University of Nottingham
BBSRC Wheat Research Centre

REPUBLIC OF TURKEY
MINISTRY OF AGRICULTURE
AND FORESTRY

Isolation and Characterisation of Wheat/*Thinopyrum elongatum* Introgressions

Sevgi HEREK BESLER

M.Sc.

Thesis submitted to the University of Nottingham for the

Degree of Doctor of Philosophy

The University of Nottingham

School of Biosciences

Sutton Bonington Campus

Loughborough, LE12 5RD

May 2022

WRC
Nottingham BBSRC Wheat Research Centre

Acknowledgements

First, I would like to express my gratitude to the Republic of Turkey Ministry of Agriculture and Forestry Scholarship programme for sponsoring my study and making this possible.

I would like to express my deepest appreciation to my supervisors, Prof Julie King, and Prof Ian King, for accepting me into the King group, which includes many successful exemplary scientists. Thanks to them for their valuable suggestions and guidance during my PhD.

Special thanks to Dr Surbhi Grewal for her advice and expertise and all the help with genotyping, and Dr Caiyun Yang for her cytogenetic expertise as well as her support. I am also grateful to Dr Stella Hubbart-Edwards, Dr Aleyda Sierra-Gonzalez, and Dr Manel Othmeni for their friendly support and advice during my PhD. Many thanks to Stephen Ashling and Duncan Scholefield for their great help with the technical support.

I am always grateful to Associate Prof Ziya Dumlupınar for encouraging me on this journey and believing in me to the end. His friendly support and invaluable advice always helped me to draw my route.

I am extremely grateful to my spouse, Ibrahim Besler, and my parents, for their love and support throughout this journey.

Lastly, but the most I want to thank my little daughter Nil, who came like sunshine to our family in the darkest time of the Covid pandemic. I get all my power from your love.

Abstract

The world population is predicted to increase to approximately 10 billion by 2050 and due to this increase 50% more food will need to be produced. Hexaploid bread wheat (*Triticum aestivum*, AABBDD, 2n=6x=42) is a major food source as cereals play an important role in the human diet. Wheat breeders have had considerable success in increasing production during the last century but now yields are starting to plateau. Bread wheat evolved nearly 10,000 years ago but domestication has reduced its genetic diversity making wheat vulnerable to everchanging abiotic and biotic stresses. To widen wheat's gene pool, new alleles from landraces and wild relatives of wheat have been transferred through wide-crossing programmes.

Thinopyrum species have been shown to be tolerant to many biotic and abiotic stresses such as drought and salt tolerance and leaf rust, yellow rust and Fusarium head blight resistance. As one of the Thinopyrum species, wheatgrass, *Thinopyrum elongatum* ($E^eE^eE^bE^bE^xE^x$StStStSt, 2n=10x=70), is an excellent reservoir for genetic diversity and can be used for the development of new wheat varieties.

116 wheat/*Th. elongatum* introgression lines were taken forward in this work with 191 crosses producing 1637 seeds and self-fertilisation of 801 of the backcrossed lines generating an estimated 60,441 seeds. Each subsequent generation showed higher levels of fertility as the wheat genome was restored.

In this thesis, the production of the wheat-*Th. elongatum* introgression lines, containing segments of *Th. elongatum* of varying sizes, and their characterisation through Kompetitive Allele Specific PCR (KASP^{TM}) markers is described. These lines were simultaneously examined with multi-colour and single-colour genomic *in situ* hybridisation (GISH) to determine the genomic constitution of the *Th. elongatum* chromosome(s) present. The KASP markers detected 530 introgressions across all seven linkage groups. In the majority of lines, the genotyping results were validated in the mc-GISH analysis. Genotyping with KASP markers allowed the rapid identification of introgressed segments and also the distinction between heterozygous and homozygous introgressions in the segregating populations.

Twenty homozygous introgression lines were identified with the molecular characterisation, but more markers are required for the identification of small segments in those chromosome regions not as yet covered by the markers.

Table of Contents

Chapter 1 Introduction 2

1.1 Wheat Production 2

1.2 The Evolution of Wheat 3

1.3 Food Security 6

1.4 Alien Introgression 8

1.5 *Ph1* Locus 10

1.6 Increasing Genetic Diversity with Pre-breeding 12

1.7 Molecular Markers: an Introduction and Their Applications 14

1.8 Genomic *in Situ* Hybridization Technique (GISH) 19

1.9 *Thinopyrum elongatum* 20

1.10 Thesis Outline and Objectives 22

Chapter 2 General Materials and Methods 23

2.1 DNA extraction 23

2.2 Genomic *in situ* hybridization 24

2.2.1 Root tips and nitrous oxide treatment 24

2.2.2 Making Probes 24

2.2.2.1 Nick translation reaction 25

2.2.2.2 Ethanol Precipitation of the probe 25

2.2.2.3 Preparation of blocking genomic DNA 25

2.2.3 Multi-colour GISH 26

Chapter 3 Development of Germplasm 27

3.1 Introduction 27

3.2 Aims and Objectives of this chapter 30

3.3 Materials and Methods 31

3.3.1 Crossing programme 31

3.3.2 Plant material 34

3.3.3 Germination 34

3.3.4 Vernalisation and potting 35

3.3.5 Crossing Method 35

3.3.5.1 Emasculation 35

3.3.5.2 Pollination 36

3.3.6 Harvesting and threshing 37

3.3.7 Selection of seeds 37

3.3.8 Observations 38

3.4 Results 39

3.5 Discussion 41

Chapter 4 Generation of introgression lines and their characterisation with KASP markers and GISH (single, multi (A, B, D, E)) 43

4.1 Introduction 43

4.1.1 Genome Constitution of the Thinopyrum Genus 43

4.1.2 Previous Work using Molecular Markers for *Th. elongatum* 46

4.1.3 SNP Discovery and KASP marker development at the WRC 49

4.1.4 KASP validation and characterization 50

4.2 Aims and Objectives 52

4.3 Material and method 53

4.3.1 Development of KASP markers 53

4.3.1.1 Selection of markers 53

4.3.1.2 Genotyping analysis via a KASP Marker 53

4.3.2 Cytogenetic analysis of the Wheat/*Th. elongatum* introgression lines 55

4.4 Result 56

4.4.1 Development of KASP markers 56

4.4.2 Genotyping of wheat-*Th. elongatum* introgression lines with KASP markers 56

4.4.3 Combining the KASP marker analysis with GISH analysis of wheat-*Th. elongatum* introgression lines 61

4.5 Discussion 72

Chapter 5 Cytogenetic and Molecular Characterization of wheat-*Th. elongatum* homozygous lines 79

5.1 Introduction 79

5.2 Aim and Objectives 81

5.3 Materials and Methods 82

5.3.1 Plant material and Crossing programme 82

5.3.2 Genotyping Analysis of Homozygous Lines via KASP 82

5.3.3 Cytogenetic Analysis of the Homozygous Lines with GISH 83

5.4 Results 84

5.5 Discussion 102

Chapter 6 General Discussion 106

Future Work 111

References 112

Appendices 125

Appendix 1 126

Appendix 2 129

List of Figures

Figure 1. 1: The pie chart representation of top 10 wheat producer in the world (Source; http://faostat.fao.org) 3

Figure 1. 2: Near East map illustrating the Fertile Crescent (Israel, Jordan, Turkey, Syria, Iran, Iraq) (Source; Feuillet et al., 2008). 4

Figure 1. 3: Diagrammatic representation of the evolution of bread wheat *T. aestivum*. After each hybridisation event, subsequent chromosome doubling occurred. 5

Figure 1. 4: The figure shows that the world production and area harvested of wheat from 2000 to 2020 inclusive (Source; http://faostat.fao.org) 7

Figure 1. 5: Diagram representation of the hexaploid wheat genomes. The red box homologous chromosomes, while the blue box homoeologous chromosomes (Moore, 2014). 10

Figure 1. 6: Wheat gene pools representing primary, secondary, and tertiary gene pools with example species in each 14

Figure 1. 7: Identification of SNPs between two DNA sequences (Source; (Tiwari *et al.*, 2014). 17

Figure 3. 1: Development of F1 interspecific hybrid and wheat/*Th. elongatum* introgressions. The green, blue, and red chromosomes represent the A, B, and D genomes of wheat respectively and the yellow chromosomes represent the E genome of *Th. elongatum*. This crossing programme was used by Dr Lauren Baker. 32

Figure 3. 2: Crossing diagram to generate single segment and homozygous wheat/*Th. elongatum* introgression lines (used in this work). The green, blue, and red chromosomes represent the A, B, and D genomes of wheat respectively and the yellow chromosomes represent the E genome of *Th. elongatum*. 33

Figure 3. 3: An example of back-crossed lines germination and sowing into modular trays. 34

Figure 3. 4: Transferring backcross lines from the vernalization to the glasshouse as an example. 35

Figure 3. 5: The left side photo has green anthers collected with forceps. Each floret must contain three anthers, so after emasculation, it was checked that there were no anthers left in the spikelet and the removed anthers were counted. An example of a glassine bag to avoid unwanted cross-pollination in the right-side photo. 36

Figure 3. 6: The image showed that the mature stigma is 'fluffy'/feathered and fanned. 37

Figure 4. 1: Mc-GISH of a metaphase spread of *Th. elongatum* accession 401007 probed with total genomic DNAs from *Th. bessarabicum* (green) and *P. strigosa* (red) (from Baker et al., 2020)……46

Figure 4. 2: An example of a homozygous introgression. In this line the introgression from *Th. elongatum* has replaced both short arms of wheat chromosome 1A. Green represents homozygous, red represents heterozygous……51

Figure 4. 3: The distribution of the 267 KASP markers across the 21 wheat chromosomes (Chr1A-7D). Figure created using RIdeogram_0.2.2. The 189 working KASP markers are shown with purple circles. The 78 KASP markers which did not work are shown with red triangles……58

Figure 4. 4: Molecular characterization of wheat–*Th. elongatum* introgression lines showing genotyping data (left), using KASP markers, and sc-GISH and mc-GISH analysis (right). The A-genome chromosomes are shown in green, the B-genome chromosomes in blue, the D-genome chromosomes in red and E-genome chromosome/introgressions in gold. Whole introgressed chromosomes are labelled using a yellow arrow while segments are labelled with a white arrow. All heterozygous calls are shown in red while wheat calls are shown in blue and where no KASP markers were available at the time of genotyping is shown in white……63

Figure 4. 5*:* Molecular analysis (KASP genotyping and GISH) for (A) a BC_5 plant and (B) one of its subsequent BC_5F_2 offspring showing the loss of a whole *Th. elongatum* chromosome (yellow arrow) in BC_5F_2-10-8 as compared to BC5-74C. (Sc-GISH for BC_5-74E from Dr Lauren Baker). Both plants carry the centric fusion 3ES.3AL (white arrows). In (B), the A-genome is shown in green, the B-genome in blue, the D-genome in red and the E-genome in gold. In the coloured bars, all heterozygous calls are shown in red, while all wheat calls are shown in blue, and white spaces indicate where there are no markers……64

Figure 4. 6: Sc-GISH and mc-GISH images of metaphase spreads and KASP genotyping data of 8 wheat-*Th. elongatum* introgression lines. (A) BC_4-118E (image from Dr Lauren Baker) and its self-fertilised progeny BC_4F_1-88-1 (no GISH is available for BC_4F_1-88-1). (B) Three of the BC5 progeny produced by crossing BC_4F_1-88-1 (C) Three of the BC_4F_2 progeny produced by self-fertilising BC_4F_1-88-1. In the mc-GISH images, the A-genome is shown in green, the B-genome in blue, the D-genome in red and the E-genome in

gold while whole *Th. elongatum* chromosomes are shown with a yellow arrow and wheat- *Th. elongatum* introgressions with a white arrow. The coloured bars represent genotyping result, wheat is shown in blue, heterozygous is shown blue and white spaces indicate where there are no markers……66

Figure 4. 7: KASP genotyping and sc-GISH/mc-GISH analyses of 3 generations of wheat–*Th. elongatum* introgression lines. (A) BC_5-85A, (B) BC_5F_1-12-1, (C) BC_5F_2-7-3 and (D) BC_5F_2-7-4. The A-genome is shown in green, the B-genome in blue, the D-genome in red and the E-genome in gold. Whole *Th. elongatum* chromosomes are shown with a yellow arrow and wheat- *Th. elongatum* introgressions with a white arrow. The coloured bars represent genotyping result. All heterozygous calls are shown in red, all wheat calls are shown in blue, and white spaces indicate where there are no markers……68

Figure 4. 8: Characterization of wheat-*Th. elongatum* introgression segments using mc-GISH and KASP markers from seven linkage groups. In all segments, two images are shown; the first one is merged colour (the yellow represents *Th. elongatum*) and the second one is Red, Green, and Blue (RGB) colour (the purple represents *Th. elongatum*). In both images, green represents the A genome, grey/blue represents the B genome, and red represents the D genome of wheat. In the genotyping result, black represents the wheat allele and green represents a heterozygous call for the *Th. elongatum* allele……70

Figure 5. 1: Molecular characterization of wheat-*Th. elongatum* homozygous lines showing genotyping analysis (left), using chromosome-specific KASP markers, and root metaphase spreads (right), using mc-GISH. Homozygous calls are shown in green heterozygous calls are shown in red, while all wheat calls are shown in blue. **(a)** BC_4F_3-51, **(b)** BC_4F_2-121, **(c)** BC_5F_2-17, and **(d)** BC_2F_3-111. Arrows indicate recombinant chromosomes…… 87

Figure 5. 2: Molecular characterization of wheat-*Th. elongatum* addition homozygous lines showing genotyping analysis (left), using chromosome-specific KASP markers, and root metaphase spreads (right), using mc-GISH. Heterozygous calls are shown in red, while all wheat calls are shown in blue. **(a)** BC_4F_3-47, **(b)** BC_4F_3-44, and **(c)** BC_5F_2-18. Whole chromosomes of *Th. elongatum* are indicated by arrows…… 89

Figure 5. 3: Molecular characterization of wheat-*Th. elongatum* homozygous lines showing genotyping analysis (left), using chromosome-specific KASP markers, and root

metaphase spreads (right), using mc-GISH. Homozygous calls are shown in green, heterozygous calls are shown in red, and all wheat calls are shown in blue. White arrows indicate recombinant chromosomes while orange arrows indicate whole/partly chromosome of E. **(a)** BC_3F_2-156-3, and **(b)** BC_3F_3-56…….. 91

Figure 5. 4: Molecular characterization of wheat-*Th. elongatum* homozygous line (BC_4F_3-46) with genotyping analysis (left), using chromosome-specific KASP markers, and root metaphase spread (right), using mc-GISH. Homozygous calls are shown in green, heterozygous calls are shown in red, and all wheat calls are shown in blue. White arrows indicate recombinant chromosomes……92

Figure 5. 5: Characterization of wheat-*Th. elongatum* homozygous line (BC_4F_4-7) using KASP markers assay (left), and their Identification through mc-GISH (right). Homozygous calls are shown in green, heterozygous calls are shown in red, and all wheat calls are shown in blue. White arrows indicate recombinant chromosomes…… 93

Figure 5. 6: On the right, mc-GISH photos of wheat-*Th. elongatum* homozygous introgression lines showed root metaphase spreads, whereas on the left, KASP markers were used to genotype them. Heterozygous calls are shown in red, while all wheat calls are shown in blue. **(a)** BC_5F_3-6 **(b)** BC_4F_3-48 and, **(c)** BC_5F_4-1. Recombinant chromosomes are indicated by white arrows……94

Figure 5. 7: Molecular characterization of wheat-*Th. elongatum* homozygous lines showing genotyping analysis (left), using chromosome-specific KASP markers, and root metaphase spreads (right), using mc-GISH. Heterozygous calls are shown in red, while all wheat calls are shown in blue. **(a)** BC_4F_2-89-2, and **(b)** BC_4F_2-89-5. Recombinant chromosomes were indicated by white while whole chromosome of *Th. elongatum* was indicated orange arrows…….. 95

Figure 5. 8: Molecular characterization of wheat-*Th. elongatum* homozygous introgression lines (BC_4F_3-43) using KASP markers (left) and mc-GISH (right). Homozygous calls are shown in green, heterozygous calls are shown in red, and all wheat calls are shown in blue. White arrows indicate recombinant chromosomes……96

Figure 5. 9: Molecular characterization of wheat-*Th. elongatum* homozygous lines showing genotyping analysis (left), using chromosome-specific KASP markers, and root

metaphase spreads (right), using mc-GISH. **(a)** BC_4F_3-49 and **(b)** BC_5F_3-7. White arrows indicate recombinant chromosomes…….. 97

Figure 5. 10: Characterization of wheat-*Th. elongatum* homozygous line (BC_3F_2-158-5) showing KASP markers assay (left), and their Identification through mc-GISH (right). Homozygous calls are shown in green, heterozygous calls are shown in red, and all wheat calls are shown in blue. White arrows indicate recombinant chromosomes…… 98

Figure 5. 11: mc-GISH analysis of wheat-*Th. elongatum* homozygous lines showing the different J and J^{St} segments present. **(a)** BC_2F_3-111, **(b)** BC_4F_2-121, **(c)** BC_3F_3-56, **(d)** BC_4F_3-51**(e)** BC_5F_2-17, **(f)** BC_4F_3-44, **(g)** BC_4F_3-46, **(h)** BC_4F_4-7, **(i)** BC_5F_3-6, **(j)** BC_4F_3-48, **(k)** BC_5F_2-18, **(l)** BC_4F_3-43, **(m)** BC_5F_4-1, **(n)** BC_4F_3-49, **(o)** BC_5F_3-7, **(p)** BC_3F_2-155, and **(r)** BC_6-7-3…… 100

Figure 5. 12: Spikes morphology; **a)** Paragon (*T. aestivum)*, **b)** BC_4F_2-121, **c)** BC_4F_3-46, **d)** BC_4F_3-51, **e)** BC_5F_3-6, **f)** BC_5F_4-1, and **g)** BC_5F_2-17……..101

Figure 5. 13: Spikes morphology; **a)** Paragon (*T. aestivum*), **b)** BC_4F_3-47, **c)** BC_4F_3-49, **d)** BC_4F_3-48, **e)** BC_4F_3-50, **f)** BC_5F_3-7, and **g)** BC_4F_3-44…… 101

List of Table

Table 1. 1: The world production of cereal crops in 2020 (Source; http://faostat.fao.org) 7

Table 1. 2: Comparison of characteristics of most widely used DNA marker systems in plants. (Source; Nadeem et al., 2017). 18

Table 3. 1: Number of seeds germinated and produced in relation to the number of crosses carried out and the cross fertility and the number of self-fertilized seed produced for each generation of the introgression programme for *Th. elongatum* into wheat. The number of self-fertilised seeds represents an estimated seed number. . 39

Table 4. 1 : A brief summary of work looking at the close relationship of the J and E genomes. 44

Table 4. 2: The number of KASP markers tested, and the number of KASP markers that worked to detect *Th. elongatum* introgressions for each wheat chromosome..... 57

Table 4. 3: The number of plants of each generation genotyped and GISHed and the number of plants with one or more introgressions from *Th. elongatum*. 60

Table 5. 1: List of wheat-*Th. elongatum* homozygous introgression lines obtained in this chapter, the number of the genomes and the total number of chromosomes (2n) in each line is homozygous for just one introgression. Lines marked with an asterisk (*) are sister lines from same BC1 parent with same genotype. 85

Table 5. 2: The table shows one diagnostic marker per segment. Lines marked with an asterisk (*) are addition lines and therefore markers for all wheat genomes show heterozygous so one marker is shown per genome. 99

Table 1: KASP assay primer details for all selected SNPs. 129

Table 2: BLASTN and BLASTX results for all chromosome-specific SNPs. 140

Table 3: PCR layout for the first genotyping analysis in September 2018, and 84 KASP markers were used for this set. 151

Table 4: PCR layout for the second genotyping analysis in February 2019, and 203 KASP markers were used. 153

Table 5: PCR layout for the Set 1 genotyping analysis in May 2019, and 235 KASP markers were used for Set 1. 154

Table 6: PCR layout for the Set 2 genotyping analysis in June 2019, and 235 KASP markers were used for Set 2. 155

Table 7: PCR layout for the Set 3 genotyping analysis in July 2019, and 235 KASP markers were used for Set 3. 156

Table 8: PCR layout for the Set 4 genotyping analysis in October 2019, and 198 KASP markers were used for Set 4.....157
Table 9: PCR layout for the Set 5 genotyping analysis in November 2019, and 198 KASP markers were used for Set 5.....158
Table 10: PCR layout for the Set 6 (homozygous lines) genotyping analysis in November 2020, and 198 KASP markers were used for Set 6.....159
Table 11: The table shows the list of markers used for the first genotyping analysis in September 2018. 160
Table 12: The table shows the list of markers used for the second genotyping analysis in February 2019. 161
Table 13: The table shows the list of markers used for the Set 1, Set 2, and Set 3 genotyping analyses. 163
Table 14: The table shows the list of markers used for the Set 4, Set 5, and Set 6 genotyping analyses. 166

Abbreviations

List of Abbreviations	Description
Ae.	*Aegilops*
Acc.	Accession
AFLP	Amplified Fragment Length Polymorphism
Am.	*Amblyopyrum*
BBSRC	Biotechnology and Biological Sciences Research Council
Bp	Base pair
BYDV	Barley Yellow Dwarf Virus
cM	Centimorgan
COS	Conserved orthologous set
Cs	Chinese Spring
Cv	Cultivar
DArT	Diversity Array Technology
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribo-nucleotide triphosphates
EDTA	Ethylene-diamine-tetraacetic acid
FAO	Food and Agriculture Organisation of the United Nations
FHB	Fusarium Head Blight
FISH	Fluorescent *in situ* hybridisation
GISH	Genomic *in situ* hybridisation
KASP	Kompetitive allele-specific PCR
MAS	Marker assisted selection
mc-GISH	Multi-colour Genomic *in situ* hybridisation
PCR	Polymerase chain reaction
Ph1	Pairing homoeologous 1
PLUG	PCR-based landmark unique genes
RAPD	Random Amplification of Polymorphic DNA
RFLP	Restriction fragment length polymorphism
sc-GISH	Single-colour Genomic *in situ* hybridisation
SCAR	Sequence characterized amplified region
SLAF-seq	Specific length amplified fragment sequencing
SNP	Single nucleotide polymorphism
SSR	Single Sequence Repeat
Tris-HCl	Tris hydrochloride
Mg	Microgram
Ml	Microlitre

Chapter 1 Introduction

1.1 Wheat Production

The tribe Triticeae of wheat belongs to Gramineae, the Poaceae family. This tribe of the family, which gave rise to the domesticated cereals wheat, barley, and rye, and several important kinds of grass, is the economically most significant tribe (Feldman and Levy, 2015). Cereal crops provide around 60 % of human calories (FAO, 2007), while wheat is the largest cereal provider 19 % of the food calories and 20 % of the protein consumed by the world population (Braun *et al*., 2010). Wheat is counted mainly a source of energy and it also includes an important amount of nutrients such as proteins, minerals, vitamins, lipids and phytochemicals which might provide a healthy diet (Shewry and Hey, 2015).

Cereals that contribute 50% of the total crop production are a significant source for food, feed, and industrial materials. After the production of maize, wheat is the second most important cereal crop with a harvest of over 760 million tonnes in 2020 (Table 1.1) (http://faostat.fao.org). Bread wheat (*T. aestivum* ssp. Aestivum L.) which makes up about 95% of total wheat is used to make bread, cakes, cookies, pastries, crackers, and noodles, while durum wheat (*T. turgidum ssp. durum L.*) makes up about 5% and is used to make pasta and semolina (Faris, 2014).

Wheat is cultivated on the largest scale from latitudes of 30° and 60°N to 27° and 40°S, mostly in East and South Asia, North America, and Eastern Europe. It can be cultivated in a wide environment partly because several different types of wheat are grown and most of these species have winter, facultative and spring species (Bedő and Láng, 2015). Currently, the available pie chart shows that the countries producing wheat on the largest producer are

China, followed by India, Russia, USA, and France (Figure 1.1) (http://faostat.fao.org).

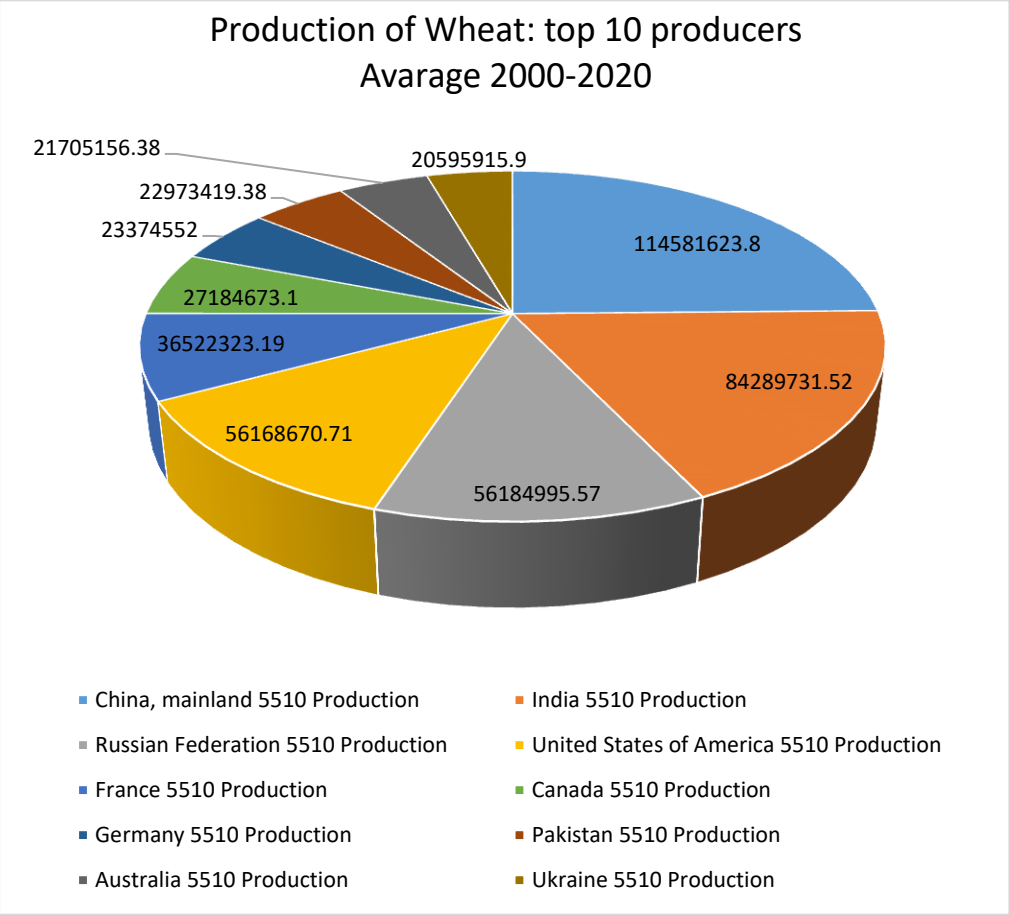


Figure 1. 1: The pie chart representation of top 10 wheat producer in the world (Source; http://faostat.fao.org)

1.2 The Evolution of Wheat

Before the Neolithic Revolution, people lived by hunting animals and gathering plants, but after the revolution, this was replaced by an agricultural lifestyle. The inception of the Neolithic Revolution occurred in the Levantine Corridor and spread throughout the Fertile Crescent of the Middle East (Figure 1.2). This area was also the birthplace of *Triticum monococcum L.* (einkorn wheat), and *Triticum turgidum ssp. dicoccum L.* (emmer wheat), that were the immediate predecessor of modern wheat (Faris, 2014).

Diploid einkorn and tetraploid emmer wheat originated 10,000 years ago in the southeast Turkey and were probably widely cultivated in the region (Feuillet *et al.*, 2008; Shewry, 2009).

Hexaploid common wheat (*Triticum aestivum*, AABBDD, 2n=6x=42) is composed of three different genomes. The A genome was derived from *Triticum urartu* (2n = 2x = 14, genomic code AA), while the D genome was derived from *Aegilops tauschii* (2n=2x=14, genomic code DD) (Shewry, 2009; Feldman and Levy, 2015), and a more recent study showed that *Aegilops speltoides* (2n=2x= 14, genomic code SS) is likely to be the direct donor of chromosomes of the wheat B genome (Miki *et al*., 2019).



Figure 1. 2: Near East map illustrating the Fertile Crescent (Israel, Jordan, Turkey, Syria, Iran, Iraq) (Source; Feuillet *et al*., 2008).

The evolution and formation of modern-day bread wheat was the consequence of two evolutionary events that involved two amphiploidization events demonstrated in Figure 1.3. After the hybridization of two different species followed by spontaneous chromosome doubling of the F1 hybrid (Faris, 2014).

The first evolutionary event occurred at 500,000-150,000 years before present (BP) between *T. urartu* (2n=2x=14, genomic code AA), and an *Ae. speltoides* like, from the *Sitopsis* section (Shewry, 2009; Feldman, 2001; Feuillet *et al.,* 2008) (Figure 1.3). This was followed by the second evolutionary event was the creation of hexaploid wheat which occurred about 10,000 BP, as a result of the hybridization of tetraploid wheat (2n=4x=28, genomic code AABB), which had immigrated to the south of the Caspian Sea, and to the Fertile

Crescent, and a diploid wild species, *Ae. tauschii*, (2n=2x=14, genomic code DD) (Feuillet *et al.,* 2008). This hybridization presumably occurred a few times independently with the new hexaploid wheat being chosen by farmers for its superior traits (Shewry, 2009).

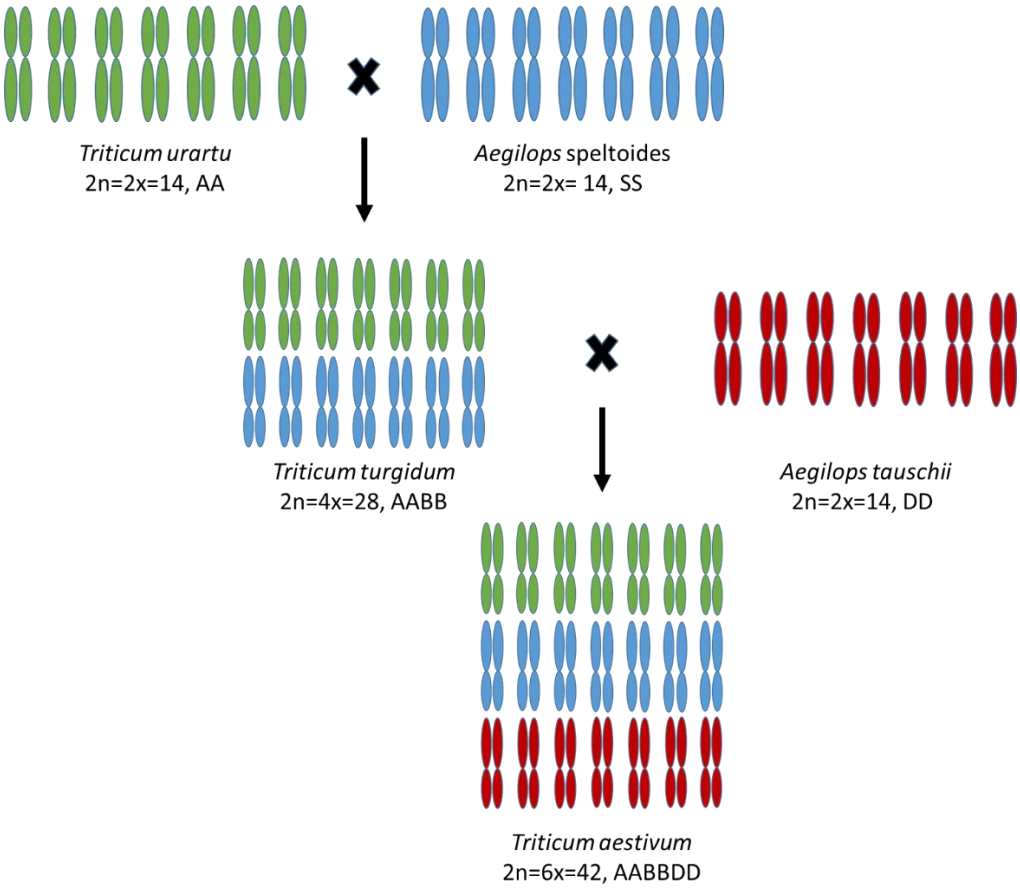


Figure 1. 3: Diagrammatic representation of the evolution of bread wheat *T. aestivum*. After each hybridisation event, subsequent chromosome doubling occurred.

The domestication of wheat produced the genetic bottlenecks, and it resulted in an increase in adapted alleles whereas potential alleles were lost. Also, this limited variety reduces the potential of wheat to adapt more to changing environments (Faris, 2014). Without sufficient alleles to develop superior adapted varieties are not possible for plant breeders. On the other hand, wild relatives are an important source for the gene pool.

The wheat groups of the genus *Amblyopyrum*, *Aegilops*, and *Triticum* are members of the grass family Gramineae or Poaceae. Also, they are close

relatives of certain weeds, such as *Agropyron*, and different wild relatives which be able to be hybridized with wheat, such as *Leymus*, and *Thinopyrum*. Most of the wild *Triticeae* species grow in the Near East, south-western Asia, and the eastern Mediterranean, while other species, such as *Australopyrum spp.* can also be found in Australia and some, such as *Aegilops spp.*, can be found in the Maghreb region, and western Mediterranean Europe (Bozzini *et al.*, 2012).

1.3 Food Security

Wheat is a cereal crop widely grown around the world for human consumption. Today, bread or common wheat (*T. aestivum ssp. Aestivum L.*) and durum wheat (*T. turgidum ssp. durum L.*) are the main forms produced, providing approximately a ninety of the calories consumed worldwide.

Currently, the available table shows that wheat is produced at 760 million tonnes in 2020 all over the world (Table 1.1). This makes it the second most important crop in terms of global production, the comparative values for the production of the other major crop that is maize over the same period (Table 1) (http://faostat.fao.org). However, maize is used more for animal feed and, in recent years, fuel production (Curtis and Halford, 2014).

Table 1. 1: The world production of cereal crops in 2020 (Source; http://faostat.fao.org)

Cereal Crops	Area harvested (ha)	Yield (hg/ha)	Production (tonnes)
Wheat	219,006,893	34,744	760,925,831
Maize	201,983,645	57547	1,162,352,997
Rice, paddy	164,192,164	46089	756,743,722
Barley	51,601,372	30432	157,030,764
Rye	4,446,927	33,781	15,022,273
Triticale	3,812,724	40,290	15,361,341
Oats	9,772,003	25,769	25,181,805

Over the last 20 years, global production of wheat has increased gradually despite the fact that the area harvested has slightly declined at the same time (Figure 1.4) (http://faostat.fao.org). However, wheat yields do not currently increase at rates comparable to those achieved in previous decades.

