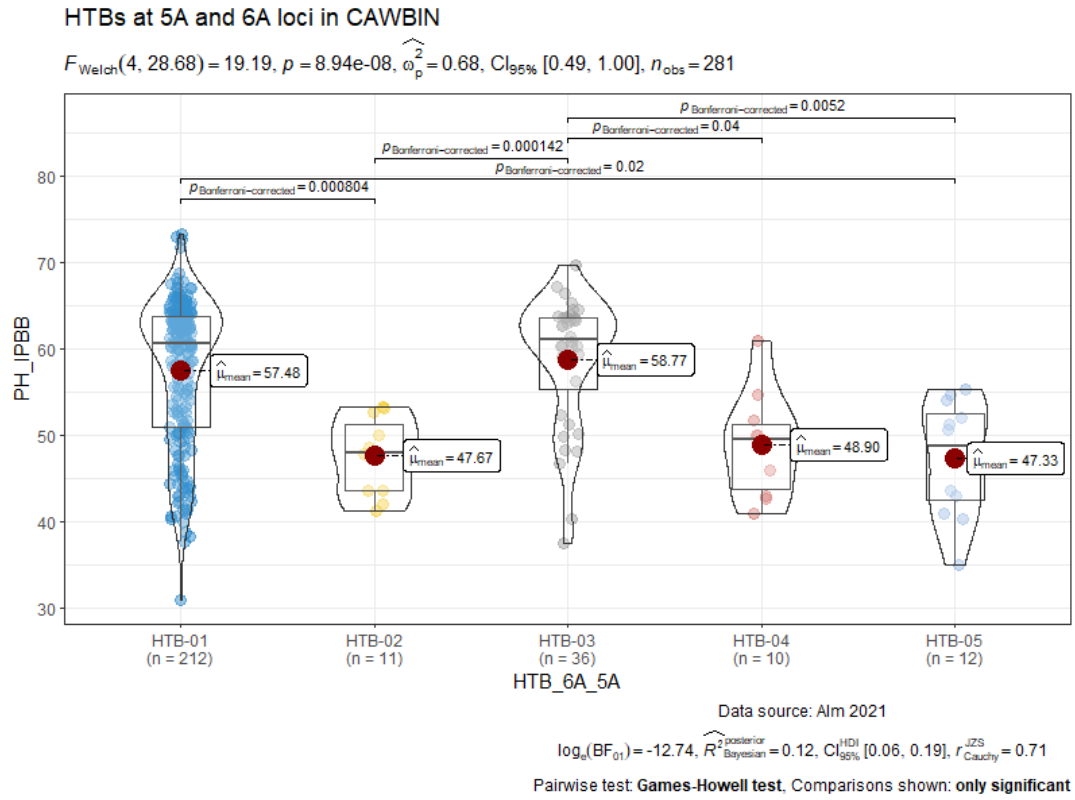


Figure 5.16 Significance test of HTBs at 5A and 6A loci in CAWBIN

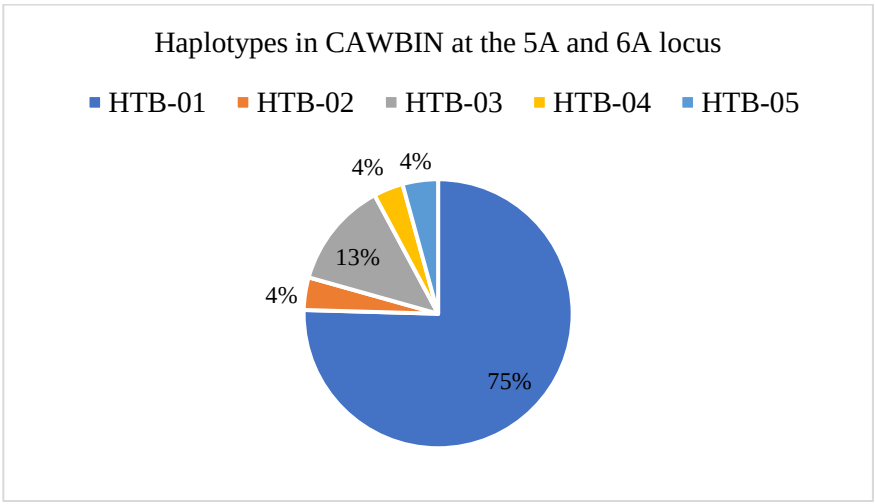


Figure 5.17 The relative percentage of haplotype groups in CAWBIN at the 5A and 6A locus

Looking at the HTBs with their PH data and comparing it with ones in GEDIFLUX suggested two possible acting modes of 5A locus:

1. The 5A locus is totally silenced in KZ. The reason for breeding for the allele in KZ is that it might have a positive effect on grain yield components. That means breeders breed for 5A PH QTL not for the sake of increased height but to achieve a gain in grain yield. However, it is important to stress the ineffectiveness of the Pamyati Azieva 5A allele in the

southern part of KZ as it significantly decreased the yield (p = 0.04) in a critical season of Alm 2021.

2. The 5A locus might be dependent in its action on 6A locus. When 6A "Pam" allele is present, it increased the height of HTB-01 and HTB-03 (Figure 5.15). When 6A "Pam" locus is absent as in HTB-05, the height effect of the 5A locus is silenced. In case of HTB-02 and HTB-04, these haplogroups basically might be missing the Pamyati Azieva's "tall" allele at the locus.

Having said all of these, the question "if the 5A and 6A act together why is the additivity half what we discovered in the UK?" led us to develop two hypotheses:

1. The height effect of both loci is suppressed by environmental stresses such as drought and heat.
2. Although 5A locus increases height, it has less additive effect on the trait in KZ.

Besides the acting mode of 5A locus, putting genetic data of two loci together seemed to provide further evidence for shrinking the 5A QTL mapped interval from ~12.5Mb (528.2 – 540.5Mb) to ~1.2Mb between the physical locations of 536.7 and 537.9 on CS genome. This was achieved comparing the heights of HTB for CAWBIN and GEDIFLUX (Figure 5.18).

Shrinking 5A gene location to 536.7 and 537.9

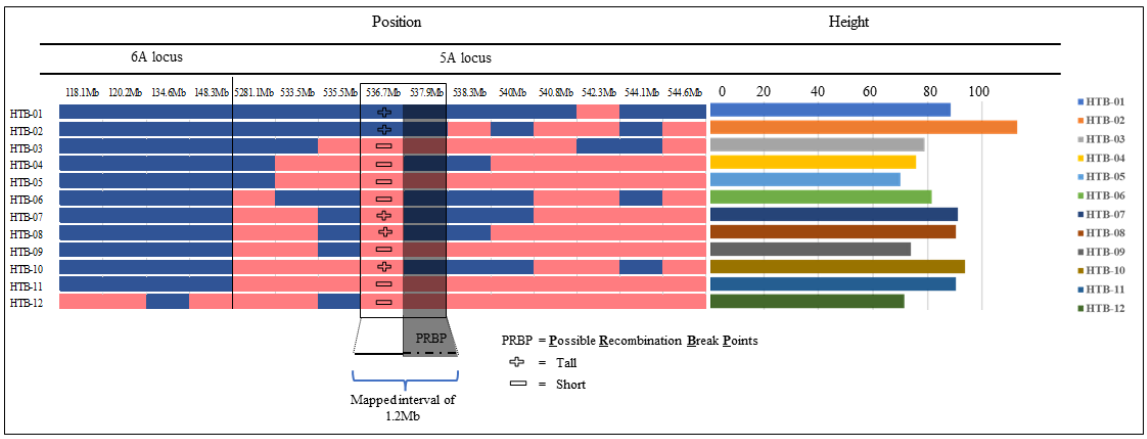


Figure 5.18 The 5A locus fine-mapped interval

Coral and blue correspond to Paragon and Pamyati Azieva respectively. PH of each HTB is given on the right of the map as horizontal bar plots (cm).

There is only one haplotype group (HTB-11) which is behaved like "tall", although it missed Pamyati Azieva allele. Because of the large height variation within this group, we investigated the lines it contains (Figure 5.14). The group had short and tall lines. Almost of the tall varieties were

released before Green Revolution. In contrast, the short genotypes, like Pernel, Piccadilly and Parade, were developed after the Green Revolution (Figure 5.19).

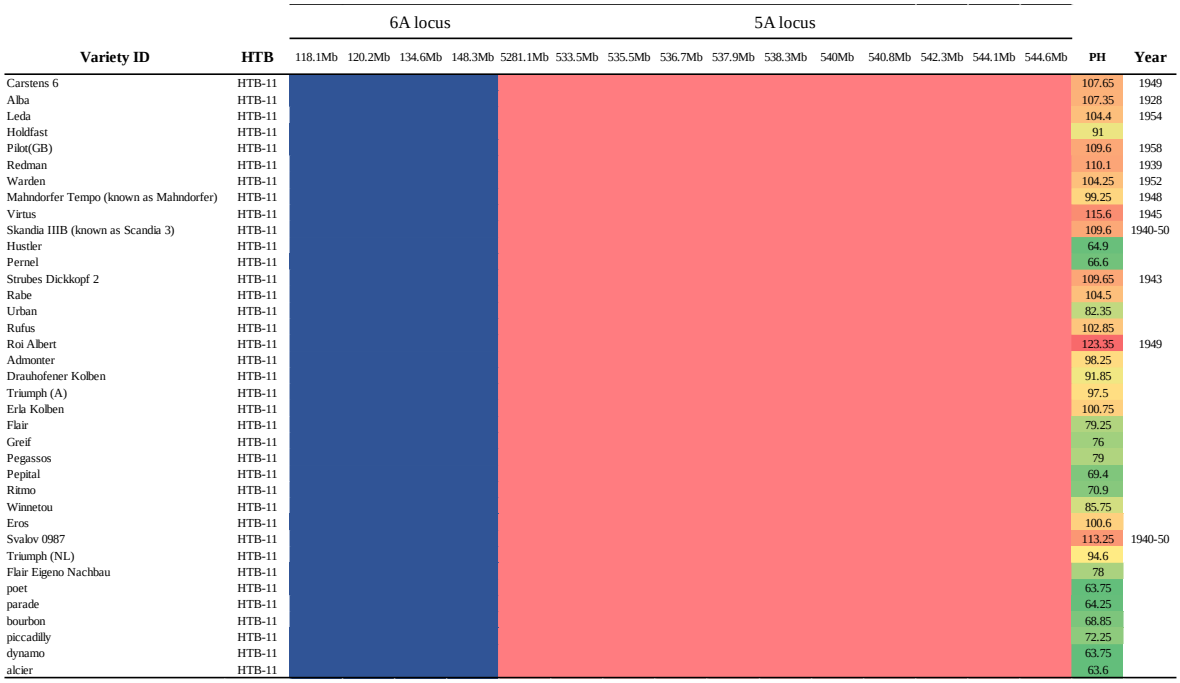


Figure 5.19 HTB-11's tall and short lines

Coral and blue correspond to Paragon and Pamyati Azieva respectively. Absolute values of the plant height (in cm) of each genotype within the HTB-11 are given on the right of the pattern as a heatmap. Plant heights of the tall pre-green revolution developed lines are shown in red. Plant heights of the short post-green revolution developed lines are shown in green. The year of release is given for pre-green revolution developed lines only.

The further investigation of the mapped 1.2Mb region between physical locations - 536.7 and 537.9Mb - showed that the genomic region contains 23 genes in total based on the IWGSC Refseqv1.0 assembly (Figure 5.20). Two of these genes were related to plant height, but none of them had missense mutations within the coding regions based on the exome capture data of Pamyati Azieva compared to Paragon or vice versa. Therefore, all 23 genes were investigated regardless of what trait they are associated with. Looking at each gene individually, we found four genes, TRAESCS5A02G327300, TRAESCS5A02G328500, TRAESCS5A02G328600 and TRAESCS5A02G328700, to which we paid much attention for because the first and last genes possessed 3bp and 6bp insertions respectively, in Pamyat Azieva, and thus these regions were deleted in CS and Paragon. However, the exome capture data of parents for TRAESCS5A02G327300 gene did not match well the IWGSC Refseqv1.0 annotation. The whole genome exome sequencing data of Pamayti Azieva and Paragon had two exons compared to one single exon in the current annotation version of CS (data is not shown). This was caused due to the fact that the exome capture baits (biotinylated oligonucleotide probes) we used were based on previous TGAC (not IWGSC) gene sets (Clavijo et al., 2017). Nevertheless, the gene which has 6bp (ATGAGG) insertion in Pamyati Azieva well matched with CS gene annotation.

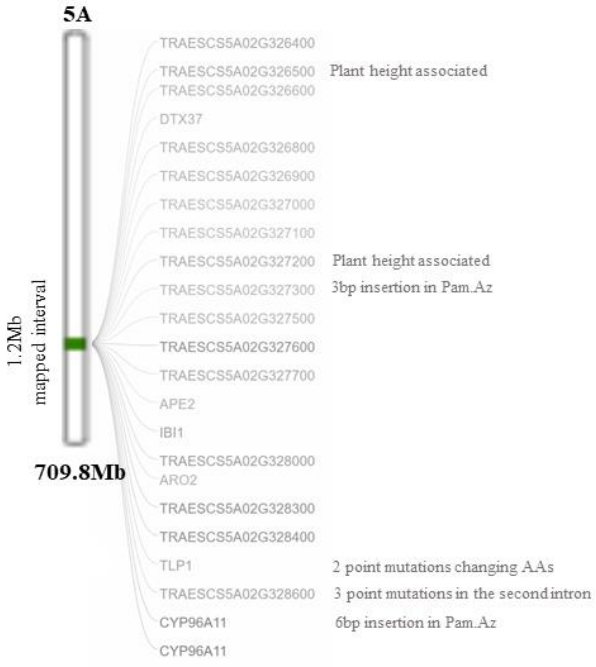


Figure 5.20 The 5A QTL 1.2Mb mapped interval gene content

The list of candidate genes within the 1.2Mb mapped interval. Two genes were associated with plant height, but no missense mutations were identified between Pamyati Azieva and Paragon in those genes. The nature of mutations are given on the right-hand side of the most promising gene candidates at this stage. Pam.Az = Pamyati Azieva, AAs = Amino acids, bp = base pair and Mb = Mega base.

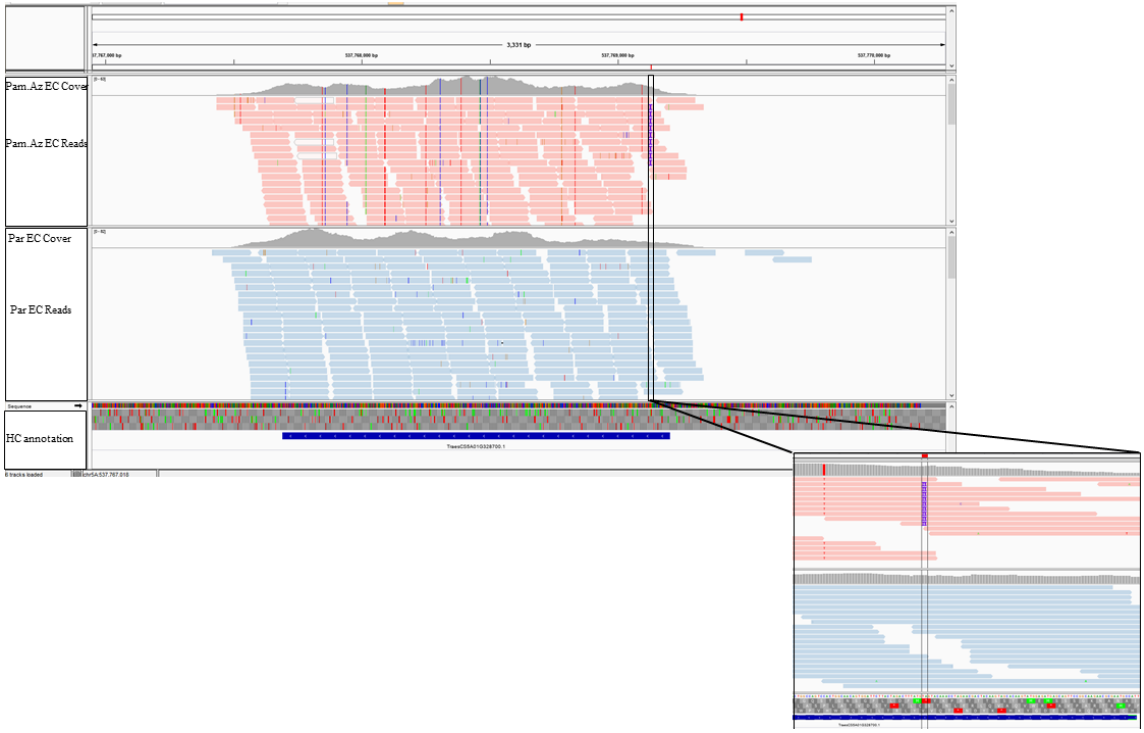


Figure 5.21 The 6bp insertion region in Pamyati Azieva

The visualisation of EC data of Pamyati Azieva and Paragon for the gene TRAESCS5A02G328700 in IGV tool (Integrative Genomics Viewer). The EC coverage and reads for Pam.Az (red) and Paragon (blue) are given on the top and bottom respectively. The region containing the 6bp insertion in Pam.Az is zoomed in (small figure below). Pam.Az = Pamyati Azieva, Par = Paragon, EC = Exome capture, Cover = Coverage and HC = High confidence.

The third candidate gene - TRAESCS5A02G328600 – has three SNPs, but all are located in the second intron (Figure 5.22).

Figure 5.22 The 3-point mutations in the third intron of the TRAESCS5A02G328600 gene

The visualisation of EC data of Pamyati Azieva and Paragon for the gene TRAESCS5A02G328600 in IGV tool (Integrative Genomics Viewer). The EC coverage and reads for Pam.Az (red) and Paragon (blue) are given on the top and bottom respectively. The region containing the 3 SNPs between Pam.Az and Par is labelled as red and shown below the physical coordinates. Pam.Az = Pamyati Azieva, Par = Paragon, EC = Exome capture, Cover = Coverage and HC = High confidence.

The last gene - TRAESCS5A02G328500; possessing two missense point mutations both altering the codon in which they are located, was the most interesting (Figure 5.23). The first "C" to "T" single base pair substitution altered the genetic code of the gene by changing the alanine to valine. The next "G" to "A" point mutation which also was of interest altered the genetic makeup of the gene through changing the arginine to histidine (Figure 5.24). Importantly, the "C" to "T" single base pair substitution was also highlighted using the whole exome sequencing of 890 diverse wheat landraces and cultivars (He et al., 2019). The C/T highest minor allele frequency observed in this population accounted for 0.32 showing that the presence of the Pamyati Azieva allele was less than that of Paragon. Interestingly, these findings were in direct contradiction to what we have found in GEDIFLUX and CAWBIN. The genotyping results from both Western European and Central Asian germplasm panels using the newly designed marker for the C/T substitution at the SNP site surprisingly revealed that among 318 GEDIFLUX lines, 294 are carrying Pamyati Azieva allele compared to only 24 Paragon alleles (Figure 5.25). In CAWBIN, the proportion was similar for GEDIFLUX and thus far higher; 288 to 18, in favour of Pamyati Azieva allele (Figure 5.26).

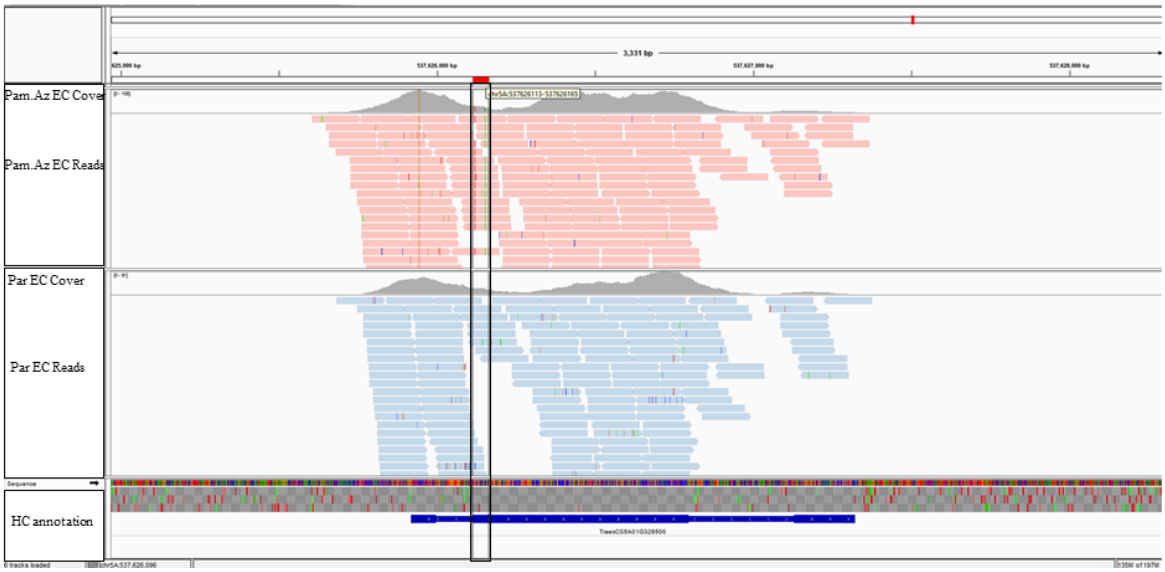


Figure 5.23 The 2 missense point mutations of the TRAESCS5A02G328500

The visualisation of EC data of Pamyati Azieva and Paragon for the gene TRAESCS5A02G328500 in IGV tool (Integrative Genomics Viewer). The EC coverage and reads for Pam.Az (red) and Paragon (blue) are given on the top and bottom respectively. The region containing the 2 SNPs between Pam.Az and Par is labelled as red and shown below the physical coordinates. Pam.Az = Pamyati Azieva, Par = Paragon, EC = Exome capture, Cover = Coverage and HC = High confidence.

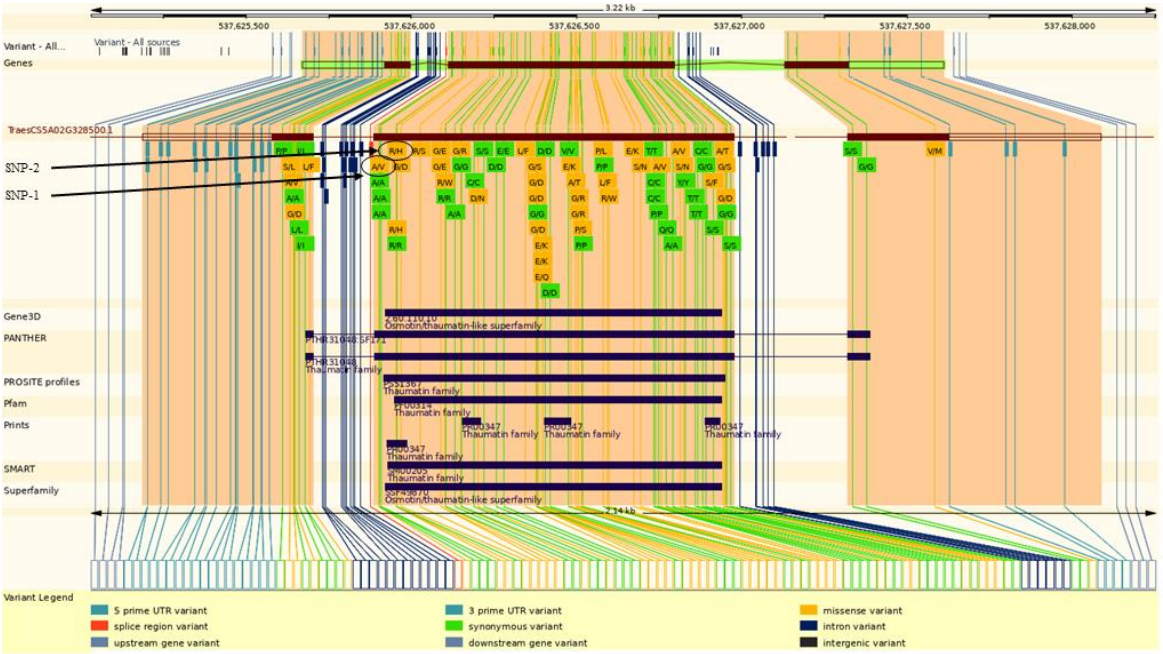


Figure 5.24 The variant image of 2 missense point mutations in the TraesCS5A02G328500 gene

In the figure, the two, SNP-1 and SNP2, missense point mutations are shown. SNP-1 and SNP-2 are the “C” to “T” and “G” to “A” single base pair substitutions changing the alanine to valine and arginine to histidine respectively.

The designed KASP marker for the G/A substitution failed as did the markers designed for 3 SNPs in the second intron, 3bp and 6bp insertions. Therefore, next steps on the agenda are either designing the new informative markers or optimising the PCR conditions. Nonetheless, the allelic frequency of the "Pamyati Azieva" and "Paragon" groups for the C/T substitution indicated that Pamyat Azieva allele indeed increases the plant height significantly in the UK and southern regions of Kazakhstan (p=0.001) (Figure 5.25 and Figure 5.26). However, from the field experiments on

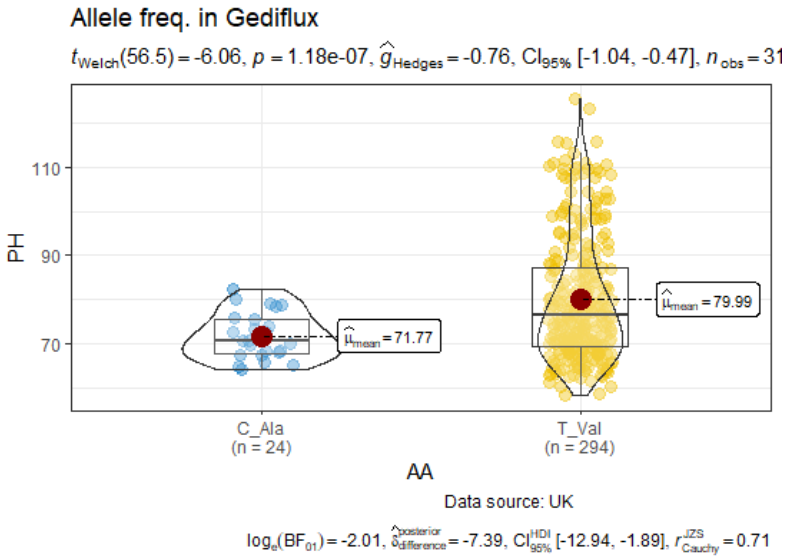


Figure 5.25 Allelic frequency of the C/T substitution in GEDIFLUX

PH = Plant height (cm), C_Ala and T_Val are Paragon allele "C" encoding "Alanine" and Pamyati Azieva allele "T" encoding "Valine". AA = Amino acid.

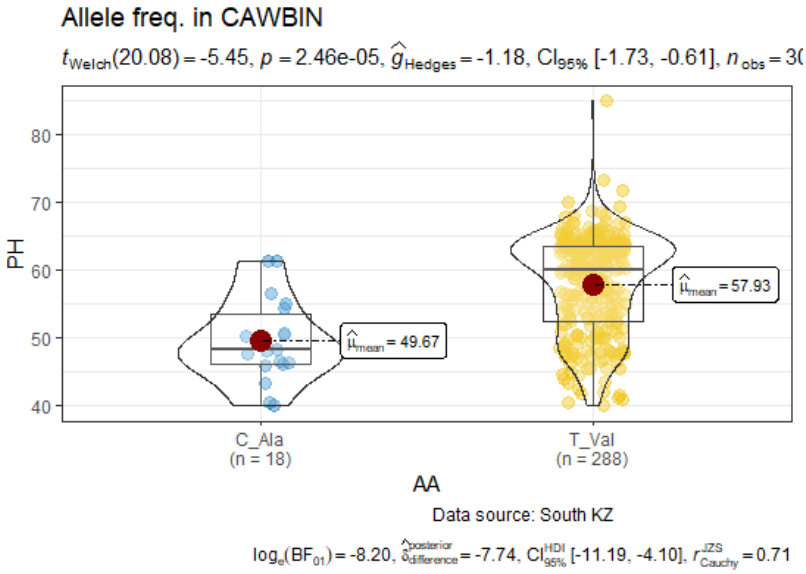


Figure 5.26 Allelic frequency of the C/T substitution in CAWBIN

PH = Plant height (cm), C_Ala and T_Val are Paragon allele "C" encoding "Alanine" and Pamyati Azieva allele "T" encoding "Valine". AA = Amino acid.

NILs, we know that the 5A QTL height effect is silenced in Kazakhstan, except Alm 2020 when the QTL was expressed as a height gene increasing the PH of NIL5A(+) significantly compared to NIL5A(-) (Figure 3.10). It is also worth to point that Pamyati Azieva allele at the 5A locus increased the plant height in RILs during the initial QTL identification in 2016 in Alm. These findings show that 5A QTL increases the plant height periodically, but in southern regions of Kazakhstan only.

Interestingly, the Paragon (C) and Pamyati Azieva (T) alleles at that SNP site were identified as "ancestral" and "derived" respectively based on the use of multiple outgroup species such as wild emmer, domesticated emmer, wheat cultivars and landraces (He et al., 2019). Presumably, these findings show that the Pamyati Azieva allele is a later derived version and might contribute to wheat improvement and local adaption in the North Kazakhstan. However, wheat breeders in Kazakhstan should be selective and cautious when breeding for the "tall" allele at the 5A locus in the South Kazakhstan due to its possible significant negative impact on grain yield and its components (Figure 3.14).

5.4 Discussion

5.4.1 The establishment of the important plant resource

Plant resource development and genetic characterisation are essential prerequisites to gain some insights into genetic structure of wheat (Wingen et al., 2017). Thus, with the joint initiation of the UK and Central Asian wheat communities the Central Asian wheat panel (CAWBIN) was established. Significant effort has been put from the Central Asian wheat community to collect and send the seed material of the panel to the UK. This valuable resource was genetically purified within the scope of this PhD research. Importantly, the panel was genetically fingerprinted based on the use of an array-based SNP genotyping platform (Burridge et al., 2018). The raw genetic data was analysed and converted into usable format to investigate the interesting genetic patterns present in CA wheat. Our initial results showed that the vast majority of wheat lines from Kazakhstan are genetically close to Russian varieties while some form clusters with other cultivars mainly from CA and AFG. These lines were also grouped closely with the historic Watkins collection reflecting the persistence of landrace farming in these countries. However, this is the only small group which outgrouped from the main cluster KZ1. This is the indication of local varieties being extensively used in plant breeding schemes in the same local area which likely caused the genetic bottleneck minimising the genetic diversification. On the other hand, the fact that Kazakh and Russian wheats are as distant from the Watkins landraces as the UK wheats are might also indicate intensive wheat selection programs in the region, although they are mainly based on traditional plant breeding methodologies compared to the UK wheat breeding which takes advantage of using the most advanced genomic techniques in plant selection.

However, in either case, the established new resource should allow us to maximise the genetic diversity in the parental selection and to fully utilise the diversity which has not been captured so far using only conventional breeding technologies.

5.4.2 Testing the SNP markers of mapped intervals for 5A and 6A loci on wider populations

Testing SNP markers flanking the mapped interval for loci showed that more than 90% of the CAWBIN and GEDIFLUX population members are fixed for Pamyati Azieva allele at the 6A locus. This suggests that the haplotype is a signature for a segment of chromosome that is identical by descent. The limited number of haplotypes across most of the 6A chromosome seems to be due to the lack of recombination in 6A chromosome of wheat, resulting in some varieties to share almost 94% identical patterns (Brinton et al., 2020). However, when it comes to 5A locus, CAWBIN and GEDIFLUX were segregating at the locus, but GEDIFLUX presented a large genetic diversity at the 5A locus compared to CAWBIN where almost every single marker showed additive effect, except 536.7Mb locus the relative plant height of which is decreased for the Paragon allele (Figure 5.6 and Figure 5.8). The highest additive effect belonged to the marker (537.9Mb) located to next to this marker (536.7Mb). This seem to add support for the reduced 1.2Mb mapped interval which is the result of the haplotype-based approach I took. The haplotype blocks were also provided to predict two modes of action of the 5A locus. The first hypothesis states that Kazakh wheat breeders might have been breeding for the locus without realising its possible height effect as it is always silenced in Kazakhstan. The second hypothesis suggests that 6A locus has a positive epistatic effect with the 5A locus without which the expression of the 5A locus is restricted.