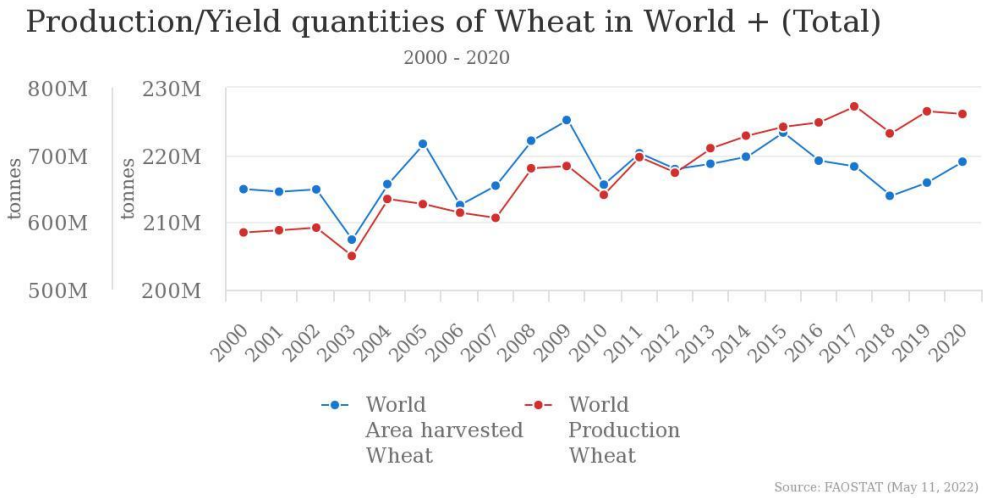


Figure 1. 4: The figure shows that the world production and area harvested of wheat from 2000 to 2020 inclusive (Source; http://faostat.fao.org)

The Food and Agriculture Organization of the United Nations has predicted that 50% more food will need to be produced by 2050, due to the anticipated increase in global population to approximately 10 billion (FAO, 2017).

As a result, the annual yield of wheat must increase by 2% by 2050, in order to ensure global food security. Besides improving farming techniques, and reducing postharvest losses, new varieties of wheat should be developed that have a high yield and are disease-resistant, as well as being able to contend with adverse environmental conditions (Rey *et al*., 2015). However, increases in production will not be easy due to lack of genetic diversity, climate change and limited land for agriculture. On the other hand, wheat breeders have had considerable successes, but new superior wheats should be created to provide food demand for the increasing population. At this point, alien introgression programs into wheat can be employed to recover the genetic diversity of wheat.

1.4 Alien Introgression

Alien introgression involves hybridising the wheat chromosomes with those of its wild relatives to expand the gene pool (Tiwari *et al*., 2014); this is an important option when aiming to recover the genetic diversity of wheat (Rey *et al*., 2015). Hybridization of wheat with its wild relatives or cultivated species has been undertaken since breeding first began (Bedő and Láng, 2015). For instance, Wilson (1876) hybridized wheat and rye, while Farrer 1904, crossed wheat and barley. Although wheat breeders have made positive developments using genetic variants of wheat in the past, future applications might include improving wheat yields via introgression of genes to enhance global food security.

Alien introgression method used to transfer genes from involved species to wheat largely depends on the domestication distance between the related species. *Triticeae* tribe contains three gene pools with an annual and perennial wild relative of wheat to produce interspecific hybrids. A piece of more in-

depth information is included in the Increasing Genetic Diversity with Pre-breeding section (see section 1.5).

The primary gene pool of wheat shares homologous genome with species belonging to this gene pool (Bedő and Láng, 2015). Transferring genes to wheat can be possible via direct hybridization.

The secondary gene pool has at least one homologous genome in wheat and also gene transfer can be achieved by homologous recombination.

The tertiary gene pool does not have a homologous genome in wheat, species belonging to this pool are more distantly related. In general, F1 plants with produced via interspecific crosses are sterile and colchicine is usually used to double chromosome for producing fertile amphidiploids (Bedő and Láng, 2015).

Alien introgression is an option for the improvement of the useful diversity that was lost and to obtain new and beneficial genes and alleles are to use wild relatives of wheat that are not subject to human selection and thus demonstrate a rich resource of diversity (Rey *et al*., 2015).

Indeed, Ceoloni *et al*., (2015) reported that wheat breeding using alien introgression represents an important avenue for progress, as it is possible to develop new lines of wheat for a variety of purposes. However, the presence of the *Pairing homoeologous 1* locus (*Ph1*) hinders introgression from the wild relative of wheat. Hybridization between a wild relative of wheat and wheat are suppressed the pairing by the *Ph1* thus recombination of wheat and alien chromosomes cannot occur (Bedő and Láng, 2015). Therefore, special techniques need to be used for successful crosses for the use of these species in alien introgression programmes. However, *ph1b* mutant having a major deletion at the *Ph1* locus, recombination of homoeologous wheat and alien chromosomes can pair (Bedő and Láng, 2015).

The aim of breeding is to transfer desirable genes from one species to alternative varieties. King *et al.* (2013) explained that alien introgression, which requires the transference of genetic diversity from the distant and near relatives of plants offers a valuable and largely untouched reserve of genetic

variation, to introduce agriculturally significant features that could enable wheat breeders to create better-adapted crops with higher yields.

Alien introgression has increased the importance of pre-breeding for the growing requirement of new genetic variations (Bedő and Láng, 2015). Pre-breeding offers a great opportunity to improve genetic diversity in wheat.

1.5 *Ph1* Locus

Common wheat consists of three genomes, A, B and D, derived from three different diploid species and therefore, bread wheat contains both homologous and homoeologous chromosomes (see Figure 1.5). The chromosomes within a homoeologous group, e.g., chromosomes 1A, 1B and 1D have a similar gene order. However, due to the *Ph1* locus, homoeologous chromosomes cannot pair each chromosome can only pair with its homologue, i.e., 1A can pair with 1A at meiosis, but not with 1B or 1D (Moore, 2002; Yousafzai *et al.*, 2010).

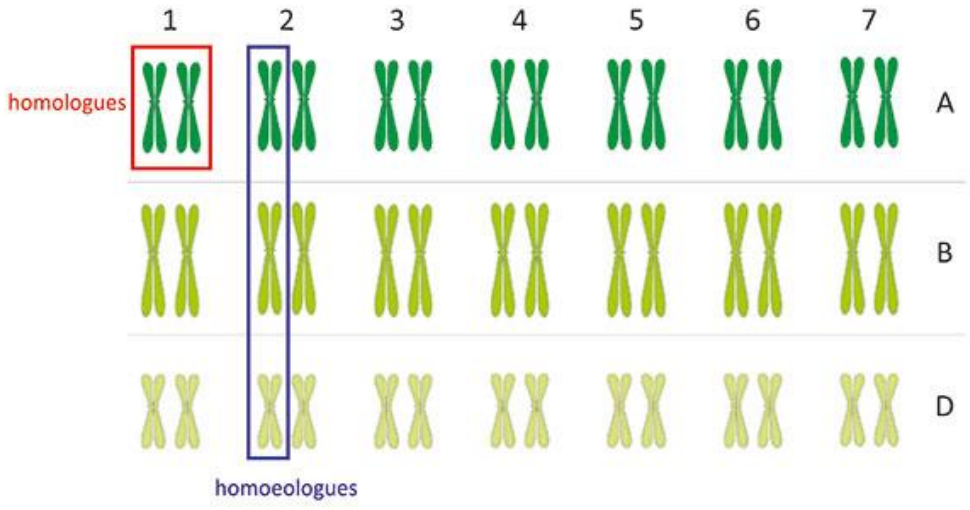


Figure 1. 5: Diagram representation of the hexaploid wheat genomes. The red box homologous chromosomes, while the blue box homoeologous chromosomes (Moore, 2014).

The *Ph1* locus is located on the long arm of chromosome 5B (5BL) (Riley and Chapman, 1958, Sears and Okamoto, 1958). A more recent study localised the *Ph1* locus to a 2.5-megabase (Mb) region on wheat chromosome 5B containing a cluster of *cdc2* (*cdk*)-related genes (Griffiths *et al.,* 2006). The *Ph1*

locus must have arisen during polyploidization because the diploid relatives of wheat do not show *Ph1* activity (Chapman and Riley, 1970).

The many distant relatives of wheat offer a rich source for developing wheat's genetic diversity (Mujeeb-Kazi *et* al., 2008). However, *Ph1* is a problem because this locus prevents recombination between homoeologous chromosomes including between wheat/alien chromosomes (Yousafzai *et al.*, 2010). Martin *et al*. (2014) proposed that *Ph1* has a dual effect during meiosis. The effect occurs place early in meiosis, when *Ph1* promotes homologous synapsis, rather than suppresses homoeologous synapsis. The effect on synapsis occurs during the telomere bouquet stage, when Ph1 promotes more efficient homologous synapsis, whereas homologous pairing begins after the telomere bouquet dispersal (Martin *et al*. 2017). In the absence of the *Ph1* locus, homologous synapsis is delayed in relation to the telomere bouquet, resulting in more synapsis. Thus, in wheat, the *Ph1* locus encourages early homologous pairing while avoiding homoeologous pairing.

One solution employed to overcome *Ph1* is to use lines where the locus has been deleted, e.g., the Chinese Spring *ph1b* mutant has a section of chromosome 5B, including the *Ph1* locus, deleted. This was first generated by Sears (1977) using X-ray irradiation on the pollen of Chinese Spring. The *ph1b* deletion is located near the middle of chromosome 5B, on the long arm about 1.0 cM from the centromere (Sears, 1984). For the last 40 years, a single *Ph1* deletion (*ph1b*) mutant in CS cultivar has been used for wheat/alien crosses. However, a recent study using new chemically induced mutant lines selected for a mutation in *TaZIP4-B2* within the *Ph1* locus which when crossed with wild relatives display high levels homoeologous crossovers (Rey *et al.,* 2017).

Nottingham BBSRC Wheat Research Centre used the Chinese Spring *ph1b* deletion to make the initial crosses between *Thinopyrum elongatum* and wheat to enable homoeologous recombination to occur between the chromosomes of *Th. elongatum* and the chromosomes of wheat.

1.6 Increasing Genetic Diversity with Pre-breeding

Common wheat (*T. aestivum*) is one of the cereals that constitutes a primary food source for human diets, and it, therefore, occupies a significant place in the world. As explained in the previous section the evolution of wheat has two hybridization events and evolved only once or twice times almost 10,000 years ago in the Fertile Crescent consequently, the domestication of wheat has reduced its genetic diversity and resulted in a narrow gene pool, which hinders the improvement of new varieties.

Plant breeding represents a way for the selection, creation, and fixation of desirable plants in the improvement of developed cultivars to needs of consumers and farmers. The main aim of plant breeding is to improve yields, quality of nutrition and other traits of commercial importance (Moose and Mumm, 2008). Using the entire existing genetic diversity can help develop new wheat varieties that will be needed, but commercial varieties currently capture only 10% of this resource (Moore, 2015).

Today, wheat varieties are higher in quality, yield, biotic and abiotic resistance, and other agronomic performances than those that were considered elites a couple of decades ago. However, if the growing population of the world will be fed, deeper progress should be made in these and other features in the near future (Faris, 2014).

Breeders have worked to recover unexploited, or lost, sources by hybridizing between wheat varieties and wild species. It is due to this pre-breeding that beneficial genetic diversity can be regained (Dempewolf *et al*., 2017). Pre-breeding aims to create a new genetic diversity, and a larger gene pool with resistance or tolerance to biotic stresses, such as diseases and insect damage, and abiotic stresses, such as heat, salinity, and drought. Landraces and wild relatives place as a rich resource to develop new varieties thus these sources place on important for pre-breeding.

Bedő and Láng (2015) reported that the variation between alleles is reduced in new elite cultivars, as a result of commercial breeding methods, thus pre-breeding programmes are crucial for developing new germplasm. Five kinds of genetic resources exist that will advance new genetic diversity:

- Landraces and cultivar population
- Adapted lines and cultivars
- Cultivated and wild relatives
- Mutant genotypes and
- Exotic germplasm.

These sources can be part of a pool depending on how related they are to wheat (Figure 1.6) (Feuillet *et al*., 2008);

- **Primary gene pool:** The primary gene pool of wheat contains landraces, wild species and early domesticates such as *T. turgidum* (AABB), *Ae. tauschii* (DD) and *T. urartu* (AA).
- **Secondary gene pool:** There are species that have at least one homologous genome in the secondary gene pool such as *T. timopheevii* (AAGG) and *Aegilops* species (S genome).
- **Tertiary gene pool:** This gene pool contains distantly related diploids and polyploids and also contains many *Triticeae* members that are not inside the primary and secondary gene pools such as *Hordeum marinum* (XX), *Secale* (RR), and *Thinopyrum* (EE). They need special strategies for gene transfer because homologous recombination cannot be used.

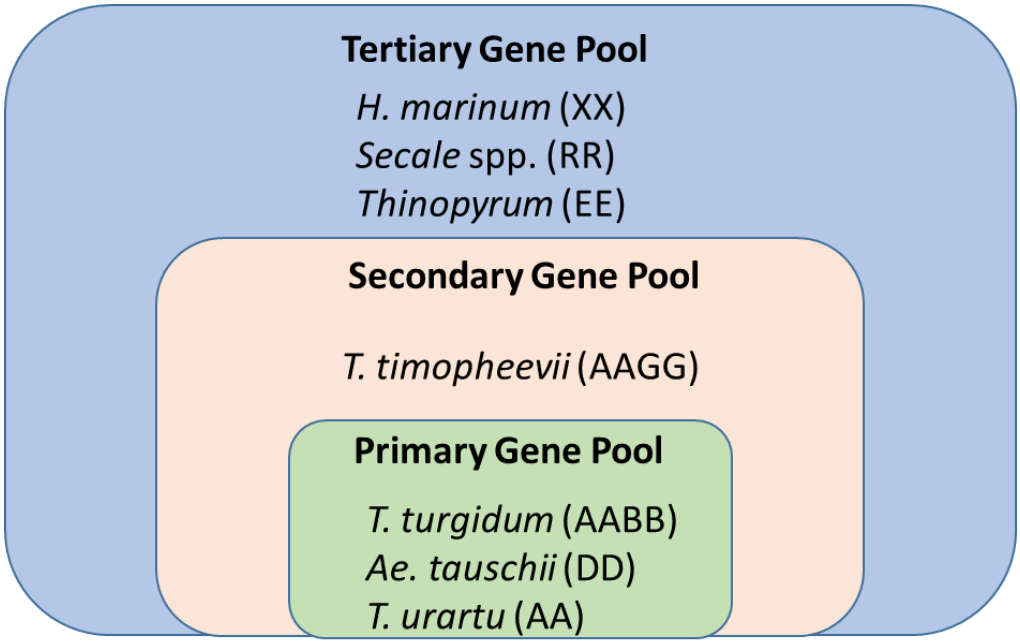


Figure 1. 6: Wheat gene pools representing primary, secondary, and tertiary gene pools with example species in each

To solve the lack of genetic diversity problem, wild relatives are a rich source of beneficial alien genes for enriching the genetic diversity of wheat. *Th. elongatum* that belong to the tertiary gene pool of wheat has always been the focus of breeders (Li *et al*., 2018).

1.7 Molecular Markers: an Introduction and Their Applications

The fundamentals of breeding in plants science for genetic improvement of crops have mainly been based on the selection of superior progenies from the population for traits such as high yield, nutritional quality, suitable colour, etc., as preferred by the user. Some traits cannot be measured directly from the crop. Instead, they are measured via markers and tags related to the traits. The specific segment of DNA which has a nucleotide between individuals is defined as a molecular marker (Boopathi, 2013). Using molecular markers is a major step for the improvement of the new germplasm and also to develop the pre-breeding efficiency in cereal breeding. It allows breeders to speed the introgression and backcrossing of genes into a variety genetic background (Bedő and Láng, 2015).

The DNA based markers that can be produced in large numbers can demonstrate to be useful for many kinds of aims related to crop improvement. For instance, these DNA markers are extensively used to prepare genetic maps. The genetic map is required both for reliable detection, mapping and prediction of gene effects of significant agronomic properties as well as for further study on the organization, evolution, structure and function of the crop genome (Kundan *et al.*, 2014).

An ideal molecular marker should have some desirable properties such as codominant, high reproducibility, higher frequency of polymorphism and also it should be easy, fast to detect and have a low cost (Kundan *et al.*, 2014).

Today there are a variety of molecular markers available such as restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), microsatellites or simple sequence repeats (SSRs), and single nucleotide polymorphism (SNP). The following section is briefly explained in the below.

RFLP markers, the first DNA-based markers, originated in the 1980s. These markers based on the digestion of genomic DNA that is digested with restriction enzymes. These enzymes are utilised to cut DNA at a specific sequence, as a result, a huge number of fragments with varying length. After digested with enzyme, agarose or polyacrylamide gel electrophoresis is applied to separate these fragments and show a characteristic pattern. Besides, these patterns are visualised using Southern blotting (Kundan *et al.*, 2014). They can be implemented in any organism and although they are still used today because of their utility in comparative mapping analysis and map-based cloning studies, they are used limitedly. The main problem with RFLP is that they are not suitable for high-throughput analysis (Tiwari *et al.*, 2014). The other limitations of RFLP markers are that they are time consuming, require quantity and high quality of DNA and expensive probes (Amom *et al.,* 2017).

The first PCR based molecular marker is RAPD and its PCR primers are short (approximately 10 nucleotides), and random. RAPD fragments generated are visualised on agarose gels stained with ethidium bromide. These markers have many advantages such as highly polymorphic, and simply applying.

However, RAPD shows dominant inheritance and hence they unable to distinguish between homozygous and heterozygous (Kundan *et al.*, 2014) and they have limited utilization in mapping studies and molecular biology owing to their low reproducibility (Tiwari *et al.*, 2014).

AFLP that is a combination of RFLP, and RAPD is the latest model of marker-assisted selection and is an extremely sensitive method (Kundan *et al*., 2014). Good quality also partially degraded DNA can be utilized for AFLP analysis, but DNA should be free of restriction enzymes and PCR inhibitors (Kundan *al.,* 2014, Amom *et al.,* 2017). After digesting the genomic DNA, followed by ligation of oligonucleotide adapters to the ends of fragments and visualisation on agarose gel (Kundan *et al.,* 2014; Tiwari *et al*., 2016).

The advantages of AFLP technology consist highly reproducible and also it has the capability to detect a large number of polymorphisms in different genomic areas at the same time (Kundan *et al.,* 2014; Tiwari *et al*., 2016) as compared to RAPD and RFLP (Kundan *et al*., 2014). Although AFLP markers that have been applied in a large range of organisms unite, it is expensive and most AFLP loci do not distinguish between homozygous and heterozygous thus it is dominant (Kundan *et al*., 2014).

SSRs are also called as microsatellites and are randomly tandem repeats motif of nucleotide (Jiang, 2013). Mono: A, T Di: AT, GA Tri: AGG Tetra: AAAC are a common motif to see (Kundan *et al*., 2014). The DNA sequences flanking these microsatellites are generally conserved and can be utilized to design primers (Jiang, 2013; Kundan *et al.,* 2014).

SSR markers have been shown to be polymorphic, co-dominant, reproducibility and they can be analysed by using PCR and detected with electrophoresis. Small DNA samples (~100 ng per individual) are enough for SSR assay. However, SSR markers have high start-up costs for automated detections, and they also require nucleotide information to design primer (Jiang, 2013).

SNPs are important and new generation molecular markers, used to define genetic diversity. These markers detect differences in a single nucleotide (A, T, C, or G) between parental lines (Figure 1.7) (Tiwari *et al.*, 2014). According

to nucleotide substitutions, SNPs may be transitions (C/T or G/A) or transversions (C/G, A/T, C/A or T/G) (Jiang, 2003).

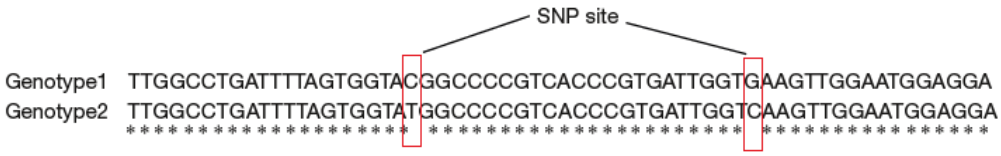


Figure 1. 7: Identification of SNPs between two DNA sequences (Source; (Tiwari *et al.*, 2014).

SNPs are codominant and highly polymorphic markers, and they are potential genetic molecular markers in crop breeding and genetic study. They also can be quickly detected and easily automated to define polymorphism with high efficiency. Hence, it can be expected that SNPs will be increasingly used for more and more species (e.g., rice, maize, soybean, etc.). However, high costs for initialisation or marker development and high technical equipment are required (Jiang, 2013).

The choice and utilize of molecular markers are still a challenge for breeders in research and breeding. The benefit or disadvantages of a molecular marker system are relevant largely to the aim of the study, available genetic resources, laboratories and equipment, personnel and funding resources (Jiang, 2013). The commonly used DNA markers discussed above are compared in Table 1.2 and this table was presented by Nadeem *et al.* (2017).

Table 1. 2: Comparison of characteristics of most widely used DNA marker systems in plants. (Source; Nadeem *et al.*, 2017).

Characteristics	**RFLP**	**RAPD**	**AFLP**	**SNP**	**SSR**
Codominant/ Dominant	Codominant	Dominant	Dominant	Codominant	Codominant
Reproducibility	High	High	Intermediate	High	High
Polymorphism level	Medium	Very high	High	High	High
Required DNA quality	High	High	High	High	Low
Required DNA quantity	High	Medium	Low	Low	Low
Marker index	Low	High	Medium	High	Medium
Genome abundance	High	Very high	Very high	Very high	Medium
Cost	High	Less	High	Variable	High
Sequencing	Yes	No	No	Yes	Yes
Status	Past	Past	Past	Present	Present
PCR requirement	No	Yes	Yes	Yes	Yes
Visualization	Radioactive	Agarose gel	Agarose gel	SNP-VISTA	Agarose gel
Required DNA (ng)	10000	20	500-1000	50	50

Thanks to developments in DNA markers and sequencing methods, comparison of the genome of wheat and its wild relatives is now possible. Improvements in DNA sequencing techniques have led to the development of high-throughput highly effective DNA marker methods, e.g., Kompetitive allele specific PCR (KASPAR), SNP, and diversity arrays technology (DArT) markers, which are suitable for the wide-scale screening of generations of distant relative hybrids. In addition, it assists the breeding of lines, which introgressed

the gene of interest with a minimum of undesirable chromatin structures (Rey *et al*., 2015).

1.8 Genomic *in Situ* Hybridization Technique (GISH)

Chromosomes were analyzed with classical cytogenetics which depended on morphological studies such as the size of the arms and position of the centromeres up to the early 1970s (Silva and Souza, 2013). Molecular cytogenetic techniques, such as in situ hybridization methods, most reliable tools for analysing the genomic structure, genome evolution, aneuploidy, and polyploidy and also alien gene introgression from the wild types into cultivated species can be analysed. These have been also used in the identification of allopolyploids (Ramzan, Younis and Lim, 2017).

Genomic *in situ* hybridization (GISH), a modification of fluorescence *in situ* hybridization (FISH) technology, first developed during the 1980s is a method which can be applied for the genome discrimination of parents from one another, using distinctive chromosome labelling methods (Ramzan, Younis and Lim, 2017). Nick translation is a molecular technique for labelling DNA with florosens dNTPs. DNase I hydrolyses DNA by creating random nicks in double-stranded DNA. This was followed by DNA polymerase replacing the nicked sites with dNTPs. For example, one of the first applications of GISH involved differentiating the genomes of an intergeneric hybrid between the *Secale africanum* and *Hordeum chilense* (Schwarzacher *et al*., 1989).

GISH is utilized to distinguish a genome of one parent from the other by using chromosome labelling methods (Ramzan, Younis and Lim, 2017). GISH follows the same protocol as in the FISH however, FISH utilises specific DNA sequences as a probe while GISH uses genomic DNA of a species as a probe (Silva and Souza, 2013).

The primary purpose of crop improvement is to transfer the genes responsible for resistance to extreme temperatures, diseases, and other factors, between wild species and cultivated ones. One of the most important features of GISH involves identifying the inserted DNA sequences of wild species in the chromosome of the cultivated species (Silva and Souza,

2013). The visualization of wild relative chromosomes in wheat is possible by GISH using labelled genomic DNA (gDNA). For example, in this project when applied single colour GISH (sc-GISH), the total genomic DNA of *Th. elongatum* was used as a probe for visualizing introgression into wheat. On the other hand, when visualizing introgressions by using multi-colour GISH (mc-GISH), the total genomic DNA of *T. urartu, Ae. tauschii, Ae. speltoides* and *Th. elongatum* labelled with four different colour probes were used. Mc-GISH thus enables chromosomes from different genomes to be visualized by using several probes at the same time. Thanks to mc-GISH can distinguish the wheat genomes and genome of wild relative of wheat.

Mc-GISH has been used to identify many different alien chromosomes into hexaploid wheat such as *Th. bessarabicum* (Grewal *et al.,* 2018a), *Triticum timopheevii* (Devi *et al*., 2019), *Th. elongatum* (Hu *et al*., 2012).

1.9 *Thinopyrum elongatum*

Th. elongatum, in the tribe *Triticeae* and genus Thinopyrum, is closely related to wheat, and is composed of four species: diploid (2n=2X= 14), tetraploid (2n=4X= 28), hexaploid (2n=6x=42) and decaploid (2n=10X= 70) (Chen *et al*., 2013). The diploid *Th. elongatum* (syn. *Lophopyrum elongatum* or *Agropyron elongatum*), known as tall wheatgrass, is a wild relative of wheat, and also an excellent reservoir for the gene pool in terms of the improvement of new superior wheat species (Jauhar, Doğramaci and Peterson, 2004; Hu *et al*., 2012; Chen *et al*., 2013).

Thinopyrum species are a domesticated species in Europe, especially the Mediterranean region, northern Africa, and western Asia (Wang, 2011). Wheatgrass possesses many excellent traits, including its grain protein content, multi-flowers, and long spikes, its fertility, and its resistance, and the fact that it grows in both cold and temperate areas (Chen *et al*., 2013). Recently, the wild species that include the E genome have been employed by breeders for improving disease-resistance, and resistance to adverse environmental conditions (Hu *et al*., 2012). *L. elongatum* (Host) Á. Löve, which is also known as *Thinopyrum*, or *Agropyron* and, which originates in the tertiary gene pool, is

tolerant to many abiotic and biotic stresses (Mujeeb-Kazi *et al*., 2008; Mullan *et al*., 2005). For instance, Dvorak *et al*. (1988) reported that *Lophopyrum* has genes for tolerating salinity.

Hu *et al*. (2012) suggested that the use of *Th. elongatum* as a resource of genetic diversity to improve wheat species would be enhanced by the development of specific molecular markers for the plant. These markers would then be used to scan many genotypes for alien introgressions. The study by Chen *et al*. (2013), which created molecular markers for the 7E chromosome of *Th. elongatum* suggested that the markers can be employed to find the location of *Th. elongatum* in wheat breeding, via marker-assisted selection.

Diseases such as rusts and Fusarium head blight are currently suffered by wheat and can cause a reduction in the volume of the wheat cultivated. Previous studies have reported that certain wild relatives of wheat possess resistance genes to these diseases including, wheatgrass which has been used extensively to enhance the disease-resistance of wheat (see Chapter 3, section 3.1). *Th. elongatum*, a wild relative used in this study, is shortly described in this section. Chapters 3 and 4 provide a more detailed introduction (see sections 3.1, and 4.1).

1.10 Thesis Outline and Objectives

Chapter 3: The study started with the step of reducing the number of *Th. elongatum* segments. The main aims and objectives of this chapter are:

- Introduce and explain the crossing programme used to reduce the number of *Th. elongatum* introgressions present in the lines carried forward in this work.
- Present the fertility rates of the different generations achieved in the backcrossing program.

Chapter 4: In this chapter, the chromosome-specific Kompetitive Allele Specific PCR (KASP^{TM}) markers and mc-GISH have been used for the characterisation of the wheat/*Th. elongatum* introgression lines. These markers have therefore been used in this work to identify *Th. elongatum* introgressions in a wheat background and characterise, as far as possible, the introgressions identified in terms of size. Track the introgressions through the self-fertilised generations. The main aims and objectives of this chapter are:

- To generate introgression lines containing single segments of varying sizes of *Th. elongatum* into wheat background.
- To characterise the introgression lines using the chromosome-specific KASP markers and validate with mc-GISH.

Chapter 5: In this work, twenty introgressions in wheat- *Th. elongatum* were generated using self-fertilization on some of the lines from the back-crossed population generated previously and molecular characterization of these lines. Then, these twenty homozygous lines were characterised with KASP markers and mc-GISH. The main aims and objectives of this chapter are:

- To self-fertilise the lines containing single segments to make the introgressions homozygous and thus stable.
- To characterise the introgression lines using KASP markers and validate with mc-GISH.

Chapter 2 General Materials and Methods

2.1 DNA extraction

Four to five cm of young leaf material was collected into 2 ml tubes on ice and dried overnight in a Christ Alpha 1-4 LD-Plus freeze-dryer (Martin Christ GmbH, Germany). One 5mm stainless steel ball was added to each tube. Samples were ground using a tissue lyser (Qiagen Tissue Lyser II, Type MM 300, Germany) at 25Hc/second for 6 min. 800 ml extraction buffer (Appendix a) was added to each sample and then incubated at 65°C for 1 hour. After incubation, samples were mixed with 400 ml 6 M ammonium acetate ($NH_4C_2H_3O$, Appendix b) on ice for 15 min. Samples were centrifuged for 10 min at 13000 rpm the supernatant transferred to a new 2 ml tube and 400 ml isopropanol added to each tube. The samples were mixed thoroughly, and the DNA allowed to precipitate for at least 20 minutes on ice. The samples were centrifuged for 10 minutes at 13000 rpm to pellet the DNA and the supernatant discarded. 200-300 µl of 70% ethanol was added to wash the pellet, samples were centrifuged for 3 min at 13000 rpm and the supernatant again discarded. This step was repeated once more. DNA was purified by using 200 ml phenol/chloroform (1:1 V/V). Samples were centrifuged for 5 min at 13000 rpm, the supernatant transferred to a new 1.5 ml tube and a 1/10th sample volume of 3M NaOAC pH5.2 (Appendix c) and 2.5 sample volume of 100% ethanol added and mixed well. The samples were kept at -20^{o}C for 1h or overnight and samples were centrifuged for 10 min at 13000 rpm to remove the supernatant. The pellet was washed three times with 200-300 µl of 70% ethanol, centrifuging for 3 min at 13000 rpm after each wash to remove the ethanol. The pellets were incubated at 37^{o}C for 10 min to dry completely and the pellet was re-suspended in 100-200 µl of vortexed briefly, and the tube kept on ice for 1 hr overnight. The DNA was briefly centrifuged and stored at -20^{o}C. DNA was quantified with either a Nanodrop spectrophotometer (Thermo Fisher Scientific, Applied Biosystems, Invitrogen, Gibco or Ion Torrent) or Qubit 2.0 fluorometer (Thermo Fisher Scientific, Applied Biosystems, Invitrogen^{TM}) following manufacturer`s instruction.

2.2 Genomic *in situ* hybridization

2.2.1 Root tips and nitrous oxide treatment

After germination, 1 cm root tips were collected from each seed into 0.5 ml Eppendorf tubes (that were sprayed with dH_2O) with a hole on the lid and treated with nitrous oxide gas at 10 bars for 2 hrs. All roots were fixed in 90% acetic acid for 10 min and then washed 3 times with dH_2O on ice. The root tip was then cut and put into 20 µl of 1% pectolyase Y23 and 2% cellulase Onozuka R-10 (Appendix e) (Yakult Pharmaceutical, Tokyo) for digestion at 37°C for 50-55 min. All samples were washed 3 times, with 70% ethanol. After each wash, the tubes were centrifuged for 2 minutes at 6000rpm, and the supernatant discarded. The root tips were then crushed in 50 µl 70% ethanol, centrifuged at 5000 rpm for 2 mins and all ethanol was removed. 15-40 µl 100% acetic acid was added to the pellet on ice for 10 mins-2 hrs, depending on the size of the root tip. Glass slides (Thermo Fisher Scientific, Applied Biosystems, Invitrogen, D-38116, 76 x 26 mm) were labelled and placed in a cardboard box with a few layers of wet tissue. Samples were mixed gently, and 7 µl of root tip suspension dropped onto a glass slide from a height of 5-10 cm. Slides were air dried slowly in a humid chamber. Slides were stored at -20°C.

2.2.2 Making Probes

Genomic DNA from *Th. elongatum* 401007 was labelled by nick translation with Chroma Tide Alexa Fluor 546-14-dUTP (Life Technologies™; C11401). Genomic DNA from *T. urartu* was labelled by nick translation with Chroma Tide Alexa Fluor 488-5-dUTP (Thermo Fisher Scientific (Invitrogen; C11397), Waltham, MA, United States). Genomic DNA from *Ae. tauschii* was labelled by nick translation with Chroma Tide Alexa Fluor 594-5-dUTP (Thermo Fisher Scientific (Invitrogen; C11400), Waltham, MA, United States). Genomic DNA from *Ae. speltoides* was labelled by nick translation with DEAC-dUTP (Jena Bioscience; NU-803-DEAC; coloured blueish purple).

2.2.2.1 Nick translation reaction

The nick translation reaction mixture for 20 µl in 1.5 ml tube as described below:

dH_2O	x µl to 20 µl
DNA (gDNA >200ng/µl)	2-3 µg
10x Nick translation buffer or 10x buffer2	2 µl
Non-labelled dNTPs (2mM each, mixed)	2 µl
Labelled dNTP (1mM)	0.5 µl
DNA polymerase I (10U/µl)	5 µl
*DNase	0.8 µl

*DNase (100 mU/µl) was diluted (5 µl of 2U/µl DNase by adding 95 µl of 50% glycerol.

The solution was mixed well by pipetting and then incubated at 16°C for 2 hours in the dark.

2.2.2.2 Ethanol Precipitation of the probe

160 µl of single stranded DNA (SS DNA) working solution (20 µg SSDNA, Appendix f) was added to the nick translation reaction and vortexed to mix well. 500 µl of ethanol and 3M Sodium acetate pH 5.2 (Appendix c) were then added and mixed well. The solution was incubated at -20°C overnight. After spinning at 12,000 rpm for 30 min at 4°C probe pellet was washed with 70% ethanol and then spun again to discard the ethanol and the pellet was air-dried for 10 min in the dark to ensure all ethanol was removed. The pellet was dissolved in 20 µl of 2x SSC+1x TE buffer pH7.6 (Appendix g), left on ice for 30 min-1h and then stored at -20°C in the dark.

2.2.2.3 Preparation of blocking genomic DNA

Genomic DNA from Chinese Spring was fragmented to 300-500 base pair in length by treating in a heat block at 100°C for 12 ~20 min according to the DNA volume.

2.2.3 Multi-colour GISH

Selected slides for GISH were treated twice with UV light at 0.125 Joules to cross-link the chromatin on the slides. 10 µl probe mix for mc-GISH (Appendix h) was added to each slide and covered by a plastic coverslip (22mm x 22mm) avoiding air bubbles. Slides were put in a tray in a water-bath at 75°C for 5 mins and then incubated at 55°C overnight. Slides were dipped into 2x SSC (2×Saline-sodium citrate) solution (Appendix g) to remove the coverslips and air dried on a tissue. One drop of 4', 6-diamidino-2-phenylindole (DAPI) was dropped onto each slide, and a glass coverslip added carefully avoiding air bubbles. Samples were observed on a fluorescence microscope (Zeiss Axio ImagerZ2, Carl Zeiss Ltd, Oberkochen, Germany) using filters for DAPI (blue), Alexa Fluor-594 (Red), Alexa Fluor-488 (Green) and Alexa Fluor-546 (Gold). Photographs were taken using a MetaSystems Coolcube 1m CCD camera and analysed using Metafer (automated metaphase image capture) and ISIS (image processing) software (Metasystems GmbH, Altlussheim, Germany).

Chapter 3 Development of Germplasm

3.1 Introduction

Thinopyrum species have attracted the attention of breeders as sources of genes for important agronomic traits (Arterburn *et al.,* 2011). There are approximately 20 species in this genus (Dewey 1984) with species native to Europe (especially the Mediterranean region), west Asia, north Africa, and North America. Species in this genus include the diploid *Thinopyrum bessarabicum*, hexaploid *Thinopyrum intermedium* and decaploid *Thinopyrum ponticum*.