6. Final discussion

6.1 Introduction

The base aim of this thesis was to use a genetic strategy specifically aimed at finding the type of genetic variation that would be fixed in the elite breeding pools of Kazakh wheat breeders. This is because the environment of the Southern Steppe Zone requires very specific adaption to a short season for rain fed spring wheat. By crossing with a UK wheat which is adapted to longer Western European growing seasons the intention was to cause the Kazakh alleles for specific adaptation to become detectable in the mapping population. An analogous situation would be to cross a UK wheat with variety bred by CIMMYT. The subsequent cross would segregate strongly for photoperiod sensitivity controlled by *Ppd-1*, but a CIMMYT x CIMMYT or UK x UK cross would generally not be useful for identifying this genetic variation which is actually essential for adaptation to short day lengths. This approach was fruitful in this study because strong segregating factors were indeed identified. There were two plant height QTL located on 5A and 6A wheat chromosomes with the height increasing alleles coming from Pamyati Azieva in both cases. The

question then was to ask whether these large genetic effects were actually beneficial, providing some adaptive advantage to wheat growing in Kazakhstan. To address this question Near Isogenic Lines were developed in the genetic background of the UK parent Paragon. The hypothesis being that even though Paragon was not bred for the region the addition of a allele with adaptive benefit would improve the performance of Paragon in some way in Kazakhstan. This required the development and multiplication of NILs in the UK and then ambitious international trials in Kazakhstan.

6.2 The identification of two plant height increasing QTL

The height increasing alleles of the two plant height QTL were donated by Pamyati Azieva. The variety is well-adapted to Kazakh environment and included into the state register. When each QTL was identified, with the large additive effects of 10cm each, it was quite surprising due to the fact that most parts of the world have seen selection towards shorter wheats. This is also evidenced by Canada and Australia with environmental conditions quite similar to that of northern Kazakhstan (Rebetzke et al., 2012; Chen et al., 2016). In both countries a considerable yield increase was experienced between 1960-2019 (FAO). So, the widespread belief amongst breeders for these environments that taller wheats perform better in rain-fed environments such as Kazakhstan should be tested and the specify genetic components controlling these height differences are important. It is a fact that farmers pay for this extra height in terms of increased vulnerability to lodging by wind and rain (Peng et al., 1999) or when high levels of fertilisers are applied (Griffiths et al., 2012), so it is very important that the pros and cons of breeders "gut feeling" are properly quantified. These lodging pressures are faced in Kazakhstan. For example, fertilizers and pesticides are used excessively which led to land degradation (Environment Division of Asian Development Bank, 1997). Strong winds in the North Kazakhstan, where the main wheat growing areas are located, is the major cause of lodging when higher rainfall is experienced. The lodging in turn complicates the wheat grain collection during harvest seasons and reduces grain quality which is the major selling point of Kazakh wheat. Considering these aspects, it seems that the widespread belief of wheat varieties with shortened stem performing badly relative to tall counterparts is something that has not been well tested and proven. Therefore, it was interesting to us to investigate the following points although initial hypotheses were that these genes should be associated with increased adaptation and yield:

- ✓ What these genes are doing in the environment where they are originated from?
- ✓ Is there any valuable association/s in which these genes are worth to keep in CA germplasm panel?
- ✓ What is the individual and combined effect of each locus? Are they independent from each other in their mode of action?
- ✓ Do they increase early vigour and establishment to any extent?
- ✓ Is it a gene or cluster of genes affecting the trait?

- ✓ What is the allele frequency of these loci in the wider Central Asian breeding pool?

6.3 Testing the Individual and combined effects in an isogenic background

Near Isogenic Lines are valuable genetic resources to address these questions. The results obtained from field experiments conducted on two sites in Kazakhstan showed possible ineffectiveness of the 6A locus in increasing any beneficial trait that was measured. However, the height increasing effect was extremely consistent. This leads to the conclusion that it would be an extremely valuable and interesting exercise in experimental breeding for the region to develop varieties in which the 6A height reducing alleles were selected instead of the Pamyati Azieva allele.

In the case of the 5A allele there is some evidence from the experiments described here that the Pamyati Azieva allele increases yield or at least yield components in the Paragon background in the experiments carried out in Northern trials locations. This QTL was also interesting in that it shows a strong interaction with the environment. The height effects were validated in UK trials but generally not in Kazakhstan. Moreover, any beneficial yield effect was restricted to the North while in the South there was actually a significant yield reduction observed. This suggests that the underlying gene/s controls some element of environmental sensitivity. Temperature response would be one obvious target for future dissection of this physiological response. When considering the relatively high p values of the putative yield increase in Pet it should be held in mind that the growing environment of the North is stressful. It is well known that stress reduces the heritability of almost all traits, especially complex polygenic traits like yield, so the statistical power of the experiments was reduced. Efforts were put in place to counter this, with the first trials in the North based on $10m^2$ plots replicated twice and the next year this was increased to four replicates but this year was particularly stressful as evidenced by the overall reduction in height and yield. There was never a significant reduction in yield in the North associated with the height increasing alleles. As a result, the case for individually selecting for or against 5A alleles is much less strong than for 6A. A further possibility is that Pamyati Azieva actually needs both alleles to achieve a beneficial effect. To move towards testing this the double NILs were developed. In the UK they expressed the phenotype of both genes with a spectacular additive height increasing effect. Time constraints meant that they were only grown in Southern Kazakhstan in the last year of the project and with difficulties of weather and the covid pandemic the NILs did not grow. In fact the seeds sent from the UK often showed poor viability when first received in Kazakhstan although they had high germination frequencies when dispatched. The reason for this was never clear, possibly the customs process involved the use of a particularly strong or prolonged X-ray machine? However, some height data for the double NILs was obtained (data not shown) and the doubles were only as tall as the 6A+ single NILs suggesting that, again, the 5A height effect was lost in Kazakhstan even in combination with 6A+.

The development and genotyping of the CAWBIN panel allowed the project to address much broader questions than the analysis of the biparental population and NILs had. It clearly showed the discrete clustering of Central Asian wheat compared to Western European (GEDIFLUX) and global landraces (Watkins). Even within the CAWBIN accessions the genomic signatures picked up the breeding streams of the region with USSR overlapping and non overlapping groups. This does not reflect pre/post Soviet influence but parallel streams which merit further study .

Coming back to the QTL identified here, the fine mapping work gave good molecular markers to tag the haploptypes carrying these genes. It showed that 5A and 6A Pamyati Azioeva alleles were present at very high frequency in Central Asian germplasm. This further validates the approach of crossing with a UK parent as it seems most likely that a Kazakhstan x Kazakhstan cross would not segregate for these alleles. The marker and candidate genes developed here provide very precise information for local breeders on how they can sample and select the haplotypic diversity of these two loci and test the proposal that 6A might be dispensable at least in terms of yield potential.

Because of the outputs of this research the approach carried out for these two genes could now be implemented systematically in Central Asian Breeding. For example the Axiom 35K genotyping allows the selection of a diversity maximised set of parents for the construction of a next generation genetic mapping resource such as a MAGIC or NAM population. This would provide a formidable resource to produce a catalogue of the genes that local breeders are using now. Wider crosses such as that used here could test the utility of genetic diversity associated with haplotypes that are not represented in CAWBIN. Here we showed the comparison with GEDIFLUX and Watkins but the same could be done for germplasm from CIMMYT, Eastern Europe, North America and China. In this way breeding for diversity in Kazakhstan can be much more precise than what would be possible if selections were simply made from pedigrees formed from these wide crosses. At a much simpler level genome wide association mapping of CAWBIN would provide a route for the identification of disease resistance and other important major gene traits.

Wheat genomics has entered a new phase with projects like the 10+ pan genome project (Walkowiak et al., 2020). The varieties chosen for this do not include any representation from Central Asian germplasm, but the varieties adapted to this part of the world are extremely important for global food security. The work of this PhD, with collaboration and support from many in Central Asia and the UK, has elevated the genetic resources of Central Asia to a higher level. Perhaps this is exactly what is needed to gain the attention of the international community so that investments such as this can be argued for to help the breeders and farmers of Kazakhstan accelerate the genetic gains for their crops.

7. References

Abugalieva, A. and Peña-Bautista, R. J. (2010). Grain quality of spring and winter wheat of Kazakhstan. *Asian Australas. J. Plant Sci. Biotechnol.* 4, 87–90.

Abugaliyeva, A. and Morgounov, A. (2016). The History of Wheat Breeding in Southern Kazakhstan. *See Discuss. stats, author profiles this Publ. https//www.researchgate.net/publication/303627899*.

Agarwal, P., Baranwal, V. K. and Khurana, P. (2019). Genome-wide Analysis of bZIP Transcription Factors in wheat and Functional Characterization of a TabZIP under Abiotic Stress. *Sci. Rep.* 9, 1–18. doi:10.1038/s41598-019-40659-7.

Alghabari, F., Lukac, M., Jones, H. E. and Gooding, M. J. (2014). Effect of Rht Alleles on the Tolerance of Wheat Grain Set to High Temperature and Drought Stress During Booting and Anthesis. *J. Agron. Crop Sci.* 200, 36–45. doi:10.1111/jac.12038.

Alqudah, A. M., Sallam, A., Stephen Baenziger, P. and Börner, A. (2020). GWAS: Fast-forwarding gene identification and characterization in temperate Cereals: lessons from Barley – A review. *J. Adv. Res.* 22, 119–135. doi:10.1016/j.jare.2019.10.013.

Amirthadevarathinam, A. (1983). Genetic variability, correlation and path analysis of yield components in upland rice. *Madras Agric. J.* 70, 781–785.

Amram, A., Fadida-Myers, A., Golan, G., Nashef, K., Ben-David, R. and Peleg, Z. (2015). Effect of GA-sensitivity on wheat early vigor and yield components under deep sowing. *Front. Plant Sci.* 6, 1–10. doi:10.3389/fpls.2015.00487.

Andreeva, Y. A. (2015). Salt resistance of common spring wheat varieties in the steppe zone of Northern Kazakhstan. *PhD thesis, Ualikhanov Univ. Kokshetau*.

Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., Pozniak, C. J., Choulet, F., Distelfeld, A., Poland, J., Ronen, G., Barad, O., Baruch, K., Keeble-Gagnère, G., Mascher, M., Ben-Zvi, G., Josselin, A. A., Himmelbach, A., Balfourier, F., Gutierrez-Gonzalez, J., Wang, L. et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science (80-.).* 361. doi:10.1126/science.aar7191.

Appleford, N. E. J., Lenton, J. R., Appleford, N. E. J. and Lenton, J. R. (1991). Linked references are available on JSTOR for this article : Planta Gibberellins and leaf expansion in near-isogenic wheat lines containing Rhtl and Rht3 dwarfing alleles. 183, 229–236.

Aradottir, G. I., Martin, J. L., Clark, S. J., Pickett, J. A. and Smart, L. E. (2017). Searching for wheat resistance to aphids and wheat bulb fly in the historical Watkins and Gediflux wheat collections. *Ann. Appl. Biol.* 170, 179–188. doi:10.1111/aab.12326.

Arora, S., Steuernagel, B., Gaurav, K., Chandramohan, S., Long, Y., Matny, O., Johnson, R., Enk, J., Periyannan, S., Singh, N., Asyraf Md Hatta, M., Athiyannan, N., Cheema, J., Yu, G., Kangara, N., Ghosh, S., Szabo, L. J., Poland, J., Bariana, H., Jones, J. D. G., Bentley, A. R., Ayliffe, M., Olson, E., Xu, S. S., Steffenson, B. J., Lagudah, E. and Wulff,

B. B. H. (2019). Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nat. Biotechnol.* 37, 139–143. doi:10.1038/s41587-018-0007-9.

Astafev, V. L. (2015). Comparison of wheat planting methods in different conditions of Northern Kazakhstan. *AIC Russ.* , 11–15.

Astafev, V. L. (2017). Accumulation and preservation of moisture in water-saving technologies and techniques for their implementation. in *Scientific support for the implementation of state programs of agro-industrial complex and rural areas (in Russian)*, 38–45.

Babu, P., Baranwal, D. K., Harikrishna, Pal, D., Bharti, H., Joshi, P., Thiyagarajan, B., Gaikwad, K. B., Bhardwaj, S. C., Singh, G. P. and Singh, A. (2020). Application of Genomics Tools in Wheat Breeding to Attain Durable Rust Resistance. *Front. Plant Sci.* 11, 1–18. doi:10.3389/fpls.2020.567147.

Bai, C., Liang, Y. and Hawkesford, M. J. (2013). Identification of QTLs associated with seedling root traits and their correlation with plant height in wheat. *J. Exp. Bot.* 64, 1745–1753. doi:10.1093/jxb/ert041.

Barayev, A., Bakayev, N. and Vedeneeva, M. (1978). Spring wheat, Moscow. *Kolos*, 1–429.

Barley, M. (2021). Genetic Variation in Grain Yield and Quality Traits of Spring.

Beales, J., Turner, A., Griffiths, S., Snape, J. W. and Laurie, D. A. (2007). A Pseudo-Response Regulator is misexpressed in the photoperiod insensitive Ppd-D1a mutant of wheat (Triticum aestivum L.). *Theor. Appl. Genet.* 115, 721–733. doi:10.1007/s00122-007-0603-4.

Becker, J., Vos, P., Kuiper, M., Salamini, F. and Heun, M. (1995). Combined mapping of AFLP and RFLP markers in barley. *MGG Mol. Gen. Genet.* 249, 65–73. doi:10.1007/BF00290237.

Bentley, A. R., Turner, A. S., Gosman, N., Leigh, F. J., Maccaferri, M., Dreisigacker, S., Greenland, A. and Laurie, D. A. (2011). Frequency of photoperiod-insensitive Ppd-A1a alleles in tetraploid, hexaploid and synthetic hexaploid wheat germplasm. *Plant Breed.* 130, 10–15. doi:10.1111/j.1439-0523.2010.01802.x.

Bernardo, R. (2009). Should maize doubled haploids be induced among F1 or F 2 plants? *Theor. Appl. Genet.* 119, 255–262. doi:10.1007/s00122-009-1034-1.

Bernardo, R. (2020). Reinventing quantitative genetics for plant breeding: something old, something new, something borrowed, something BLUE. *Heredity (Edinb).* 125, 375–385. doi:10.1038/s41437-020-0312-1.

Bespalova, L. A. (2003). Selection of common winter wheat for adaptability in Krasnodar. *Vestnik. A Reg. Netw. Introd. wheat Var. seed Prod.* No. 3 (6), 114–119.

Bhat, J. A. and Yu, D. (2021). High-throughput NGS-based genotyping and phenotyping: Role in genomics-assisted breeding for soybean improvement. *Legum. Sci.* 3. doi:10.1002/leg3.81.

Bogale, A., Tesfaye, K. and Geleto, T. (2011). Journal of Biodiversity and Environmental Sciences (JBES) ISSN : 2220-6663 (Print) 2222-3045 (Online) Morphological and physiological attributes associated to drought tolerance of Ethiopian durum wheat genotypes under water deficit condition. 1, 22–36.

Bonke, M., Thitamadee, S., Mähönen, A. P., Hauser, M. T. and Helariutta, Y. (2003). APL regulates vascular tissue identity in Arabidopsis. *Nature* 426, 181–186. doi:10.1038/nature02100.

Borlaug, N. E. (1983). Contributions of conventional plant breeding to food production. *Science (80-.).*, 689–693. doi:10.4324/9780429050329-14.

Börner, A. and Worland, A. J. (2002). Does the Chinese dwarf wheat variety "XN0004" carry Rht21? *Cereal Res. Commun.* 30, 25–29. doi:10.1007/bf03543385.

Börner, A., Worland, A. J., Plaschke, J., Schumann, E. and Law, C. N. (1993). Pleiotropic effects of genes for reduced height (Rht) and day-length insensitivity (Ppd) on yield and its components for wheat grown in middle Europe. *Plant Breed.* 111, 204–216.

Botwright, T., Rebetzke, G., Condon, T. and Richards, R. (2001). The effect of rht genotype and temperature on coleoptile growth and dry matter partitioning in young wheat seedlings. *Aust. J. Plant Physiol.* 28, 417–423. doi:10.1071/pp01010.

Braun, H. J., Rajaram, S. and Van Ginkel, M. (1996). CIMMYT's approach to breeding for wide adaptation. *Euphytica* 92, 175–183. doi:10.1007/BF00022843.

Brinton, J., Ramirez-Gonzalez, R. H., Simmonds, J., Wingen, L., Orford, S., Griffiths, S., Haberer, G., Spannagl, M., Walkowiak, S., Pozniak, C. and Uauy, C. (2020). A haplotype-led approach to increase the precision of wheat breeding. *Commun. Biol.* 3, 1–11. doi:10.1038/s42003-020-01413-2.

Broman, K. W. (2005). The genomes of recombinant inbred lines. *Genetics* 169, 1133–1146. doi:10.1534/genetics.104.035212.

Broman, K. W. (2010). Genetic map construction with R/qtl. *Univ. Wisconsin-Madison Dep. Biostat. Med. Informatics* Technical.

Broman, K. W., Wu, H., Sen, Ś. and Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19, 889–890. doi:10.1093/bioinformatics/btg112.

Bullrich, L., Appendino, M., Tranquilli, G., Lewis, S. and Dubcovsky, J. (2002). Mapping of a thermo-sensitive earliness per se gene on Triticum monococcum chromosome 1A m. *Theor. Appl. Genet.* 105, 585–593. Available at: https://cloudfront.escholarship.org/dist/prd/content/qt5p18m5rp/qt5p18m5rp.pdf?t=onjgtb.

Burridge, A. J., Wilkinson, P. A., Winfield, M. O., Barker, G. L. A., Allen, A. M., Coghill, J. A., Waterfall, C. and Edwards, K. J. (2018). Conversion of array-based single nucleotide polymorphic markers for use in targeted genotyping by sequencing in hexaploid wheat (Triticum aestivum). *Plant Biotechnol. J.* 16, 867–876. doi:10.1111/pbi.12834.

Busari, M. A., Kukal, S. S., Kaur, A., Bhatt, R. and Dulazi, A. A. (2015). Conservation tillage impacts on soil, crop and the environment. *Int. Soil Water Conserv. Res.* 3, 119–129.

Bushuk, W. (1997). "Wheat breeding for end-product use," in *Wheat: Prospects for global improvement* (Springer), 203–211.

Buss, W., Ford, B. A., Foo, E., Schnippenkoetter, W., Borrill, P., Brooks, B., Ashton, A. R.,

Chandler, P. M. and Spielmeyer, W. (2020). Overgrowth mutants determine the causal role of gibberellin GA2oxidaseA13 in Rht12 dwarfism of wheat. *J. Exp. Bot.* 71, 7171–7178. doi:10.1093/jxb/eraa443.

Cao, H. X., Wang, W. Q., Le, H. T. T. and Vu, G. T. H. (2016). The Power of CRISPR-Cas9-Induced Genome Editing to Speed Up Plant Breeding. *Int. J. Genomics*. doi:10.1155/2016/5078796.

Cha, J., Connor, K. O., Alahmad, S., Lee, J., Dinglasan, E., Kim, K., Ko, J., Hickey, L. T., Shin, D. and Dixon, L. E. (2021). A new protocol for speed vernalisation of winter cereals. *bioRxiv*, 1–23.

Chebotar, S. B. (2008). Allelic characteristics of short-stature genes in the genetic pool of varieties of winter soft wheat of Ukraine. *Genet. plant Resour.* 6, 96–103.

Chen, H., Moakhar, N. P., Iqbal, M., Pozniak, C., Hucl, P. and Spaner, D. (2016). Genetic variation for flowering time and height reducing genes and important traits in western Canadian spring wheat. *Euphytica* 208, 377–390. doi:10.1007/s10681-015-1615-9.

Chen, Y., Wang, M. and Ouwerkerk, P. B. F. (2012). Molecular and environmental factors determining grain quality in rice Grain Quality Traits in Rice. *Food Energy Secur.*, 111–132. doi:10.1002/fes3.11.

Chen, Z., Tang, D., Ni, J., Li, P., Wang, L., Zhou, J., Li, C., Lan, H., Li, L. and Liu, J. (2021). Development of genic KASP SNP markers from RNA-Seq data for map-based cloning and marker-assisted selection in maize. *BMC Plant Biol.* 21, 1–11. doi:10.1186/s12870-021-02932-8.

Chew, W., Hrmova, M. and Lopato, S. (2013). Role of Homeodomain Leucine Zipper (HD-Zip) IV Transcription Factors in Plant Development and Plant Protection from Deleterious Environmental Factors. 8122–8147. doi:10.3390/ijms14048122.

Cho, Y.-C., Suh, J.-P., Yoon, M.-R., Baek, M.-K., Won, Y.-J., Lee, J.-H., Park, H.-S., Baek, S.-H. and Lee, J.-H. (2013). QTL Mapping for Paste Viscosity Characteristics Related to Eating Quality and QTL-NIL Development in Japonica Rice (Oryza sativa L.) . *Plant Breed. Biotechnol.* 1, 333–346. doi:10.9787/pbb.2013.1.4.333.

Clavijo, B. J., Venturini, L., Schudoma, C., Accinelli, G. G., Kaithakottil, G., Wright, J., Borrill, P., Kettleborough, G., Heavens, D., Chapman, H., Lipscombe, J., Barker, T., Lu, F. H., McKenzie, N., Raats, D., Ramirez-Gonzalez, R. H., Coince, A., Peel, N., Percival-Alwyn, L., Duncan, O., Trösch, J., Yu, G., Bolser, D. M., Namaati, G., Kerhornou, A., Spannagl, M., Gundlach, H., Haberer, G., Davey, R. P., Fosker, C., Di Palma, F., Phillips, A. L., Millar, A. H., Kersey, P. J., Uauy, C., Krasileva, K. V., Swarbreck, D., Bevan, M. W. and Clark, M. D. (2017). An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations. *Genome Res.* 27, 885–896. doi:10.1101/gr.217117.116.

Collard, B. C. Y. and Mackill, D. J. (2008). Marker-assisted selection: An approach for precision plant breeding in the twenty-first century. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 557–572. doi:10.1098/rstb.2007.2170.

Concibido, V. C., Diers, B. W. and Arelli, P. R. (2004). A decade of QTL mapping for cyst nematode resistance in soybean. *Crop Sci.* 44, 1121–1131. doi:10.2135/cropsci2004.1121.

Couto, D. and Zipfel, C. (2016). Regulation of pattern recognition receptor signalling in plants. *Nat. Rev. Immunol.* 16, 537–552. doi:10.1038/nri.2016.77.

Crain, J., Mondal, S., Rutkoski, J., Singh, R. P. and Poland, J. (2018). Combining High-Throughput Phenotyping and Genomic Information to Increase Prediction and Selection Accuracy in Wheat Breeding. *Plant Genome* 11. doi:10.3835/plantgenome2017.05.0043.

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de los Campos, G., Burgueño, J., González-Camacho, J. M., Pérez-Elizalde, S., Beyene, Y., Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J. and Varshney, R. K. (2017). Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. *Trends Plant Sci.* 22, 961–975. doi:10.1016/j.tplants.2017.08.011.

Cui, F., Li, J., Ding, A., Zhao, C., Wang, L., Wang, X., Li, S., Bao, Y., Li, X., Feng, D., Kong, L. and Wang, H. (2011). Conditional QTL mapping for plant height with respect to the length of the spike and internode in two mapping populations of wheat. *Theor. Appl. Genet.* 122, 1517–1536. doi:10.1007/s00122-011-1551-6.

Czedik-Eysenberg, A., Seitner, S., Güldener, U., Koemeda, S., Jez, J., Colombini, M. and Djamei, A. (2018). The 'PhenoBox', a flexible, automated, open-source plant phenotyping solution. *New Phytol.* 2, 808–823.

Das, P., Lakra, N., Nutan, K. K., Singla-Pareek, S. L. and Pareek, A. (2019). A unique bZIP transcription factor imparting multiple stress tolerance in Rice. *Rice* 12. doi:10.1186/s12284-019-0316-8.

De Beurs, K. M. and Henebry, G. M. (2004). Land surface phenology, climatic variation, and institutional change: Analyzing agricultural land cover change in Kazakhstan. *Remote Sens. Environ.* 89, 497–509.

Díaz, A., Zikhali, M., Turner, A. S., Isaac, P. and Laurie, D. A. (2012). Copy number variation affecting the photoperiod-B1 and vernalization-A1 genes is associated with altered flowering time in wheat (Triticum aestivum). *PLoS One* 7. doi:10.1371/journal.pone.0033234.

Distelfeld, A., Li, C. and Dubcovsky, J. (2009). Regulation of flowering in temperate cereals. *Curr. Opin. Plant Biol.* 12, 178–184. doi:10.1016/j.pbi.2008.12.010.

Dong, C., Zhang, L., Chen, Z., Xia, C., Gu, Y., Wang, J., Li, D., Xie, Z., Zhang, Q., Zhang, X., Gui, L., Liu, X. and Kong, X. (2020). Combining a New Exome Capture Panel With an Effective varBScore Algorithm Accelerates BSA-Based Gene Cloning in Wheat. *Front. Plant Sci.* 11, 1–12. doi:10.3389/fpls.2020.01249.

Dorofeev, V. F., Udachin, R. A. and Semenova, L. V (1987). Wheats of the world.

Agropromizdat Leningrad, USSR.

Du, Y., Chen, L., Wang, Y., Yang, Z., Saeed, I., Daoura, B. G. and Hu, Y. G. (2018). The combination of dwarfing genes Rht4 and Rht8 reduced plant height, improved yield traits of rainfed bread wheat (Triticum aestivum L.). *F. Crop. Res.* 215, 149–155. doi:10.1016/j.fcr.2017.10.015.

Duan, S., Cui, C., Chen, L., Yang, Z. and Gang, Y. (2022). Fine mapping and candidate gene analysis of dwarf gene Rht14 in durum wheat (Triticum durum). *Funct. Integr. Genomics*. doi:10.1007/s10142-021-00825-5.

Duan, S., Zhao, Z., Qiao, Y., Cui, C., Morgunov, A., Condon, A. G., Chen, L. and Hu, Y. G. (2020). GAR dwarf gene Rht14 reduced plant height and affected agronomic traits in durum wheat (Triticum durum). *F. Crop. Res.* 248, 107721. doi:10.1016/j.fcr.2020.107721.

Dunwell, J. M. (2010). Haploids in flowering plants: Origins and exploitation. *Plant Biotechnol. J.* 8, 377–424. doi:10.1111/j.1467-7652.2009.00498.x.

Duveiller, E., Singh, P., Mezzalama, M., Singh, R. and Dababat, A. (2018). *Wheat diseases and pests: a guide for field identification*. Ankara - CIMMYT.

Eduardo, I., Arús, P. and Monforte, A. J. (2005). Development of a genomic library of near isogenic lines (NILs) in melon (Cucumis melo L.) from the exotic accession PI161375. *Theor. Appl. Genet.* 112, 139–148. doi:10.1007/s00122-005-0116-y.

Eken, C., Spanbayev, A., Tulegenova, Z. and Yechshzhanov, T. (2016). First report of Nigrospora oryzae on wheat in Kazakhstan. *Plant Dis.* 100, 861.

Ellis, M. H., Rebetzke, G. J., Chandler, P., Bonnet, D., Spielmeyer, W. and Richards, R. A. (2004). The effect of different height reducing genes on the early growth of wheat. *Funct. Plant Biol.* 31, 583–589. doi:10.1071/FP03207.

Environment Division of Asian Development Bank (1997). *Central Asian environments in transition, Pt. II. An Environmental Profile of the Republic of Kazakhstan*. Asian Development Bank.

Farshadfar, E., Rafiee, F. and Hasheminasab, H. (2013). Evaluation of genetic parameters of agronomic and morpho-physiological indicators of drought tolerance in bread wheat (Triticum aestivum L .) using diallel mating design. 7, 268–275.

Fehér, I., Lehota, J., Lakner, Z., Kende, Z., Bálint, C., Vinogradov, S. and Fieldsend, A. (2017). "Kazakhstan's Wheat Production Potential," in *The Eurasian Wheat Belt and Food Security* (Springer), 177–194.

Feldman, M., Liu, B., Segal, G., Abbo, S., Levy, A. A. and Vega, J. M. (1997). Rapid elimination of low-copy DNA sequences in polyploid wheat: a possible mechanism for differentiation of homoeologous chromosomes. *Genetics* 147, 1381–1387. Available at: http://www.genetics.org/content/genetics/147/3/1381.full.pdf.

Field, A. (2009). *Reading Field's Discovering Statistics Using SPSS*. Third Edit. Sage Publications Ltd.

Fletcher, R. S., Mullen, J. L., Yoder, S., Bauerle, W. L., Reuning, G., Sen, S., Meyer, E., Juenger, T. E. and McKay, J. K. (2013). Development of a next-generation NIL library in Arabidopsis thaliana for dissecting complex traits. *BMC Genomics* 14. doi:10.1186/1471-2164-14-655.

Fleury, D., Jefferies, S., Kuchel, H. and Langridge, P. (2010). Genetic and genomic tools to improve drought tolerance in wheat. *J. Exp. Bot.* 61, 3211–3222. Available at: https://watermark.silverchair.com/erq152.pdf?token=AQECAHi208BE49Ooan9kkhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAAcMwggG_BgkqhkiG9w0BBwagggGwMIIBrAIBADCCAaUGCSqGSIb3DQEHATAeBglghkgBZQMEAS4wEQQM9pV6NeZ3_OrpCFepAgEQgIIBdoMN2owW7DLNGBCrBnt1EErcWQpJcI5DDcQNurefxsi2QTZa.

Flintham, J. E., Börner, A., Worland, A. J. and Gale, M. D. (1997). Optimizing wheat grain yield: Effects of Rht (gibberellin-insensitive) dwarfing genes. *J. Agric. Sci.* 128, 11–25. doi:10.1017/S0021859696003942.

Ford, B. A., Foo, E., Sharwood, R., Karafiatova, M., Vrána, J., MacMillan, C., Nichols, D. S., Steuernagel, B., Uauy, C., Dolezel, J., Chandler, P. M. and Spielmeyer, W. (2018). Rht18 Semidwarfism in Wheat Is Due to Increased GA 2-oxidaseA9 Expression and Reduced GA Content. *Plant Physiol.* 177, 168–180. doi:10.1104/PP.18.00023.

Fu, D., Szűcs, P., Yan, L., Helguera, M., Skinner, J. S., Von Zitzewitz, J., Hayes, P. M. and Dubcovsky, J. (2005). Large deletions within the first intron in VRN-1 are associated with spring growth habit in barley and wheat. *Mol. Genet. Genomics* 273, 54–65. Available at: https://cloudfront.escholarship.org/dist/prd/content/qt56w8t3pt/qt56w8t3pt.pdf?t=onjfw3.

Funakawa, S., Nakamura, L., Akshalov, K. and Kosaki, T. (2004). Water dynamics in soil-plant systems under grain farming in northern Kazakhstan. *Soil Sci. Plant Nutr.* 50, 1219–1227. doi:10.1080/00380768.2004.10408597.

Gagkaeva, T. U., Gavrilova, O. P., Levtinin, M. M. and Novozhilov, K. V (2011). Fusarium of cereals. *Append. to J. "Plant Prot. Quar.*, 112.

Gale, M. D. and Marshall, G. A. (1976). The chromosomal location of Gai 1 and Rht 1, genes for gibberellin insensitivity and semi-dwarfism, in a derivative of Norin 10 wheat. *Heredity (Edinb).* 37, 283.

Gale, M. D., Youssefian, S. and Russell, G. E. (1985). Dwarfing genes in wheat. *Prog. plant Breed.* 1, 1–35.

Genievskaya, Y., Pecchioni, N., Laidò, G., Anuarbek, S., Rsaliyev, A., Chudinov, V., Zatybekov, A., Turuspekov, Y. and Abugalieva, S. (2022). Genome-Wide Association Study of Leaf Rust and Stem Rust Seedling and Adult Resistances in Tetraploid Wheat Accessions Harvested in Kazakhstan. *Plants* 11, 1–20. doi:10.3390/plants11151904.

Geshtovt, U. N. (1986). Protection of crops from pests, diseases and weeds. *Almaty: Kainar*, 65–70.

Gewin, V. (2017). Robotic field surveillance speeds crop breeding. *Front. Ecol. Environ.* 15, 122.

Gómez-Becerra, H. F., Abugalieva, A., Morgounov, A., Abdullayev, K., Bekenova, L., Yessimbekova, M., Sereda, G., Shpigun, S., Tsygankov, V. and Zelenskiy, Y. (2010). Phenotypic correlations, G× E interactions and broad sense heritability analysis of grain and flour quality characteristics in high latitude spring bread wheats from Kazakhstan and Siberia. *Euphytica* 171, 23–38.

Goncharov, N. P. (2003). Genetics of growth habit (spring vs winter) in common wheat: Confirmation of the existence of dominant gene Vrn4. *Theor. Appl. Genet.* 107, 768–772. doi:10.1007/s00122-003-1317-x.

Gossen, E. (2000). *Ecological niches for agro-biodiversity and production stability in Northern Kazakhstan*. , eds. M. Karabayev, A. Satybaldin, J. Benites, T. Friedrich, M. Pala, and T. Payne.

Goujon, T., Ferret, V., Mila, I., Pollet, B., Ruel, K., Burlat, V., Joseleau, J. P., Barrière, Y., Lapierre, C. and Jouanin, L. (2003). Down-regulation of the AtCCR1 gene in Arabidopsis thaliana: Effects on phenotype, lignins and cell wall degradability. *Planta* 217, 218–228. doi:10.1007/s00425-003-0987-6.