Th. elongatum (also known as tall wheatgrass) belongs to the genus Thinopyrum, which, along with *Lophopyrum* and *Trichopyrum*, was separated from the much larger genus *Elytrigia* in the late twentieth century (Baum and Johnson 2018). *Th. elongatum* exists at four ploidy levels: diploid (2n=2x= 14), tetraploid (2n=4x= 28), hexaploid (2n=6x=42) and decaploid (2n=10x= 70) (Chen *et al*., 2013) and synonyms include *L. elongatum*, *A. elongatum* and *Elytrigia elongatum*). *Th. elongatum* has been shown to be an excellent reservoir for the introduction of genetic variation into wheat such as salinity tolerance (Dvorak *et al*., 1988; Mujeeb-Kazi *et al*., 2008) and, in particular, to biotic stresses (Jauhar, Doğramaci and Peterson, 2004; Hu *et al*., 2012; Chen *et al*., 2013). Significant worldwide diseases of *T. aestivum L.* include fungal diseases such as stripe (yellow), leaf (brown) and stem (black) rusts, Fusarium head blight and powdery mildew as well as a number of diseases caused by viruses including barley yellow dwarf virus (BYDV) and wheat streak mosaic virus (WSMV). Resistance to all these diseases has been found in *Th. elongatum*.

Friebe *et al*. (1994) used C-banding and *in situ* hybridization methods to show that the stripe rust resistance genes (*Sr26 and Sr25/Lr19*) had been transferred from *Th. elongatum* to common wheat, with the *Sr26* gene having been transferred from chromosome 6E of *Th. elongatum* and *Sr25/Lr19* having been transferred from chromosome 7E. The transfer of *Lr19* from *Th.*

elongatum was confirmed by (Schachermayr *et al.*, 1995). Ma *et al*. (2000) tested several *T. aestivum-Th. elongatum* lines for resistance to yellow rust and demonstrated that a gene located on chromosome 3E of *Th. elongatum* carried resistance and Zu and Zheng (2001) also reported that *Th. elongatum* possesses resistance to stripe rust. Recently, Wang *et al*. (2020) generated new wheat–*Th. ponticum* substitution line that was shown to be resistant to stripe rust and powdery mildew in their research. They discovered that *Th. ponticum* chromosomes $1J^s$ may have novel resistance gene(s) for stripe rust and powdery mildew. As a result, they noted that it is a crucial material in the search for possible resistance genes in *Th. ponticum*. More recently, Gong *et al*. (2022) conducted another stripe rust and powdery mildew study with another Thinopyrum species, *Th. scirpeum* (2n = 4x = 28, EEEE). They reported that a new 4E (4D) chromosomal substitution line was characterised by chromosome-specific markers, GISH and FISH, and the results of the disease responses revealed it is resistant to stripe rust and powdery mildew (Gong *et al*., 2022).

Fusarium head blight (FHB) is a disease of wheat grown in humid conditions around the world, for which amphiploids with alien species and synthetic hexaploid lines can provide novel sources of resistance (Oliver *et al*., 2005). The diploid *Th. elongatum* has a high level of resistance to fusarium head blight and the authors reported having successfully identified these genes, located on the long arm of chromosome 7E (Shen *et al*., 2004; Fedak *et al.,* 2021). Shen *et al*. (2004) reported that lines resistant to FHB had been developed from crosses between wheat and *Th. elongatum,* while another study of FHB (Fu *et al*., 2012) demonstrated by examination with GISH and mc-GISH that some lines that included chromosome 7E of *Th. elongatum* provided resistance to the disease. A gene for powdery mildew resistance was also shown to be present in *Th. elongatum* (Li *et al*., 2017). More recently a study of diploid *Th. elongatum* showed that introgressions on wheat chromosomes 7A, 7B, and 7D were discovered using SSR markers unique for homoeologous wheat chromosome 7. FHB resistance was found in several of the *Th.*

elongatum/wheat introgression lines (Fedak *et al.,* 2021). Zhang *et al*. (2022) identified and integrated *Th. elongatum*-derived Fhb7 allele, designated Fhb7^{The2}, into the wheat B genome through a small 7B–7E translocation (7BS·7BL–7EL). Additionally, it enabled its use in breeding programmes for both common and durum wheat without the yellow flour linkage drag (Zhang *et al*., 2022). Recently, a study presented how the 7EL contributed to the CS-7EL transcriptome by direct expression and through alteration of wheat transcript profiles. Also, a list of candidate genes for the FHB resistance locus on 7EL has been generated（Konkin *et al*., 2022).

In their work with wheatgrass, Sharma *et al*. (1989) reported that *Th. elongatum* is not affected by BYVD and concluded that the species is an important source of BYDV resistance for transfer into wheat. Furthermore, three wheat streak mosaic virus (WSMV)-resistant lines have been developed by hybridizing wheat with *Th. elongatum* (Jiang *et al*., 1993).

Previous work has therefore shown the importance of *Th. elongatum* as a source of new genetic variation for transfer into wheat an important species for the production of introgression lines with wheat.

3.2 Aims and Objectives of this chapter

The main aims and objectives of this chapter are:

- Introduce and explain the crossing programme used by Dr Lauren Baker to generate the initial wheat/ *Th. elongatum* 401007 introgression lines.
- Introduce and explain the crossing programme used to reduce the number of *Th. elongatum* introgressions present in the lines carried forward in this work.
- Present the fertility rates of the different generations achieved in the backcrossing program.

3.3 Materials and Methods

3.3.1 Crossing programme

Hexaploid wheat cvs. Chinese Spring and Paragon, both with the *Ph1* gene deleted (obtained from the Germplasm Resource Unit at the John Innes Centre, Norwich) were pollinated with *Th. elongatum* (accession 401007; obtained from the USDA) to produce F_1 interspecific hybrids (see Figure 3.1) at the Nottingham WRC. The F_1 hybrids contained a single chromosome from each linkage group from the A, B and D genomes of wheat and each of the E genomes from *Th. elongatum* (Figure 3.1).

The F_1 interspecific hybrids were then backcrossed to *T. aestivum* cv. Paragon as the male parent to produce a BC_1 population. These BC_1 lines often had numerous introgressions present, and so several rounds of back-crossing with *T. aestivum* cv. Paragon were required to reduce the number of introgressions present as shown in the crossing diagram (Figure 3.1). The initial backcross generations (up to BC_2, and sometimes BC_3) were produced by Dr Lauren Baker and it was seed from these generation that was taken forward in this project.

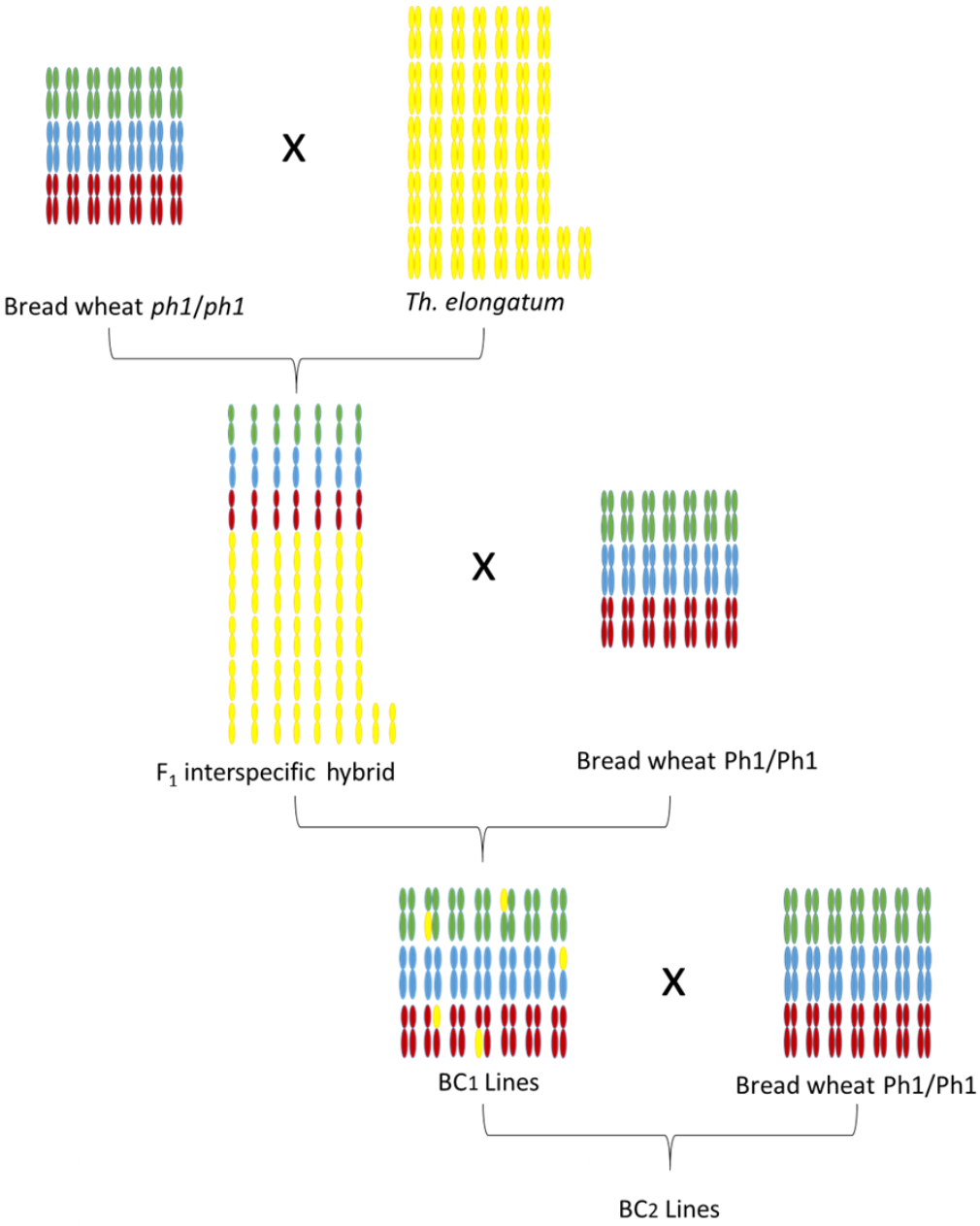


Figure 3. 1: Development of F1 interspecific hybrid and wheat/*Th. elongatum* introgressions. The green, blue, and red chromosomes represent the A, B, and D genomes of wheat respectively and the yellow chromosomes represent the E genome of *Th. elongatum*. This crossing programme was used by Dr Lauren Baker.

In this project, wheat/*Th. elongatum* introgression plants that contained more than one introgression were back-crossed as the female parent to wheat cultivar (cv.) Paragon to obtain single introgressions while wheat/*Th. elongatum* introgression plants, containing only a single introgression, were

self-fertilised to obtain stable homozygotes. This crossing programme is summarised in Figure 3.2.

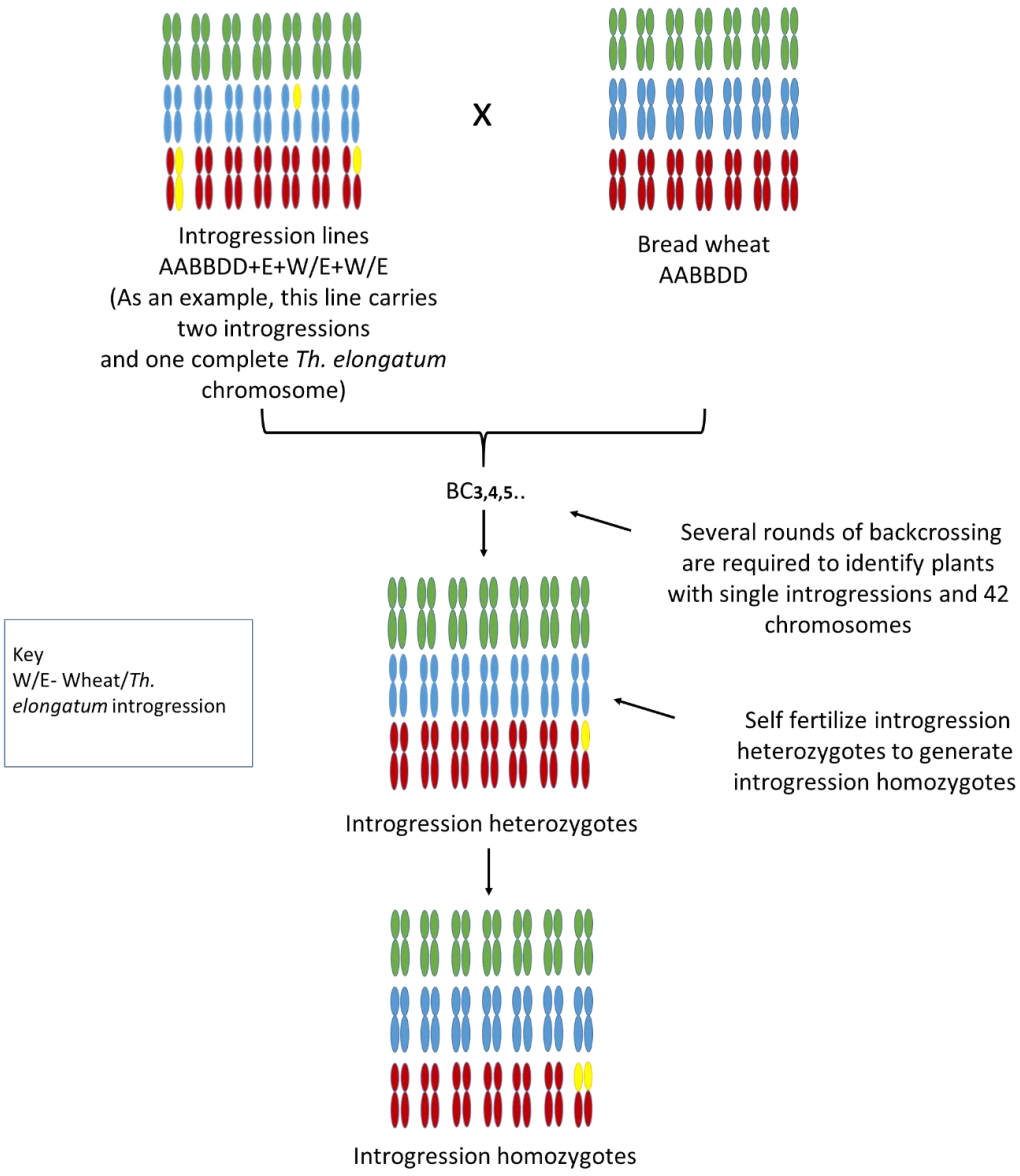


Figure 3. 2: Crossing diagram to generate single segment and homozygous wheat/*Th. elongatum* introgression lines (used in this work). The green, blue, and red chromosomes represent the A, B, and D genomes of wheat respectively and the yellow chromosomes represent the E genome of *Th. elongatum*.

3.3.2 Plant material

Wheat*/Th. elongatum* introgression lines containing segments of *Th. elongatum,* previously produced at the WRC by Dr Lauren Baker, were used in this project. While these seeds were created, two varieties of spring wheat were used *T. aestivum* Chinese Spring *ph1/ph1* was pollinated with *Th. elongatum* to produce the F_1 interspecific hybrids and then *T. aestivum* cv. Paragon was used for subsequent backcrosses at all stages of the breeding programme.

3.3.3 Germination

Nine hundred and six seed from wheat*/Th. elongatum* introgression lines (from different generations) were germinated (see table 3.1) Seeds were placed on damp, filter paper in Petri dishes for 2-3 days (the seeds were sprayed with fresh dH_2O every day to keep the environment moist). Seeds which did not germinate were embryo rescued (see below). Seed of Paragon was sown directly into modular trays containing Levington Advance F2+S compost in the glasshouse.

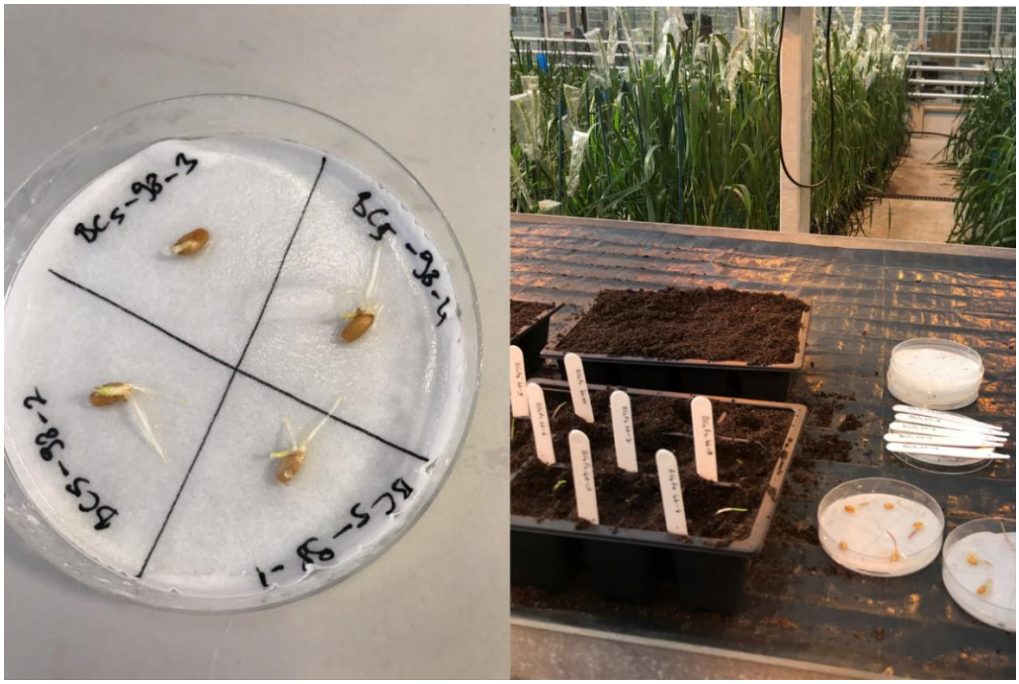


Figure 3. 3: An example of back-crossed lines germination and sowing into modular trays.

3.3.4 Vernalisation and potting

Seedling were vernalised at 6˚C with a photoperiod of 12 hours (6 am-6 pm) for four weeks. All plants were then potted individually into 2 litres pots containing John Innes compost No. 2 and transferred to the glasshouse at 25˚C with a photoperiod of 16 hours light and 8 hours dark.

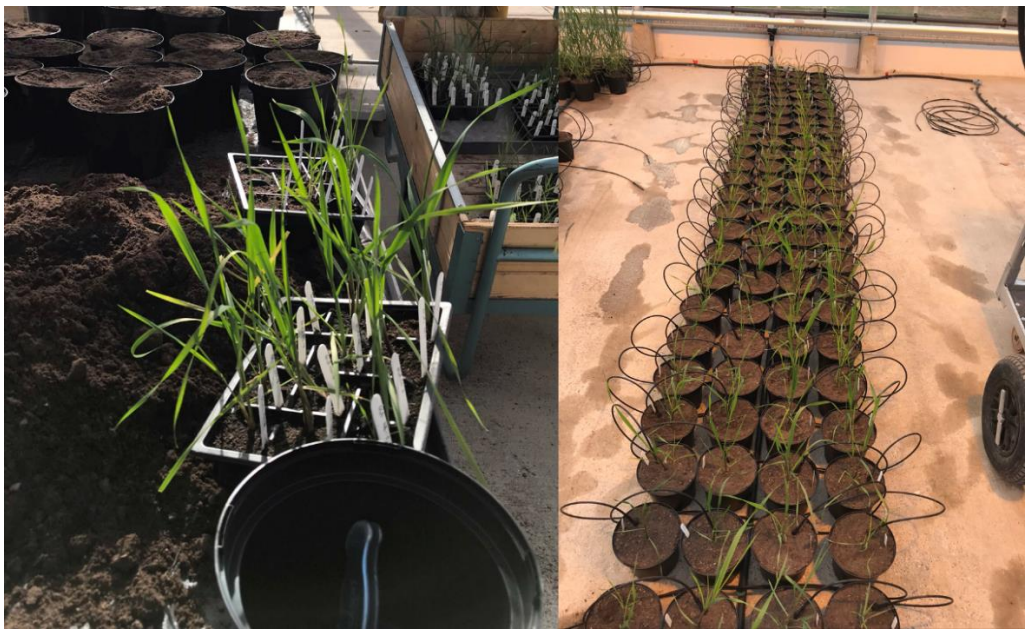


Figure 3. 4: Transferring backcross lines from the vernalization to the glasshouse as an example.

3.3.5 Crossing Method

The crossing methods are described below. Self-fertilised seeds (e.g., BC_2F_2, BC_2F_3) were produced by bagging spikes without emasculating and allowing them to self-pollinate.

3.3.5.1 Emasculation

Before the spike completely emerged from the flag leaf, the central florets of each spikelet were removed. The three green anthers were removed from the remaining florets to prevent self-pollination while taking care not to damage the stigma. After emasculation, the spike was labelled (the name of parent (male/female) and date of emasculation) and covered with a glassine bag to prevent unwanted cross-pollination (Figure 3.5). All crossing information was recorded in a crossing notebook.



Figure 3. 5: The left side photo has green anthers collected with forceps. Each floret must contain three anthers, so after emasculation, it was checked that there were no anthers left in the spikelet and the removed anthers were counted. An example of a glassine bag to avoid unwanted cross-pollination in the right-side photo.

3.3.5.2 Pollination

The stigma was checked for maturity two days after emasculation (a fluffy/feathered appearance showed the stigma was ready to be pollinated). Anthers were collected from Paragon (anthers were ready for pollination when they were a bright yellow colour) and the pollen was released onto the stigma. The spike was again covered by a glassine bag to protect it from any other pollen until threshing.



Figure 3. 6: The image showed that the mature stigma is 'fluffy'/feathered and fanned.

3.3.6 Harvesting and threshing

The crossed and selfed spikes of all plants were harvested individually and manually threshed.

3.3.7 Selection of seeds

Genotyping was available for germination in each generation, so seeds were selected based on genotyping result (See chapter 4 for KASP marker genotyping).

3.3.8 Observations

Various backcross seed set observations were made for each hybridization and recorded as follows:

$$\text{Germination rate} = \frac{\text{Number of germinated seeds}}{\text{Number of sown seeds}} \times 100$$

$$\text{Cross fertility rate} = \frac{\text{Number of successful crossings}}{\text{Number of crosses}} \times 100$$

$$\text{Seeds/cross} = \frac{\text{Number of producing seed}}{\text{Number of crosses}} \times 100$$

3.4 Results

A total of 982 seeds were selected based on genotyping results and sown either for crossing or for self-pollination. The number of seeds sown, germination rate, cross fertility, and crossed and self-fertilised seeds produced, etc., are summarized in Table 3.1.

Table 3. 1: Number of seeds germinated and produced in relation to the number of crosses carried out and the cross fertility and the number of self-fertilized seed produced for each generation of the introgression programme for *Th. elongatum* into wheat. The number of self-fertilised seeds represents an estimated seed number.

Generation	**Seed sown**	**Germination rate (%)**	**Crosses made**	**Cross fertility (%)**	**Crossed seeds produced**	**Seeds/Cross**	**Self-fertilized seeds produced**
BC_2F_2	6	83,3	-	-	-	-	160
BC_2F_3	18	72,2	-	-	-	-	402
BC_3F_1	48	58,3	23	95,5	241	10,5	2,021
BC_3F_2	48	75	9	100	76	8,4	2,264
BC_3F_3	10	90	-	-	-	-	1,090
BC_4	17	64,7	7	85,7	48	7	910
BC_4F_1	91	76,9	65	100	569	9	6,093
BC_4F_2	340	82,4	11	100	54	4,9	24,525
BC_4F_3	47	78,7	-	-	-	-	1,986
BC_4F_4	4	100	-	-	-	-	115
BC_5	62	95,1	50	92	550	11	6,582
BC_5F_1	164	82,3	26	88,4	99	3,8	7,299
BC_5F_2	97	87	-	-	-	-	4,420
BC_5F_3	18	94	-	-	-	-	1,435
BC_5F_4	4	100	-	-	-	-	275
BC_6	8	100	-	-	-	-	864
Total	982	-	191	-	1,637	-	60,441

One hundred and twenty plants were germinated for further back crossing with Paragon (some heads on each of these plants were also bagged to produce selfed seed). Pollination was missed on seven of the 120 plants and four of the plants failed to produce seed.

A total of 191 crosses were made between 113 wheat/*Th. elongatum* introgression lines and Paragon as the male parent and a total of 1637 seeds were obtained with an additional 12,864 seeds were produced via self-fertilisation. In the earlier generations worked on by Dr Lauren Baker, the fertility of the plants (based on the number of crosses setting seed) was found to be quite low particularly in the F_1 generation where only 57.3% of crosses resulted in seed set. This rose to a maximum of 90% in the BC_1, BC_4, BC_4F_1, and BC_5 generations. The BC_3F_1, BC_3F_2, BC_4, BC_4F_1, BC_4F_2, BC_5 and BC_5F_1 generations in this work showed higher levels of fertility with 95.5, 100, 85.7, 100, 100, 92 and 88.4% of crossed ears setting seed, respectively (see Table 3.1).

Eight hundred and one of the plants selected for self-fertilisation reached maturity producing an estimated 60,441 seeds. The lowest germination rate was observed in the BC_3F_1 generation where only 28 of the 48 seeds sown survived to maturity (producing 241 seed from 23 crosses and 2000 plus seed from self-fertilised heads). However, the germination rate then increased with each subsequent generation (see Table 3.1).

3.5 Discussion

Th. elongatum is a potentially important source of genetic variation for a wide range of agronomically important traits (Chen *et al*., 2013). Using *Ph1* mutant, wheat-*Th. elongatum* recombinant lines were generated in the previous study by Dr Lauren Baker to obtain F_1 interspecific hybrids which were then backcrossed three times to produce BC_3 plants (Figure 3.1). Subsequent generations have been generated firstly to obtain single introgressions and secondly to make the single introgressions homozygous (Figure 3.2). *Th. elongatum* accession 401007 was shown to be decaploid when metaphase spreads were looked at with four additional chromosomes present (2n = 10x = 70(+4)) (Baker *et al.*, 2020). A large number of *Th. elongatum* chromosomes were therefore present in the F_1 hybrids and hence several rounds of backcrossing were required to reduce the number present to single introgressions.

The cross fertility was generally above 90% with the exceptions of the BC_4 and BC_5F_1 generations (Tale 3.1). The lower cross fertility of these two generations was not expected as cross fertility had increased in all the previous generations (results from Dr Lauren Baker). The reduced cross fertility in these two generations was possibly linked to glasshouse problems when each of these generations was being grown. The glasshouse heater did not work properly for approximately three months before and during flowering of the BC_4 plants resulting in reduced fertility of the plants, e.g., a reduction in pollen production was observed. The BC_5F_1 plants were badly affected by thrips in the glasshouse. These small insects living in flowers can prevent pollination as they eat both pollen and stigmas.

The reduction in the number of introgressions through the backcrossing programme would be one reason for the increasing fertility of the later generations. Large introgressions in particular, would introduce not only useful genetic variation but also deleterious genes via linkage drag and these deleterious genes could have an effect on the overall viability of the gametes and progeny. Strategies for reducing the number of deleterious genes

introduced once single introgressions have been obtained will be discussed in chapter 4.

The backcrossing programme also enabled the recovery of the wheat genome (which partly explained the lower fertility of the earlier generations). The F_1 hybrids were expected to contain only 21 wheat chromosomes resulting in gametes with different numbers of chromosomes. Correct chromosomal pairing is a need for proper meiosis in allopolyploids and that pairing is confined to homologues. Aneuploidy is caused by anomalies in homologous chromosome pairing in meiosis and also by the potential of polyploid meiosis to create aneuploid cells. It is also widespread in natural allohexaploid wheat, owing to its hexaploid nature, which makes it more tolerant of a chromosome or chromosomal arm deletion or addition (Sears, 1954). Thus, in addition to reducing the number of *Th. elongatum* chromosomes present, each backcross helped to restore the number of wheat chromosomes with the final aim of obtaining 42 chromosome plants. The use of *ph1* also resulted in a number of inter-genomic wheat translocations which the backcrossing also helped to eliminate. Both uneven numbers of wheat chromosomes (additions or deletions) and inter-genomic translocations would cause problems at meiosis as the chromosomes would not be able to pair properly leading to the formation of unbalanced gametes. The absence of certain wheat chromosomes, if not compensated for by a *Th. elongatum* chromosome, might also have resulted in reduced fertility. For example, chromosome 4B of wheat carries a male fertility gene on the short arm (Endo *et al*., 1991). Therefore, a lack of chromosome 4B can result in plants that are completely male sterile.

In conclusion, the study started with the step of reducing the number of *Th. elongatum* segments described in the crossing programme (Figure 3.2). In the crossing programme within this aim, 1637 seeds were generated from 191 crossings with 116 wheat/*Th. elongatum* introgression lines. Also, an estimated 60,441 seeds were produced by self-fertilization with 801 backcrossing lines (see Table 3.1).

Chapter 4 Generation of introgression lines and their characterisation with KASP markers and GISH (single, multi (A, B, D, E))

4.1 Introduction

4.1.1 Genome Constitution of the Thinopyrum Genus

The wheatgrass genus Thinopyrum contains different levels of ploidy with species ranging from diploid to decaploid and therefore, this genus exhibits a remarkable complexity of evolution at the genome level. However, anatomically, these species are very similar and thus these wild grasses have sometimes grouped different ploidy levels into the same species. The genus Thinopyrum members have been counted in the genera *Lophopyrum*, *Elymus*, *Elytrigia*, *Pseudoroegneria* and *Agropyron* at different times (Arterburn, M. *et al*., 2011). The identity of some wheatgrass species is frequently confused in the literature, possibly because of the use of multiple synonyms or the requirement for accurate differentiation via cytogenetic analysis. For example, diploid *Th. elongatum* (Host) D.R. Dewey is synonymous to *A. elongatum* (Host) P. Beauv, *Triticum elongatum* Host and *L. elongatum* (Host) A. Löve (Wang, R.R.C., 2011). *Th. intermedium* (Host) Barkworth and D.R. Dewey is synonymous to *Elytrigia intermedia, Agropyron intermedium*), and the decaploid *Th. ponticum* (Popd.) Barkworth and D.R. Dewey can also be referred to as *A. elongatum*, *Lophopyrum ponticum* and *Elytrigia pontica* (Ceoloni *et al*., 2017).

The Thinopyrum genus has been an important challenge for plant systematists owing to the effects of polyploidization on the evolution of this group. The current classification for the members of the Thinopyrum genus is mostly contingent on anatomical and karyotypic homology, and the pairing behaviours of their chromosomes in F_1 interspecific hybrids. Polyploid members of this genus are thought to have occurred from polyploidization events containing three supposed ancestors of the genus: *Th. elongatum* Dewey (2n = 2x = 14, E^eE^e), *Th. bessarabicum* (Savul & Rayss) A. Löve (2n = 2x =

14, E^bE^b), and either *Pseudoroegneria strigosa* (Bieb) A. Löve (2n = 2x = 14, StSt) or the closely related *Ps. spicata* (Pursh) Á. Löve (2n = 2x = 14, StSt) (Dewey 1984; Arterburn *et al*., 2011). The genomes of *Th. bessarabicum* and *Th. elongatum* were previously represented as J and E respectively although they are now usually represented by E^b and E^e as according to Wang (1992). The J and E genomes were described as very closely related, initially through the work of Dewey (1984) and Wang (1985), and also by later studies using different methodologies (Table 4.1).

Table 4. 1 : A brief summary of work looking at the close relationship of the J and E genomes.

Studies using different methodologies	Reference
Chromosome pairing	Wang 1985 Pienaar 1988 Wang and Hsiao 1989 Forster and Miller 1989
RAPD and STS markers	Wei and Wang 1995 Li *et al*., 2007
GISH	Kosina and Heslop-Harrison 1996 Chen *et al.,* 1998, 2003
DNA sequence of chloroplast	Mason-Gamer *et al.,* 2002 Liu *et al*., 2008
A gene encoding plastid acetyl-CoA carboxylase sequence	Fan *et al.,* 2007
nuclear rDNA internal transcribed spacer (ITS) sequences	Hsiao *et al.,* 1995. Yu *et al.,* 2008 Liu *et al.* 2008

There is also considerable confusion in the literature in distinguishing between the two decaploid *Th. elongatum* and another Thinopyrum species, *Th. ponticum*. Previously, these two species were often included under the same name, *Agropyrum elongatum* (Shepherd and Islam 1988; Li et al. 2017). *Th. ponticum* (Podp.) Barkworth and D.R. Dewey has been described with three different genome formulas, each proposed by different studies. Muramatsu (1990) and Wang *et al*. (1991) proposed the genome composition $E^bE^bE^bE^bE^bE^bE^bE^bE^bE^b$, Zhang *et al.* (1996) proposed $E^eE^eE^bE^bE^xE^x$StStStSt, with the St genome derived from either *P. strigosa* or the closely related *P. spicata* (see above) while Chen *et al.* (1998) proposed $E^bE^bE^bE^bE^bE^b$ $J^SJ^SJ^SJ^S$. The *Th. elongatum* accession 401007, used in this work was shown to be a decaploid accession with added monosomy, i.e., a chromosome count showed a total of 74 chromosomes and Baker *et al*. (2020) suggested this accession showed many similarities to another decaploid member of the Thinopyrum genus, *Th. ponticum*. Mc-GISH analysis of a metaphase spread of the 401007 accession by Baker *et al.* (2020) suggested the correct genome formula was $EEEEEEE^{St}E^{St}E^{St}E^{St}$ or $JJJJJJJ^{St}J^{St}J^{St}J^{St}$ (Figure 4.1). It can be seen quite clearly that while there are whole E or J chromosomes, there are no complete St chromosomes. Instead, the St genome is found mainly in centromeric regions although a few chromosomes also show small regions interstitially and at the telomeres. It is interesting to note that the genome of *Th. intermedium* (2n=6x=42) also has both the J and St genomes. However, in *Th. intermedium*, one whole genome is St (Wang *et al*., 2015; Cseh *et al*., 2019) which is not seen in the *Th. elongatum* accession in this work.



Figure 4. 1: Mc-GISH of a metaphase spread of *Th. elongatum* accession 401007 probed with total genomic DNAs from *Th. bessarabicum* (green) and *P. strigosa* (red) (from Baker *et al*., 2020).

4.1.2 Previous Work using Molecular Markers for *Th. elongatum*

Th. elongatum exhibits a wide range of genetic diversity which could be used to improve wheat (see Chapter 3, Introduction). Thus, many introgressions from *Th. elongatum* into wheat have been produced by wheat breeders (Friebe *et al*., 1994; Schachermayr *et al*., 1995; Ma *et al*., 2000; Fu *et al*., 2012; Fedak *et al.,* 2021). However, in order to use introduced alien genetic variation, it is necessary to be able to detect the introgressions in a wheat background and also track the introgressions through a breeding programme. Due to the potential of *Th. elongatum* as a source of genetic variation, a number of studies have developed molecular markers for this detection.

Chromosome 7E of *Th. elongatum* was shown to carry resistance to both FHB and wheat rust and therefore, Chen *et al.* (2013) developed 89 7E chromosome-specific molecular markers using Specific Length Amplified Fragment Sequencing (SLAF-seq). These markers could be utilised to detect the 7E chromosome of *Th. elongatum* in a wheat background and also for marker-assisted selection (MAS) in wheat breeding (Chen *et al.,* 2013).

In work by Dr Lauren Baker at the Nottingham BBSRC Wheat Research Centre (WRC), 202 introgressions lines were identified from 352 back-crossed lines using the 35K Axiom® Wheat-Relative Genotyping Array (many lines contained more than one introgression) enabling the production of a genetic linkage map of *Th. elongatum*. The genetic linkage map allowed a comparison of the synteny between *Th. elongatum* and wheat and showed that overall macro-synteny between the two was good although *Th. elongatum* does not show the 4A/5A/7B translocations found in wheat.