Griffiths, S., Simmonds, J., Leverington, M., Wang, Y., Fish, L., Sayers, L., Alibert, L., Orford, S., Wingen, L. and Herry, L. (2009). Meta-QTL analysis of the genetic control of ear emergence in elite European winter wheat germplasm. *Theor. Appl. Genet.* 119, 383–395.

Griffiths, S., Simmonds, J., Leverington, M., Wang, Y., Fish, L., Sayers, L., Alibert, L., Orford, S., Wingen, L. and Snape, J. (2012). Meta-QTL analysis of the genetic control of crop height in elite European winter wheat germplasm. *Mol. Breed.* 29, 159–171. doi:10.1007/s11032-010-9534-x.

Guan, P., Lu, L., Jia, L., Kabir, M. R., Zhang, J., Lan, T., Zhao, Y., Xin, M., Hu, Z., Yao, Y., Ni, Z., Sun, Q. and Peng, H. (2018). Global QTL analysis identifies genomic regions on chromosomes 4A and 4B harboring stable loci for yield-related traits across different environments in wheat (Triticum aestivum L.). *Front. Plant Sci.* 9. doi:10.3389/fpls.2018.00529.

Gulnaz, S., Sajjad, M., Khaliq, I., Khan, A. S. and Khan, S. H. (2011). Relationship among coleoptile length, plant height and tillering capacity for developing improved wheat varieties. *Int. J. Agric. Biol.* 13, 130–133.

Guo, H., Bao, A., Liu, T., Jiapaer, G., Ndayisaba, F., Jiang, L., Kurban, A. and De Maeyer, P. (2018). Spatial and temporal characteristics of droughts in Central Asia during 1966-2015. *Sci. Total Environ.* 624, 1523–1538. doi:10.1016/j.scitotenv.2017.12.120.

Guo, Z., Song, Y., Zhou, R., Ren, Z. and Jia, J. (2010). Discovery, evaluation and distribution of haplotypes of the wheat Ppd-D1 gene. *New Phytol.* 185, 841–851. doi:10.1111/j.1469-8137.2009.03099.x.

Gustafson, P., Raskina, O., Ma, X. and Nevo, E. (2009). Wheat evolution, domestication, and improvement. *Wheat Sci. trade*, 3–30.

Habib, A., Powell, J. J., Stiller, J., Liu, M., Shabala, S., Zhou, M., Gardiner, D. M. and Liu, C. (2018). A multiple near isogenic line (multi-NIL) RNA-seq approach to identify candidate genes underpinning QTL. *Theor. Appl. Genet.* 131, 613–624. Available at: https://link.springer.com/content/pdf/10.1007%2Fs00122-017-3023-0.pdf.

Hafeez, A. N., Arora, S., Ghosh, S., Gilbert, D., Bowden, R. L. and Wulff, B. B. H. (2021). Creation and judicious application of a wheat resistance gene atlas. *Mol. Plant* 14, 1053–1070. doi:10.1016/j.molp.2021.05.014.

Halder, T., Liu, H., Chen, Y., Yan, G. and Siddique, K. H. M. (2021). Identification of candidate genes for root traits using genotype–phenotype association analysis of near-isogenic lines in hexaploid wheat (Triticum aestivum l.). *Int. J. Mol. Sci.* 22. doi:10.3390/ijms22073579.

Haque, M., Martinek, P., Watanabe, N. and Kuboyama, T. (2011). Genetic mapping of gibberellic acid-sensitive genes for semi-dwarfism in durum wheat. *Cereal Res. Commun.* 39, 171–178. doi:10.1556/CRC.39.2011.2.1.

Hassani-Pak, K., Singh, A., Brandizi, M., Hearnshaw, J., Parsons, J. D., Amberkar, S., Phillips, A. L., Doonan, J. H. and Rawlings, C. (2021). KnetMiner: a comprehensive approach for supporting evidence-based gene discovery and complex trait analysis across species. *Plant Biotechnol. J.* 19, 1670–1678. doi:10.1111/pbi.13583.

Hautsalo, J., Novakazi, F., Jalli, M., Göransson, M., Manninen, O. and Isolahti, M. (2021). Pyramiding of scald resistance genes in four spring barley MAGIC populations. 3829–3843. doi:10.1007/s00122-021-03930-y.

He, F., Pasam, R., Shi, F., Kant, S., Keeble-gagnere, G., Kay, P., Forrest, K., Fritz, A., Hucl, P., Wiebe, K., Knox, R., Cuthbert, R., Pozniak, C., Akhunova, A., Morrell, P. L., Davies, J. P., Webb, S. R., Spangenberg, G., Hayes, B., Daetwyler, H., Tibbits, J., Hayden, M. and Akhunov, E. (2019). Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome. *Nat. Genet.* 51. doi:10.1038/s41588-019-0382-2.

Hedden, P. (2003). The genes of the Green Revolution. *Trends Genet.* 19, 5–9. doi:10.1016/S0168-9525(02)00009-4.

Hobo, T., Kowyama, Y. and Hattori, T. (1999). A bZIP factor, TRAB1, interacts with VP1 and mediates abscisic acid- induced transcription. *Proc. Natl. Acad. Sci. U. S. A.* 96, 15348–15353. doi:10.1073/pnas.96.26.15348.

Holloway, J. and Knapp, S. (1993). Holloway, J. L., and S. J. Knapp. "GMendel 3.0 users guide. *Dep. Crop Soil Sci.*

Hongwei, Y. T. and Zhonghua, W. (1993). Chromosomal Arm Location of A Dominant Dwarfing Gene Rht21 in Common Wheat Variety-XN0004 [J]. *J. Northwest Sci-Tech Univ. Agric. For.*

Hoogendoorn, J. (1985). The physiology of variation in the time of ear emergence among wheat-

varieties from different regions of the world. *Euphytica* 34, 559–571. doi:10.1007/bf00022955.

Howe, K. L., Contreras-Moreira, B., De Silva, N., Maslen, G., Akanni, W., Allen, J., Alvarez-Jarreta, J., Barba, M., Bolser, D. M., Cambell, L., Carbajo, M., Chakiachvili, M., Christensen, M., Cummins, C., Cuzick, A., Davis, P., Fexova, S., Gall, A., George, N., Gil, L., Flicek, P. et al. (2020). Ensembl Genomes 2020-enabling non-vertebrate genomic research. *Nucleic Acids Res.* 48, D689–D695. doi:10.1093/nar/gkz890.

Hucl, P. and Baker, J. (1987). A study of ancbstral and modern canadian any , have occurred in yield-. 97, 87–97.

Hussain, M. I. (2002). Growth , Yield and Quality Response of Three Wheat (Triticum aestivum L .) Varieties to Different Levels of N , P and K.

Iizumi, T., Yokozawa, M., Sakurai, G., Travasso, M. I., Romanernkov, V., Oettli, P., Newby, T., Ishigooka, Y. and Furuya, J. (2014). Historical changes in global yields: major cereal and legume crops from 1982 to 2006. *Glob. Ecol. Biogeogr.* 23, 346–357. doi:10.1111/geb.12120.

Imaizumi, T., Schultz, T. F., Harmon, F. G., Ho, L. A. and Kay, S. A. (2005). Plant science: FKF1 F-box protein mediates cyclic degradation of a repressor of CONSTANS in Arabidopsis. *Science (80-.).* 309, 293–297. doi:10.1126/science.1110586.

Iorgansky, A. (2000). *Agro-ecological conservation tillage in Southeastern Kazakhstan*. , eds. M. Karabayev, A. Satybaldin, J. Benites, T. Friedrich, M. Pala, and T. Payne.

Iqtidar, H., Ayyaz, K. M. and Ahmad, K. E. (2006). Bread wheat varieties as influenced by different nitrogen levels. 7, 70–78. doi:10.1631/jzus.2006.B0070.

Islam, M. A., Obour, A. K., Saha, M. C., Nachtman, J. J., Cecil, W. K. and Baumgartner, R. E. (2014). Grain Yield , Forage Yield , and Nutritive Value of Dual-Purpose Small Grains in the Central High Plains of the USA. doi:10.1094/CM-2012-0154-RS.

Itoh, H., Ueguchi-Tanaka, M., Sato, Y., Ashikari, M. and Matsuoka, M. (2002). The gibberellin signaling pathway is regulated by the appearance and disappearance of slender rice1 in nuclei. *Plant Cell* 14, 57–70. doi:10.1105/tpc.010319.

Jakoby, M., Weisshaar, B., Dröge-Laser, W., Vicente-Carbajosa, J., Tiedemann, J., Kroj, T. and Parcy, F. (2002). bZIP transcription factors in Arabidopsis. *Trends Plant Sci.* 7, 106–111. doi:10.1016/S1360-1385(01)02223-3.

Jia, M., Guan, J., Zhai, Z., Geng, S., Zhang, X., Mao, L. and Li, A. (2018). Wheat functional genomics in the era of next generation sequencing: An update. *Crop J.* 6, 7–14. doi:10.1016/j.cj.2017.09.003.

Jiang, J., Yang, D. and Ali, J. (2015). Molecular marker-assisted pyramiding of broad-spectrum disease resistance genes , Pi2 and Xa23 , into GZ63-4S , an elite thermo-sensitive genic male-sterile line in rice. doi:10.1007/s11032-015-0282-9.

Jordan, K. W., Wang, S., He, F., Chao, S., Lun, Y., Paux, E., Sourdille, P., Sherman, J.,

Akhunova, A., Blake, N. K., Pumphrey, M. O., Glover, K., Dubcovsky, J., Talbert, L. and Akhunov, E. D. (2018). The genetic architecture of genome-wide recombination rate variation in allopolyploid wheat revealed by nested association mapping. *Plant J.* 95, 1039–1054. doi:10.1111/tpj.14009.

Juliana, P., Singh, R. P., Singh, P. K., Poland, J. A., Bergstrom, G. C., Huerta-Espino, J., Bhavani, S., Crossa, J. and Sorrells, M. E. (2018). Genome-wide association mapping for resistance to leaf rust, stripe rust and tan spot in wheat reveals potential candidate genes. *Theor. Appl. Genet.*, 1–18.

Karimov, A., Qadir, M., Noble, A., Vyshpolsky, F. and Anzelm, K. (2009). Development of magnesium-dominant soils under irrigated agriculture in southern Kazakhstan. *Pedosphere* 19, 331–343. Available at: https://www.sciencedirect.com/science/article/pii/S1002016009601247?via%3Dihub.

Kaskarbayev, Z. A. (1998). Research on spring wheat technology in Northern Kazakstan. in *Proceedings of the Kazakstan-CIMMYT Conference, Shortandy, Akmola,(Kazakstan), 22-24 Sep 1997* (CIMMYT).

Kaskarbayev, Z. A. and Kenzhebekov, A. (2004). Soil moisture is the basis of high yields. *Akmolinskaya pravda*, 17–19.

Kato, K., Miura, H. and Sawada, S. (1999). Comparative mapping of the wheat Vrn-A1 region with the rice Hd-6 region. *Genome* 42, 204–209.

Kaut, A. H. E. E., Mason, H. E., Navabi, A., O'Donovan, J. T. and Spaner, D. (2009). Performance and stability of performance of spring wheat variety mixtures in organic and conventional management systems in western Canada. *J. Agric. Sci.* 147, 141–153. doi:10.1017/S0021859608008319.

Kippes, N., Chen, A., Zhang, X., Lukaszewski, A. J. and Dubcovsky, J. (2016). Development and characterization of a spring hexaploid wheat line with no functional VRN2 genes. *Theor. Appl. Genet.* 129, 1417–1428. doi:10.1007/s00122-016-2713-3.

Kippes, N., Debernardi, J. M., Vasquez-Gross, H. A., Akpinar, B. A., Budak, H., Kato, K., Chao, S., Akhunov, E. and Dubcovsky, J. (2015). Identification of the VERNALIZATION 4 gene reveals the origin of spring growth habit in ancient wheats from South Asia. *Proc. Natl. Acad. Sci. U. S. A.* 112, E5401–E5410. doi:10.1073/pnas.1514883112.

Kippes, N., Zhu, J., Chen, A., Vanzetti, L., Lukaszewski, A., Nishida, H., Kato, K., Dvorak, J. and Dubcovsky, J. (2014). Fine mapping and epistatic interactions of the vernalization gene VRN-D4 in hexaploid wheat. *Mol. Genet. Genomics* 289, 47–62. Available at: https://cloudfront.escholarship.org/dist/prd/content/qt8wh2f2pr/qt8wh2f2pr.pdf?t=o6zthv.

Koishybayev, M., Boltybayeva, L. A. and Kopirova, G. I. (2008). Wheat germplasm with group resistance to diseases with airborne infection. *Agromeridian* 3, 34–42.

Kokhmetova, A., Kremneva, O., Volkova, G., Atishova, M. and Sapakhova, Z. (2017). Evaluation of wheat cultivars growing in kazakhstan and russia for resistance to tan spot. *J.*

Plant Pathol. 99, 161–167. doi:10.4454/jpp.v99i1.3812.

Kokhmetova, A., Sharma, R. C., Rsaliyev, S., Galymbek, K., Baymagambetova, K., Ziyaev, Z. and Morgounov, A. (2018). Evaluation of Central Asian wheat germplasm for stripe rust resistance. *Plant Genet. Resour. Util.* 16, 178–184. doi:10.1017/s1479262117000132.

Kowalski, A. M., Gooding, M., Ferrante, A., Slafer, G. A., Orford, S., Gasperini, D. and Griffiths, S. (2016). Agronomic assessment of the wheat semi-dwarfing gene Rht8 in contrasting nitrogen treatments and water regimes. *F. Crop. Res.* 191, 150–160. doi:10.1016/j.fcr.2016.02.026.

Kulkarni, M., Soolanayakanahally, R., Ogawa, S., Uga, Y., Selvaraj, M. G. and Kagale, S. (2017). Drought response in wheat: Key genes and regulatory mechanisms controlling root system architecture and transpiration efficiency. *Front. Chem.* 5, 1–13. doi:10.3389/fchem.2017.00106.

Kulwal, P. (2018). Trait Mapping Approaches Through Linkage Mapping in Plants. *Adv Biochem Eng Biotechnol* 164, 53–82. doi:DOI: 10.1007/10_2017_49.

Kurishbayev, A. (2003). *A strategy for soil conservation farming in Northern Kazakhstan*. , eds. L. GarciaTorres, J. Benites, A. MartinezVilela, and A. HolgadoCabrera.

Kurkiev, K. U., Tyryshkin, L. G., Kolesova, M. A., and Kurkiev, U. K. (2008). Identification of *rht2* and *rht8* genes for semidwarfness in hexaploid triticale with use of dna markers. *Vavilovskii Zhurnal Genet. Selektsii* 12, 372–377.

Kuzmin, V. P. (1970). Breeding of spring wheat for drought resistance in Northern Kazakhstan. *MIncrease drought Toler. Cereal. M*, 6a17.

Lamari, L., Strelkov, S. E., Yahyaoui, A., Amedov, M., Saidov, M., Djunusova, M. and Koichibayev, M. (2005). Virulence of Pyrenophora tritici-repentis in the countries of the Silk Road. *Can. J. Plant Pathol. Can. Phytopathol.* 27, 383–388. Available at: https://www.tandfonline.com/doi/abs/10.1080/07060660509507236.

Lander, E., Green, P., Abrahamson, J., Barlow, A., Daly, M., Lincoln, S. and Newburg, L. (1987). MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1, 174–181.

Lanning, S. P., Martin, J. M., Stougaard, R. N., Guillen-Portal, F. R., Blake, N. K., Sherman, J. D., Robbins, A. M., Kephart, K. D., Lamb, P., Carlson, G. R., Pumphrey, M. and Talbert, L. E. (2012). Evaluation of near-isogenic lines for three height-reducing genes in hard red spring wheat. *Crop Sci.* 52, 1145–1152. doi:10.2135/cropsci2011.11.0625.

Li, A., Yang, W., Guo, X., Liu, D., Sun, J. and Zhang, A. (2012). Isolation of a gibberellin-insensitive dwarfing gene, Rht-B1e, and development of an allele-specific PCR marker. *Mol. Breed.* 30, 1443–1451. doi:10.1007/s11032-012-9730-y.

Li, D., Wang, L., Wang, M., Xu, Y. Y., Luo, W., Liu, Y. J., Xu, Z. H., Li, J. and Chong, K. (2009). Engineering OsBAK1 gene as a molecular tool to improve rice architecture for high yield. *Plant Biotechnol. J.* 7, 791–806. doi:10.1111/j.1467-7652.2009.00444.x.

Liu, R. (2021). 16th ICC Cereal and Bread Congress. in *Study on the quality characteristics and applicability of wheat from Kazakhstan on Chinese traditional staple food*, 40. Available at: https://research.tees.ac.uk/ws/portalfiles/portal/25382425/Book_of_Abstracts_ICBC2021_.pdf.

Liu, S., Zhang, X., Pumphrey, M. O., Stack, R. W., Gill, B. S. and Anderson, J. A. (2006). Complex microcolinearity among wheat, rice, and barley revealed by fine mapping of the genomic region harboring a major QTL for resistance to Fusarium head blight in wheat. *Funct. Integr. Genomics* 6, 83–89. doi:10.1007/s10142-005-0007-y.

Locascio, A., Blázquez, M. A. and Alabadí, D. (2013). Genomic analysis of della protein activity. *Plant Cell Physiol.* 54, 1229–1237. doi:10.1093/pcp/pct082.

Lorieux, M. (2012). MapDisto: Fast and efficient computation of genetic linkage maps. *Mol. Breed.* 30, 1231–1235. doi:10.1007/s11032-012-9706-y.

Manly, K. F., Cudmore, R. H. and Meer, J. M. (2001). Map Manager QTX, cross-platform software for genetic mapping. *Mamm. Genome* 12, 930–932. doi:10.1007/s00335-001-1016-3.

Mansur, L. M., Orf, J. H., Chase, K., Jarvik, T., Cregan, P. B. and Lark, K. G. (1996). Genetic mapping of agronomic traits using recombinant inbred lines of soybean. *Crop Sci.* 36, 1327–1336. doi:10.2135/cropsci1996.0011183X003600050042x.

Masle, J., Gilmore, S. R. and Farquhar, G. D. (2005). The ERECTA gene regulates plant transpiration efficiency in Arabidopsis. *Nature* 436, 866–870. doi:10.1038/nature03835.

Mayer, K. F. X., Marcussen, T., Sandve, S. R., Heier, L., Pfeifer, M., Kugler, K. G., Zhan, B., Spannagl, M., Pfeifer, M., Jakobsen, K. S., Wulff, B. B. H., Steuernagel, B. and Olsen, O.-A. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome Ancient hybridizations among the ancestral genomes of bread wheat Genome interplay in the grain transcriptome of hexaploid bread wheat Structural and functional pa. *Science* 345, 1250092. Available at: http://www.sciencemag.org/content/345/6194/1250092.abstract.

McFadden, E. S. and Sears, E. R. (1946). The origin of Triticum spelta and its free-threshing hexaploid relatives. *J. Hered.* 37, 107–116.

McIntosh, R. A., Dubcovsky, J., Rogers, W. J., Morris, C., Appels, R., Xia, X. C. and Azul, P. de B. A. (2000). CATALOGUE OF GENE SYMBOLS FOR WHEAT: 2009 SUPPLEMENT. *Wheat Inf. Serv*, 33–70.

McIntosh, R. A., Dubcovsky, J., Rogers, W. J., Morris, C., Xia, X. C. and Designators, L. (2017). Catalogue of Gene Symbols for Wheat: 2017 Supplement. *Komugi Wheat Genet. Resour. Database* Comm. Natl, Kyoto, Japan, updated Jan. 11. Available at: https://shigen.nig.ac.jp/wheat/komugi/genes/macgene/supplement2017.pdf.

McIntosh, R. and Yamazaki, Y. (2013). Catalogue of gene symbols for wheat. *Int. Wheat Genet. Symp.* 4, 1–197. Available at:

http://www.cabdirect.org/abstracts/20083014739.html%5Cnhttp://scabusatest1.com/ggpages/wgc/2003/Catalogue.pdf.

McWhorter (1981). Johnsongrass as a Weed. *Farmers' Bull. United States Dep. Agric. Sci. Educ. Adminstration* , 13. Available at: https://books.google.co.uk/books?id=CLfcf-OoOBgC&pg=PA13&dq=disk+cultivation&hl=en&sa=X&ved=0ahUKEwiB5OrI3-vaAhWlB8AKHXnUB0wQ6AEILDAB#v=onepage&q=disk cultivation&f=false.

Milec, Z., Sumíková, T., Tomková, L. and Pánková, K. (2013). Distribution of different Vrn-B1 alleles in hexaploid spring wheat germplasm. *Euphytica* 192, 371–378. doi:10.1007/s10681-013-0863-9.

Milec, Z., Tomková, L., Sumíková, T. and Pánková, K. (2012). A new multiplex PCR test for the determination of Vrn-B1 alleles in bread wheat (Triticum aestivum L.). *Mol. Breed.* 30, 317–323. doi:10.1007/s11032-011-9621-7.

Miura, H., Nakagawa, M. and Worland, A. J. (1999). Control of ear emergence time by chromosome 3A of wheat. *Plant Breed.* 118, 85–87.

Mizina, S. V, Eserkepova, I. B., Pilifosova, O. V, Dolgih, S. A. and Gossen, E. F. (1996). "Model-based climate change vulnerability and adaptation assessment for wheat yields in Kazakhstan," in *Adapting to Climate Change* (Springer), 148–163.

Mo, Y., Vanzetti, L. S., Hale, I., Spagnolo, E. J., Guidobaldi, F., Al-Oboudi, J., Odle, N., Pearce, S., Helguera, M. and Dubcovsky, J. (2018). Identification and characterization of Rht25, a locus on chromosome arm 6AS affecting wheat plant height, heading time, and spike development. *Theor. Appl. Genet.* 131, 2021–2035. doi:10.1007/s00122-018-3130-6.

Mohan, A., Grant, N. P., Schillinger, W. F. and Gill, K. S. (2021). Characterizing reduced height wheat mutants for traits affecting abiotic stress and photosynthesis during seedling growth. *Physiol. Plant.* 172, 233–246. doi:10.1111/ppl.13321.

Morgounov, A., Akin, B., Demir, L., Keser, M., Kokhmetova, A., Martynov, S., Orhan, S., Ozdemir, F., Ozseven, I., Sapakhova, Z. and Yessimbekova, M. (2015). Yield gain due to fungicide application in varieties of winter wheat (Triticum aestivum) resistant and susceptible to leaf rust. *Crop Pasture Sci.* 66, 649–659. doi:10.1071/cp14158.

Morgounov, A., Gómez-Becerra, H. F., Abugalieva, A., Dzhunusova, M., Yessimbekova, M., Muminjanov, H., Zelenskiy, Y., Ozturk, L. and Cakmak, I. (2007a). Iron and zinc grain density in common wheat grown in Central Asia. *Euphytica* 155, 193–203.

Morgounov, A., Rosseeva, L. and Koyshibayev, M. (2007b). Leaf rust of spring wheat in Northern Kazakhstan and Siberia: Incidence, virulence, and breeding for resistance. *Aust. J. Agric. Res.* 58, 847–853. doi:10.1071/AR07086.

Morgounov, A., Zykin, V. A., Sereda, G. A. and Urazaliev, R. A. (2001). Siberian and North Kazakhstan wheat pool. *World Wheat B. A Hist. Wheat Breeding. A. Bonjean W. Angus (eds.). Lavoisier Publ. Paris*, 755–772.

Movchan, V. (1998). Spring wheat breeding in Kazakstan. in *Proceedings of the Kazakstan-*

CIMMYT Conference, Shortandy, Akmola,(Kazakstan), 22-24 Sep 1997 (CIMMYT).

Mu, Q., Guo, T., Li, X. and Yu, J. (2021). Phenotypic plasticity in plant height shaped by interaction between genetic loci and diurnal temperature range. *New Phytol.* doi:10.1111/nph.17904.

Munoz-Amatriain, M., Moscou, M. J., Bhat, P. R., Svensson, J. T., Bartos, J., Suchankova, P., Simkova, H., Endo, T. R., Fenton, R. D., Lonardi, S., Castillo, A. M., Chao, S. M., Cistue, L., Cuesta-Marcos, A., Forrest, K. L., Hayden, M. J., Hayes, P. M., Horsley, R. D., Makoto, K., Moody, D., Sato, K., Valles, M. P., Wulff, B. B. H., Muehlbauer, G. J., Dolezel, J. and Close, T. J. (2011). An Improved Consensus Linkage Map of Barley Based on Flow-Sorted Chromosomes and Single Nucleotide Polymorphism Markers. *Plant Genome* 4, 238–249. doi:10.3835/plantgenome2011.08.0023.

Mursalova, J., Akparov, Z., Ojaghi, J., Eldarov, M. and Yuma, A. (2015). Evaluation of drought tolerance of winter bread wheat genotypes under drip irrigation and rain-fed conditions. 817–824. doi:10.3906/tar-1407-152.

Muterko, A., Kalendar, R., Cockram, J. and Balashova, I. (2015). Discovery, evaluation and distribution of haplotypes and new alleles of the Photoperiod-A1 gene in wheat. *Plant Mol. Biol.* 88, 149–164. doi:10.1007/s11103-015-0313-2.

Nagaoka, T. and Ogihara, Y. (1997). Applicability of inter-simple sequence repeat polymorphisms in wheat for use as DNA markers in comparison to RFLP and RAPD markers. *Theor. Appl. Genet.* 94, 597–602. doi:10.1007/s001220050456.

Nakamichi, N., Kita, M., Niinuma, K., Ito, S., Yamashino, T., Mizoguchi, T. and Mizuno, T. (2007). Arabidopsis clock-associated pseudo-response regulators PRR9, PRR7 and PRR5 coordinately and positively regulate flowering time through the canonical CONSTANS-dependent photoperiodic pathway. *Plant Cell Physiol.* 48, 822–832. doi:10.1093/pcp/pcm056.

Nantawan, U., Kanchana-Udomkan, C., Bar, I. and Ford, R. (2019). Linkage mapping and quantitative trait loci analysis of sweetness and other fruit quality traits in papaya. *BMC Plant Biol.* 19, 1–11. doi:10.1186/s12870-019-2043-0.

Naoura, G., Sawadogo, N., Atchozou, E. A., Emendack, Y., Tabo, R. and Laza, H. (2019). Assessment of agro-morphological variability of dry-season sorghum cultivars in Chad as novel sources of drought tolerance. 1–12. doi:10.1038/s41598-019-56192-6.

Nesbitt, M. and Samuel, D. (1998). Wheat domestication: archaeobotanical evidence. *Science (80-.).* 279, 1431.

Nishida, H., Yoshida, T., Kawakami, K., Fujita, M., Long, B., Akashi, Y., Laurie, D. A. and Kato, K. (2013). Structural variation in the 5′ upstream region of photoperiod-insensitive alleles Ppd-A1a and Ppd-B1a identified in hexaploid wheat (Triticum aestivum L.), and their effect on heading time. *Mol. Breed.* 31, 27–37. doi:10.1007/s11032-012-9765-0.

Omirbekova, N., Zhussupova, A., Askanbayeva, B., Kenzhebaeva, S., Zhunusbayeva, Z. and Sgem (2016). Study of storage proteins in endosperm and antioxidant enzyme activity of soft

wheat and brachypodium distachyon infected by puccinia recondita, in *Ecology, Economics, Education and Legislation Conference Proceedings, Sgem 2016, Vol Ii*, 767–774.

Pachikin, K., Erokhina, O. and Funakawa, S. (2014). "Soils of Kazakhstan, their distribution and mapping," in *Novel Measurement and Assessment Tools for Monitoring and Management of Land and Water Resources in Agricultural Landscapes of Central Asia* (Springer), 519–533.

Patil, H. B., Chaurasia, A. K., Azeez, A., Krishna, B., Subramaniam, V. R., Sane, A. P. and Sane, P. V (2017). Research paper Characterization of two TERMINAL FLOWER1 homologs PgTFL1 and PgCENa from pomegranate (Punica granatum L .). 772–784. doi:10.1093/treephys/tpx154.

Pearce, S., Saville, R., Vaughan, S. P., Chandler, P. M., Wilhelm, E. P., Sparks, C. A., Al-Kaff, N., Korolev, A., Boulton, M. I., Phillips, A. L., Hedden, P., Nicholson, P. and Thomas, S. G. (2011). Molecular characterization of Rht-1 dwarfing genes in hexaploid wheat. *Plant Physiol.* 157, 1820–1831. doi:10.1104/pp.111.183657.

Peng, J., Richards, D. E., Hartley, N. M., Murphy, G. P., Devos, K. M., Flintham, J. E., Beales, J., Fish, L. J., Worland, A. J. and Pelica, F. (1999). 'Green revolution'genes encode mutant gibberellin response modulators. *Nature* 400, 256. Available at: https://www.nature.com/articles/22307.pdf.

Peng, Z. S., Li, X., Yang, Z. J. and Liao, M. L. (2011). A new reduced height gene found in the tetraploid semi-dwarf wheat landrace Aiganfanmai. *Genet. Mol. Res.* 10, 2349–2357. doi:10.4238/2011.October.5.5.

Peters, J. L., Cnudde, F. and Gerats, T. (2003). Forward genetics and map-based cloning approaches. *Trends Plant Sci.* 8, 484–491. doi:10.1016/j.tplants.2003.09.002.

Pilifosova, O. V, Eserkepova, I. B. and Dolgih, S. A. (1997). Regional climate change scenarios under global warming in Kazakhstan. *Clim. Change* 36, 23–40.

Prescott, J. M., Burnett, P. A., Saari, E. E., Ransom, J. K., Bowman, J., De Milliano, W. A. J., Singh, R. P. and Bekele, G. T. (2002). Wheat Diseases and Pests: a guide for field identification. *CIMMYT*.

Pysh, L. D., Wysocka-Diller, J. W., Camilleri, C., Bouchez, D. and Benfey, P. N. (1999). The GRAS gene family in Arabidopsis: Sequence characterization and basic expression analysis of the SCARECROW-LIKE genes. *Plant J.* 18, 111–119. doi:10.1046/j.1365-313X.1999.00431.x.

R Scarth and Law, C. (1983). The control of day-length response in wheat by the group 2 chromosome 2B of wheat. *Heredity (Edinb).* 51, 607–619.

Ragab, R. and Prudhomme, C. (2002). Sw—soil and Water: climate change and water resources management in arid and semi-arid regions: prospective and challenges for the 21st century. *Biosyst. Eng.* 81, 3–34. Available at: https://www.sciencedirect.com/science/article/pii/S1537511001900133?via%3Dihub.

Ramazanov, O. (2006). Proposed research in irrigated agriculture and water resources management. *Proc. Conf. Transit. Towar. Mark. Econ. Water Resour. Manag. L. Amelior.*, 4–9.

Ramírez-González, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., Davey, M., Jacobs, J., Van Ex, F., Pasha, A., Khedikar, Y., Robinson, S. J., Cory, A. T., Florio, T., Concia, L., Juery, C., Schoonbeek, H., Steuernagel, B., Xiang, D., Ridout, C. J., Chalhoub, B., Mayer, K. F. X., Benhamed, M., Latrasse, D., Bendahmane, A., Wulff, B. B. H., Appels, R., Tiwari, V., Datla, R., Choulet, F., Pozniak, C. J., Provart, N. J., Sharpe, A. G., Paux, E., Spannagl, M., Bräutigam, A. and Uauy, C. (2018). The transcriptional landscape of polyploid wheat. *Science (80-.).* 361. doi:10.1126/science.aar6089.

Rebetzke, G. J., Appels, R., Morrison, A. D., Richards, R. A., McDonald, G., Ellis, M. H., Spielmeyer, W. and Bonnett, D. G. (2001). Quantitative trait loci on chromosome 4B for coleoptile length and early vigour in wheat (Triticum aestivum L.). *Aust. J. Agric. Res.* 52, 1221–1234. doi:10.1071/ar01042.

Rebetzke, G. J., Bonnett, D. G. and Ellis, M. H. (2012). Combining gibberellic acid-sensitive and insensitive dwarfing genes in breeding of higher-yielding, sesqui-dwarf wheats. *F. Crop. Res.* 127, 17–25. doi:10.1016/j.fcr.2011.11.003.

Rebetzke, G. J. and Richards, R. A. (2000). Gibberellic acid-sensitive dwarfing genes reduce plant height to increase kernel number and grain yield of wheat. *Aust. J. Agric. Res.* 51, 235–246.

Rebetzke, G. J., Richards, R. A., Fettell, N. A., Long, M., Condon, A. G., Forrester, R. I. and Botwright, T. L. (2007). Genotypic increases in coleoptile length improves stand establishment, vigour and grain yield of deep-sown wheat. *F. Crop. Res.* 100, 10–23. Available at: https://www.sciencedirect.com/science/article/pii/S0378429006001092?via%3Dihub.