Lou *et al.* (2017) aimed to develop E-genome-specific SNPs between *Th. elongatum* and Chinese Spring (CS) and found 35,193 candidate SNPs between CS and *Th. elongatum.* Primers were designed for 420 of the 35,193 SNPs, 373 of which were validated on the parental material. When tested on 78 wheat-*Th. elongatum* hybrids, 14 of the SNPs were able to reliably detect the E genome in the CS background. They also suggested that it would be possible to use some of the SNPs for studies between wheat and other Thinopyrum species containing the E and/or St genomes.

Kompetitive Allele-Specific PCR (KASP) assays (a method of SNP genotyping) have been reported as flexible and cost-effective in different studies (Semagn *et al*., 2013; Allen *et al*., 2013; Neelam *et al*., 2013; Yu *et al*., 2017; Tan *et al*., 2017; Grewal *et al*., 2020) and can be performed in basic molecular laboratories (Chen *et al*., 2021). It is also possible to design the KASP assays so that they are able to distinguish between individuals that are homozygous and heterozygous for a particular introgression (Grewal *et al*., 2020; Makhoul *et al.*, 2020). KASP markers have been used in a number of different studies.

KASP, a proprietary genotyping technology from LGC Genomics, has been widely used by plant breeders for marker-assisted selection over the past decade. KASP genotyping, which can be carried out in 96-, 384-, and 1,536-well plate formats, enables biallelic evaluation of SNPs or inserts and deletions (InDels). The KASP primers are designed around a target SNP (e.g., for wheat/wild relative introgression a target SNP would need to be polymorphic between wheat and the wild relative). The KASP assay mix will contain two allele-specific forward primers and one common reverse primer, with the allele-specific forward primers specific for the polymorphic SNP. The allele-specific primers each contain a unique tail sequence that corresponds to a universal fluorescence resonant energy transfer (FRET) cassette; one labelled with FAM™ dye and the other with HEX™ dye which can be detected in the PCR amplified products using a plate reader or real-time PCR machine. Thus, in a homozygous genotype, only one of the fluorescent dyes will be incorporated, while in a heterozygous genotype, both dyes will be incorporated.

KASP assays have been increasingly used during the last decade in many studies on crop species. In wheat, Chandra *et. al*. (2017) reported the successful validation of 83 SNPs tested in 47 wheat genotypes with different rust and mildew resistance genes, resulting in the identification of SNPs associated with leaf rust resistance. In work by Qureshi *et al*. (2018), KASP markers were developed to genotype a population of recombinant inbred lines, derived from a cross between a landrace and the cultivar Avocet S (which carries the stem rust resistance gene Sr26). In this study, two KASP markers were identified as linked to the stem rust resistance. In another recent study, KASP markers were tested for marker-assisted selection of wheat curl mite (WCM) resistance. Two KASP markers were developed and tested in various germplasm with results suggesting that the two markers could be used in MAS to develop WCM resistance in some wheat genetic backgrounds (Dhakal *et al*., 2018).

In maize, KASP markers were developed from the RNA-Seq data of 368 inbred lines to meet the requirements of genotyping fewer loci in many

samples (Chen *et. al*., 2021). The results showed that the KASP markers identified were accurate and polymorphic. Another study on different crops was done on sorghum by Burow *et. al*., (2019). Using variants in the *Bmr*6 and *Bmr*12 genes that control the brown midrib (bmr) phenotype, the aim was to analyse single-nucleotide variants using KASP markers and to evaluate the benefit of the DNA markers for allelic discrimination in breeding lines. Validation of 15 KASP markers in an F_2 population, showed that the markers could be used for effective marker-assisted selection for the bmr trait in sorghum.

KASP markers have been developed at the WRC for use in wheat/wild relative introgression and the KASP markers developed validated for ten wild relative species including *Th. elongatum*.

4.1.3 SNP Discovery and KASP marker development at the WRC

SNPs can be transformed into KASP assays, which are more cost-effective and efficient than Affymetrix for genotyping introgression lines at low densities. Due to wheat's polyploid nature, KASP analyses for homoeologous SNPs may struggle to differentiate between heterozygous and homozygous hybrid lines in a backcross population. The KASP markers used in this project were developed by Grewal *et al*. (2020) at the WRC. They PCR-amplified and sequenced genomic DNA from putative single-copy areas of the wheat genome and compared it to orthologous copies from various wild relatives to find co-dominant SNPs that can distinguish between heterozygotes and homozygotes.

The KASP genotyping technology was used to validate a selection of 1000 putative SNPs, of which 620 were possibly chromosome-specific and evenly spread throughout the wheat chromosomes. The target SNP sequences were given to LGC Genomics for KASP assay design, along with any annotations for chromosomal specificity. In addition, a set of 90 chromosome-nonspecific assays was developed for genotyping introgression lines Grewal *et al*. (2020).

4.1.4 KASP validation and characterization

Grewal *et al*. (2020) tested the 1000 primers for PCR amplification using genomic DNA from two wheat types, Paragon and Chinese Spring, as well as ten wild relatives of wheat, *Th. elongatum*, *Th. ponticum*, *Th. bessarabicum*, *Th. intermedium*, *T. urartu*, *Amblyopyrum muticum*, *Ae. speltoides*, *Aegilops caudata*, *T. timopheevii*, and *Secale cereale*. Introgression lines from each of these wild relative species were included in the genotyping set because they were previously genotyped on the Axiom® Wheat-Relative Genotyping Array and were known to contain segments from each of the wild relative species' linkage groups. From the validated assays, 710 markers were found polymorphic between the four main wheats and at least one wild relative species with 303 of the markers polymorphic between hexaploid wheat and *Th. elongatum* 401007 (Grewal *et al.,* 2020).

Genotyping with KASP markers enabled distinction between homozygous and heterozygous lines as well as the wheat recombined with their wild relative segment can be rapidly identified. Markers on each of the three homologous subgenomes generate a heterozygous call when a single wild relative segment from an orthologous group is present. Where the orthologous segment is homozygous in the introgression lines, the recombination markers on the wheat chromosome will call homozygous (owing to the lack of both wheat alleles), but the markers on the other two subgenomes will call heterozygous. In Figure 4.2, line replaced both short arms of wheat chromosome 1A with *Th. elongatum*. As a result, 1B and 1D chromosomes didn't replace with a wild relative segment, so they showed heterozygous calls even if the line had a homozygous introgression.

Each SNP is designed to be polymorphic between *Th. elongatum* and one of the wheat chromosomes. All three wheat chromosomes get lighting up because the markers of all linkage group 1 chromosomes detect 1E. They are still detecting both copy of 1B and 1D genomes, they are also detecting 1E which is present in the 1A genome. Thus, they all come up heterozygous initially because they will say 1E is present. Grewal *et al*. (2020) explained all these situations with a nice diagram in their KASP assay design studies.



Figure 4. 2: An example of a homozygous introgression. In this line the introgression from *Th. elongatum* has replaced both short arms of wheat chromosome 1A. Green represents homozygous, red represents heterozygous.

Some markers can detect the presence of the wild relative segment but cannot distinguish between heterozygous and homozygous lines for this segment. Because for example, if there is a heterozygous introgression from the wild relative into chromosome 2A, the result shows a heterozygous call. If there is a homozygous introgression in 2A, then due to the absence of the wheat call entirely, a homozygous call is indicated. However, if introgression has gone into 2A is homozygous and chromosome 2A KASP marker is still viewed, a heterozygous call will be seen for the 2A marker. Because it is still detecting 2A chromosome(s) and wild relative segments. However, the mc-GISH analysis validates whether it is homozygous or heterozygous. This is also explained in more detail with the examples the Chapter 5.

These markers have therefore been utilised in this work to: (i) Identify *Th. elongatum* introgressions in a wheat background (ii) Characterise, as far as possible, the introgressions identified in terms of size (iii) Track the introgressions through the self-fertilised generations.

4.2 Aims and Objectives

The main aims and objectives of this chapter are:

- To generate introgression lines containing single segments of varying sizes of *Th. elongatum* in a wheat background.
- To characterise the introgression lines using chromosome specific KASP markers and validate with mc-GISH.

4.3 Material and method

4.3.1 Development of KASP markers

4.3.1.1 Selection of markers

KASP markers were screened on a set of test plants including the wheat accessions used to make the wheat/*Th. elongatum* introgression lines (Chinese Spring and Paragon) and *Th. elongatum* 401007. Markers were removed from further analysis if they showed no call, were heterozygous for either parent or if they showed no polymorphism between the parents.

For the earliest set, 84 KASP markers for picking up each linkage group were chosen in September 2018. These markers were taken to test to make a genotype quicker because to screen as many plants possible as quickly as possible. After analysing the genotyping results, it was decided to increase the number of markers. Because genotyping results showed gaps so to fill in the gaps the number of markers needed to increase. Some segments can be missed due to a lack of markers, so the number of markers was increased.

For the next set in 2019, 203 KASP markers were tested. In those 203 markers, more information had been got about introgression lines. Then, some of the markers were developed for other wild species and tested on *Th. elongatum*. So, for Set 1, Set 2, and Set 3, 235 KASP markers were used, and for Set 4 and Set 5, 198 KASP markers were used. After many rounds of analysis, some of them were removed, and the final 189 markers were chosen as robust markers. These markers were run for 8 different PCR layouts (Appendix 2) but not always 267 markers were run for each layout. As mentioned above, in the beginning, genotyping was started with 84 markers and the number of markers was increased, and in each analysis, markers were added and removed according to genotyping results.

4.3.1.2 Genotyping analysis via a KASP Marker

In this chapter, 267 KASP markers polymorphic between hexaploid wheat and *Th. elongatum* 401007 were used, these markers were previously produced and tested by Grewal *et al.* (2020). Leaf tissue of 10 days old seedlings were collected in a 96 well plate for DNA extraction. Paragon and *Th.*

elongatum were used as a control. Genotyping reactions were performed as described in below using a ProFlex PCR system (Applied Biosystems by Life Technology).

DNA extraction and quantification and preparation of the DNA samples for use in the sc-GISH, mc-GISH and KASP genotyping was as described in chapter 2, (section 2.1.). DNA was diluted (1:30) and transferred to a 384 PCR plate prepared with a PIPETMAX® (Automated Liquid Handling Platform) (Gilson, UK). One well was filled with de-ionised water as a control. The plates were centrifuged for 1 minute at 1000 rpm. The plates were then dried at 80°C for 20-30 min and 5µl RCR master mix added to each well. The PCR master mix is shown below:

Reagent	**1x**	**384x**
dH_2O	2.43µl	653.8 µl
Rox PACE mix (see Appendix j)	2.50µl	672 µl
74x Primers (see Appendix 2)	0.068µl	18.2 µl
Total volumes	5µl	1344 µl

PCRs were carried out using the ProFlex PCR system (Applied Biosystems by Life Technology) with the following programme:

Stage	**Step**	**Temperature (°C)**	**Time (minutes)**	**No. of cycles**
Stage 1	Denaturation	94°C	15	1x
Stage 2		94°C	0.25	10x
	Annealing	65°C	1	
Stage 3	Elongation	94°C	0.25	30x
		57°C	1	
Stage 4	Final hold	4°C		1x

The PCR results were analysed with a QuantStudio™ 5 Real-Time PCR System and QuantStudio™ Design and Analysis Software V1.5.0 (Applied Biosystems, California).

4.3.2 Cytogenetic analysis of the Wheat/*Th. elongatum* introgression lines

Wheat/*Th. elongatum* introgression plants produced from the crossing programme (Figure 3.2) were characterised using sc-GISH and mc-GISH to determine the genomic constitution of the *Th. elongatum* chromosome(s) present. Root tips were collected after seed germination for metaphase chromosome spreads and the sc-GISH and mc-GISH protocol applied as described in chapter 2 (section 2.2.). Genomic DNA from *Th. elongatum*, *T. urartu*, *Ae. speltoides*, and *Ae. tauschii* were extracted to make probes. Genomic DNA from *Th. elongatum* 401007 was labelled by nick translation with Chroma Tide Alexa Fluor 546-14-dUTP (Thermo Fisher Scientific (Invitrogen; C11401), Waltham, MA, United States). Genomic DNA from *T. urartu* was labelled by nick translation with Chroma Tide™ Alexa Fluor™ 488-5-dUTP (Thermo Fisher Scientific (Invitrogen; C11397), Waltham, MA, United States), *Ae. speltoides* was labelled by nick translation with DEAC-dUTP (Jena Bioscience; NU-803-DEAC), Ae. tauschii was labelled with nick translation with Chroma Tide Alexa Fluor 594-5-dUTP (Thermo Fisher Scientific (Invitrogen; C11400), Waltham, MA, United States).

4.4 Result

4.4.1 Development of KASP markers

One thousand KASP markers had previously been generated at the WRC with the aim of using the markers to detect introgressions from 10 different wild relatives into wheat. When tested on wheat and the wild relatives a panel of 710 wheat chromosome-specific markers were selected, with each marker polymorphic between at least one wild relative and wheat (Grewal *et al*., 2020) with 303 of the markers polymorphic between common wheat and *Th. elongatum* 401007 (Grewal *et al.,* 2020).

Previous work by Dr Lauren Baker suggested some of the 303 markers should be removed as they did not work/did not give a clear signal. Therefore, after examination of her genotyping results 35 KASP markers were not selected for use in the work described her and genotyping started with 84 KASP markers to get quick results. After analysing the genotyping results, in order to fill in the gaps the number of markers increased and 203 KASP markers were used for the second set in 2019. Then, for Set 1, Set 2, and Set 3, 235 KASP markers were used and for Set 4 and Set 5, 198 KASP markers were used. To sum up, in each analysis, markers were added and removed according to genotyping results. In total 267 KASP markers were used to genotype the *Th. elongatum* chromosome segments present in the introgression lines and after many rounds of genotyping 189 markers were able to distinguish the wheat and *Th. elongatum* alleles, thereby detecting introgression plants (Table 4.2).

4.4.2 Genotyping of wheat-*Th. elongatum* introgression lines with KASP markers

A further 78 markers were removed after rounds of analysis because they could not generate a useful amplification signal, i.e., these markers worked only on the wheat varieties or were not polymorphic with *Th. elongatum*. Thus, out of the 267 chromosome-specific KASP markers tested, 189 markers were found to be polymorphic between wheat and the *Th. elongatum* accession used in this work (Table 4.2). Table 4.2 also shows the spread of the KASP markers across the wheat genome.

Table 4. 2: The number of KASP markers tested, and the number of KASP markers that worked to detect *Th. elongatum* introgressions for each wheat chromosome.

Chromosome	KASP markers used to detect *Th. elongatum*	KASP markers work to *Th. elongatum*
1A	9	5
1B	17	10
1D	13	10
2A	12	8
2B	19	12
2D	10	8
3A	8	6
3B	16	13
3D	14	10
4A	17	8
4B	7	3
4D	15	13
5A	11	11
5B	12	10
5D	18	11
6A	11	9
6B	13	13
6D	13	8
7A	10	5
7B	14	11
7D	8	5
Total	267	189

The 267 chromosome-specific KASP markers used for the analysis were visualized across each wheat chromosome with RIdeogram_0.2.2 (Figure 4.3). The distribution of the 189 working KASP marker are shown as purple circles across the seven linkage groups, while the physical locations of the 78 eliminated KASP markers are shown as a red triangle.

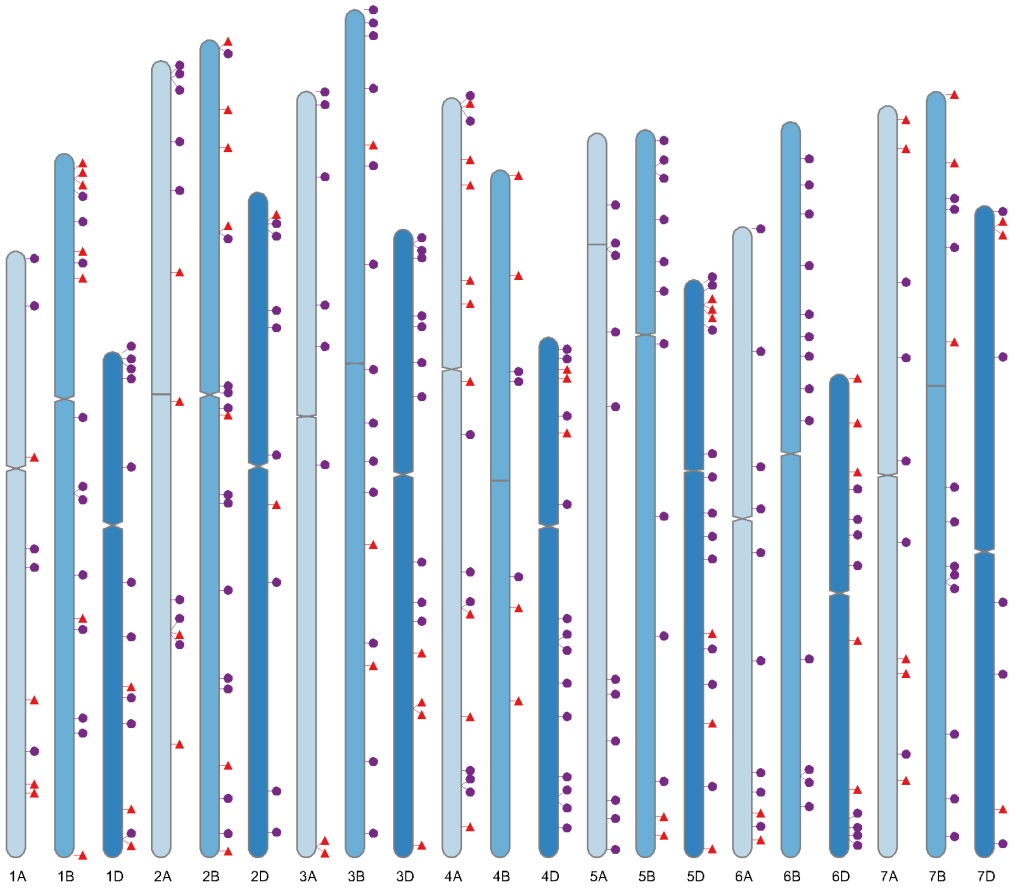


Figure 4. 3: The distribution of the 267 KASP markers across the 21 wheat chromosomes (Chr1A-7D). Figure created using RIdeogram_0.2.2. The 189 working KASP markers are shown with purple circles. The 78 KASP markers which did not work are shown with red triangles.

The elimination of some of the KASP markers that were not working properly markers results in the markers not being evenly distributed throughout the 21 chromosomes. The fewest KASP markers were tested on

chromosome 4B (7), and testing resulted in the removal of four of the markers leaving only three KASP markers polymorphic with *Th. elongatum.* Chromosomes 1A, 7A and 7D had the next fewest with five markers that were polymorphic with *Th. elongatum*, while 6 KASP markers produced a useful amplification signal for chromosome 3A (Table 4.2, Figure 4.3). With 13 markers, chromosomes 3B and 6B possessed the most working KASP markers, followed by 2B, 4D, and 5D chromosomes with 12 markers each. Although the 6B chromosome has 13 KASP markers, Figure 4.3 showed that more markers on the top of the long arm were required. Despite the fact that some of the chromosomes have the same number of markers, some of them need more markers to fill the gaps as some markers were well dispersed, while others tended to be clustered. For example, chromosomes 1B, 1D, 3D and 5B each have 10 markers. The 10 markers were well distributed on chromosome 1D and therefore marker coverage was enough to detect the introgressed segments. However, chromosomes 1B, 3D and 5B all have significant gaps between the markers in both the short and long arms for 1B and 3D and the long arm for chromosome 5B. Likewise, chromosomes 2A, 2D, 4A, and 6D all had eight markers with a good distribution over chromosome 2D but with markers more clustered on the other three chromosomes.

The 267 KASP markers were used to detect *Th. elongatum* segments in a wheat background. To ensure that the KASP markers were robust for use in breeding programmes, all the KASP assays were also used to genotype 796 backcrossed and self-fertilised lines (Table 4.3). In addition to the KASP marker genotyping, GISH analyses were also carried out to confirm the genotyping results as the GISH analysis was able to visualise the presence of a *Th. elongatum* segment (Table 4.3). The genotyping and GISH analyses showed that 530 lines from the 796 tested contained one or more introgressions from *Th. elongatum*.

Table 4. 3: The number of plants of each generation genotyped and GISHed and the number of plants with one or more introgressions from *Th. elongatum.*

Generation	Number of plants genotyped	Number of plants GISHed	Plants with *Th. elongatum introgressions*	Total number of plants germinated
BC_2F_2	5	2	4	5
BC_2F_3	13	9	12	13
BC_3F_1	28	10	25	28
BC_3F_2	34	8	24	36
BC_3F_3	9	2	9	9
BC_4	11	6	8	11
BC_4F_1	69	19	31	70
BC_4F_2	280	40	187	280
BC_4F_3	37	14	35	37
BC_4F_4	4	1	4	4
BC_5	58	10	38	59
BC_5F_1	134	52	80	135
BC_5F_2	85	20	45	85
BC_5F_3	17	3	17	17
BC_5F_4	4	2	4	4
BC_6	8	4	7	8
Total	796	202	530	801

4.4.3 Combining the KASP marker analysis with GISH analysis of wheat-*Th. elongatum* introgression lines

GISH analysis is relatively lower throughput than the KASP marker analyses. However, it provides a very useful tool to not only validate the marker work but also, on occasion, to help work out what the markers are actually characterising. This is particularly true in wild relative introgression work with a complex species such as the decaploid accession of *Th. elongatum* used in this work. Marker analysis enabled genotypes containing introgressions from *Th. elongatum* to be distinguished from those genotypes with no *Th. elongatum* segments present. Genotypes containing introgressions were then selected for GISH analysis. Root metaphase spreads were made for a total of 202 out of the 233 wheat-*Th. elongatum* self-fertilizing and backcrossed populations were selected for GISH analysis (Table 4.3). The GISH analysis allowed the detection of all wheat chromosomes as well as the *Th. elongatum* introgressions. GISH analysis was not carried out on the 266 lines where marker analysis showed no *Th. elongatum* introgressions present.

In this chapter, selected sc-GISH analyses performed by Dr Lauren Baker were also used as a guide to show lines containing several *Th. elongatum* introgressions and hence lines where further crossing in order to reduce the number of *Th. elongatum* introgressions per plant was required, i.e., the segregation of segments into different individual plants.

A relatively small number of KASP markers were available at the start of this project. However, more markers became available through the course of the project. Therefore, in the genotyping part of the figures below (Figure 4.4 to 4.7), the positions where no KASP markers were available at the time of genotyping is shown in white. Coloured regions represent different genotype calls. All heterozygous calls are shown in red while wheat calls are shown in blue. Differences between genotyping for related plants are therefore due to the increased number of markers available for the genotyping of the later generations.

Figures 4.4 to 4.7 also show the GISH analyses of plants from different generations of four different wheat/*Th. elongatum* introgression lines. These figures illustrate the separation of different *Th. elongatum* chromosomes/introgressions into different plants through the crossing programme.

Figure 4.4A-D shows the analysis of a BC_4 parent (which was both backcrossed to wheat and self-fertilised), one BC_4F_1 offspring and two BC_5 offspring. The sc-GISH image of BC_4-90E lines (from Dr Lauren Baker) showed the presence of three whole *Th. elongatum* chromosomes and two introgressions. This was partially supported by the initial KASP analysis which suggested one 2E, 5E and one 7E whole chromosomes and an introgression from group 3E. GISH analysis is not available for BC_4F_1-96-2 although the genotyping result suggested the same whole chromosomes and introgression from groups 2E, 5E and 3E. The genotyping of BC_4F_1-96-2 also showed that this plant had lost the whole 7E chromosome that was present in BC_4-90E. Genotyping with KASP markers showed a single introgression from 3E in BC_5-104-6 and one whole *Th. elongatum* group 5 chromosome in BC_5-104-3. Mc-GISH analysis of these two plants validated the genotyping in both cases while demonstrating that the introgression present in BC_5-104-6 was actually a centric fusion between 3E and 3A. Additional KASP markers were used to genotype the two BC_5 plants compared to the BC4 and BC_4F_1 plants giving fewer gaps along the chromosomes.

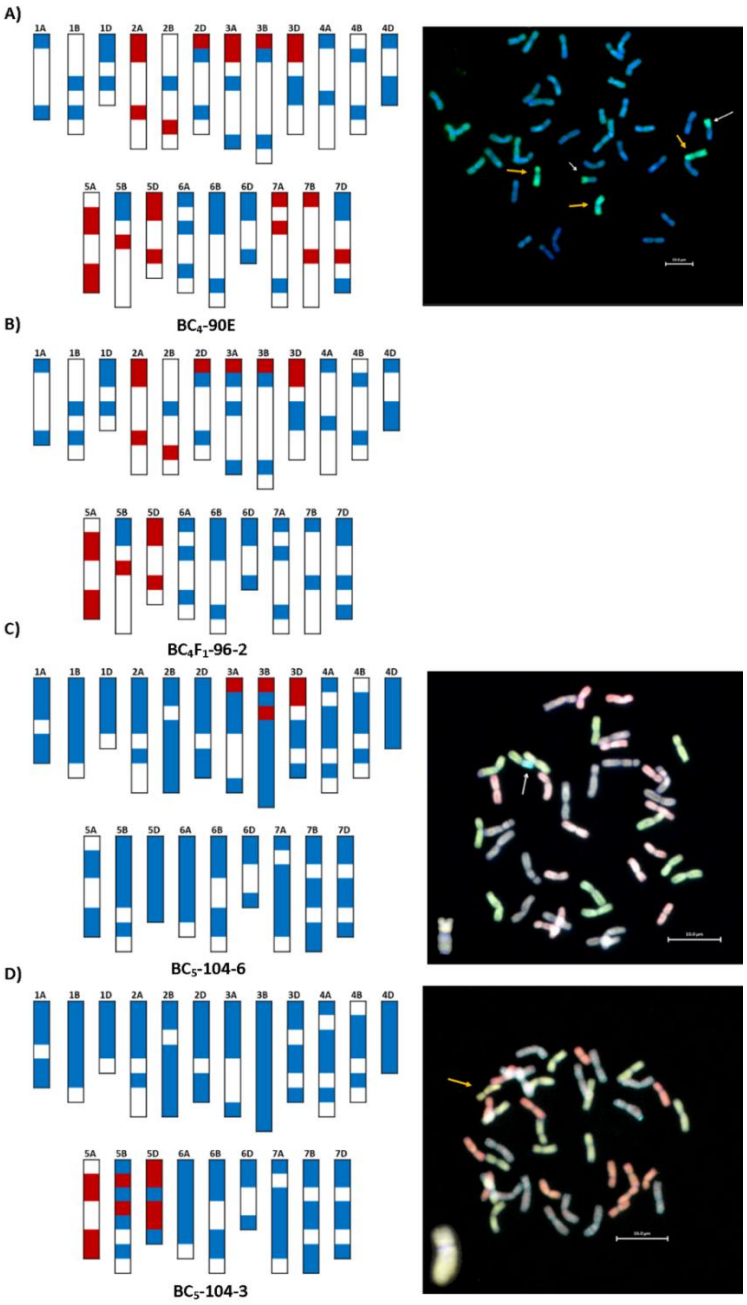


Figure 4. 4: Molecular characterization of wheat–*Th. elongatum* introgression lines showing genotyping data (left), using KASP markers, and sc-GISH and mc-GISH analysis (right). The A-genome chromosomes are shown in green, the B-genome chromosomes in blue, the D-genome chromosomes in red and E-genome chromosome/introgressions in gold. Whole introgressed chromosomes are labelled using a yellow arrow while segments are labelled with a white arrow. All heterozygous calls are shown in red while wheat calls are shown in blue and where no KASP markers were available at the time of genotyping is shown in white.

Figure 4.5 (A and B) shows the genotyping and GISH analysis of a BC_5 plant and a BC_5F_2 plant generated by two rounds of self-fertilisation. Marker analysis of BC_5-74C showed the presence of a whole group 2E chromosome and an introgression from 3E confirming the sc-GISH analysis carried out by Dr Lauren Baker. Genotyping of two plants from the BC_5F_1 generation showed them to have exactly the same *Th. elongatum* chromosomes/chromosome segments present as the BC_5. However, the next generation, BC_5F_2-10-8 was found to have lost the whole 2E chromosome from *Th. elongatum* while still carrying the introgressed segment identified as a centric fusion 3ES.3AL.

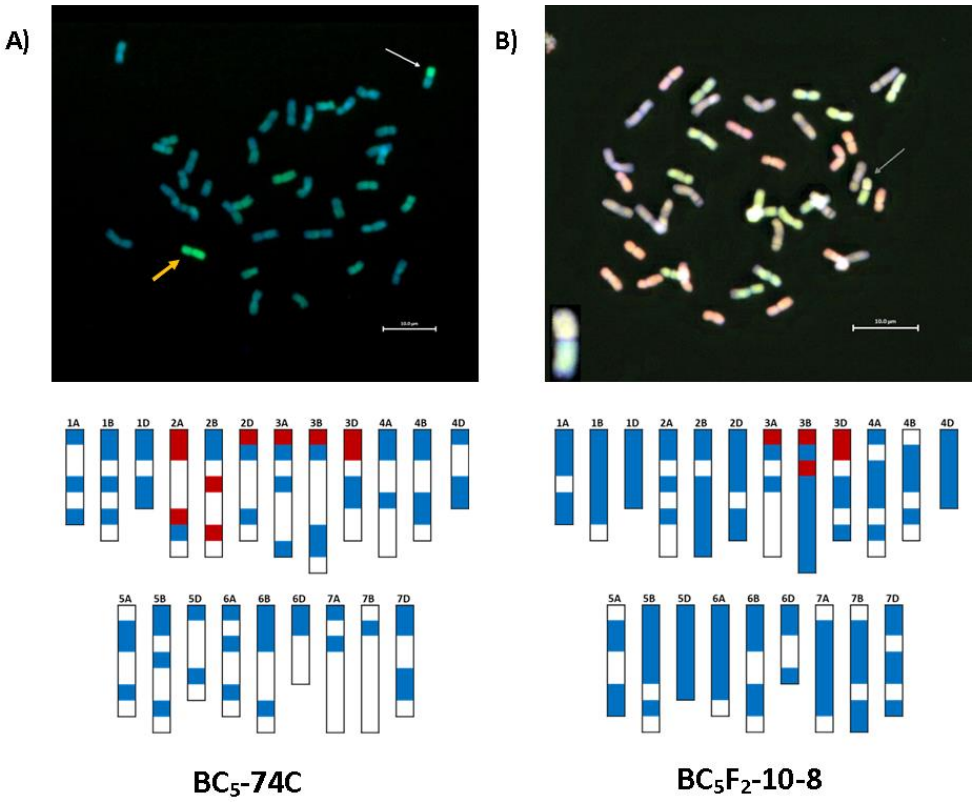


Figure 4. 5*:* Molecular analysis (KASP genotyping and GISH) for (A) a BC_5 plant and (B) one of its subsequent BC_5F_2 offspring showing the loss of a whole *Th. elongatum* chromosome (yellow arrow) in BC_5F_2-10-8 as compared to BC5-74C. (Sc-GISH for BC_5-74E from Dr Lauren Baker). Both plants carry the centric fusion 3ES.3AL (white arrows). In (B), the A-genome is shown in green, the B-genome in blue, the D-genome in red and the E-genome in gold. In the coloured bars, all heterozygous calls are shown in red, while all wheat calls are shown in blue, and white spaces indicate where there are no markers.

The original genotyping result for BC_4-118E (Figure 4.6A) indicated that it contained three whole chromosomes from groups 2, 3 and 5 from *Th. elongatum* and one introgression from linkage group 7E, a result confirmed with sc-GISH (Figure 4.6A). The KASP marker analysis was not conclusive as some markers showed a wheat call instead of the expected *Th. elongatum* call which might have been the result of recombined chromosomes rather than whole chromosomes, e.g., the top marker on group 2A (Figure 4.6A). The KASP marker results for BC_4F_1-88-1 (no GISH analysis available) showed that this plant had lost both the whole 2E and 3E chromosomes (Figure 4.6A). BC_4F_1-88-1 was then both back crossed to Paragon and also self-fertilised in order to separate the remaining E-genome chromosomes/segments. The analysis for three of the BC_5 plants (Figure 6B) shows the different segregation of segments (Figure 6B) in the three plants. BC_5-99-8 has the same whole chromosome (5E) and segment (7E) as its parent BC_4F_1-88-1. The other two BC_5 plants, however, had each lost one of the *Th. elongatum* chromosomes. BC_5-99-9 had lost the whole 5E chromosome while retaining the introgression segment from group 7 (subsequently validated by mc-GISH). BC_5-99-10 had lost the group 7 introgression while retaining the whole group 5 chromosome. In the self-fertilized plants (Figure 4.6C) the same segregation patterns were seen as in the BC_5 plants with BC_4F_1-43-8 containing both segments, BC_4F_1-43-1 containing the group 7 chromosome segment and BC_4F_1-43-2 containing the whole chromosome from group 5. This result demonstrates that both backcrossing and self-fertilisation can be used to separate chromosomes/introgressions from a wild relative in a wheat background into different plants for use in breeding programmes (Figure 4.6A-C).

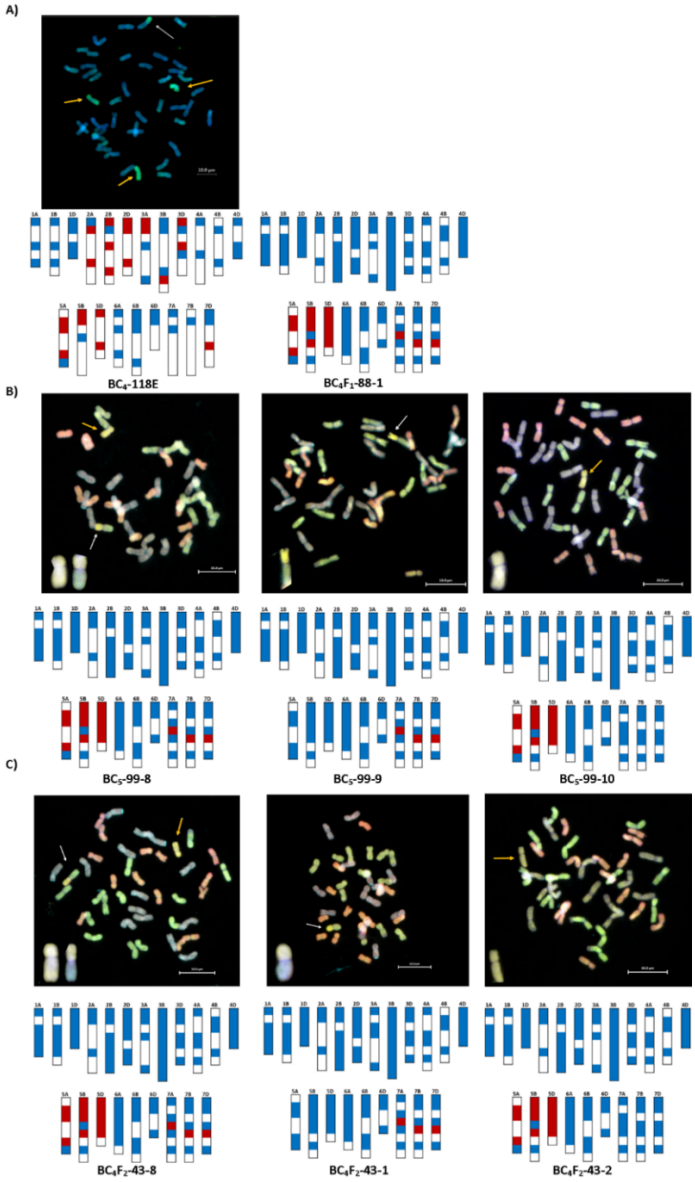


Figure 4. 6: Sc-GISH and mc-GISH images of metaphase spreads and KASP genotyping data of 8 wheat-*Th. elongatum* introgression lines. (A) BC_4-118E (image from Dr Lauren Baker) and its self-fertilised progeny BC_4F_1-88-1 (no GISH is available for BC_4F_1-88-1). (B) Three of the BC5 progeny produced by crossing BC_4F_1-88-1 (C) Three of the BC_4F_2 progeny produced by self-fertilising BC_4F_1-88-1. In the mc-GISH images, the A-genome is shown in green, the B-genome in blue, the D-genome in red and the E-genome in gold while whole *Th. elongatum* chromosomes are shown with a yellow arrow and wheat- *Th. elongatum* introgressions with a white arrow. The coloured bars represent genotyping result, wheat is shown in blue, heterozygous is shown blue and white spaces indicate where there are no markers.

Figure 4.7 (A-C) shows the KASP genotyping and GISH analyses of three generations of plants starting with BC_5-83A. The KASP genotyping for BC_5-83A showed the presence of one wheat-*Th. elongatum* introgression from group 6 and a whole genome *Th. elongatum* identified as a group 3 chromosome. Confirmation of the genotyping was made using the sc-GISH result (from Dr Lauren Baker). BC_5F_1-12-1 was found to have the same segments, with this result also subsequently validated by sc-GISH. Analysis of two BC_5F_2 plants (7-3 and 7-4) showed that BC_5F_2-7-3 had lost the introgression from group 6 while retaining the whole group 3 chromosome and BC_5F_2-7-4 had retained the group 6 introgression, but the group 3 chromosome had broken at the centromere and only the short arm was retained in this plant.

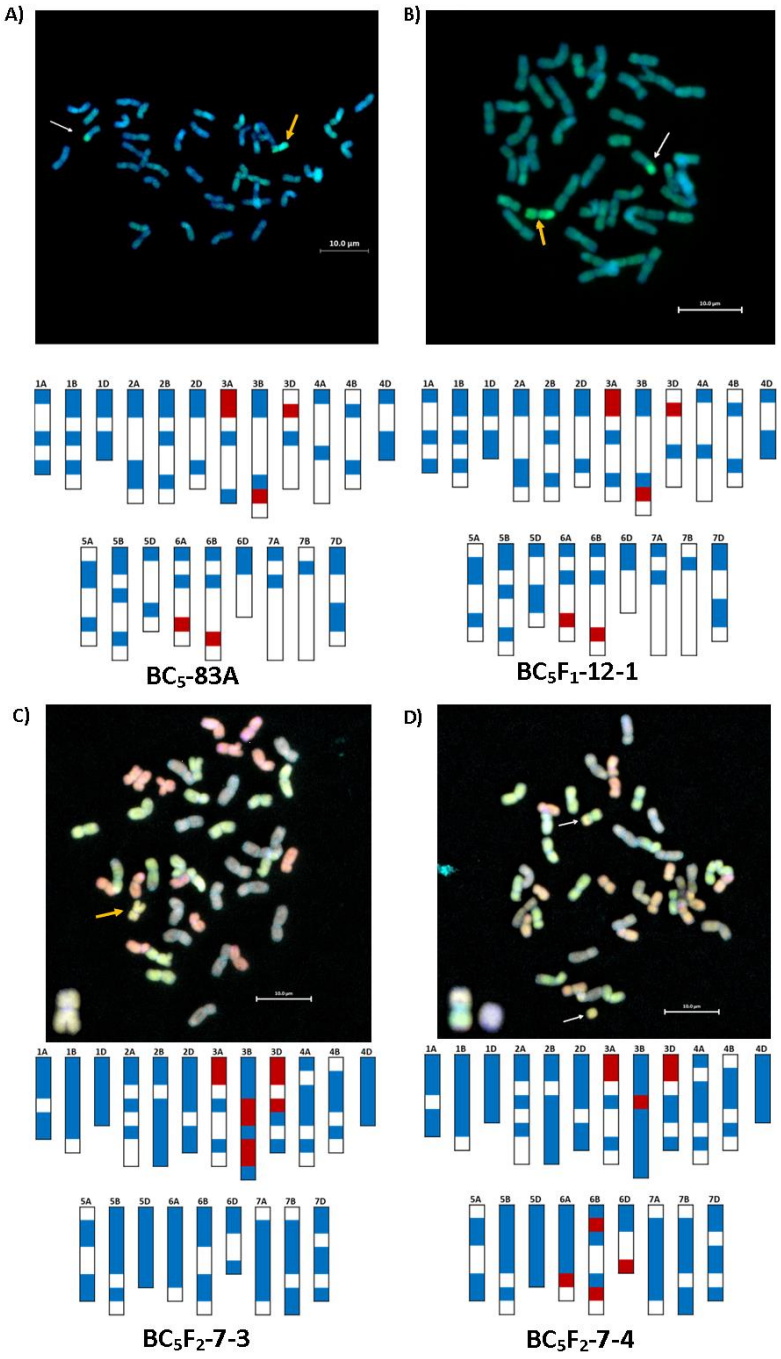


Figure 4. 7: KASP genotyping and sc-GISH/mc-GISH analyses of 3 generations of wheat–*Th. elongatum* introgression lines. (A) BC_5-85A, (B) BC_5F_1-12-1, (C) BC_5F_2-7-3 and (D) BC_5F_2-7-4. The A-genome is shown in green, the B-genome in blue, the D-genome in red and the E-genome in gold. Whole *Th. elongatum* chromosomes are shown with a yellow arrow and wheat- *Th. elongatum* introgressions with a white arrow. The coloured bars represent genotyping result. All heterozygous calls are shown in red, all wheat calls are shown in blue, and white spaces indicate where there are no markers.

Figures 4.4 to 4.7 above show examples of the segregation of different *Th. elongatum* segments via both back crossing and self-fertilisation. Many more introgressions from *Th. elongatum* were actually generated in this programme than those shown in these figures. Several of the introgressions generated are now homozygous and thus stable (see chapter 5) enabling the lines to be multiplied and distributed. However, many introgressions are still heterozygous and hence further work will be needed to generate homozygous lines from these introgressions. The introgressions generated from each of the linkage groups of *Th. elongatum* are shown below (Figure 4.8) although this is not a complete list as only those introgressions characterised by both KASP markers and GISH are shown. In the genotyping shown for each introgression, the KASP markers have been ordered against the published *Th. elongatum* sequence instead of against the wheat sequence as for figures 4.4 to 4.7.

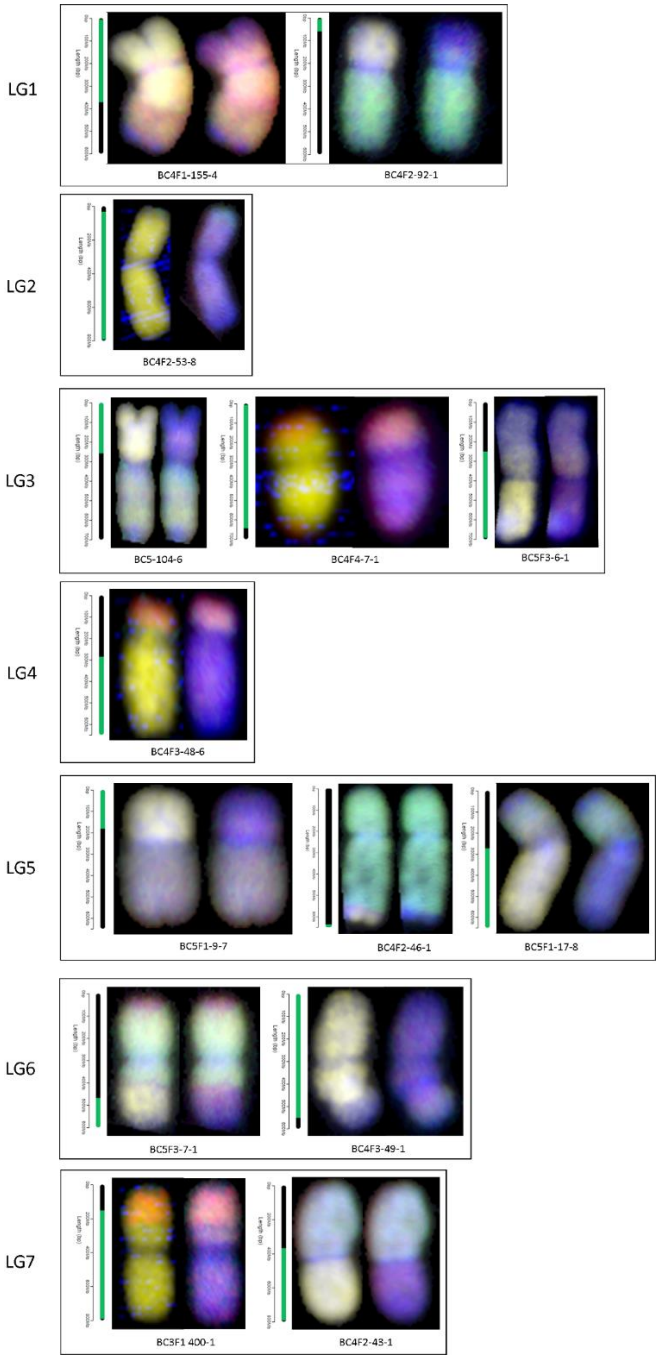


Figure 4. 8: Characterization of wheat-*Th. elongatum* introgression segments using mc-GISH and KASP markers from seven linkage groups. In all segments, two images are shown; the first one is merged colour (the yellow represents *Th. elongatum*) and the second one is Red, Green, and Blue (RGB) colour (the purple represents *Th. elongatum*). In both images, green represents the A genome, grey/blue represents the B genome, and red represents the D genome of wheat. In the genotyping result, black represents the wheat allele and green represents a heterozygous call for the *Th. elongatum* allele.

To sum up, mc-GISH analysis was performed in order to further characterisation and confirmation of the introgressions being detected from KASP markers in this project. While the performed mc-GISH to examine whether the *Th. elongatum* genomes were transferred into wheat, total genomic DNA probes of the E genomes were utilized. In addition, the A, B, and D genomic DNA probes were utilised to define which wheat genomes the *Th. elongatum* introgressions were recombining with.

4.5 Discussion

In this work, we have utilised the chromosome-specific KASP markers and mc-GISH for the characterisation of the wheat/*Th. elongatum* introgression lines. The 303 markers used were selected from a panel of 710 KASP markers polymorphic between wheat and all ten wild relatives by the Nottingham WRC (Grewal *et al*., 2020). 267 KASP markers were finally used, as previous work suggested removing 36 markers.

There are a number of previous studies using different types of molecular markers designed to characterize *Th. elongatum* chromosomes/introgressions in a wheat background such as microsatellite markers or SSR markers (You *et al*., 2003), expressed sequence tag-derived simple sequence repeat markers (EST-SSR) markers (Mullan *et al*., 2005) PCR-based landmark unique genes (PLUG) (Hu *et al*., 2012), sequence characterized amplified region (SCAR), conserved orthologous set (COS) markers (Gaál *et al*., 2018) and SNP (Lou *et al*., 2017) and specific length amplified fragment sequencing (SLAF-seq) markers (Chen *et al*., 2013). However, the above studies mainly concentrated on diploid *Th. elongatum* and the application of markers is still limited for breeding using decaploid *Th. elongatum*.

The early generations of the wheat/*Th. elongatum* introgression lines were originally genotyped with the Axiom® Wheat Relative Genotyping Array by Dr Lauren Baker. However, subsequent generations produced in this study have been characterised with chromosome-specific KASP markers. KASP offers a cost-effective system, with flexibility and efficiency for genotyping in wheat. One of the main advantages of the chromosome-specific KASP markers is that in the backcrossed and self-fertilized populations the markers not only detect the *Th. elongatum* introgression in the wheat background but also whether the introgression is homozygous or heterozygous and which wheat chromosome has recombined with *Th. elongatum*. SNPs can be converted into KASP markers, although identification and characterization of SNPs are especially challenging in polyploid species due to the existence of similar homoeologous and paralogous genome sequences (Makhoul *et al*., 2020). A study by Rasheed *et*

al. (2016) successfully converted a panel of PCR markers for genes in wheat into KASP markers. Grewal *et al*. (2018a, b) also reported the successful conversion of array-based, putative co-dominant SNPs between a variety of wild relatives and wheat to KASP assays.

However, genotyping the wheat/*Th. elongatum* introgression lines in this project was challenging for two main reasons: (1) Accession 401007 is a decaploid accession of *Th. elongatum* with additional aneuploidy. (2) Decaploid *Th. elongatum* is an allopolyploid which appears to have undergone further re-arrangements within the polyploid.

Accession 401007 of *Th. elongatum* contains five separate genomes and therefore, five copies of each linkage group. The cluster calling by the Quant Studio programme was checked manually in this programme as multiple copies of each linkage group can often result in incorrect KASP cluster calling. KASP markers are expected to separate into three clusters, i.e., one cluster for the wheat genome, one cluster for the *Th. elongatum* genome and one cluster for heterozygotes between the two genomes. However, incorrect calls observed included the wrong number of clusters caused partly by the decaploid genome. Heterozygous or homozygous lines were also found to produce unexpected KASP clusters. A case of incorrect clustering, frequently reported in the literature, is a hemi-SNP, in which all homozygous calls for one allele cluster in one 'heterozygous' cluster due to the binding of the common primer to more than one sub-genome (Makhoul *et al.,* 2020). False KASP clustering calls might also explain why some of the markers can detect the presence of the wild relative genome but cannot distinguish between heterozygous and homozygous lines for a segment in a wheat background. Makhoul *et al*. (2020) recently reported the same false clustering calls problem. The LGC group have published a KASP troubleshooting guide where they suggest two possible reasons: (1) Contamination in the KASP assay mix or DNA samples, (2) polymorphism in the primer binding site. In this work, some genotyping was carried out twice on two different samples with the same result, suggesting contamination of the KASP assay mix or DNA samples was not the problem.

In the genotyping analysis some markers were found to detect introgressions in some families but were not able to detect introgressions in other families. This was probably due to the different genomes present in the decaploid *Th. elongatum*. There are five genomes present in decaploid *Th. elongatum* and some of the genomes are definitely different, but it is possible for two genomes to be the same, for example. The real genome constitution is not clear for the decaploid *Th. elongatum*. For example, some markers did not work for the 3E chromosome in Figure 4.7 while those markers worked for the 3E chromosome in Figure 4.5. When markers A/T were run wheat and Th elongatum accession for example A for wheat and T for *Th. elongatum*, but it is unknown whether the T allele is on all five genomes of the *Th. elongatum* or two of them or four of them. Thus, for Figure 4.5, 3E is still there, just for those markers in that genome, the *Th. elongatum* allele just depends on which 3E comes from, so it works differently in different families. Because the same homoeologous group in *Th. elongatum* can have variation and only be between SNP present/absence.

As outlined in the Introduction to this chapter, there is considerable confusion in the literature concerning both the distinction between decaploid *Th. elongatum* and a second decaploid species *Th. ponticum* and also their genomic composition. Baker *et al*. (2020) concluded from mc-GISH analysis that the genome composition of accession 401007 was very similar to that found for *Th. ponticum* by Molnar-Lang and Kruppa (2016) and was actually $EEEEEEE^{St}E^{St}E^{St}E^{St}$ or $JJJJJJ^{St}J^{St}J^{St}J^{St}$ although it is also not clear if the E/J genomes are autopolyploid or allopolyploid as suggested by Zhang *et al*. (1996) who concluded the genome composition to be $E^{e}E^{e}E^{b}E^{b}E^{x}E^{x}StStStSt$. The mc-GISH analysis showed that a high level of translocation had occurred between the genomes of *Th. bessarabicum/Th. elongatum* and *P. strigosa*. Out of the 74 chromosomes, 32 chromosomes were observed to have the St genome at the centromere while the other 42 chromosomes either did not show any St genome or a very weak non-centromeric St genome fluorescence (see Figure 4.1) (Baker *et al*, 2020). The KASP markers would not all be expected to work

for all the genomes present in *Th. elongatum* 401007. For example, an introgression from linkage group 1 could be from the E or E^{St} genomes or even the E^{b}, E^{e}, E^{x} or E^{St} genomes. In future work, it would be possible to show this if genotyping and mc-GISH were carried out on lines containing 1 whole chromosome. The mc-GISH would show any E^{St} regions present on the chromosome and the markers could then be specifically assigned to those regions.

As shown in Table 4.2, 267 chromosome-specific KASP markers were able to detect 530 lines out of 796 that contained introgressions and/or whole chromosomes of *Th. elongatum* across all seven linkage groups. The markers were not evenly distributed throughout the 21 chromosomes with significant gaps on most of the wheat chromosomes. These gaps were made more extensive due to the removal of some of the markers as early genotyping analyses showed they were not reliable. Table 4.1 shows a comparison of the number of KASP markers tested on the wheat/*Th. elongatum* introgression lines and the number that were found to be robust in genotyping the introgression lines. In total, 189 markers out of the 267 tested were found to detect the *Th. elongatum* genome in a wheat background. The significant gaps in the KASP marker distribution meant that while larger introgressions and whole chromosome additions/substitutions were detected by the markers it was possible that smaller segments were not being detected. With the development of additional KASP markers, the genotyping analysis was able to detect smaller introgression segments in some regions of the genome. However, Figure 4.3, which shows the physical locations of the 189 working KASP markers and the 78 removed markers, does indicate that more markers are still needed to fill the remaining gaps on the linkage groups. This applies particularly to chromosomes 2A, 3A, 3D, 4A, 4B, 6B, 6D and 7A. For example, there is only one marker on the long arm of chromosome 3A relatively close to the centromere which would result in even large introgressions in this arm being missed by the genotyping. The spread of markers on chromosomes 1D,2B, 2D, 3B, 4D, 5A, 5B, 5D, and 6A, however, is better than on the

chromosomes listed above but even on these chromosomes a higher density of markers would provide increased confidence in the completeness of introgression identification.

In order to validate the genotyping data, lines with red representing heterozygous segments *Th. elongatum* alleles, were also analysed with mc-GISH, and the KASP marker results were compared with the cytogenetic results (see Figure 4.4 to 4.7). Two hundred and two of the 530 plants identified as containing introgressions via genotyping were validated with GISH analysis. Although this appears to be under half, it was possible to reduce the number of plants undergoing GISH analysis by selecting one plant from a set of sister lines where they showed the same genotyping. GISH was also not used on progeny when genotyping and GISH had confirmed the segments in a parent plant. For example, if genotyping of the parent plant and offspring showed that although the offspring had lost one or more segments compared to the parent, the offspring still had more than one segment present, the offspring was immediately backcrossed or self-fertilised to reduce the number of segments present. GISH is technically demanding and of much lower throughput than genotyping, and hence the need to reduce as far as possible the number of plants being analysed. A comparison of the genotyping and cytogenetic analyses showed that they matched for most of the introgression lines. Where GISH showed additional segments to be present, the introgression was usually relatively small.

The genotyping initially showed that 266 plants did not contain any *Th. elongatum* chromosomes/introgressions and therefore these plants were not taken forward. However, due to the gaps in the KASP genotyping, particularly for the earlier generations, if time had allowed, selecting a number of these lines at random for GISH analysis might have identified the presence of smaller introgressions not detected by the genotyping. With the increased density of markers, it would be interesting to germinate progeny from some of these lines to see if introgressions were, in fact, missed.

In some lines, genotyping suggested that a segment was heterozygous while GISH showed it to be homozygous and vice versa. Therefore, mc-GISH was used for a second reason in this study and that was for the simultaneous visualization of the A, B and D chromosomes of wheat as well as the E chromosomes from *Th. elongatum* to establish the chromosome composition of the different lines. Thus, the number of A, B and D genome chromosomes could also be counted for each line. This was important as in those cases where the genotyping called an introgression as heterozygous, but the GISH showed the segment to be homozygous, the mc-GISH usually showed extra wheat chromosomes present in the background. Although mc-GISH cannot identify individual chromosomes, it was assumed that the extra wheat chromosomes belonged to the same linkage group as the introgressions resulting in the heterozygous genotyping call. In those cases where the genotyping showed a homozygous introgression, but the GISH showed a heterozygous segment, mc-GISH usually showed a missing wheat chromosome(s), again assumed to be from the same linkage group as the introgression.

Visualisation of the wheat chromosomes was also necessary as it is important to have a homozygous introgression in a normal wheat background. Missing wheat chromosomes can result in essential genes being missing. This in turn can result in weak or even inviable plants in which the true value of the introgression cannot be assessed. For example, the *Ph1* gene, which controls homologous pairing, is located on the long arm of chromosome 5B. Therefore, in plants without chromosome 5B, homoeologous chromosomes pairing can occur (Sears, 1976) causing inter-genomic translocations, which in turn can lead to problems at meiosis.

In conclusion, although there were problems with the genotyping, the KASP markers showed the potential to identify the presence of heterozygous and homozygous *Th. elongatum* segments in a wheat background. KASP markers were found to be more flexible than Affymetrix as a genotyping platform. The KASP technology converts SNPs from fixed chip systems to a stand-alone format, allowing hundreds to thousands of samples to be

genotyped with fewer markers as well as the flexibility to customise the genotyping run with various combinations of SNP markers and sample numbers. The genotyping and GISH analyses were able to identify plants of interest containing whole chromosomes or segments from *Th. elongatum*. These plants were taken forward to reduce the number of introgressions and also to obtain plants homozygous for the introgressions which would be stable for multiplication.

Chapter 5 Cytogenetic and Molecular Characterization of wheat-*Th. elongatum* homozygous lines

5.1 Introduction

In pre-breeding programmes, many wild relatives are utilised to improve wheat genetic diversity. Introgression of new genes from progenitors and related species has been frequently utilised to enhance bread wheat agronomic traits. However, this strategy has been hindered by a lack of markers (Winfield *et al*., 2016). For example, SSR markers enable the detection of introgression segments (Zhao *et al*., 2013) but are tedious and time-consuming to utilise and therefore, they have limited use in wheat breeding (Grewal *et al.,* 2020). SNP markers, on the other hand, are now commonplace in the genotyping of wheat (Akhunov *et al*., 2009; van Poecke *et al*., 2013).

Wheat breeders need tools that can help them detect and characterise wild relative segments in a wheat background. Breeders may carry out this procedure with the use of molecular markers. However, developing molecular markers in bread wheat has been challenging until recently, owing to the existence of homoeologous and paralogous copies of the different genes (Barker and Edwards, 2009; Kaur, Francki and Forster, 2012). Allen *et al*. (2011) previously employed Next Generation Sequencing technology to find and map relatively large numbers of gene-based SNP loci to overcome the many bottlenecks and issues associated with SNP production, characterization, and, most critically, validation in wheat.

The Axiom® Wheat-Relative Genotyping SNP Array was created using genome-specific SNPs from exome sequencing of different wild wheat species (Winfield *et al*., 2016; King *et al*., 2017). The array was then used to identify introgressions from a variety of wild relatives in a wheat background (King *et al*., 2018; Grewal *et al*., 2018a; Grewal *et al*., 2018b; Devi *et al*., 2019; Cseh *et al*., 2019), and they were previously used in this project by Baker *et al*., (2020). Recently, Grewal *et al*. (2020; 2022) addressed the fact that these markers are unable to effectually distinguish between homozygous and heterozygous

individuals, despite the fact that these genotyping arrays have the potential to be very high-throughput and efficient.

For genotyping of introgression lines utilising wild relative species genome-specific SNPs, the KASP technology has been shown to be a cost-effective, efficient, and flexible approach (Grewal *et al*., 2020; Bansal *et al*., 2020; Han *et al*., 2020). However, this technology has problems due to the polyploid genome of bread wheat, which makes it difficult to generate co-dominant interspecific SNPs. The first barrier is discerning between interspecific the abundant paralogous SNPs present in the wheat genome. The scoring of interspecific SNPs in polyploid segregating populations when the SNP contains three homoeologous copies in the wheat genome is the second obstacle. Thus, in polyploid species such as common wheat, it is difficult to develop chromosome-specific KASP markers that can differentiate heterozygous from homozygous individuals, and validation is necessary (Allen *et al*., 2011; Makhoul *et al*., 2020; Grewal *et al*., 2020). Recently, this problem was addressed by Grewal *et al*. (2020) and wild relative genome-interspecific SNPs were converted to KASP assays using PCR amplicon-based sequencing. This work was further developed with the design of new sets of KASP markers to provide a higher density of SNPs between the selected wild relative and wheat and hence fill the gaps (Grewal *et al*., 2021 and 2022).

In conclusion, for the first time, wheat homozygous introgressions and recombination sites have been identified thanks to the development of chromosome-specific KASP markers.

5.2 Aim and Objectives

The main aims and objectives of this chapter are:

- To self-fertilise the lines containing single segments to make the introgressions homozygous and thus stable.
- To characterise the introgression lines using KASP markers and validate with multi-colour genomic *in situ* hybridisation.

5.3 Materials and Methods

5.3.1 Plant material and Crossing programme

Seed materials with single introgression segments or whole E chromosome lines were selected from materials that had been produced so far for this chapter. The crossing diagram to generate homozygous wheat/*Th. elongatum* introgression lines in this work is shown in Figure 3.2 (Chapter 3, section 3.3.1). Plant material were self-fertilised to obtain stable homozygotes lines. Then, selected homozygous seeds were germinated (Chapter 3, section 3.3.3).

5.3.2 Genotyping Analysis of Homozygous Lines via KASP

DNA extraction, quantification and preparation of the DNA samples for use in the mc-GISH and KASP genotyping was as described in Chapter 2, (section 2.1). Plant materials produced from the crossing programme (Figure 3.2) were genotyped using chromosome-specific KASP markers.

Genotyping was performed as described in Chapter 4 (section 4.3.1.2) using a PCR reaction containing 1 ng of genomic DNA, 2.5 µL KASP reaction mix (ROX), 0.068 µL primer mix and 2.43 µL nuclease free water in a final volume of 5 µL set up using the automated PIPETMAX® 268 (Gilson, UK). PCR programs were run for denaturation step as 15 mins at 94°C; 10 touchdown cycles of 10 s at 94°C, then annealing for 1 min at 65–57°C; and elongation step was 30 cycles of 10 s at 94°C, 1 min at 57°C and performed in a ProFlex PCR system (Applied Biosystems by Life Technology). Fluorescence detection of the reactions was performed using a QuantStudio™ 5 Real-Time PCR System (Applied Biosystems) and the data was analyzed using QuantStudio™ Design and Analysis Software V1.5.0.

5.3.3 Cytogenetic Analysis of the Homozygous Lines with GISH

Preparation of the root tips chromosome spreads protocol for mc-GISH was described in Chapter 2. The metaphase spreads of 20 selected lines in the homozygous introgression lines were analysed via two different combinations of mc-GISH. The first detected both wheat genomes and *Th. elongatum* genomes. Genomic DNA from *T. urartu*, *Ae. speltoides*, *Ae. tauschii*, and *Th. elongatum* were extracted to make probes. For the second, *Th. bessarabicum* and *P. strigosa* genomic DNA as probes were used to detect the presence of the J genome and St genomes respectively in homozygous lines as indicated in Chapter 2 (section 2.2).

5.4 Results

Seed materials with single introgression or whole *Th. elongatum* chromosome lines, genotyped with KASP markers and confirmed by mc-GISH, were selected from materials obtained so far for further breeding programs. These selected lines were self-fertilized to obtain homozygous lines. After self-fertilization lines were characterized using KASP markers and mc-GISH analysis. Twenty BCxFy backcross lines were found to contain homozygous introgression. Examination of these lines showed there were difference in the *Th. elongatum* segment present, e.g., recombinant chromosomes, substitution lines or addition lines (Table 5.1).

Figures 5.1 to 5.10 show the KASP analysis of wheat-*Th. elongatum* homozygous lines (left) and mc-GISH images of metaphase spreads. The genotyping shows the markers ordered on the *Th. elongatum* chromosomes. Mc-GISH was used for confirmation of the genotyping results, with both the *Th. elongatum* and wheat genomes visualised at the same time. In the GISH photos, E genome segments (yellow) recombined with the A (green), B (grey/blue) or D (red) genome chromosomes of wheat.

Chromosome counts of each genome together with the total chromosome counts were made for each of these 20 homozygous introgression lines (Table 5.1). In addition, both genotyping and mc-GISH were able to detect which wheat genome the E genome had recombined with (see below). Most of these homozygous lines have one introgressed segment, with the exception of some lines that carry a part or whole chromosome of E (Table 5.1).

When a line had a homozygous introgression, the KASP markers on the wheat chromosome involved in the recombination indicated homozygous calls (green) as shown in Figure 5.1. This is due to the genome-specificity of the KASP markers in wheat. Thus, when two copies of the *Th. elongatum* segment replaced both copies of the homoeologous wheat regions in the homozygous recombinant chromosomes, it resulted in no wheat allele for the genome-

specific KASP markers in those regions. All other KASP markers detecting the E genome segment but not physically located on the recombinant chromosome showed a heterozygous call (red) to indicate the presence of *Th. elongatum* in a wheat background.

Table 5. 1: List of wheat-*Th. elongatum* homozygous introgression lines obtained in this chapter, the number of the genomes and the total number of chromosomes (2n) in each line is homozygous for just one introgression. Lines marked with an asterisk (*) are sister lines from same BC1 parent with same genotype.

WRC Distribution Code	Introgression lines	Recombinant chromosomes	Whole chromosomes	Chromosome number of genomes	Total chromosome number
WRC21_Elo1	BC_2F_3-111		2	14A 14B 12D 2E	42
WRC22_Elo21	BC_3F_2-156-3	1: E with B	2	14A 10B 14D 2E/B 2E	42
Line lost	BC_3F_2-158-5	1: E with B		14A 12B 14D 2E/B	42
WRC21_Elo2	BC_3F_3-56	1: E with B	A part	14A 12B 14D 2E/B	42
WRC21_Elo11	BC_4F_2-89-2	1: E with B	2	14A 13B 14D 2E/B 2E	45
WRC22_Elo22	BC_4F_2-89-5	1: E with B		14A 14B 14D 2E/B	44
WRC21_Elo3	BC_4F_2-121	1: E with B		15A 13B 13D 2E/B	43
WRC21_Elo4	BC_4F_3-43	1: E with A		12A 14B 14D 2E/A	42
WRC21_Elo5	BC_4F_3-44		2	14 A 14B 14D 2E	44
WRC21_ Elo6	BC_4F_3-45*	1: E with A		12A 14B 14D 2E/A	42
WRC21_ Elo7	BC_4F_3-46	1: E with A		12A 14B 14D 2E/A	42

WRC21_Elo8	BC_4F_3-47		2	14 A 14B 14D 2E	44
WRC21_Elo9	BC_4F_3-48	1: E with D		13A 14B 13D 2E/D	42
WRC21_Elo10	BC_4F_3-49	1: E with B		14A 12B 14D 2E/B	42
WRC21_Elo12	BC_4F_3-51	1: E with A		12A 14B 14D 2E/A	42
WRC21_Elo13	BC_4F_4-7	1: E with D		14A 14B 12D 2E/D	42
WRC21_Elo14	BC_5F_2-17	1: E with D		14A 14B 12D 2E/D	42
WRC21_Elo15	BC_5F_2-18		2	14A 14B 13D 2E	43
WRC21_Elo18	BC_5F_3-6	1: E with B		14A 12B 14D 2E/B	42
WRC21_Elo19	BC_5F_3-7	1: E with D		14A 14B 12D 2E/D	42
WRC21_Elo20	BC_5F_4-1	1: E with B		14A 13B 14D 2E/B	43

The analysis of genotyping in Figure 5.1a showed the E-genome segment had recombined with the A genome of wheat, in Figure 5.1b, the markers indicated the E genome had recombined with the B genome of wheat and in Figure 5.1c the E genome had recombined with the D genome of wheat. Figure 5.1d, however, shows a substitution line where two whole *Th. elongatum* chromosomes replaced both copies of the homoeologous D genome of the wheat. The mc-GISH analysis, shown in Figure 5.1a, 5.1b, 5.1c, and 5.1d confirmed markers results. In addition, an unbalanced chromosome number was observed in BC_4F_2-121 using mc-GISH.

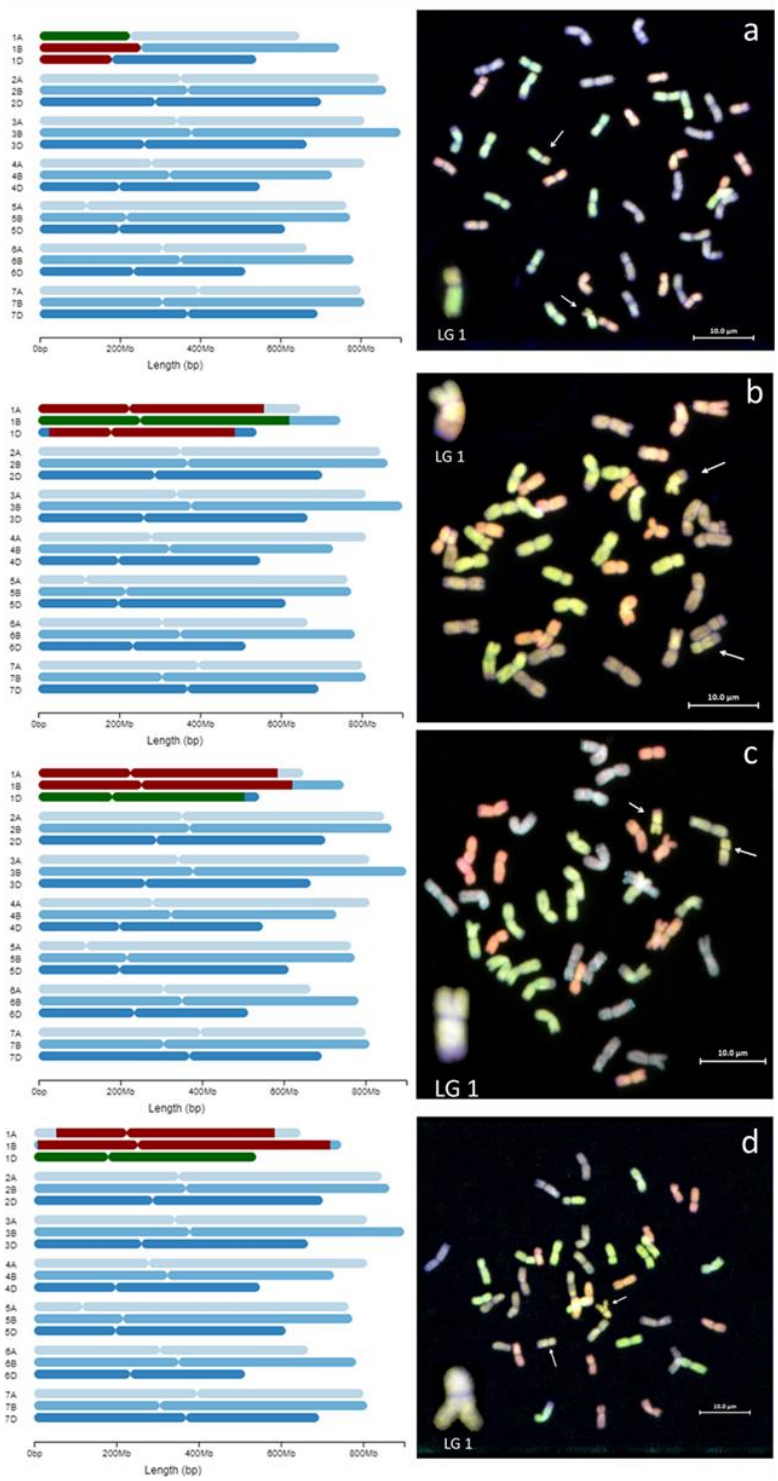


Figure 5. 1: Molecular characterization of wheat-*Th. elongatum* homozygous lines showing genotyping analysis (left), using chromosome-specific KASP markers, and root metaphase spreads (right), using mc-GISH. Homozygous calls are shown in green heterozygous calls are shown in red, while all wheat calls are shown in blue. **(a)** BC_4F_3-51, **(b)** BC_4F_2-121, **(c)** BC_5F_2-17, and **(d)** BC_2F_3-111. Arrows indicate recombinant chromosomes.