Ren, R., Yao, M. and Guan, M. (2022). Regional association and transcriptome analysis revealed candidate genes controlling plant height in Brassica napus. *Res. Sq.*, 1–19. doi:https://doi.org/10.21203/rs.3.rs-1460223/v1.

Rivera-Amado, C., Trujillo-Negrellos, E., Molero, G., Reynolds, M. P., Sylvester-Bradley, R. and Foulkes, M. J. (2019). Optimizing dry-matter partitioning for increased spike growth, grain number and harvest index in spring wheat. *F. Crop. Res.* 240, 154–167. doi:10.1016/j.fcr.2019.04.016.

Rsaliyev, A., Yskakova, G., Maulenbay, A., Zakarya, K. and Rsaliyev, S. (2020). Virulence and race structure of puccinia graminis f. Sp. tritici in kazakhstan. *Plant Prot. Sci.* 56, 275–284. doi:10.17221/85/2020-PPS.

Ruxton, G. D. and Beauchamp, G. (2008). Time for some a priori thinking about post hoc testing. *Behav. Ecol.* 19, 690–693. doi:10.1093/beheco/arn020.

Sanchez-Martin, J., Steuernagel, B., Ghosh, S., Herren, G., Hurni, S., Adamski, N., Vrana, J., Kubalakova, M., Krattinger, S. G., Wicker, T., Dolezel, J., Keller, B. and Wulff, B. B. H. (2016). Rapid gene isolation in barley and wheat by mutant chromosome sequencing. *Genome Biol.* 17. doi:10.1186/s13059-016-1082-1.

Santra, M., Wang, H., Seifert, S. and Haley, S. (2017). *Doubled Haploid Laboratory Protocol for Wheat Using Wheat–Maize Wide Hybridization*. doi:10.1007/978-1-4939-7337-8_6.

Schiex, T. and Gaspin, C. (1997). CARTHAGENE: constructing and joining maximum likelihood genetic maps. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 5, 258–267.

Shamanin, V. P., Morgunov, A. I., Manes, J., Zelensky, Y. I., Chursin, A. S., Levshunov, M. A., IV, P., Likhenko, I. E., Ianko, T. A. and Karakoz, I. I. (2012). Breeding and genetic estimation of spring bread wheat populations of the Siberian shuttle breeding nursery of CIMMYT.

Shamanin, V., Salina, E., Wanyera, R., Zelenskiy, Y., Olivera, P. and Morgounov, A. (2016). Genetic diversity of spring wheat from Kazakhstan and Russia for resistance to stem rust Ug99. *Euphytica* 212, 287–296. doi:10.1007/s10681-016-1769-0.

Shaw, L. M., Turner, A. S. and Laurie, D. A. (2012). The impact of photoperiod insensitive Ppd-1a mutations on the photoperiod pathway across the three genomes of hexaploid wheat (Triticum aestivum). *Plant J.* 71, 71–84.

Sherman, J. D., Yan, L., Talbert, L. and Dubcovsky, J. (2004). A PCR marker for growth habit in common wheat based on allelic variation at the VRN-A1 gene. *Crop Sci.* 44, 1832–1838. doi:10.2135/cropsci2004.1832.

Shewry, P. R. (2009). Wheat. *J. Exp. Bot.* 60, 1537–1553. doi:10.1093/jxb/erp058.

Shinada, H., Iwata, N., Sato, T. and Fujino, K. (2014). QTL pyramiding for improving of cold tolerance at fertilization stage in rice. 488, 483–488. doi:10.1270/jsbbs.63.483.

Shingala, M. C. (2015). Comparison of Post Hoc Tests for Unequal Variance. 2, 22–33.

Shirvanov, R. B. and Rybalko, A. G. (2001). Snow retention as one of the elements of resource conservation in the cultivation of crops. *САРАТОВСКОГО ГОСАГРОУНИВЕРСИТЕТА им. НИ ВАВИЛОВАVAVILOV NI Saratov State Agrar. Univ.*, 70.

Singh, R. P., Huerta-Espino, J. and William, H. M. (2005). Genetics and breeding for durable resistance to leaf and stripe rusts in wheat. *Turkish J. Agric. For.* 29, 121–127.

Singla, B., Chugh, A., Khurana, J. P. and Khurana, P. (2006). An early auxin-responsive Aux/IAA gene from wheat (Triticum aestivum) is induced by epibrassinolide and differentially regulated by light and calcium. *J. Exp. Bot.* 57, 4059–4070. doi:10.1093/jxb/erl182.

Sivasakthi, K., Thudi, M., Tharanya, M., Kale, S. M., Kholová, J., Halime, M. H., Jaganathan, D., Baddam, R., Thirunalasundari, T., Gaur, P. M., Varshney, R. K. and Vadez, V. (2018). Plant vigour QTLs co-map with an earlier reported QTL hotspot for drought tolerance while water saving QTLs map in other regions of the chickpea genome.

BMC Plant Biol. 18, 1–18. doi:10.1186/s12870-018-1245-1.

Smith, C. W. (1995). *Crop production: evolution, history, and technology*. John Wiley & Sons.

Snape, J. W., Law, C. N., Parker, B. B. and Worland, A. J. (1985). Genetical analysis of chromosome 5A of wheat and its influence on important agronomic characters. *Theor. Appl. Genet.* 71, 518–526. doi:10.1007/BF00251199.

Sommer, R., Glazirina, M., Yuldashev, T., Otarov, A., Ibraeva, M., Martynova, L., Bekenov, M., Kholov, B., Ibragimov, N., Kobilov, R., Karaev, S., Sultonov, M., Khasanova, F., Esanbekov, M., Mavlyanov, D., Isaev, S., Abdurahimov, S., Ikramov, R., Shezdyukova, L. and de Pauw, E. (2013). Impact of climate change on wheat productivity in Central Asia. *Agric. Ecosyst. Environ.* 178, 78–99. doi:10.1016/j.agee.2013.06.011.

Sonder, K. (2016). Global map of wheat mega-environments. *CIMMYT Res. Data Softw. Repos.* version 4.

Song, P., Wang, J., Guo, X., Yang, W. and Zhao, C. (2021). High-throughput phenotyping: Breaking through the bottleneck in future crop breeding. *Crop J.* 9, 633–645. doi:10.1016/j.cj.2021.03.015.

Sreekanta, S., Bethke, G., Hatsugai, N., Tsuda, K., Thao, A., Wang, L., Katagiri, F. and Glazebrook, J. (2015). The receptor-like cytoplasmic kinase PCRK1 contributes to pattern-triggered immunity against Pseudomonas syringae in Arabidopsis thaliana. *New Phytol.* 207, 78–90. doi:10.1111/nph.13345.

Stam, P. (1993). Construction of integrated genetic linkage maps by means of a new computer package: JOINMAP. *Plant J.* 3, 739–744. doi:10.1111/j.1365-313X.1993.00739.x.

Steuernagel, B., Periyannan, S. K., Hernandez-Pinzon, I., Witek, K., Rouse, M. N., Yu, G. T., Hatta, A., Ayliffe, M., Bariana, H., Jones, J. D. G., Lagudah, E. S. and Wulff, B. B. H. (2016). Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nat. Biotechnol.* 34, 652–655. doi:10.1038/nbt.3543.

Sturtevant, A. H. (1913). The linear arrangement of six sex-linked factors in Drosophila, as shown by their mode of association. *J. Exp. Zool.* 14, 43–59. doi:10.1002/jez.1400140104.

Suiter, K. A., Wendel, J. F. and Case, J. S. (1983). LINKAGE-1: A PASCAL computer program for the detection and analysis of genetic linkage. *J. Hered.* 74, 203–204. doi:10.1093/oxfordjournals.jhered.a109766.

Suleimenova, N. (2000). *Agro-ecological soil cultivation for mountainous farming in Southeastern Kazakhstan*. , eds. M. Karabayev, A. Satybaldin, J. Benites, T. Friedrich, M. Pala, and T. Payne.

Sun, L., Yang, W., Li, Y., Shan, Q., Ye, X., Wang, D., Yu, K., Lu, W., Xin, P., Pei, Z., Guo, X., Liu, D., Sun, J., Zhan, K., Chu, J. and Zhang, A. (2019). A wheat dominant dwarfing line with Rht12, which reduces stem cell length and affects gibberellic acid synthesis, is a 5AL terminal deletion line. *Plant J.* 97, 887–900. doi:10.1111/tpj.14168.

Swinnen, J., Burkitbayeva, S., Schierhorn, F., Prishchepov, A. V and Müller, D. (2017).

Production potential in the “bread baskets” of Eastern Europe and Central Asia. *Glob. Food Sec.* 14, 38–53.

Szalma, S. J., Hostert, B. M., LeDeaux, J. R., Stuber, C. W. and Holland, J. B. (2007). QTL mapping with near-isogenic lines in maize. *Theor. Appl. Genet.* 114, 1211–1228. doi:10.1007/s00122-007-0512-6.

Tang, T., Botwright Acuña, T., Spielmeyer, W. and Richards, R. A. (2021). Effect of gibberellin-sensitive Rht18 and gibberellin-insensitive Rht-D1b dwarfing genes on vegetative and reproductive growth in bread wheat. *J. Exp. Bot.* 72, 445–458. doi:10.1093/jxb/eraa481.

Tanio, M. and Kato, K. (2007). Development of near-isogenic lines for photoperiod-insensitive genes, Ppd-B1 and Ppd-D1, carried by the Japanese wheat cultivars and their effect on apical development. *Breed. Sci.* 57, 65–72.

Taylor, J. and Butler, D. (2017). R package ASMap: Efficient genetic linkage map construction and diagnosis. *J. Stat. Softw.* 79. doi:10.18637/jss.v079.i06.

Tee, T. S. and Qualset, C. O. (1975). Bulk populations in wheat breeding: comparison of single-seed descent and random bulk methods. *Euphytica* 24, 393–405. doi:10.1007/BF00028206.

Thind, A. K., Wicker, T., Šimková, H., Fossati, D., Moullet, O., Brabant, C., Vrána, J., Doležel, J. and Krattinger, S. G. (2017). Rapid cloning of genes in hexaploid wheat using cultivar-specific long-range chromosome assembly. *Nat. Biotechnol.* 35, 793–796. doi:10.1038/nbt.3877.

Tian, X., Wen, W., Xie, L., Fu, L., Xu, D., Fu, C., Wang, D., Chen, X., Xia, X., Chen, Q., He, Z. and Cao, S. (2017). Molecular mapping of reduced plant height gene Rht24 in bread wheat. *Front. Plant Sci.* 8, 1–9. doi:10.3389/fpls.2017.01379.

Tóth, B., Galiba, G., Fehér, E., Sutka, J. and Snape, J. W. (2003). Mapping genes affecting flowering time and frost resistance on chromosome 5B of wheat. *Theor. Appl. Genet.* 107, 509–514. Available at: https://link.springer.com/content/pdf/10.1007%2Fs00122-003-1275-3.pdf.

Trethowan, R. M., Morgunov, A., He, Z., De Pauw, R., Crossa, J., Warburton, M., Baytasov, A., Zhang, C., Mergoum, M. and Alvarado, G. (2006). The global adaptation of bread wheat at high latitudes. *Euphytica* 152, 303–316.

Trevaskis, B. (2015). Wheat gene for all seasons. *Proc. Natl. Acad. Sci. U. S. A.* 112, 11991–11992. doi:10.1073/pnas.1516398112.

Trevaskis, B., Hemming, M. N., Dennis, E. S. and Peacock, W. J. (2007). The molecular basis of vernalization-induced flowering in cereals. *Trends Plant Sci.* 12, 352–357. Available at: https://www.sciencedirect.com/science/article/pii/S1360138507001562?via%3Dihub.

Turner, A., Beales, J., Faure, S., Dunford, R. P. and Laurie, D. A. (2005). The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. *Science (80-.).* 310, 1031–1034.

Turuspekov, Y., Baibulatova, A., Yermekbayev, K., Tokhetova, L., Chudinov, V., Sereda, G.,

Ganal, M., Griffiths, S. and Abugalieva, S. (2017a). GWAS for plant growth stages and yield components in spring wheat (Triticum aestivum L.) harvested in three regions of Kazakhstan. *BMC Plant Biol.* 17. doi:10.1186/s12870-017-1131-2.

Turuspekov, Y., Sariev, B., Chudinov, V., Sereda, G., Tokhetova, L., Ortaev, A., Tsygankov, V., Doszhanov, M., Volis, S. and Abugalieva, S. (2013). Genotype x environment interaction patterns for grain yield of spring barley in different regions of Kazakhstan. *Russ. J. Genet.* 49, 196–205. doi:10.1134/s1022795413020129.

Turuspekov, Y., Yermekbayev, K. and Griffiths, S. (2017b). GENOTYPING OF SPRING HEXAPLOID WHEAT COLLECTION FROM KAZAKHSTAN USING VRN, PPD AND EPS GENES. in *"Plant Genetics, Genomics, Bioinformatics and Biotechnology": Materials of the 4th International Conference PlantGen. Poster presentation* (Almaty), 94. Available at: http://primerdigital.com/PlantGen2017/files/Abstracts_book.pdf.

Tyagi, S., Mir, R. R., Kaur, H., Chhuneja, P., Ramesh, B., Balyan, H. S. and Gupta, P. K. (2014). Marker-assisted pyramiding of eight QTLs / genes for seven different traits in common wheat (Triticum aestivum L .). 167–175. doi:10.1007/s11032-014-0027-1.

Uffelmann, E., Huang, Q. Q., Nchangwi, S. M., Jantina, de V., Yukinori, O., Alicia, R. M., Hilary, C. M., Tuuli, L. and Danielle, P. (2021). Genome-Wide Association Studies. *Nat. Rev. Methods Prim.*, 33–41. doi:10.1016/B978-0-12-804182-6.00003-4.

Van De Velde, K., Ruelens, P., Geuten, K., Rohde, A. and Van Der Straeten, D. (2017). Exploiting DELLA Signaling in Cereals. *Trends Plant Sci.* 22, 880–893. doi:10.1016/j.tplants.2017.07.010.

Van Os, H., Stam, P., Visser, R. G. F. and Van Eck, H. J. (2005). RECORD: A novel method for ordering loci on a genetic linkage map. *Theor. Appl. Genet.* 112, 30–40. doi:10.1007/s00122-005-0097-x.

Varshney, R. K., Sinha, P., Singh, V. K., Kumar, A., Zhang, Q. and Bennetzen, J. L. (2020). 5Gs for crop genetic improvement. *Curr. Opin. Plant Biol.* 56, 190–196. doi:10.1016/j.pbi.2019.12.004.

Véronique, L., Paul, M., Paul-Henry, C., Philippe, D. R. and Brigitte, C. (2008). Quantitative Genetics and Functional-Structural Plant Growth Models: Simulation of Quantitative Trait Loci Detection for Model Parameters and Application to Potential Yield Optimization. *Ann. Bot.* 101, 1243–1254.

Vikhe, P., Venkatesan, S., Chavan, A., Tamhankar, S. and Patil, R. (2019). Mapping of dwarfing gene Rht14 in durum wheat and its effect on seedling vigor, internode length and plant height. *Crop J.* 7, 187–197. doi:10.1016/j.cj.2018.11.004.

Viurkov, V. V (2005). Technologies of cultivation of winter crops on black and coulomb pairs in farms of the Urals.

Walkowiak, S., Gao, L., Monat, C., Haberer, G., Kassa, M. T., Brinton, J., Ramirez-Gonzalez, R. H., Kolodziej, M. C., Delorean, E., Thambugala, D., Klymiuk, V., Byrns,

B., Gundlach, H., Bandi, V., Siri, J. N., Nilsen, K., Aquino, C., Himmelbach, A., Copetti, D., Ban, T., Pozniak, C. J. et al. (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature* 588, 277–283. doi:10.1038/s41586-020-2961-x.

Wang, H., Zhang, Y., Norris, A. and Jiang, C. Z. (2022). S1-bZIP Transcription Factors Play Important Roles in the Regulation of Fruit Quality and Stress Response. *Front. Plant Sci.* 12, 1–13. doi:10.3389/fpls.2021.802802.

Wang, J., Luo, M., Chen, Z., You, F. M., Wei, Y., Zheng, Y. and Dvorak, J. (2013). Aegilops tauschii single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat. *New Phytol.* 198, 925–937. Available at: https://nph.onlinelibrary.wiley.com/doi/pdf/10.1111/nph.12164.