The marker analysis in Figure 5.2 indicated the presence of a heterozygous *Th. elongatum* introgression. However, cytological analysis showed that the introgression was homozygous and thus the recombined chromosome in this line was present as a disomic addition. When a line had the recombined chromosomes as an addition, the KASP markers on the wheat chromosome indicated heterozygous calls (red) as shown in Figure 5.2. Because two copies of the *Th. elongatum* segment did not replace both copies of the wheat genome in the homozygous recombinant chromosomes, it resulted in a wheat allele for the KASP markers in those regions. However, the detection of homozygous *Th. elongatum* was possible via mc-GISH, although the markers indicate that these lines were heterozygous in Figure 5.2.

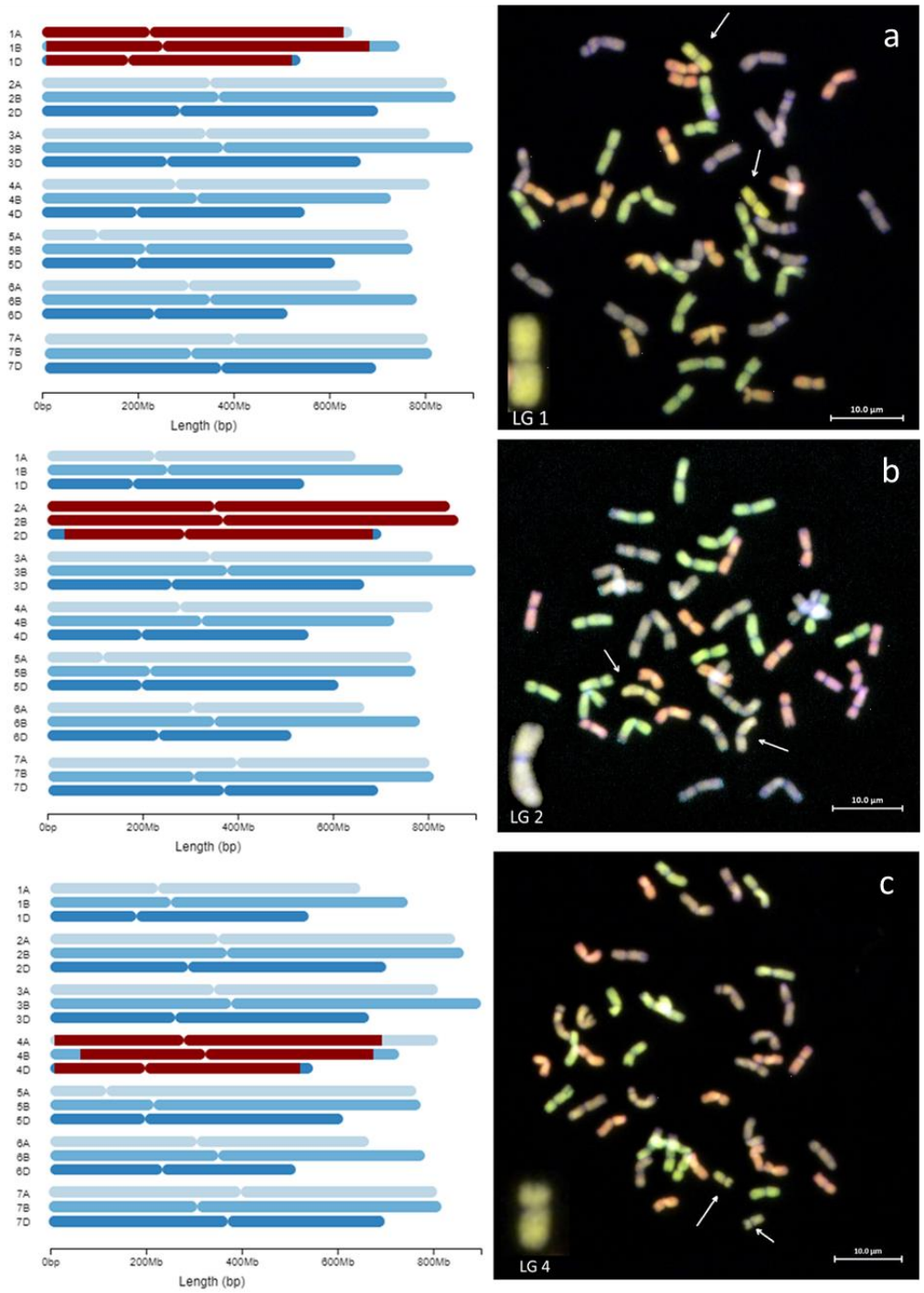


Figure 5. 2: Molecular characterization of wheat-*Th. elongatum* addition homozygous lines showing genotyping analysis (left), using chromosome-specific KASP markers, and root metaphase spreads (right), using mc-GISH. Heterozygous calls are shown in red, while all wheat calls are shown in blue. **(a)** BC_4F_3-47, **(b)** BC_4F_3-44, and **(c)** BC_5F_2-18. Whole chromosomes of *Th. elongatum* are indicated by arrows.

In Figure 5.3, the marker detected the *Th. elongatum* segment in two different linkage groups (Figures 5.3a LG 1 and 6, 5.3b LG 1 and 5). The genotyping results showed that homozygous calls for linkage group 1 and LG 6, while LG 5 was found to be heterozygous. In Figure 5.3a, KASP markers showed a segment covering the short arm and approximately half of the proximal region of the long arm.

When this line was examined with mc-GISH, it was confirmed and showed the segment was homozygous and recombined with the B genome of wheat. The GISH result also indicated the presence of two whole E genomes LG6 chromosomes. The genotyping diagram in Figure 5.3b shows a small homozygous E introgression at the end of the long arm that recombines with the wheat genome B in LG1 as well as there was a part of the E genome in the LG5. The KASP markers` results were confirmed by mc-GISH.

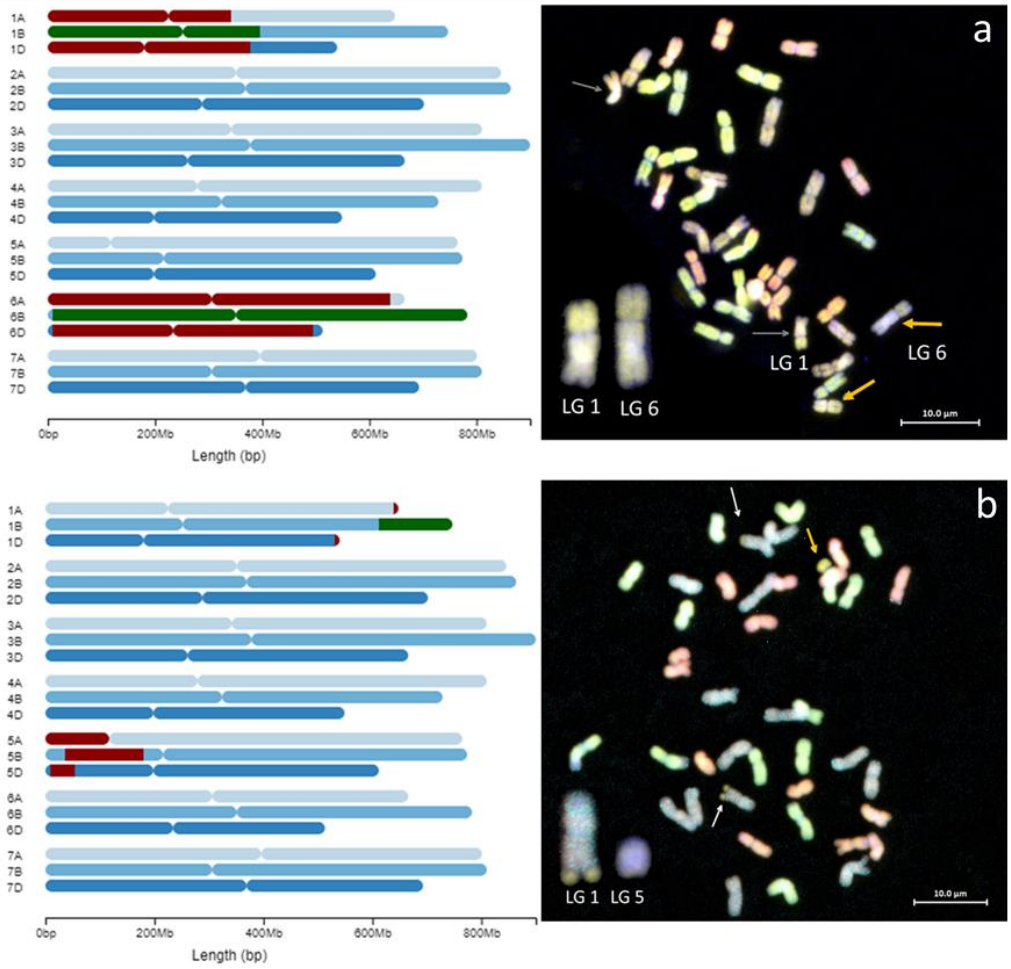


Figure 5. 3: Molecular characterization of wheat-*Th. elongatum* homozygous lines showing genotyping analysis (left), using chromosome-specific KASP markers, and root metaphase spreads (right), using mc-GISH. Homozygous calls are shown in green, heterozygous calls are shown in red, and all wheat calls are shown in blue. White arrows indicate recombinant chromosomes while orange arrows indicate whole/partly chromosome of E. **(a)** BC_3F_2-156-3, and **(b)** BC_3F_3-56

In Figure 5.4, the markers detected the presence of a large segment of 2E via heterozygous calls on chromosome 2B and 2D (in red) and its presence as a homozygous introgression on chromosome 2A, due to homozygous calls on chromosome 2A (in green). The KASP markers` results were subsequently confirmed by mc-GISH analysis that showed a small part of the A genome on the top of the short arm.

Also, the same segment was observed on the homozygous introgression plant BC4F3-45. These two lines are sister lines from the same BC1 parent with the same genotype. Thus, their GISH and genotyping results were exactly the same.

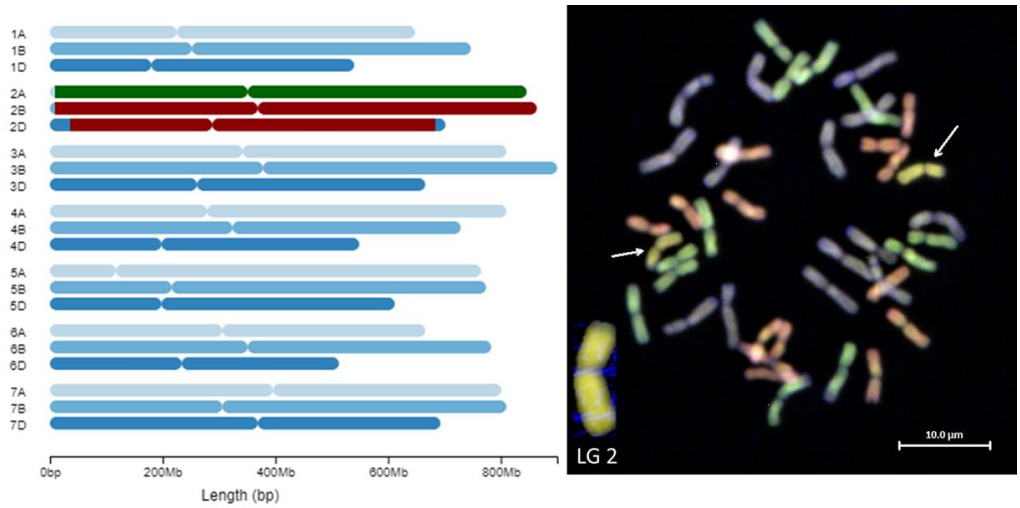


Figure 5. 4: Molecular characterization of wheat-*Th. elongatum* homozygous line (BC_4F_3-46) with genotyping analysis (left), using chromosome-specific KASP markers, and root metaphase spread (right), using mc-GISH. Homozygous calls are shown in green, heterozygous calls are shown in red, and all wheat calls are shown in blue. White arrows indicate recombinant chromosomes.

KASP markers detected homozygous lines in LG3. In Figure 5.5a, one large segment was detected by markers on the D genome. Mc-GISH analysis of the introgression showed that the segment is from half the short arm until the end of the long arm.

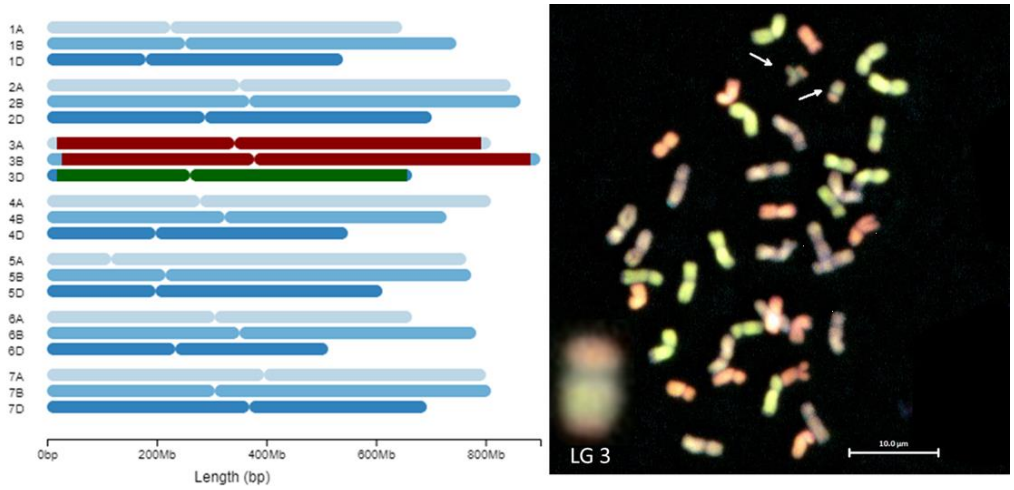


Figure 5. 5: Characterization of wheat-*Th. elongatum* homozygous line (BC_4F_4-7) using KASP markers assay (left), and their Identification through mc-GISH (right). Homozygous calls are shown in green, heterozygous calls are shown in red, and all wheat calls are shown in blue. White arrows indicate recombinant chromosomes.

Genotyping results showed heterozygous calls for a total of five homozygous introgression lines which was confirmed with GISH analysis (see Figure 5.6 and 5.7). However, the total chromosome counts showed all but one of these lines had an unbalanced chromosomes number (the exception was BC_5F_3-6) (Table 5.1) in that the total number of chromosomes was more than 42 and hence the recombined chromosome was present as an addition. Thus, for the genotyping, even when an introgression was homozygous, the homoeologous wheat chromosome was present, resulting in a heterozygous call.

In Figure 5.6 there were three different homozygous introgressions that were identified with mc-GISH after the genotyping showed them as heterozygous. In Figure 5.6a, the segment was the whole of the long arm and was recombinant with the wheat B genome while in Figure 5.6c, the E segment located on the whole of the short arm. Mc-GISH analysis also showed introgression was on the whole of the long arm and that there was a small part D genome on the short arm (Figure 5.6b).

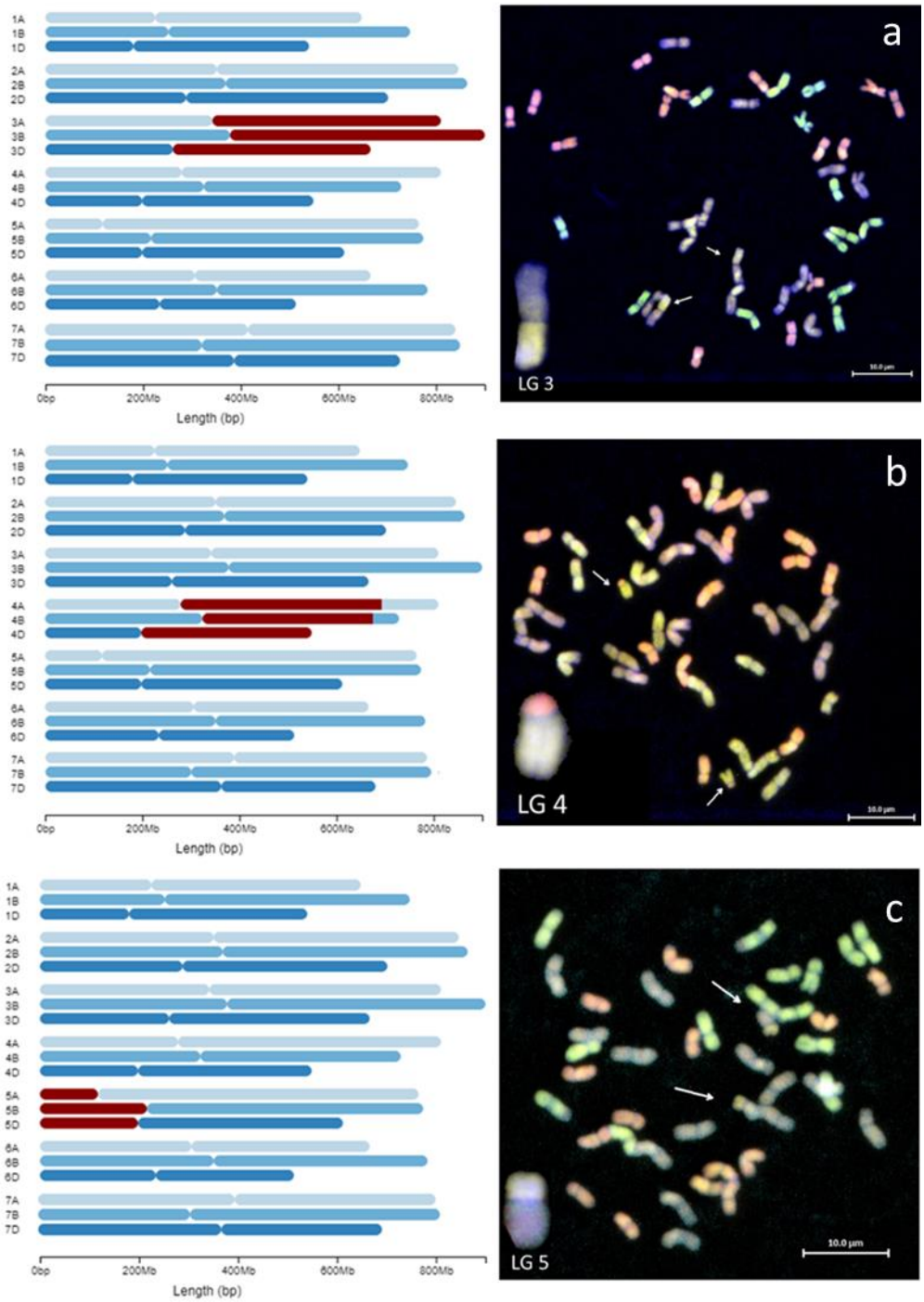


Figure 5. 6: On the right, mc-GISH photos of wheat-*Th. elongatum* homozygous introgression lines showed root metaphase spreads, whereas on the left, KASP markers were used to genotype them. Heterozygous calls are shown in red, while all wheat calls are shown in blue. **(a)** BC_5F_3-6 **(b)** BC_4F_3-48 and, **(c)** BC_5F_4-1. Recombinant chromosomes are indicated by white arrows.

In sister lines, introgression had recombined with chromosome 5B of wheat as indicated by mc-GISH analysis in Figures 5.7a and b, and the whole chromosome of E was examined in linkage group 1 for BC_4F_2-89-2. While BC_4F_2-89-5 line (14A, 14B, 14D, and 2E/B) was losing the whole chromosome of E, BC_4F_2-89-2 line has (14A, 13B, 14D, 2E/B, and 2E) both recombinant and whole chromosomes.

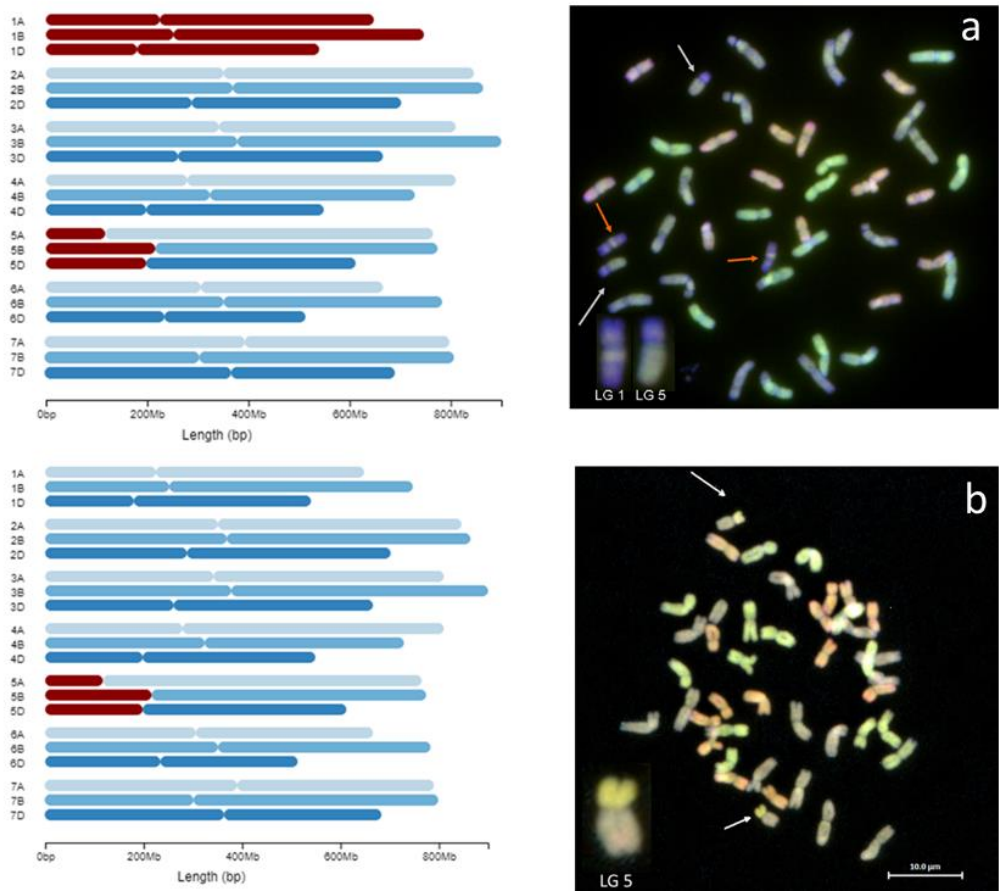


Figure 5. 7: Molecular characterization of wheat-*Th. elongatum* homozygous lines showing genotyping analysis (left), using chromosome-specific KASP markers, and root metaphase spreads (right), using mc-GISH. Heterozygous calls are shown in red, while all wheat calls are shown in blue. **(a)** BC_4F_2-89-2, and **(b)** BC_4F_2-89-5. Recombinant chromosomes were indicated by white while whole chromosome of *Th. elongatum* was indicated orange arrows.

The presence of *Th. elongatum* in wheat was easily detected by the markers, as shown in Figures 5.8, where the homozygous segments from the E chromosome was recombined with chromosome 5A of wheat. Mc-GISH analysis showed a small E chromosome located at the end of the long arm.

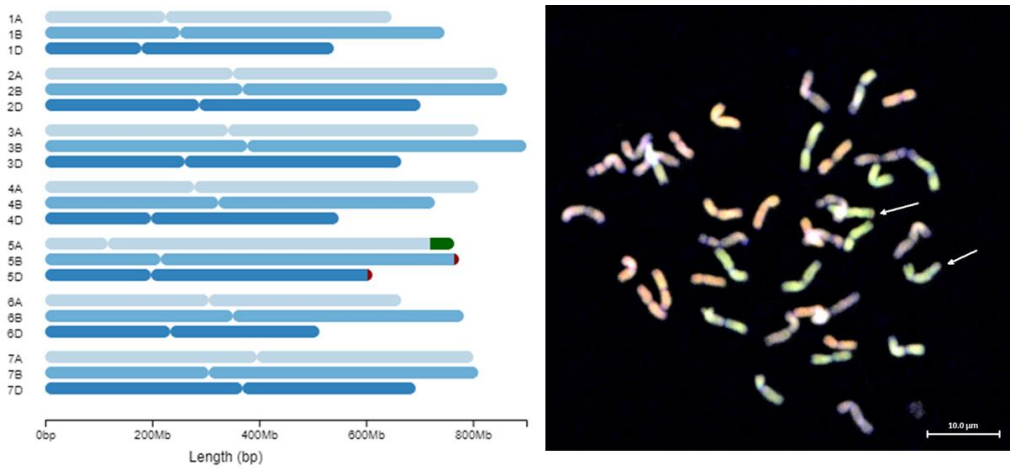


Figure 5. 8: Molecular characterization of wheat-*Th. elongatum* homozygous introgression lines (BC_4F_3-43) using KASP markers (left) and mc-GISH (right). Homozygous calls are shown in green, heterozygous calls are shown in red, and all wheat calls are shown in blue. White arrows indicate recombinant chromosomes.

Figures 5.9a and 5.9b show two homozygous lines with different sized segments on LG 6 and their genotyping result of KASP markers and validation of the result with mc-GISH. The homozygous calls for KASP markers on chromosome 6B of wheat indicated the presence of a large segment from the E genome in Figure 5.9a, while a small introgression was indicated on chromosome 6D in Figure 5.9b. When GISH results of BC_4F_3-49 were examined, it was found that the segment included the short arm and most of the long arm except for the tip of the long arm (see Figure 5.9a). Figure 5.9b mc-GISH results indicated an E genome on the half part of the long arm and a very small D genome region at the distal end of the short arm. Moreover, that photo showed that between the D and E genomes there was a B genome of wheat.

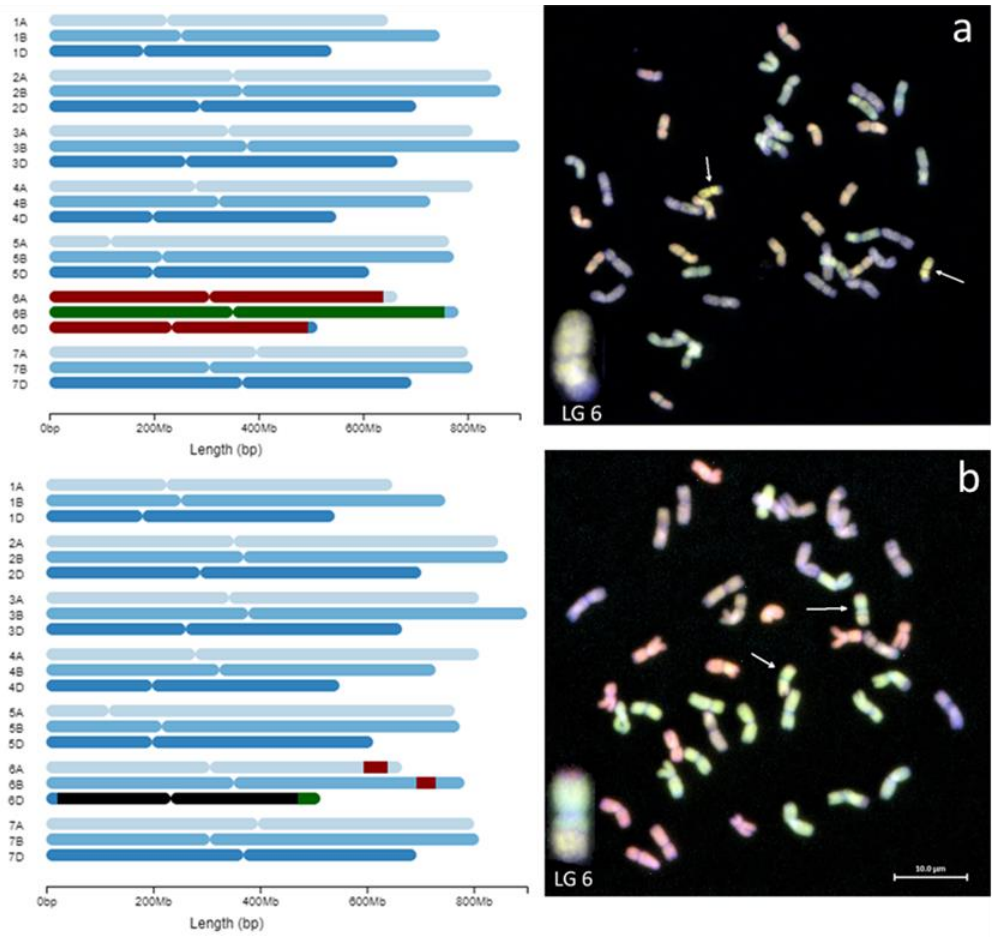


Figure 5. 9: Molecular characterization of wheat-*Th. elongatum* homozygous lines showing genotyping analysis (left), using chromosome-specific KASP markers, and root metaphase spreads (right), using mc-GISH. **(a)** BC_4F_3-49 and **(b)** BC_5F_3-7. White arrows indicate recombinant chromosomes.

Figure 5.10 shows the genotyping of the wheat-*Th. elongatum* homozygous line with KASP markers and validation of the result with GISH. The large segment of E was detected, and a very small segment of the B genome of wheat was visualized at the distal end of the short arm (see below). Also, only one homozygous line was generated in LG7, however, unfortunately, this line was lost.

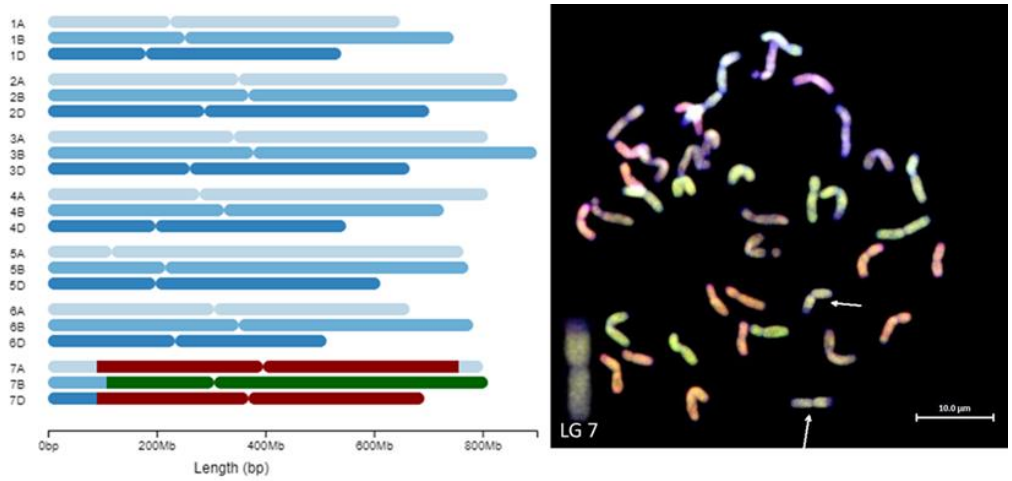


Figure 5. 10: Characterization of wheat-*Th. elongatum* homozygous line (BC_3F_2-158-5) showing KASP markers assay (left), and their Identification through mc-GISH (right). Homozygous calls are shown in green, heterozygous calls are shown in red, and all wheat calls are shown in blue. White arrows indicate recombinant chromosomes.

These lines will be multiplied and released free of IP for trait analyses. It is therefore important that the introgressions are able to be tracked within a breeding programme. Table 5.2, therefore, lists one KASP marker for each of the segments present in the homozygous lines, which can be used to detect that segment when it is both heterozygous and homozygous.

Table 5. 2: The table shows one diagnostic marker per segment. Lines marked with an asterisk (*) are addition lines and therefore markers for all wheat genomes show heterozygous so one marker is shown per genome.

Introgression Lines	Linkage Group	Recombinant chromosomes	Diagnostic Markers
BC_2F_3-111	LG1	Chr D	WRC0135
BC_3F_2-156-3	LG1	Chr B	WRC0076
BC_3F_2-156-3	LG6	Chr B	WRC0802
BC_3F_2-158-5	LG7	Chr B	WRC0908
BC_3F_3-56	LG1	Chr B	WRC0091
BC_4F_2-121	LG1	Chr B	WRC0074
BC_4F_2-89-2*	LG1	Chr A	WRC0006
		Chr B	WRC0074
		Chr D	WRC0135
BC_4F_2-89-2	LG5	Chr B	WRC0620
BC_4F_2-89-5	LG5	Chr B	WRC0620
BC_4F_3-43	LG5	Chr A	WRC1137
BC_4F_3-44*	LG2	Chr A	WRC0184
		Chr B	WRC0218
		Chr D	WRC0267
BC_4F_3-46	LG2	Chr A	WRC0184
BC_4F_3-47*	LG1	Chr A	WRC0037
		Chr B	WRC0074
		Chr D	WRC0128
BC_4F_3-48	LG4	Chr D	WRC0536
BC_4F_3-49	LG6	Chr B	WRC0763
BC_4F_3-51	LG1	Chr A	WRC0002
BC_4F_4-7	LG3	Chr D	WRC0391
BC_5F_2-17	LG1	Chr D	WRC0128
BC_5F_2-18*	LG4	Chr A	WRC0439
		Chr B	WRC1364
		Chr D	WRC0536

BC_5F_3-6	LG3	Chr B	WRC0363
BC_5F_3-7	LG6	Chr D	WRC0852
BC_5F_4-1	LG5	Chr B	WRC0620

Fifteen wheat-*Th. elongatum* homozygous lines (Figure 5.11a to o) and two heterozygous lines (Figure 5.11p and r) were analysed using mc-GISH (green represents *Th. bessarabicum* (J) while red represent *Pseudoroegneria spicata* (St)) and J and J^{St} of genomes were observed from different linkage groups. The size of introgression was also visualised with mc-GISH analysis (Figure 5.11). Due to the complicated constitution of the decaploid *Th. elongatum* genome, some of the introgression lines were analysed to determine whether there was a transfer of J or J^{St} genomes into wheat.

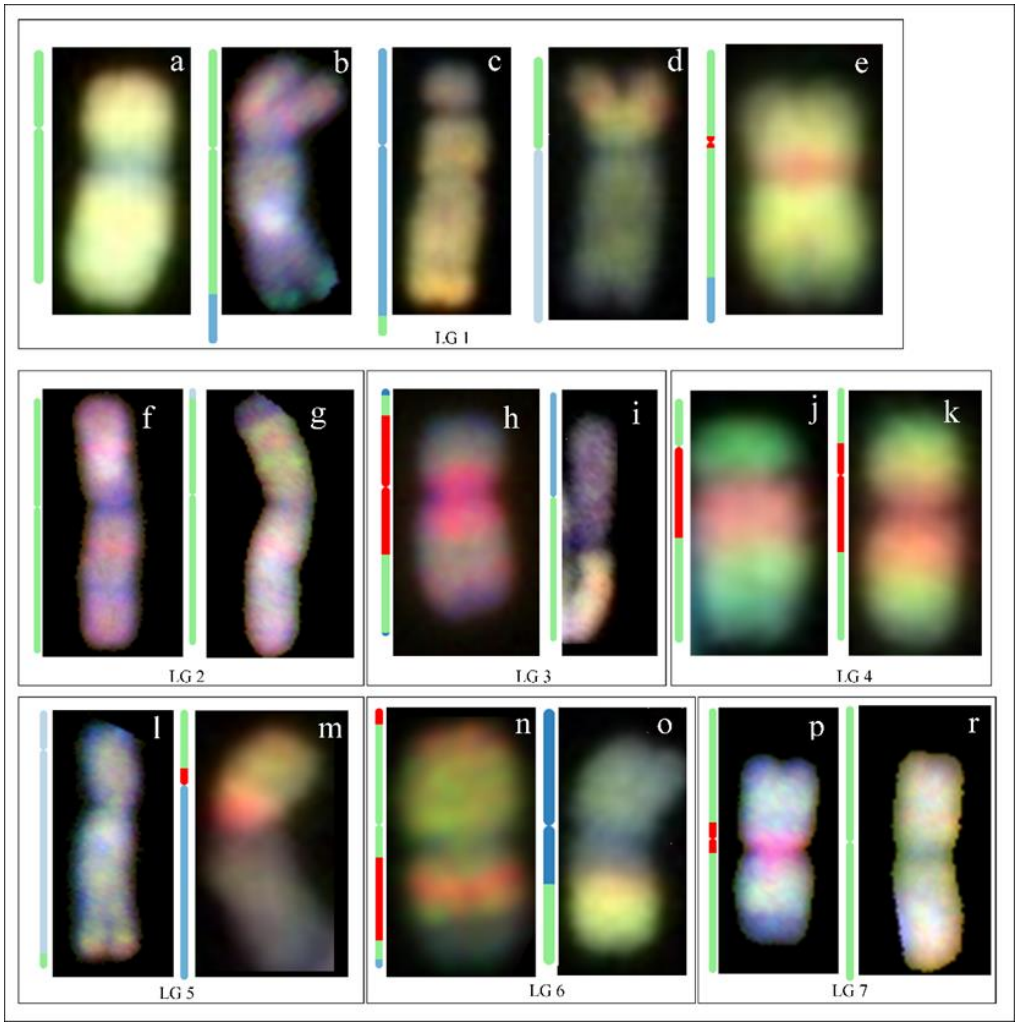


Figure 5. 11: mc-GISH analysis of wheat-*Th. elongatum* homozygous lines showing the different J and J^{St} segments present. **(a)** BC_2F_3-111, **(b)** BC_4F_2-121, **(c)** BC_3F_3-56, **(d)** BC_4F_3-51**(e)** BC_5F_2-17, **(f)** BC_4F_3-44, **(g)** BC_4F_3-46, **(h)** BC_4F_4-7, **(i)**

BC_5F_3-6, **(j)** BC_4F_3-48, **(k)** BC_5F_2-18, **(l)** BC_4F_3-43, **(m)** BC_5F_4-1, **(n)** BC_4F_3-49, **(o)** BC_5F_3-7, **(p)** BC_3F_2-155, and **(r)** BC_6-7-3.



Figure 5. 12: Spikes morphology; **a)** Paragon (*T. aestivum*), **b)** BC_4F_2-121, **c)** BC_4F_3-46, **d)** BC_4F_3-51, **e)** BC_5F_3-6, **f)** BC_5F_4-1, and **g)** BC_5F_2-17.



Figure 5. 13: Spikes morphology; **a)** Paragon (*T. aestivum*), **b)** BC_4F_3-47, **c)** BC_4F_3-49, **d)** BC_4F_3-48, **e)** BC_4F_3-50, **f)** BC_5F_3-7, and **g)** BC_4F_3-44.

All heads shown were after self-pollination. The morphology clearly became more wheat-like. However, if Figure 5.13 were compared to Figure 5.14, in Figure 5.13, six of the homozygous lines had lax spikes compared to wheat, and in Figure 5.14, homozygous lines had dense spikes.

5.5 Discussion

The main objective of this chapter was to generate stable homozygous wheat-*Th. elongatum* introgression lines and to characterise their segments. Hence, introgression lines with only one or a maximum of two introgressions were selected for further work. In this work, twenty introgressions in wheat-*Th. elongatum* were generated using self-fertilization on some of the lines from the back-crossed population generated previously and molecular characterization of these lines. Then, these twenty homozygous lines were characterised with KASP markers and mc-GISH.

PCR primers for a high level of polymorphism between wheat and its wild relatives are difficult to design as mentioned above (Section 5.1). Thus, the Nottingham BBSRC WRC`s aim was to design KASP assays that would detect only one of the genomes in wheat and would be polymorphic with the wild relative genome. These chromosome-specific KASP markers are able to distinguish between homozygous and heterozygous individuals, and therefore, genotyping analysis was performed with these KASP markers in this project.

According to the genotyping results, this project needed more markers. This was firstly as a greater density of markers was needed to ensure even small segments were detected. However, adding to the insufficient density was the fact that several markers did not work well, resulting in even larger gaps between markers. For example, two homozygous lines (BC_5F_5-18 and BC_4F_3-48) were obtained from LG4. KASP marker analysis did not detect the telomere end of the LG4 long arm with the A and B genome specific markers whereas mc-GISH (Figures 5.2c and 5.6) validated the introgression as being the whole of the long arm. Another example where more markers are needed is the short arm of LG1 and particularly the A genome specific markers. The LG1 A marker set still had gaps with no segments detected whereas the original array genotyping had shown segments present. A number of working markers were identified but more markers were needed for this location (see Figure 5.1).

When the genotyping results were examined, markers gave heterozygous calls even though some of the lines were homozygous (Figure

5.2). If the wheat genome wasn't completely replaced by the E genome, the genotyping result could give a heterozygous call due to the presence of at least one copy of the wheat allele, i.e., the plants carried all of the wheat chromosomes plus the E genome chromosomes/introgressions. Thus, after genotyping analysis, GISH analysis is suggested to confirm the result as besides the detection of recombinant or whole chromosomes of *Th. elongatum*, mc-GISH also allowed the observation of the wheat chromosomes. For example, mc-GISH analysis of the BC_4F_3-48 and BC_5F_2-18 homozygous lines (respectively 13A, 14B, 13D, 2E/D and 14A, 14B, 13D, 2E) demonstrated that the total chromosome number of the wheat genome (A, B, and D) was lower than expected, i.e., the plants carried less than the 14 A, 14B, and 14D chromosomes expected (Table 5.1). Plants with high numbers of chromosomes have been shown to be very unstable during both mitosis and meiosis (Comai, 2005), possibly due to the potentially disruptive effects of nuclear and cell enlargement and also the tendency of polyploid meiosis to produce aneuploid cells (Mestiri *et al*., 2010). Table 5.1 shows that the BC_4F_2-121 homozygous line has an extra A genome of wheat and one missing B and D genomes (15 A, 13B, 13D, and 2E/B).

Mc-GISH also revealed that there were introgressions not only between wheat and *Th. elongatum* but also between wheat genomes. For example, BC_5F_3-7 was shown to be homozygous and have a small introgression from the genotyping results. The mc-GISH results showed that, in contrast to the KASP result, this homozygous line was more complex than expected. The A genome of wheat occupied a large space between a very small segment of the D genome of wheat on the distal end of the short arm and the E genome of *Th. elongatum* on the distal end of the long arm. At this point, the importance of the GISH analysis was once again demonstrated (Figure 5.9b).

The production of stable homozygous introgression lines takes time because it needs a number of rounds of self-fertilisation to remove heterozygosity. BC_3F_3-56 (Figure 5.3b) needs more work in that it has a small part of telomer of *Th. elongatum* in addition to a homozygous recombinant

chromosome. The self-fertilised progeny of this line therefore needs to be taken forward and screened for lines that have retained the homozygous recombinant chromosome but that have lost the single telomere (Figure 5.3b). Genotyping with KASP markers allowed the distinction of homozygous lines from heterozygous lines in segregated populations and allowed fast characterization of the wheat chromosome recombined with segments of wild relatives (Grewal *et al*., 2020).

Mc-GISH was used to examine fifteen wheat-*Th. elongatum* homozygous lines and two heterozygous lines, with green representing *Th. bessarabicum* (J) and red representing *Pseudoroegneria spicata* (St), and J and J^{St} of genomes observed from seven linkage groups. The size of the introgression was also readily seen using mc-GISH analysis (Figure 5.11). If it has been seen as a faded red, this segment counts as a J/E genome. As indicated in Figure 5.11e, h, j, k, m, n, and p, the St genome should be bright red. More work on the J and St is required in this project because these genomes are closely related to *Th. elongatum* as mentioned in the KASP chapter (section 4.1).

The KASP and GISH analyses have some limitations according to the results. The challenge is that the genotyping results indicate heterozygous calls, while the GISH analysis indicates homozygous (Figure 5.6 and 5.7). Only the presence of wheat genomes makes this possible. When the wheat genome is gone or replaced by an E genome, the marker calls homozygous. However, if one of the wheat genome pairs is present, its presence is indicated by heterozygous calls for the markers.

In order to get heterozygous calls, the homozygous introgression line should include any of the three wheat genomes. For example, BC_4F_3-48 (13A, 14B, 13D, and 2E/D) recombined with the D genome of wheat in LG4. If lines contain at least one additional 4D chromosome, the genotyping result will show heterozygous calls (Figure 5.6b). Because these markers only show up as homozygous calls if a segment has completely replaced the equivalent linkage group/linkage group segment in one of the genomes. BC_5F_4-1 (14A, 13B, 14D, and 2E/B) is a second example. KASP results indicated a heterozygous LG5

segment which GISH showed to be recombined with a B genome chromosome. This line possessed an additional B chromosome which was probably LG5 B given the heterozygous KASP result (Figure 5.6c).

The normal chromosome number of bread wheat is 42 with 14A, 14B, and 14D chromosomes. For example, the BC_5F_3-6 homozygous introgression line has 14A, 12B, 14D, and 2 recombined LG3 E/B chromosomes (Table 5.1). GISH analysis showed that chromosome numbers were as expected, i.e., 14A, 12B, 14D and 2E/B chromosomes. However, KASP analysis showed heterozygous calls (Figure 5.6a). It is therefore possible that the line has lost either 2B, 4B, or 5B and then there is a duplicated copy of the 3B chromosomes.

BC_4F_2-121 had an unbalanced chromosome number (15A, 13B, 13D, 2E/B LG1 chromosomes) but the KASP results showed homozygous calls (confirmed with mc-GISH analysis) (Table 5.1, Figure 5.1b). Again, it is also possible that in this line the E genome has recombined with LG1 B but there is a duplicated copy of either 2B, 6B, or 7B then lost one of them so that there are three of a particular chromosome instead of two. GISH analysis would not show which linkage group's chromosome was lost or duplicated.

The only way to definitely prove that certain chromosomes are missing or added would be to do fluorescent *in situ* hybridisation (FISH) analysis, as FISH analysis allows individual chromosomes to be identified by their FISH signal and a complete karyotype can be obtained. Mc-GISH is a very useful technique but is unable to identify individual chromosomes.

Chapter 6 General Discussion

The evolution of wheat following its domestication and inbreeding has reduced the genetic variation available for wheat breeding. However, genetic diversity will be required to develop new wheat varieties in the future adapted to the challenges of global warming and an increasing population. The wild relatives of wheat provide a rich source of genetic diversity that might be used to improve diversity in wheat breeding programs. One such wild relative, *Th. elongatum*, has been shown to carry useful genetic diversity that can be used for the development of new wheat varieties.

Th. elongatum is a perennial species with many excellent traits, such as long spikes, multi-flowers, high grain protein content, strong adaptability, and reproductive ability, as mentioned in Section 1.9. Thinopyrum species have always been used by breeders to use their potential, but in general, breeders focus on the resistance to biotic and abiotic stresses. In this project, there were not any observed perennial introgression plants. However, Thinopyrum species can also be used for the development of perennial wheat species. Perennial wheats have been developed from crossing between common wheat and diploid *Th. elongatum*, hexaploid *Th. intermedium*, and decaploid *Th. ponticum* (Hayes *et al*., 2012; Gazza *et al*., 2016). Agricultural schemes based on annual crops have been recognized as among the main causes of soil corruption (Glover 2005), and each year those crops require to be reseeded, resulting in ongoing soil tillage (Baronti *et al*., 2022). Perennial crops do not need intensive soil cultivation; thus, they might contribute to reducing erosion and disruption and increasing the organic material of soil content (Ryan *et al*. 2018). However, there are some problems to be solved because the perennial wheats grown so far produce an average of 30% less yield than their annual counterparts (Hayes *et al*. 2012; Newell and Hayes 2017). Perennial cereal crops were developed in Russia as early as the 1920s through the wide hybridization of wheat and wheatgrass species (Tsitsin and Lubimova, 1959; Wagoner, 1990). Hybrids created between 1920 and 1980 had issues such as being meiotically unstable,

lacking fertility, and obtaining a variety of undesirable agronomic traits. Later, The Land Institute (Cox *et al*., 2010) and Washington State University (Murphy *et al*., 2010), both in the USA, began efforts to develop perennial cereal crops using wheat derivatives. Also, more recently, in Central Italy, four perennial wheat genotypes were grown, and agronomic traits were compared by Baronti *et al*. (2022).

In this project, wheat/*Th. elongatum* introgression lines were backcrossed to obtain lines containing single introgressions which were then self-fertilised to generate stable homozygous lines. The introgression lines were characterised with chromosome specific KASP markers and mc-GISH. Each subsequent generation in this work showed higher levels of fertility as the wheat genome was gradually restored. Babu *et al*. (2004) reported that the recurrent parent genome could be expected to be restored to 99% with subsequently six generations. However, the lines require analysis at every generation as the introgressed chromosomes are often not as stable as the wheat chromosomes and can thus be lost. Once seen the introgression is homozygous, every time selfed the plant should remain homozygous. So, the seed can be bulked and can be phenotyped, and this needs a stable background. However, phenotyping experiments were not possible within the scope of this programme. The production of stable homozygous introgression lines takes time because it needs several rounds of self-fertilisation to remove heterozygosity and until the lines are stable, they cannot be multiplied. However, it was without a doubt evident that *Th. elongatum* 401007's phenotypic traits had been transferred into the introgression lines. Some of the homozygous introgression lines' spike shapes and structures were observed to be lax and long like *Th. elongatum*, while those of other lines' spike shapes were dense and stockier like wheat (see Figures 5. 12 and 5. 13).

The chromosome-specific KASP markers were successfully validated and shown to be very efficient in identifying *Th. elongatum* DNA in a wheat background. A total of 267 KASP markers were tested in the genotyping of 796 backcrossed and self-fertilized lines and of these, 189 chromosome-specific KASP markers were found to be polymorphic between wheat and the *Th.*

elongatum accession used in this study (see Chapter 4). Genotyping results suggested that 530 backcrossing lines from the 796 lines included one or more introgressed segments. The results of the *Th. elongatum* 401007 genomic study raised concerns about the molecular marker work done on the introgression lines. Firstly, some of the KASP markers worked for some families but not for others for example, when a marker with SNP A/T was run on introgression lines (the A allele for wheat and the T allele for *Th. elongatum*), the T allele indicating presence of *Th. elongatum* was only detected in some families and not in others. Decaploid *Th. elongatum* has 5 genomes, and according to literature, at least 3 of its progenitors are known as *Th. elongatum*, *Th. bessarabicum*, and *P. strigosa*. It also is more complicated because there are no whole St chromosomes. Thus, for the example above, it is unknown whether all five genomes of *Th. elongatum* contain A/T SNP at homoeologous positions, only two of them, or four of them. Families that have a segment from homoeologous *Th. elongatum* chromosomes containing the SNP will detect the T allele for the marker and those that have a segment without this SNP will not detect it because these KASP markers are chromosome-specific in wheat, but they aren't necessarily chromosome-specific in *Th. elongatum*. The idea was to design KASP markers based on SNPs that spanned across all 5 homoeologous genomes of *Th. elongatum* but as can be seen in the genotyping, they work for some families but not for others, indicating that certain markers are subgenomes-specific in *Th. elongatum*. This results in some markers indicating the presence of the wheat allele only despite the presence of *Th. elongatum.*

For example, the line has a large *Th. elongatum* segment containing a large portion of the St genome but some of the KASP markers are not detecting the *Th. elongatum* alleles, only the wheat alleles. This is probably due to those markers being polymorphic between wheat and the J genome or wheat and/or the E genome but not wheat and the St genome. Alternatively, it would be possible that the St genome has the same allele as wheat.

GISH analysis was carried out to validate the genotyping results, and the mc-GISH analysis detected and characterised the presence of a *Th. elongatum*

segment in 202 lines (Table 4.2). In most cases, the results of the KASP markers were found to be similar to those found in the mc-GISH analysis.

Ten of the WRC 's wild relative species were sequenced last year by Anthony Hall's group at Earlham Institute. And initially, they focused on *Am. muticum* for SNP discovery (Grewal et al. 2022). Short read sequences were aligned to the CS Refseq1 assembly (IWGSC et al, 2018) and called homozygous SNPs for *Am. muticum*. SNPs that had other SNPs in the 50bp flanking region were filtered out, as these can interfere with KASP primer design. And then, finally, and most importantly, the sequence flanking of the remaining SNPs was used in a BLAST search against RefSeq1, and those with significant multiple hits were also removed. Thus, only those SNPs that belonged to potential single-copy regions in the wheat genome were selected. Then KASP markers were developed for *Am. muticum*, and they were also tested on *Th. elongatum* resulting in potentially more KASP markers available for genotyping in the future. The *Th. elongatum* species was also one of the ten species that were sequenced. However, because of its decaploid nature, it has been very difficult to call SNPs between wheat and *Th. elongatum* that are chromosome-specific without the availability of a chromosome-level genome assembly for *Th. elongatum*. This is something the Nottingham BBSRC Wheat Research Centre is working towards and hopefully, will be a valuable resource for the wheat breeding community.

At the beginning of this study, there was a relatively small number of KASP markers accessible, so the genotyping results were sometimes difficult to interpret. As a result, the places where no KASP markers were available at the time of genotyping are represented in white in the genotyping figures (in Chapter 4, Section 4.4). The mc-GISH analyses were therefore particularly important in these early generations to validate the genotyping results. Later during the project, more markers became available, and subsequent genotyping analyses detected the introgressions better. For example, in Figure 4. 7 the top marker diagrams have more white areas than the bottom marker diagrams. One of the major advantages of the chromosome specific KASP markers is that they are able to detect not only the *Th. elongatum* introgression

in the wheat background but also whether the introgression is heterozygous or homozygous and thus the KASP markers enabled rapid identification of the wheat chromosome that had recombined with the wild relative segment. The genotyping results showed that more markers were needed for both the full characterisation of the introgressions and the detection of smaller introgressions (Chapter 5, section 5.4).

In this thesis, the work performed using mc-GISH to detect *Th. elongatum* segments in a wheat background was successful. A benefit of the mc-GISH analyses was that all the wheat chromosomes, as well as the *Th. elongatum* chromosomes, could be visualised and counted. This visualization of chromosome numbers is important in the characterization of stable homozygous lines in the breeding program. Missing wheat chromosomes can result in essential genes being missing as mentioned in chapters 3 and 4. The mc-GISH analysis also revealed the presence of inter-genomic translocations within the wheat genome, e.g., in BC_5F_3-7 (Figure 5.9b) an A genome chromosome has recombined with a D genome chromosome. In contrast to the genotyping result, the mc-GISH results revealed that this homozygous line was more complicated than predicted.

After examining the segregating population with KASP markers and GISH analyses, twenty seeds from lines with a single introgression or whole E chromosome were selected for a subsequent breeding programme (see Figure 3.2, Chapter 3). These seeds were self-fertilized to generate stable homozygous lines, with KASP markers and mc-GISH analyses again used to identify the homozygous segment. A total of twenty lines were identified as containing different homozygous introgressions. It is important that these lines are screened for as many traits as possible.

Future Work

Molecular characterisation of introgressed *Th. elongatum* segments using chromosome-specific KASP markers will be ongoing and thus more markers will need to be developed to fill the gaps that are currently present. These new markers will firstly need to be assigned to a specific linkage group, and then secondly to specific genomes, e.g., J^J or J^E. One of the reasons for the extra markers is that the target of the overall programme is to transfer as much as possible of *Th. elongatum* genome into a wheat background. Thus, a higher density of markers will enable those areas of the genome not already transferred to be targeted. Thinopyrum species have been shown to be tolerant to many biotic and abiotic stresses such as drought and salt tolerance and leaf rust, yellow rust, and Fusarium head blight resistance. Stable homozygous lines obtained through this thesis should be screened for as many traits as possible by breeders and other researchers in as many different environments as possible. The lines generated will remain as just seeds in a packet if they are not screened for a wide a range of traits.

The *Th. elongatum* accession (401007) used in this work to generate the introgression lines, should also be screened for resistance to biotic and abiotic stresses. If accession 401007 does not have a resistance of interest, other accessions could be screened, and an introgression programme started on the accession shown to have the resistance of interest. This would now be a more feasible possibility as many of the tools necessary for such a programme have been put in place via this research, e.g., the chromosome-specific KASP markers have been shown to work across different accessions of the same species and also across different species.

References

Akhunov, E., Nicolet, C. and Dvorak, J. (2009) 'Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay', *Theoretical and Applied Genetics*, 119(3), pp.507-517.

Allen, A., Barker, G., Berry, S., Coghill, J., Gwilliam, R., Kirby, S., Robinson, P., Brenchley, R., D'Amore, R., McKenzie, N., Waite, D., Hall, A., Bevan, M., Hall, N. and Edwards, K. (2011) 'Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (*Triticum aestivum* L.) ', *Plant Biotechnology Journal*, 9(9), pp.1086-1099.

Arterburn, M., Kleinhofs, A., Murray, T. and Jones, S. (2011) ' Polymorphic nuclear gene sequences indicate a novel genome donor in the polyploid genus Thinopyrum', *Hereditas*, 148(1), pp.8-27.

Amom, T., Nongdam, P. (2017) 'The use of molecular marker methods in plants': A review *International Journal of Current Research and Review*, pp.1-7.

Babu, R., Nair, SK., Prasann,a BM., Gupta, HS. (2004) 'Integrating marker assisted selection in crop breeding - prospects and challenges', *Curr Sci*.;87: pp. 607-619.

Baker, L., Grewal, S., Yang, C., Hubbart-Edwards, S., Scholefield, D., Ashling, S., Burridge, A., Przewieslik-Allen, A., Wilkinson, P., King, I. and King, J. (2020) 'Exploiting the genome of *Thinopyrum elongatum* to expand the gene pool of hexaploid wheat', *Theoretical and Applied Genetics*, 133(7), pp.2213-2226.

Bansal, M., Adamski, N., Toor, P., Kaur, S., Molnár, I., Holušová, K., Vrána, J., Doležel, J., Valárik, M., Uauy, C. and Chhuneja, P. (2020) '*Aegilops umbellulata* introgression carrying leaf rust and stripe rust resistance genes Lr76 and Yr70 located to 9.47-Mb region on 5DS telomeric end through a combination of chromosome sorting and sequencing', *Theoretical and Applied Genetics*, 133(3), pp.903-915.

Barker, G. and Edwards, K. (2009) 'A genome-wide analysis of single nucleotide polymorphism diversity in the world's major cereal crops', *Plant Biotechnology Journal*, 7(4), pp.318-325.

Baronti, S., Galassi, E., Ugolini F., Miglietta F., Genesio L., Vaccari F., Cacciatori P., and Gazza L. (2021) 'Agronomic and ecophysiological evaluation of an early establishment of perennial wheat lines in Central Italy', *Genetic Resources and*

Crop Evolution, 69, pp. 619–633Bedő, Z. and Láng, L. (2015) 'Wheat Breeding: Current Status and Bottlenecks', *Alien Introgression in Wheat*, pp.77-101.

Boopathi, N.M. (2013) 'Genetic Mapping and Marker Assisted Selection', *Springer*, *India*. Pp.173-186.

Bozzini, A., David, J. and Natoli, V. (2012) 'Origin and Distribution of Durum Wheat Genetic Diversity in the World', *Durum Wheat*, pp.1-14.

Braun HJ, Atlin G, Payne T. (2010) 'Multi-location testing as a tool to identify plant response to global climate change', *Climate change and crop production*, pp.115–138.

Burow, G., Chopra, R., Sattler, S., Burke, J., Acosta-Martinez, V. and Xin, Z. (2019) 'Deployment of SNP (CAPS and KASP) markers for allelic discrimination and easy access to functional variants for brown midrib genes *bmr6* and *bmr12* in *Sorghum bicolor*', *Molecular Breeding*, 39(8), pp.115.

Chapman, V. and Riley, R. (1970) 'Homoeologous Meiotic Chromosome Pairing in *Triticum aestivum* in which Chromosome 5B is replaced by an Alien Homoeologue', *Nature*, 226(5243), pp.376-377.

Ceoloni, C., Kuzmanovic, L., Forte, P., Virili, M. and Bitti, A. (2015) 'Wheat-Perennial Triticeae Introgressions: Major Achievements and Prospects', *Alien Introgression in Wheat*, pp.273-313.

Chandra, S., Singh, D., Pathak, J., Kumari, S., Kumar, M., Poddar, R., Balyan, H., Prabhu, K., Gupta, P. and Mukhopadhyay, K. (2017) 'SNP discovery from next-generation transcriptome sequencing data and their validation using KASP assay in wheat (*Triticum aestivum* L.)', *Molecular Breeding*, 37(7):92.

Chen, S., Huang, Z., Dai, Y., Qin, S., Gao, Y., Zhang, L., Gao, Y. and Chen, J. (2013) 'The Development of 7E Chromosome-Specific Molecular Markers for *Thinopyrum elongatum* Based on SLAF-seq Technology', *PLoS ONE* 8(6): p.e65122.

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 Biology*, 21(1), pp.1-11.

Comai, L. (2005) 'The advantages and disadvantages of being polyploid', *Nature Reviews Genetics*, 6(11), pp.836-846.

Cox TS., Van Tassel DL., Cox CM., DeHaan LR. (2010) 'Progress in bree
grains', *Crop Pasture Science*, 61, pp. 513–521.

Cseh, A., Yang, C., Hubbart-Edwards, S., Scholefield, D., Ashling, S., Burridge, A., Wilkinson, P., King, I., King, J. and Grewal, S. (2019) 'Development and validation of an exome-based SNP marker set for identification of the St, Jr and Jvs genomes of *Thinopyrum intermedium* in a wheat background', *Theoretical and Applied Genetics*, 132(5), pp.1555-1570.

Curtis, T. and Halford, N. (2014) 'Food security: the challenge of increasing wheat yield and the importance of not compromising food safety', *Annals of Applied Biology*, 164(3), pp.354-372.

Dempewolf, H., Baute, G., Anderson, J., Kilian, B., Smith, C. and Guarino, L. (2017) 'Past and Future Use of Wild Relatives in Crop Breeding', *Crop Science*, 57(3), pp.1070-1082.

Devi, U., Grewal, S., Yang, C., Hubbart-Edwards, S., Scholefield, D., Ashling, S., Burridge, A., King, I. and King, J. (2019) 'Development and characterisation of interspecific hybrid lines with genome-wide introgressions from Triticum timopheevii in a hexaploid wheat background', *BMC Plant Biology*, 19(1), pp.1-14.

Dewey, D. R. (1984) 'The genomic system of classification as a guide to intergeneric hybridization with the perennial Triticeae', *Stadler genetics symposia* (USA), pp.209 – 279.

Dhakal, S., Tan, C., Anderson, V., Yu, H., Fuentealba, M., Rudd, J., Haley, S., Xue, Q., Ibrahim, A., Garza, L., Devkota, R. and Liu, S. (2018) 'Mapping and KASP marker development for wheat curl mite resistance in “TAM 112” wheat using linkage and association analysis', *Molecular Breeding*, 38(10), pp. 1-13.

Dvorák, J., Edge, M., Ross, K. (1988) 'On the evolution of the adaptation of *Lophopyrum elongatum* to growth in saline environments', *Proceedings of the National Academy of Sciences*, 85(11), pp.3805–3809.

Endo, T.R., Mukai, Y., Yamamoto, M., Gill, B.S. (1991) 'Physical mapping of a male-fertility gene of common wheat', *The Japanese Journal of Genetics* 66, pp.291–295.

Fan, X., Zhang, H., Sha, L., Zhang, L., Yang, R., Ding, C. and Zhou, Y. (2007) 'Phylogenetic analysis among Hystrix, Leymus and its affinitive genera (Poaceae: Triticeae) based on the sequences of a gene encoding plastid acetyl-CoA carboxylase', *Plant Science*, 172(4), pp.701-707.

Fao.org. (2017) [online] Available at: http://www.fao.org/home/en/ [Accessed 30 Jan. 2018].

Fao.org. (2022) [online] *FAOSTAT*. Available at: https://www.fao.org/faostat/en/#data [Accessed 20 May 2022].

Faris, J. (2014) 'Wheat Domestication: Key to Agricultural Revolutions Past and Future', *Genomics of Plant Genetic Resources*, pp.439-464.

Fedak, G., Chi, D., Wolfe, D., Ouellet, T., Cao, W., Han, F. and Xue, A. (2021) 'Transfer of fusarium head blight resistance from *Thinopyrum elongatum* to bread wheat cultivar Chinese Spring' *Genome*, 64(11), pp.997-1008.

Feldman, M. (2001) 'Origin of Cultivated Wheat', In: A. P. Bonjean and W. J. Angus, Ed., The World Wheat Book: A History of Wheat Breeding, Intercept Ltd., London, 2001, pp.3-56.

Feldman, M. and Levy, A. A. (2015) 'Origin and Evolution of Wheat and Related Triticeae Species', in *Alien Introgression in Wheat*. Cham: *Springer* International Publishing. pp.21–76.

Feuillet, C., Langridge, P. and Waugh, R. (2008) 'Cereal breeding takes a walk on the wild side', *Trends in Genetics*, 24(1), pp.24-32.

Forster, BP., Miller, TE. (1989) 'Genome relationship between *Thinopyrum bessarabicum* and *Thinopyrum elongatum*', *Genome* 32: pp.930–931.

Fu, S., Lv, Z., Qi, B., Guo, X., Li, J., Liu, B. and Han, F. (2012) 'Molecular cytogenetic characterisation of wheat *Thinopyrum elongatum* addition, substitution and translocation lines with a novel source of resistance to wheat Fusarium head blight', *Journal of Genetics and Genomics*, 39, pp.103-110.

Friebe, B., Jiang, J., Knott, D. and Gill, B. (1994) 'Compensation Indices of Radiation-Induced Wheat-*Agropyron elongatum* Translocations Conferring Resistance to Leaf Rust and Stem Rust', *Crop Science*, 34(2), p.400.

Gaál, E., Valárik, M., Molnár, I., Farkas, A., Linc, G. (2018) 'Identification of COS markers for *Thinopyrum elongatum* chromosomes preliminary revealed high level of macro syntenic relationship between the wheat and *Th. el* genomes', *PLoSONE* 13: e0208840.

Gazza, L., Galassi, E., Ciccoritti, R. (2016) 'Qualitative traits of perennial wheat lines derived from different Thinopyrum species', *Genetic Resources and Crop Evolution*, 63, pp. 209–219.

Glover JD. (2005) 'The necessity and possibility of perennial grain production systems', Renew Agric Food Syst 20, pp. 1–4.

Gong B, Zhang H, Yang Y, Zhang J, Zhu W, Xu L, Wang Y, Zeng J, Fan X, Sha L, Zhang H, Wu D, Chen G, Zhou Y, Kang H. (2022) 'Development and Identification of a Novel Wheat-*Thinopyrum scirpeum* 4E (4D) Chromosomal Substitution Line with Stripe Rust and Powdery Mildew Resistance', *The American Phytopathological Society*; 106(3, pp. 975-983.

Grewal, S., Yang, C., Edwards, S., Scholefield, D., Ashling, S., Burridge, A., King, I. and King, J. (2018a) 'Characterisation of *Thinopyrum bessarabicum* chromosomes through genome-wide introgressions into wheat', *Theoretical and Applied Genetics*, 131(2), pp.389-406.

Grewal, S., Hubbart-Edwards, S., Yang, C., Scholefield, D., Ashling, S., Burridge, A., Wilkinson, P., King, I. and King, J. (2018b) 'Detection of *T. urartu* Introgressions in Wheat and Development of a Panel of Interspecific Introgression Lines', *Frontiers in Plant Science*, 9, 1565.

Grewal, S., Hubbart-Edwards, S.,Yang, C.,Devi,U., Baker, L., Heath, J., Ashling, S., Scholefield, D., Howells, C., Yarde, J., Isaac, P., King, I. P., and King, J. (2020) 'Rapid identification of homozygosity and site of wild relative introgressions in wheat through chromosome-specific KASP genotyping assays', *Plant Biotechnology Journal*, 18, pp.743–755.

Grewal, S., Guwela, V., Newell, C., Yang, C. Y., Ashling, S., Scholefield, D., Hubbart-Edwards, S., Burridge, A., Stride, A., King, I. P., and King, J. (2021) 'Generation

of doubled haploid wheat–*Triticum urartu* introgression lines and their characterisation using chromosome-specific KASP markers', *Frontiers in Plant Science*, *12*, pp.643-636.

Grewal, S., Coombes, B., Joynson, R., Hall, A., Fellers, J., Yang, C., Scholefield, D., Ashling, S., Isaac, P., King, I.P., King, J. (2022) 'Chromosome-specific KASP markers for detecting *Amblyopyrum muticum* segments in wheat introgression lines', *The Plant Genome,* 15: e220193.

Griffiths S., Sharp R., Foote TN., Bertin I., Wanous M., Reader S., Colas I., Moore G. (2006) 'Molecular characterization of Ph1 as a major chromosome pairing locus in polyploid wheat', *Nature,* 439, pp.749-752.

Han, G., Liu, S., Jin, Y., Jia, M., Ma, P., Liu, H., Wang, J. and An, D. (2020) 'Scale development and utilization of universal PCR-based and high-throughput KASP markers specific for chromosome arms of rye (*Secale cereale* L.) ', *BMC Genomics*, 21(1), pp. 1-14.

Hayes RC., Newell MT., DeHaan LR., Murphy KM., Crane S., Norton MR., Wade LJ., Newberry M., Fahim M., Jones SS., Cox TS., Larkin PJ. (2012) ' Perennial cereal crops: an initial evaluation of wheat derivatives', *Field Crops Research* 133, pp. 68–89.

Hu, L., Liu, C., Zeng, Z., Li, G., Song, X. and Yang, Z. (2012) 'Genomic rearrangement between wheat and *Thinopyrum elongatum* revealed by mapped functional molecular markers', *Genes Genomics*, 34(1), pp.67-75.

Hsiao, C., Chatterton, N., Asay, K. and Jensen, K. (1995) 'Phylogenetic relationships of the monogenomic species of the wheat tribe, Triticeae (Poaceae), inferred from nuclear rDNA (internal transcribed spacer) sequences', *Genome*, 38(2), pp.211-223.

Jauhar, P., Doğramaci, M. and Peterson, T. (2004) 'Synthesis and cytological characterization of trigeneric hybrids of durum wheat with and without Ph1', *Genome*, 47(6), pp.1173-1181.

Jiang, J., Friebe, B., Dhaliwal, H., Martin, T. and Gill, B. (1993) 'Molecular cytogenetic analysis of *Agropyron elongatum* chromatin in wheat germplasm specifying

resistance to wheat streak mosaic virus', *Theoretical and Applied Genetics*, 86(1), pp.41-48.

Jiang, G. (2013) 'Molecular Markers and Marker-Assisted Breeding in Plants', *Plant Breeding from Laboratories to Fields*, pp.45-83.

Kaur, S., Francki, M. and Forster, J. (2012) 'Identification, characterization and interpretation of single-nucleotide sequence variation in allopolyploid crop species', *Plant Biotechnology Journal*, 10(2), pp.125-138.