Wang, X., Liu, H., Mia, M. S., Siddique, K. H. M. and Yan, G. (2018). Development of near-isogenic lines targeting a major QTL on 3AL for pre-harvest sprouting resistance in bread wheat. *Crop Pasture Sci.* 69, 864–872. doi:10.1071/CP17423.

Wang, Z., Wu, F., Chen, X., Zhou, W., Shi, H., Lin, Y., Hou, S., Yu, S., Zhou, H., Li, C. and Liu, Y. (2021). Fine mapping of the tiller inhibition gene TIN4 contributing to ideal plant architecture in common wheat. *Theor. Appl. Genet.* doi:10.1007/s00122-021-03981-1.

Watson, A., Ghosh, S., Williams, M. J., Cuddy, W. S., Simmonds, J., Rey, M. D., Hatta, M. A. M., Hinchliffe, A., Steed, A., Reynolds, D., Adamski, N. M., Breakspear, A., Korolev, A., Rayner, T., Dixon, L. E., Riaz, A., Martin, W., Ryan, M., Edwards, D., Batley, J., Raman, H., Carter, J., Rogers, C., Domoney, C., Moore, G., Harwood, W., Nicholson, P., Dieters, M. J., DeLacy, I. H., Zhou, J., Uauy, C., Boden, S. A., Park, R. F., Wulff, B. B. H. and Hickey, L. T. (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. *Nat. Plants* 4, 23–29. doi:10.1038/s41477-017-0083-8.

Wilhelm, E. P., Turner, A. S. and Laurie, D. A. (2009). Photoperiod insensitive Ppd-A1a mutations in tetraploid wheat (Triticum durum Desf.). *Theor. Appl. Genet.* 118, 285–294. doi:10.1007/s00122-008-0898-9.

Wingen, L. U., Orford, S., Goram, R., Leverington-Waite, M., Bilham, L., Patsiou, T. S., Ambrose, M., Dicks, J. and Griffiths, S. (2014). Establishing the A. E. Watkins landrace cultivar collection as a resource for systematic gene discovery in bread wheat. *Theor. Appl. Genet.* 127, 1831–1842. doi:10.1007/s00122-014-2344-5.

Wingen, L. U., West, C., Leverington-Waite, M., Collier, S., Orford, S., Goram, R., Yang, C.-Y., King, J., Allen, A. M. and Burridge, A. (2017). Wheat landrace genome diversity. *Genetics* 205, 1657–1676. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5378120/pdf/1657.pdf.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V. and Provart, N. J. (2007). An "electronic fluorescent pictograph" Browser for exploring and analyzing large-scale biological data sets. *PLoS One* 2, 1–12. doi:10.1371/journal.pone.0000718.

Wojciechowski, T., Gooding, M. J., Ramsay, L. and Gregory, P. J. (2009). The effects of

dwarfing genes on seedling root growth of wheat. *J. Exp. Bot.* 60, 2565–2573. doi:10.1093/jxb/erp107.

Woodward, A. W., Ratzel, S. E., Woodward, E. E., Shamoo, Y. and Bartel, B. (2007). Mutation of E1-CONJUGATING ENZYME-RELATED1 decreases RELATED TO UBIQUITIN conjugation and alters auxin response and development. *Plant Physiol.* 144, 976–987. doi:10.1104/pp.107.100404.

Worland, A. J., Borner, A., Korzun, V., Li, W. M., Petrovic, S. and Sayers, E. J. (1998a). The influence of photoperiod genes on the adaptability of European winter wheats (Reprinted from Wheat: Prospects for global improvement, 1998). *Euphytica* 100, 385–394. doi:10.1023/a:1018327700985.

Worland, A. J., Korzun, V., Röder, M. S., Ganal, M. W. and Law, C. N. (1998b). Genetic analysis of the dwarfing gene Rht8 in wheat. Part II. The distribution and adaptive significance of allelic variants at the Rht8 locus of wheat as revealed by microsatellite screening. *Theor. Appl. Genet.* 96, 1110–1120.

Worland, A. J., Law, C. N. and Shakoor, A. (1980). The genetical analysis of an induced height mutant in wheat. *Heredity (Edinb).* 45, 61.

Worland, A. J., Petrovic, S. and Law, C. N. (1988). Genetic analysis of chromosome 2D of wheat. *Plant Breed.* 100, 247–259.

Worthen, J. M., Yamburenko, M. V., Lim, J., Nimchuk, Z. L., Kieber, J. J. and Schaller, G. E. (2019). Type-B response regulators of rice play key roles in growth, development and cytokinin signaling. *Dev.* 146, 1–11. doi:10.1242/dev.174870.

Wu, Y., Bhat, P. R., Close, T. J. and Lonardi, S. (2008). Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph. *PLoS Genet.* 4. doi:10.1371/journal.pgen.1000212.

Würschum, T., Langer, S. M., Longin, C. F. H., Tucker, M. R. and Leiser, W. L. (2017). A modern Green Revolution gene for reduced height in wheat. *Plant J.* 92, 892–903. doi:10.1111/tpj.13726.

Würschum, T., Rapp, M., Miedaner, T., Longin, C. F. H. and Leiser, W. L. (2019). Copy number variation of Ppd-B1 is the major determinant of heading time in durum wheat. *BMC Genet.* 20, 1–8. doi:10.1186/s12863-019-0768-2.

Xu, S. (2013). *Principles of statistical genomics*. doi:10.1007/978-0-387-70807-2.

Xu, Y., Li, P., Yang, Z. and Xu, C. (2017). Genetic mapping of quantitative trait loci in crops. *Crop J.* 5, 175–184. doi:10.1016/j.cj.2016.06.003.

Xue, H., Gao, X., He, P. and Xiao, G. (2021). Origin, evolution, and molecular function of DELLA proteins in plants. *Crop J.* doi:10.1016/j.cj.2021.06.005.

Xue, S., Xu, F., Li, G., Zhou, Y., Lin, M., Gao, Z., Su, X., Xu, X., Jiang, G., Zhang, S., Jia, H., Kong, Z., Zhang, L. and Ma, Z. (2013). Fine mapping TaFLW1, a major QTL controlling flag leaf width in bread wheat (Triticum aestivum L.). *Theor. Appl. Genet.* 126, 1941–1949.

doi:10.1007/s00122-013-2108-7.

Xue, W., Xing, Y., Weng, X., Zhao, Y., Tang, W., Wang, L., Zhou, H., Yu, S., Xu, C., Li, X. and Zhang, Q. (2008). Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. *Nat. Genet.* 40, 761–767. doi:10.1038/ng.143.

Ya-wen, Z. E. N. G., Shi-quan, S. H. E. N., Lu-xiang, W. A. N. G., Jia-fu, L. I. U., Juan, D. U. and Min, Q. I. U. (2005). Correlation of Plant Morphological and Grain Quality Traits with Mineral Element Contents in Yunnan Rice. 12, 101–106.

Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., Sanchez, A., Valarik, M., Yasuda, S. and Dubcovsky, J. (2006). The wheat and barley vernalization gene VRN3 is an orthologue of FT. *Proc. Natl. Acad. Sci.* 103, 19581–19586. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1748268/pdf/zpq19581.pdf.

Yan, L., Helguera, M., Kato, K., Fukuyama, S., Sherman, J. and Dubcovsky, J. (2004). Allelic variation at the VRN-1 promoter region in polyploid wheat. *Theor. Appl. Genet.* 109, 1677–1686. Available at: https://link.springer.com/content/pdf/10.1007%2Fs00122-004-1796-4.pdf.

Yanai, J., Mishima, A., Funakawa, S., Akshalov, K. and Kosaki, T. (2005). Spatial variability of organic matter dynamics in the semi-arid croplands of northern Kazakhstan. *Soil Sci. Plant Nutr.* 51, 261–269.

Yang, J., Chen, F., Yu, O. and Beachy, R. N. (2011). Controlled silencing of 4-coumarate:CoA ligase alters lignocellulose composition without affecting stem growth. *Plant Physiol. Biochem.* 49, 103–109. doi:10.1016/j.plaphy.2010.10.004.

Yang, Y., Wang, W., Xu, T., Liu, N., Wang, H. and Feng, D. (2017). Heterologous expression of wheat TaRUB1 gene enhances disease resistance in Arabidopsis thaliana. *Plant Cell Rep.* 36, 1985–1994. doi:10.1007/s00299-017-2221-1.

Yang, Z., Zheng, J., Liu, C., Wang, Y., Condon, A. G., Chen, Y. and Hu, Y. G. (2015). Effects of the GA-responsive dwarfing gene Rht18 from tetraploid wheat on agronomic traits of common wheat. *F. Crop. Res.* 183, 92–101. doi:10.1016/j.fcr.2015.07.028.

Yasir, T. A., Ali, H., Wasaya, A., Aatif, H. M., Hussain, M., Farooq, M., Baloch, A. W. and Hu, Y. G. (2018). Single nucleotide polymorphisms in TaER genes and their association with carbon isotope discrimination in wheat genotypes under drought. *Biol. Plant.* 62, 703–710. doi:10.1007/s10535-018-0818-z.

Youssefian, S., Kirby, E. J. M. and Gale, M. D. (1992). Pleiotropic effects of the GA-insensitive Rht dwarfing genes in wheat. 2. Effects on leaf, stem, ear and floret growth. *F. Crop. Res.* 28, 191–210. doi:10.1016/0378-4290(92)90040-G.

Yu, H., Luo, C., Fan, Y., Zhang, X., Huang, F., Li, M. and He, X. (2020a). Isolation and characterization of two APETALA1-Like genes from mango (Mangifera indica L.). *Sci. Hortic. (Amsterdam).* 259, 108814. doi:10.1016/j.scienta.2019.108814.

Yu, M., Liu, Z. H., Yang, B., Chen, H., Zhang, H. and Hou, D. Bin (2020b). The contribution of photosynthesis traits and plant height components to plant height in wheat at the individual

quantitative trait locus level. *Sci. Rep.* 10, 1–10. doi:10.1038/s41598-020-69138-0.

Yutshenko, N., Yutshenko, D., Wall, P. and Karabaev, M. (2005). Zero Tillage Grain Crop Rotation in Central Kazakhstan Preserving Soils from Water Erosion. *Increasing wheat Prod. Cent. Asia through Sci. Int. Coop.*, 167.

Zhang, J., Wang, Y., Wu, S., Yang, J., Liu, H. and Zhou, Y. (2012). A single nucleotide polymorphism at the Vrn-D1 promoter region in common wheat is associated with vernalization response. *Theor. Appl. Genet.* 125, 1697–1704. doi:10.1007/s00122-012-1946-z.

Zhang, J., Zhang, P., Dodds, P. and Lagudah, E. (2020a). How Target-Sequence Enrichment and Sequencing (TEnSeq) Pipelines Have Catalyzed Resistance Gene Cloning in the Wheat-Rust Pathosystem. *Front. Plant Sci.* 11, 1–13. doi:10.3389/fpls.2020.00678.

Zhang, N., Fan, X., Cui, F., Zhao, C., Zhang, W., Zhao, X., Yang, L., Pan, R., Chen, M., Han, J., Ji, J., Liu, D., Zhao, Z., Tong, Y., Zhang, A., Wang, T. and Li, J. (2017a). Characterization of the temporal and spatial expression of wheat (Triticum aestivum L.) plant height at the QTL level and their influence on yield-related traits. *Theor. Appl. Genet.* 130, 1235–1252. doi:10.1007/s00122-017-2884-6.

Zhang, S., Shen, J., Li, D. and Cheng, Y. (2020b). Strategies in the delivery of Cas9 ribonucleoprotein for CRISPR/Cas9 genome editing. *Theranostics* 11, 614–648. doi:10.7150/thno.47007.

Zhang, Y., Yu, C., Lin, J., Liu, J., Liu, B., Wang, J., Huang, A., Li, H. and Zhao, T. (2017b). OsMPH1 regulates plant height and improves grain yield in rice. *PLoS One* 12, 1–17. doi:10.1371/journal.pone.0180825.

Zhao, K., Xiao, J., Liu, Y., Chen, S., Yuan, C., Cao, A., You, F. M., Yang, D., An, S., Wang, H. and Wang, X. (2018). Rht23 (5Dq′) likely encodes a Q homeologue with pleiotropic effects on plant height and spike compactness. *Theor. Appl. Genet.* 131, 1825–1834. doi:10.1007/s00122-018-3115-5.

Zhao, Z., Duan, S., Hao, J., Cui, C., Yang, Y., Condon, A. G., Chen, F., Hu, Y. G. and Chen, L. (2021). The dwarf gene Rht15 improved lodging resistance but differentially affected agronomic and quality traits in durum wheat. *F. Crop. Res.* 263, 108058. doi:10.1016/j.fcr.2021.108058.

Zheleva, D., Todorovska, E., Jacquemin, J.-M., Atanassov, A., Christov, N., Panayotov, I. and Tsenov, N. (2006). Allele distribution at microsatellite locus xgwm 261 marking the dwarfing gene rht8 in hexaploid wheat from bulgarian and belgian gene bank collections and its application in breeding programs. *Biotechnol. Biotechnol. Equip.* 20, 45–56.

Zheng, J., Yang, Z., Madgwick, P. J., Carmo-Silva, E., Parry, M. A. J. and Hu, Y. G. (2015). TaER expression is associated with transpiration efficiency traits and yield in bread wheat. *PLoS One* 10, 1–17. doi:10.1371/journal.pone.0128415.

Zhong, C., Sun, S., Li, Y., Duan, C. and Zhu, Z. (2018). Next-generation sequencing to identify

candidate genes and develop diagnostic markers for a novel Phytophthora resistance gene, RpsHC18, in soybean. *Theor. Appl. Genet.* 131, 525–538. Available at: https://link.springer.com/content/pdf/10.1007%2Fs00122-017-3016-z.pdf.

Zhou, C., Xiong, H., Li, Y., Guo, H., Xie, Y., Zhao, L., Gu, J., Zhao, S., Ding, Y., Song, X. and Liu, L. (2020). Genetic analysis and QTL mapping of a novel reduced height gene in common wheat (Triticum aestivum L.). *J. Integr. Agric.* 19, 1721–1730. doi:10.1016/S2095-3119(20)63224-5.

Zhou, J., Reynolds, D., Websdale, D., Le Cornu, T., Gonzalez-Navarro, O., Lister, C., Orford, S., Laycock, S., Finlayson, G. and Stitt, T. (2017). CropQuant: An automated and scalable field phenotyping platform for crop monitoring and trait measurements to facilitate breeding and digital agriculture. *bioRxiv*, 161547.

Zhou, R., Zhu, Z., Kong, X., Huo, N., Tian, Q., Li, P., Jin, C., Dong, Y. and Jia, J. (2005). Development of wheat near-isogenic lines for powdery mildew resistance. *Theor. Appl. Genet.* 110, 640–648. doi:10.1007/s00122-004-1889-0.

Zhu, J., Fan, Y., Shabala, S., Li, C., Lv, C., Guo, B., Xu, R. and Zhou, M. (2020). Understanding mechanisms of salinity tolerance in barley by proteomic and biochemical analysis of near-isogenic lines. *Int. J. Mol. Sci.* 21, 1–13. doi:10.3390/ijms21041516.

Zikhali, M., Leverington-Waite, M., Fish, L., Simmonds, J., Orford, S., Wingen, L. U., Goram, R., Gosman, N., Bentley, A. and Griffiths, S. (2014). Validation of a 1DL earliness per se (eps) flowering QTL in bread wheat (Triticum aestivum). *Mol. Breed.* 34, 1023–1033.

Zikhali, M., Wingen, L. U. and Griffiths, S. (2015). Delimitation of the Earliness per se D1 (Eps-D1) flowering gene to a subtelomeric chromosomal deletion in bread wheat (Triticum aestivum). *J. Exp. Bot.* 67, 287–299. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4682435/pdf/erv458.pdf.

Ziyaev, Z. M., Sharma, R. C., Nazari, K., Morgounov, A. I., Amanov, A. A., Ziyadullaev, Z. F., Khalikulov, Z. I. and Alikulov, S. M. (2011). Improving wheat stripe rust resistance in Central Asia and the Caucasus. *Euphytica* 179, 197–207. doi:10.1007/s10681-010-0305-x.

Zong, W., Tang, N., Yang, J., Peng, L., Ma, S., Xu, Y., Li, G. and Xiong, L. (2016). Feedback regulation of ABA signaling and biosynthesis by a bZIP transcription factor targets drought-resistance-related genes. *Plant Physiol.* 171, 2810–2825. doi:10.1104/pp.16.00469.

Zubairov, O. Z. (2002). Amelioration condition of irrigated soils in Kyzylorda oblast. *Bastau*, 385–387.