King, J., Armstead, I., Harper, J., Ramsey, L., Snape, J., Waugh, R., James, C., Thomas, A., Gasior, D., Kelly, R., Roberts, L., Gustafson, P. and King, I. (2013) 'Exploitation of interspecific diversity for monocot crop improvement', *Heredity*, 110(5), pp.475-483.

King, J., Grewal, S., Yang, C.Y., Hubbart, S., Scholefield, D., Ashling, S., Edwards, K.J. et al. (2017) 'A step change in the transfer of interspecific variation into wheat from *Ambylopyrum muticum*', *Plant Biotechnology Journal,* 15, pp.217–226.

King, J., Grewal, S., Yang, C., Hubbart Edwards, S., Scholefield, D., Ashling, S., Harper, J., Allen, A., Edwards, K., Burridge, A. and King, I. (2018) 'Introgression of *Aegilops speltoides* segments in *Triticum aestivum* and the [illegible] gametocidal genes', *Annals of Botany*, 121(2), pp.229-240.

Kundan, M.K., Fougat, Ballani, Vinita, T., Yachana, J., and Madhumati, B. (2014) 'Potential and application of molecular markers techniques for plant genome analysis', *International Journal of Pure & Applied Bioscience*, 2 (1): pp.169-188.

Konkin D., Hsueh YC., Kirzinger M., Kubaláková M., Haldar A., Balcerzak M., Han F., Fedak G., Doležel J., Sharpe A., Ouellet T. (2022) 'Genomic sequencing of Thinopyrum elongatum chromosome arm 7EL, carrying fusarium head blight resistance, and characterization of its impact on the transcriptome of the introgressed line CS-7EL', *BMC Genomics,* 23;23(1), p.228.

Kosina, R. (1996) 'Molecular Cytogenetics of an Amphiploid Trigeneric Hybrid between *Triticum durum*, *Thinopyrum distichum* and *Lophopyrum elongatum*', *Annals of Botany*, 78(5), pp.583-589.

Li, D., Long, D., Li, T., Wu, Y., Wang, Y., Zeng, J., Xu, L., Fan, X., Sha, L., Zhang, H., Zhou, Y. and Kang, H. (2018) 'Cytogenetics and stripe rust resistance of wheat–

Thinopyrum elongatum hybrid derivatives', *Molecular Cytogenetics*, 11(1), pp. 1-9.

Li, X., Jiang, X., Chen, X., Song, J., Ren, C., Xiao, Y., Gao, X. and Ru, Z. (2017) 'Molecular cytogenetic identification of a novel wheat-*Agropyron elongatum* chromosome translocation line with powdery mildew resistance', *PLOS ONE*, 12(9), p.e0184462.

Li XM, Lee BS, Mammadov AC, Koo BC, Mott IW, Wang RRC. (2007) 'CAPS markers specific to Eb, Ee and R genomes in the tribe Triticeae', *Genome* 50: pp.400–411

Liu, Z., Chen, Z., Pan, J., Li, X., Su, M., Wang, L., Li, H. and Liu, G. (2008) 'Phylogenetic relationships in Leymus (Poaceae: Triticeae) revealed by the nuclear ribosomal internal transcribed spacer and chloroplast trnL-F sequences', *Molecular Phylogenetics and Evolution*, 46(1), pp.278-289.

Lou, H., Dong, L., Zhang, K., Wang, D., Zhao, M., Li, Y., Rong, C., Qin, H., Zhang, A., Dong, Z. and Wang, D. (2017) 'High-throughput mining of E-genome-specific SNPs for characterizing *Thinopyrum elongatum* introgressions in common wheat', *Molecular Ecology Resources*, 17(6), pp.1318-1329.

Ma, J., Zhou, R., Dong, Y. and Jia, J. (2000) 'Control and inheritance of resistance to yellow rust in *Triticum aestivum- Lophopyrum elongatum* chromosome substitution lines', *Euphytica*, 111(1), pp.57-60.

Makhoul, M., Rambla, C., Voss-Fels, K., Hickey, L., Snowdon, R. and Obermeier, C. (2020) 'Overcoming polyploidy pitfalls: a user guide for effective SNP conversion into KASP markers in wheat', *Theoretical and Applied Genetics*, 133(8), pp.2413-2430.

Martin, A.C. Rey, M., Shaw, P., and Moore G., (2017) `Dual effect of the wheat Ph1 locus on chromosome synapsis and crossover`. *Chromosoma*, pp.1–12.

Martin A.C., Shaw P., Phillips D., Reader S., Moore G., (2014) `Licensing MLH1 sites for crossover during meiosis`. *Nat Commun* 5, pp.4580.

Mason-Gamer, RJ., Orme, NL., Anderson, CM. (2002) 'Phylogenetic analysis of North American Elymus and the monogenomic Triticeae (Poaceae) using three chloroplast DNA data sets', *Genome* 45: pp.991–1002.

Mestiri, I., Chagué, V., Tanguy, A., Huneau, C., Huteau, V., Belcram, H., Coriton, O., Chalhoub, B. and Jahier, J. (2010) 'Newly synthesized wheat allohexaploids display progenitor-dependent meiotic stability and aneuploidy but structural genomic additivity', *New Phytologist*, 186(1), pp.86-101.

Miki, Y., Yoshida, K., Mizuno, N., Nasuda, S., Sato, K. and Takumi, S. (2019) 'Origin of wheat B-genome chromosomes inferred from RNA sequencing analysis of leaf transcripts from section Sitopsis species of *Aegilops*', *DNA Research*, 26(2), pp.171-182.

Molnar-Lang, M. and Kruppa, K. (2016) 'Simultaneous visualization of different genomes (J, JSt and St) in a *Thinopyrum intermedium* × *Thinopyrum ponticum* synthetic hybrid (Poaceae) and in its parental species by multicolour genomic in situ hybridization (mcGISH)', *Comparative Cytogenetics*, 10(2), pp.283-293.

Moore, G. (2002) 'Meiosis in allopolyploids – the importance of 'Teflon' chromosomes', *Trends in Genetics*, 18(9), pp.456-463.

Moore, G. (2014) 'Wheat meiosis and the Ph1 locus', *John Innes Centre - The Moore Lab*. Available at: https://www.jic.ac.uk/staff/graham-moore/Wheat_meiosis.htm

Moore, G. (2015) 'Strategic pre-breeding for wheat improvement', *Nature Plants*, 1(3), p.15018.

Mujeeb-Kazi, A., Cortes, A., Gul, A., Farooq, M., Majeed, F., Ahmad, I., Bux, H., William, M., Rosas V. and Delgado, R. (2008) 'Production and cytogenetics of a new *Thinopyrum elongatum*/ *Triticum aestivum* hybrid, its amphiploid and backcross derivatives', *Pakistan Journal of Botany*, 40, pp.565-579.

Mullan, D., Platteter, A., Teakle, N., Appels, R., Colmer, T., Anderson, J. and Francki, M. (2005) 'EST-derived SSR markers from defined regions of the wheat genome to identify *Lophopyrum elongatum* specific loci', *Genome*, 48(5), pp.811-822.

Muramatsu, M. (1990) 'Cytogenetics of decaploid *Agropyron elongatum* (*Elytrigia elongata*) (2n = 70). I. Frequency of decavalent formation', *Genome* 33: pp.811 – 817.

Murphy KM., Lyon SR., Balow KA., Jones SS. (2010) 'Post-sexual cycle regrowth and grain yield of *Thinopyrum elongatum* x *Triticum aestivum* amphiploids', *Plant Breeding*, 129, pp. 480–483.

Neelam, K., Brown-Guedira, G.L., and Huang, L. (2012) 'Development and validation of a breeder friendly KASPar marker for wheat leaf rust resistance locus Lr21', *Molecular Breeding*, 31, pp.233-237.

Newell MT., Hayes RC. (2017) 'An initial investigation of forage production and feed quality of perennial wheat derivatives', *Crop Pasture Sci,* 68, pp. 1141–1148.

Oliver, R., Cai, X., Xu, S., Chen, X. and Stack, R. (2005) 'Wheat-Alien Species Derivatives: A Novel Source of Resistance to Fusarium Head Blight in Wheat', *Crop Science*, 45(4), pp.1353-1360.

Pienaar, RV. (1988) 'Hexaploid 'wheats' with novel third genomes', In: *Proceedings of 7th international wheat genet symposium*, Cambridge, UK, pp.403–407

Qureshi, N., Kandiah, P., Gessese, M., Nsabiyera, V., Wells, V., Babu, P., Wong, D., Hayden, M., Bariana, H. and Bansal, U. (2018) 'Development of co-dominant KASP markers co-segregating with Ug99 effective stem rust resistance gene Sr26 in wheat', *Molecular Breeding*, 38(8), pp. 1-9.

Ramzan, F., Younis, A. and Lim, K. (2017) 'Application of Genomic in Situ Hybridization in Horticultural Science', *International Journal of Genomics*, pp.1-12.

Rasheed, A., Wen, W., Gao, F., Zhai, S., Jin, H., Liu, J., Guo, Q. et al. (2016) 'Development and validation of KASP assays for genes underpinning key economic traits in bread wheat', *Theoretical and Applied Genetics,* 129(10), pp.1843–1860.

Rey, E., Molnár, I. and Doležel, J. (2015) 'Genomics of Wild Relatives and Alien Introgressions', *Alien Introgression in Wheat*, pp.347-381.

Rey, MD., Martín, A.C., Higgins, J. *et al.* (2017) 'Exploiting the *ZIP4* homologue within the wheat *Ph1* locus has identified two lines exhibiting homoeologous crossover in wheat-wild relative hybrids', *Molecular Breeding,* 37(8), pp. 1-11.

Riley R. and Chapman V. (1958) 'Genetic control of the cytological diploid behaviour of hexaploid wheat', *Nature,* 182(4637), pp.712–715.

Ryan MR., Crews TE., Culman SW., DeHaan LR., Hayes RC., Jungers JM., Bakker MG. (2018) 'Managing for multifunctionality in perennial grain crops', *Bioscience,* 68, pp. 294–304.

Schachermayr, GM., Messmer, MM., Feuillet, C., Winzeler, H., Keller, B. (1995) 'Identification of molecular markers linked to the *Agropyron elongatum*-derived leaf rust resistance gene *Lr24* in wheat', *Theoretical and Applied Genetics,* 90(7), pp.982–990.

Schwarzacher, T., Leitch, A. R., Bennett, M. D. and Heslop-Harrison, J. S. (1989) 'In situ localization of parental genomes in a wide hybrid', *Annals of Botany*, 64(3), pp.315–324.

Sears, E. and Okamoto, M. (1958) 'Intergenomic chromosome relationships in hexaploid wheat', *In Proceedings of the 10th International Congress of Genetics*. pp.258–259.

Sears, E. (1976) 'Genetic Control of Chromosome Pairing in Wheat', *Annual Review of Genetics*, 10(1), pp.31-51.

Sears, ER. (1977) 'An induced mutant with homoeologous chromosome pairing', *Canadian Journal of Genetics and Cytology* 19, pp.585-593.

Sears, E. (1984) 'Mutations in Wheat that Raise the Level of Meiotic Chromosome Pairing', *Gene Manipulation in Plant Improvement*, pp.295-300.

Semagn, K., Babu, R., Hearne, S. and Olsen, M. (2013) 'Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement', *Molecular Breeding*, 33(1), pp.1-14.

Sharma, H., Ohm, H., Lister, R., Foster, J., and Shukle, R. (1989) 'Response of wheatgrasses and wheat x wheatgrass hybrids to barley yellow dwarf virus', *Theoretical and Applied Genetics*, 77(3), pp.369-374.

Shen, X., Kong, L., and Ohm, H. (2004) 'Fusarium head blight resistance in hexaploid wheat (*Triticum aestivum*)-Lophopyrum genetic lines and tagging of the alien chromatin by PCR markers', *Theoretical and Applied Genetics*, 108: pp.808-813.

Shewry, P. (2009) 'Wheat', *Journal of Experimental Botany*, 60(6), pp.1537-1553.

Shewry PR, Hey SJ. (2015) 'The contribution of wheat to human diet and health', *Food Energy Secur*ity, 4(3): pp.178-202.

Silva, G., and Souza, M. (2013) 'Review Genomic in situ hybridization in plants', *Genetics and Molecular Research*, 12(3), pp.2953-2965.

Moose, S. and Mumm, R. (2008) 'Molecular Plant Breeding as the Foı
21st Century Crop Improvement', *Plant Physiology*, 147(3), pp.969-977.

Nadeem, M., Nawaz, M., Shahid, M., Doğan, Y., Comertpay, G., Yıldız, M., Hatipoğlu, R., Ahmad, F., Alsaleh, A., Labhane, N., Özkan, H., Chung, G. and Baloch, F. (2017) 'DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing', *Biotechnology and Biotechnological Equipment*, 32(2), pp.261-285.

Tan, C., Yu, H., Yang, Y., Xu, X., Chen, M., Rudd, J., Xue, Q., Ibrahim, A., Garza, L., Wang, S., Sorrells, M. and Liu, S. (2017) 'Development and validation of KASP markers for the greenbug resistance gene Gb7 and the Hessian fly resistance gene H32 in wheat', *Theoretical and Applied Genetics*, 130(9), pp.1867-1884.

Tsintin NV., Lubinova VF. (1959) 'New species and form of cereals derived from hybridization between wheat and couch grass'. *American Naturalist,* 93, pp. 181–191.

Tiwari, V. K., Wang, S., Sehgal, S., Vrána, J., Friebe, B., Kubaláková, M., Chhuneja, P., Doležel, J., Akhunov, E., Kalia, B., Sabir, J. and Gill, B. S. (2014) 'SNP Discovery for mapping alien introgressions in wheat', BMC Genomics, 15(1), pp. 1-11.

Tiwari, V.K., Faris, J.D., Friebe, B., Gill, B.S. (2016) 'Genome mapping', *Encyclopedia of Grain Science*, pp.365-375.

Wagoner P. (1990) 'Perennial grain new use for intermediate wheatgrass', *Journal of Soil and Water Conservation*, 45, pp. 81–82.

Wang, R., Marburger, J. and Hu, C. (1991) 'Tissue-culture-facilitated production of aneupolyhaploid *Thinopyrum ponticum* and amphidiploid *Hordeum violaceum* x *H. bogdanii* and their uses in phylogenetic studies', *Theoretical and Applied Genetics*, 81(2), pp.151-156.

Wang, R.R.C. (1985) 'Genome analysis of *Thinopyrum bessarabicum* and *T. elongatum*', *Canadian Journal of Genetics and Cytology*, 27, pp.722–728.

Wang, R.R.C. and Hsiao, C. (1989) 'Genome relationships between *Thinopyrum bessarabicum* and *T. elongatum*: revisited', *Genome*, 32(1984), pp.804–809.

Wang, R. R. C. (1992) 'Genome relationships in the perennial Triticeae based on diploid hybrids and beyond', *Hereditas* 116: pp.133 – 136

Wang, R.RC. (2011) '*Agropyron* and *Psathyrostachys*', Wild Crop Relatives: Genomic and Breeding Resources, *Springer,* pp 77–108.

Wang Y, Cao Q, Zhang J, Wang S, Chen C, Wang C, Zhang H, Wang Y, Ji W. (2020) 'Cytogenetic Analysis and Molecular Marker Development for a New Wheat-*Thinopyrum ponticum* 1Js (1D) Disomic Substitution Line with Resistance to Stripe Rust and Powdery Mildew', *Front Plant Science,* 11, 1282.

Wei, JZ., Wang, RRC. (1995) 'Genome and species-specific markers and genome relationships of diploid perennial species in Triticeae based on RAPD analyses', *Genome,* 38, pp.1230–1236.

Wilson, A.S. (1876) 'Wheat and rye hybrids', *Transactions of the Botanical Society of Edinburgh,* 12(1973), pp.286-288.

Yousafzai, F., Al-Kaff, N. and Moore, G. (2010) 'The molecular features of chromosome pairing at meiosis: the polyploid challenge using wheat as a reference', *Functional and amp; Integrative Genomics*, 10(2), pp.147-156.

Winfield, M. O., Allen, A. M., Burridge, A. J., Barker, G. L. A., Benbow, H. R., Wilkinson, P. A., Coghill, J., Waterfall, C., Davassi, A., Scopes, G., Pirani, A., Webster, T., Brew, F., Bloor, C., King, J., West, C., Griffiths, S., King, I., Bentley, A. R. and Edwards, K. J. (2016) 'High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool', *Plant Biotechnology Journal,* 14(5), pp.1195–1206.

Yu, H., Fan, X., Zhang, C., Ding, C., Wang, X. and Zhou, Y. (2008) 'Phylogenetic relationships of species in Pseudoroegneria (Poaceae: Triticeae) and related genera inferred from nuclear rDNA ITS (internal transcribed spacer) sequences', *Biologia*, 63(4), pp.498-505.

Yu, L., Chao, S., Singh, R. and Sorrells, M. (2017) 'Identification and validation of single nucleotide polymorphic markers linked to Ug99 stem rust resistance in spring wheat', *PLOS ONE*, 12(2), p.e0171963.

Zhang, W., Danilova, T., Zhang, M., Ren S., Zhu X., Zhang Q., Zhong S., Dykes L., Fiedler J., Xu S., Frels K., Wegulo S., Boehm J., & Cai X. (2022) 'Cytogenetic and genomic characterization of a novel tall wheatgrass-derived Fhb7 allele integrated into wheat B genome', *Theoretical and Applied Genetics.*

Zhang, X., Dong, Y. and Wang, R.R. (1996) 'Characterization of genomes and chromosomes in partial amphiploids of the hybrid *Triticum aestivum* x *Thinopyrum ponticum* by in situ hybridization, isozyme analysis, and RAPD', *Genome*, 39(6), pp.1062–1071.

Zu-Jun, Y. and Zheng-Long, R. (2001) 'Chromosomal distribution and genetic expression of *Lophopyrum elongatum* (Host) A. Löve genes for adult plant resistance to stripe rust in wheat background', *Genetic Resources and Crop Evolution*, 48(2), pp.183-187.

Appendices

Appendix 1

The protocol of Buffers and Solutions

a) DNA extraction buffer (To make 100ml)

0.1M Tris-HCl (pH 7.5)	10ml
0.05M EDTA (pH 8.0)	10ml
10% SDS	12.5ml
Water	67.5ml

b) 6M ammonium acetate ($NH_4C_2H_3O_2$) (for 50ml)

6M NH4C2H3O2	23.124g
Water	50ml

Store at 4°C

c) 3M Sodium Acetate (NaOAc, pH 5.2) for 100 ml

Sodium acetate (m.w.-136.08)	40.8 g
Water	100 ml

Adjust pH to 5.2

d) 90% Ethanol + 10% 3M NaOAc (pH 5.2)

To make 100ml:

Add 90ml of 100% Ethanol to 10ml of 3M Sodium Acetate*.

Store at room temperature.

e) Cell digestion enzyme (for 100ml)

Enzyme solution made on ice by mixing the enzyme in 1X Citric buffer*

Cellulase Onozuka R-10 (2% w/w)	0.2 g
Pectolyase Y-23 (1% w/w)	0.1 g
1X Citric Buffer (pH 5.5)*	9.7 g

The enzyme solution was divided into 20 µl aliquots into each 0.5ml tube on ice and stored at -20°C.

***For 100ml of 1x Citric Buffer**

5x Citric Buffer**	20 ml
Water	80 ml

****For 50ml of 5x Citric Buffer**

Sodium Citrate (50mM)	0.735 g
0.5M EDTA (50mM)	5 ml

Water 30 ml

Adjust the solution to pH 5.5 using citric acid

f) Salmon Sperm (SS) DNA (140 µg/ml) for 50 ml

50X TAE (pH 6.3)	5 ml
SS DNA (10mg/ml)	0.7 ml
Water	44.3 ml

Store at 4°C in a fridge

g) 2xSSC+1×TE solution for 20 ml

2×SSC*	2 ml
10×TE**	2 ml
Water	16 ml

***2× SSC (for 1l)**

20x SSC*	100 ml
Water	900 ml

***20x SSC (for 1l)**

NaCl 175.3g

Sodium citrate 88.2g

pH 7.0

Water upto 1l

****10 × T.E.**

100mM Tris

10mM EDTA

pH7.5

****1 × T.E.**

10xT.E	10 ml
Water	90 ml

h) Single colour-GISH probe mixture 10 µl per slide

2xSSC in 1x TE	X µl make up to 10 µl
Probe gDNA (green) (100ng/ µl)	0.5 - 1 µl
Blocking DNA	2 µl

i) Multi colour-GISH probe mixture 10 µl per slide

2xSSC in 1x TE	X µl make up to 10 µl
Probe gDNA (green) (100ng/ µl)	0.5 - 1 µl
Probe gDNA (aqua) (100ng/ µl)	0.5 - 1 µl
Probe gDNA (red) (100ng/ µl)	0.5 - 1 µl
Probe gDNA (gold) (100ng/ µl)	0.5 - 1 µl

j) Rox PACE mix

Make 500µl of 1:10 dilution of the Rox stock (Applied Biosystems)

Rox Stock	50µl
dH_2O	450µl

Store at -20C.

Appendix 2

Table 1: KASP assay primer details for all selected SNPs.

Type	**Primer_AlleleFAM**	**Primer_AlleleHEX**	**Primer_Common**
WRC0002	GGACAGCCAAAGAGTAAAAGGAGT	GGACAGCCAAAGAGTAAAAGGAGA	CAATTGATTCTGAACCTCACTGCAAGTAA
WRC0006	TGATTTGTAGCCAATGATTTAATTGTTCAT	GATTTGTAGCCAATGATTTAATTGTTCAG	GTTGTAGCCCCTAATAYGATATTAGTAGTT
WRC0022	AACATAACATGATAGATCAACCTGGGA	CATAACATGATAGATCAACCTGGGC	CGTTGCAACTTGCAGGACTCTTGTA
WRC0023	TCGTCGAACACTATTTGTCGTTC	CTTCGTCGAACACTATTTGTCGTTG	TGGTACCCTGTTTACAGCCCACTT
WRC0037	GCAGGATGAGCGTGATGGCG	GCAGGATGAGCGTGATGGCC	CGGAGGCGACGCTGGTGC
WRC0052	GAAGTTGCAGYGGCGGCGA	GAAGTTGCAGYGGCGGCGG	AGGGTGYCGCCCACATCCAC
WRC0057	AGAAAAATCTAAATCTAGCTAAATGTAATCTG	AGAAAAATCTAAATCTAGCTAAATGTAATCTC	CCTTTTATTTCAGCTTCATTAGRTGCACAT
WRC0069	TGATCTATGTTGTAGCCCCACAC	GTTGATCTATGTTGTAGCCCCACAT	CCCATATRAACTTCAAAATGAAGTAAGCAT
WRC0074	AGGAGCTGCGATTGTTCGTCC	CAGGAGCTGCGATTGTTCGTCT	GGAACCGCCTCACACAAGGGTA
WRC0076	CCAAGCAAGCGGTGTTGATACTG	CCAAGCAAGCGGTGTTGATACTC	CCAGCTCGCTCAACGACTTGCAAA
WRC0077	CAACACCACATTCAACTATATAACTCG	CCAACACCACATTCAACTATATAACTCA	CTTTTTACCGAATGTTGATCATTTGTGCCA
WRC0083	CAAATGGTACATGCTGGAGCGG	CAAATGGTACATGCTGGAGCGA	TCCATCCATGTGTTTGCCTGKGGT
WRC0089	CATTTAATATATGGATTCTGCAAAATAAGAAGA	TAATATATGGATTCTGCAAAATAAGAAGG	GATACCCATATTGCTGACGCRCTTATT
WRC0091	CCTTGGTCATCTGCTGGATGTTC	CCTTGGTCATCTGCTGGATGTTA	GATGCCGAAGAACGGAAAGATGAGT
WRC1323	GACGGGAGTGGATATTCATCG	GACGGGAGTGGATATTCATCA	AAGTGGAACCGTAACCACCG

WRC0102	AGAGGCAACATCGATGGAGCTC	AAGAGGCAACATCGATGGAGCTA	GGTCGATCTTAGGATTGACCAGAAATAAA
WRC0103	GGTCATCGTTGGCAACATCAACTTT	GTCATCGTTGGCAACATCAACTTG	GRAGCTCATGATCTCCGGAAGCTT
WRC0105	CAATCAGACACCAGAAGCCAGC	CCAATCAGACACCAGAAGCCAGA	CAGTGGTATGGTCTACTTGAATCTGCA
WRC0114	GGTCCTATTATGGATCTGATCCTA	GGTCCTATTATGGATCTGATCCTG	WTACCCACTGTCTGTCACGAAACCT
WRC0122	CGAGGAGAACGAGATGCTACAC	ACGAGGAGAACGAGATGCTACAT	ATTTTCTGCTCGTATTCTAGCCAC
WRC0125	GCAGGGATGGCCACAGTTCCA	CAGGGATGGCCACAGTTCCC	GTCGTCAAGGACAGYGATGTCTTC
WRC0128	GCTTAATTAACCTTTCTTGTACC	CTGCTGCTTAATTAACCTTTCTTGTACT	CTCTATAGGTACAGTAACATCTGGTAGT
WRC0131	TGCTTCATCAATGCCCCTTTCCTT	GCTTCATCAATGCCCCTTTCCTG	CCGGGTCGAAAGCWATAGATTCAAGAA
WRC0135	ACCACACATGGATACCAAGCTATCA	CACACATGGATACCAAGCTATCG	CTCATCAAGCTCATCAGGGTGTGTA
WRC0143	AGTAGCCATAGTATTGATGCTAGTTTC	CAGTAGCCATAGTATTGATGCTAGTTTT	GCGCCAGGAGGTGGCCCAA
WRC0150	AGTTGAGGAAGAGGAAGCATGTTG	AGTTGAGGAAGAGGAAGCATGTTC	CTGATCTTGGGTGTTTTGGGTATGATAAA
WRC0151	ATTGAAGAGTTCAAATACTTTAAGATGAAATCT	GAAGAGTTCAAATACTTTAAGATGAAATCA	CTKCATTAGTGATTGTGGACCATCCAT
WRC0152	GCCCGTCCAAGCTTTGTACTCA	CCCGTCCAAGCTTTGTACTCG	CAGGAACTCCATGACCGATGCAG
WRC0155	GGCTTGCTTCCAGATGCTTCG	CTGGCTTGCTTCCAGATGCTTCA	GTCATAGCTTCATGGGAGCTAGCTT
WRC0159	GTTGTYTTCGACGAGGAGGCAG	GTTGTYTTCGACGAGGAGGCAA	CCGRTGCCTCCCAGCTCCAA
WRC0179	CCCAAAGCAAAGGGATGTCACG	ACCCAAAGCAAAGGGATGTCACA	GCAACAGCTGTGGTTGACCAAGAAT
WRC0182	CAGGGCTATCGGAATTGTCTCG	ACAGGGCTATCGGAATTGTCTCA	GACCCTGACGATCAGAAGAGGCA
WRC0184	CAGGAGCTGCGCGGCCC	CAGGAGCTGCGCGGCCG	GGTGGGAAGCCAGTCTGCAAC

WRC0199	ATTCAGCAAATGAATATTTGAAGCTTCTTG	CAGCAAATGAATATTTGAAGCTTCTTA	TCATGTGCCGCTGGGATCACAAATA
WRC0208	ACCCTGCTTGTGCAAGGCGATT	CCCTGCTTGTGCAAGGCGATC	ATCAAGGAGGATGCTACCAACACTC
WRC0217	CACTTGTTGCCTGATTAAACAATAATATTCT	ACTTGTTGCCTGATTAAACAATAATATTCC	GTGGTTCTTTTAATCTAGAATAGCAGCCAT
WRC0218	AAATTTTGCTACCTTCGTTCTGTGGA	AATTTTGCTACCTTCGTTCTGTGGC	ATAGTTCAACAGACATGCCTGGACATAAA
WRC0219	AATGCTTCTCCGGGAGAACGTC	AAATGCTTCTCCGGGAGAACGTT	ATGAACTGGCCGTTGCGGTCG
WRC0225	GCCTTGGCATGGATWGGATCC	GGCCTTGGCATGGATWGGATCT	AGTGGCCCTRGCAACATGGAGTA
WRC0226	ATTGCCAGTATAAAGGCAGGCAC	GCCAGTATAAAGGCAGGCAG	TAGTATTCCACCAGCCAAGGGGTAA
WRC0230	TAYTTCAGCTTTTGATAGACTGGT	CTTAYTTCAGCTTTTGATAGACTGGC	AACTTATGTTTACAATTARAAGATGCACAT
WRC0236	GATCMAGCAGAGCAAGAACTCG	GATCMAGCAGAGCAAGAACTCC	GACGGCGAAGGGTTGCAGCAA
WRC0239	TCTTCTCTTGCAGCAGATTCATGG	CTTCTCTTGCAGCAGATTCATGC	SGGTGTCTCCAAAGGGAGGCTT
WRC0244	AGATGCGAGAGACTGGATCCGA	GATGCGAGAGACTGGATCCGG	CTCGTGCAYGAGTGTCAGGAATTG
WRC0248	AATGRTAAAAAGAAGCAACTCAATCAATCAAT	AATGRTAAAAAGAAGCAACTCAATCAATCAAA	GGCCTRATTGAGGATGACATATATTGTG
WRC0254	CCAAGGTAATATACAAATATTCACAGGGT	CAAGGTAATATACAAATATTCACAGGGC	TTKACGTGGCCTCTCTTCTGCAAG
WRC0255	GTGCTTACAGTTACAGTTATAGGGG	GGTGCTTACAGTTACAGTTATAGGGA	CATCAGAACCAGGACCTTTTGTCCAT
WRC0258	ATCAGTGAATCAAATACAGCAACGTCA	CAGTGAATCAAATACAGCAACGTCG	GAGCGGGTGCTTAAACAGATGACC
WRC0260	ATCTTGTCTCCGTCTCCTTYTTGAAT	CTTGTCTCCGTCTCCTTYTTGAAC	CGGGTACTCAAGCAAACTCGAATTGAT
WRC0267	AAGCACAATCCCACTAAACTCACAG	AAGCACAATCCCACTAAACTCACAA	CTGGATCTCATTAAGTCTGGCACGTT
WRC0275	GTATACTGATGTGATCCCATCATAGA	GTATACTGATGTGATCCCATCATAGC	GATGGACAGCAAGACTGCTACATAC

WRC0288	CATCAGAAAGTTCACAGGGATCTGA	ATCAGAAAGTTCACAGGGATCTGG	ATAGTTGGCATAATTACMAAAGTGATCCG
WRC0292	GGTGCACTAAGTTGAGGACACTG	GGTGCACTAAGTTGAGGACACTA	CAAGTGAGAGGATCAACAGGGATCT
WRC0293	AGCAATCTGCAGTATCTTGCCATTTA	GCAATCTGCAGTATCTTGCCATTTG	GCTGTCATTTGGCCAGGTGATTCAAT
WRC0300	CCTTTCTTGTTTYGCTTGAACTTAGG	TCCTTTCTTGTTTYGCTTGAACTTAGT	AGCTCAAGCTCCACGGGGGWAA
WRC0311	GAAACAAGCAGAGCTAAATAGACGAAAT	AAACAAGCAGAGCTAAATAGACGAAAC	AACTTTTCGGGGTGTAAATTTGTACCG
WRC0312	GRCTTCTTTATCCTGTGAAATTGATCA	CTTCTTTATCCTGTGAAATTGATCC	CCAGTCAGCACTACAATGTCACGAA
WRC0316	ATGCAAAGTCTACGACTTCGCTG	CTATGCAAAGTCTACGACTTCGCTA	CACCCCTAATTAGGTTCCATATATACAAG
WRC1339	CATCGCTACGAGCATCCTAG	CATCGCTACGAGCATCCTAC	GGCAGTAATCCTATGGTTCG
WRC0340	CCAGCGATGCCACTTCGAGC	CCAGCGATGCCACTTCGAGA	CCGCGGTGGATATGGTCGTGTT
WRC0345	CCGGCAGAATCCATGGGACAC	CCGGCAGAATCCATGGGACAT	CTGTAGTGATGAAATGGTGTTTCCTCAAA
WRC0350	GAGGACCTGAAGCCGAAACTC	CGAGGACCTGAAGCCGAAACTT	GCGAACCAGAGTTTCAGCAGYGA
WRC0355	GGACTGGTTCGCGAACGTCAG	GGACTGGTTCGCGAACGTCAA	CTTYGATCTGACGGCGTCGTGA
WRC0360	GTGCGAACCTCAAGGAGATAACC	GTGCGAACCTCAAGGAGATAACT	CCTCTGCGTCCTCAGTGCCT
WRC0363	GAAGCAAGGCTCTCACATCGG	GAAGCAAGGCTCTCACATCGC	CAATGCGTTGATGATAAAAAGGAAGTACAC
WRC0367	AAGGGTACATTGCCCTTCACAAGA	GGGTACATTGCCCTTCACAAGG	CCAAATTGGAGACAATAGYATTAGACGAG
WRC0370	ATGGAGCCATTCTTGTTGAAGGTG	CATGGAGCCATTCTTGTTGAAGGTA	CTGGAGTTCTCAAGCATCCGATCC
WRC0371	CCAGGAAAGAGCTTGCACTAATGATT	CAGGAAAGAGCTTGCACTAATGATC	CCTAAAACCCRAATGATCAACCATAGAC
WRC0377	GAAGGATCCGTCACTGAACACG	AGAAGGATCCGTCACTGAACACA	CAAGCGCTACATCGGGAGCATC

WRC0379	GGAAYGTTTGGCAGCTAACCCA	GGAAYGTTTGGCAGCTAACCCT	GAATTGRTCTGCTTCCTGTTGATCCTT
WRC0384	GCCTTTGCACGGRTCATTGGTG	GCCTTTGCACGGRTCATTGGTA	AAGTTTGCTGTTATGACACAACCTSTGA
WRC0389	TCCTAGCTATTAATTCCAAGAGCACA	CCTAGCTATTAATTCCAAGAGCACT	GACGTCTTACCCCAAAGAGTGATGTA
WRC0391	AACACAATCAGAACCAAACATGTGAAG	AACACAATCAGAACCAAACATGTGAAC	AATTATGTGACTTATGTAATATRGTACCTG
WRC0392	TSAAATTTAACATTATCCTCCCTGG	CACTTSAAATTTAACATTATCCTCCCTGA	GGATGTTGSACAAGTCACTCCCG
WRC0396	ATTCTTTTRTGACGAAGGGCACATC	CATTCTTTTRTGACGAAGGGCACATT	TCCTCRGAGATTTCAGTCAACTCA
WRC0399	GCAGCTTCCTTCATCATGTCCG	GCAGCTTCCTTCATCATGTCCC	AAGGGGCTCTTCTTGGATAAAWCATCATA
WRC0401	TTAGGCCTTATGTTTGTCAAGTGTCA	AGGCCTTATGTTTGTCAAGTGTCG	TGACTACACATTTTACCCCGCCAAAAAAT
WRC0403	AAGCATTATGAATGCACAGACGACG	AAGCATTATGAATGCACAGACGACC	TGTATACAGAGAATACATGGGTTTCCGTA
WRC0405	CAAAAGTACTGCGTTTTACCCTCCAT	AAAAGTACTGCGTTTTACCCTCCAC	GACARGACAGTCTCGGTCAAAGCAT
WRC0412	ACGARATCGAGCGTGAGAAAGATT	CGARATCGAGCGTGAGAAAGATC	ACCGGCCTCGCCCTRCTCTTT
WRC0417	ACTTTGTCACATGACGTGGAAACG	GACTTTGTCACATGACGTGGAAACA	ATGTTCTAGACATCAAGGTTATCAGTACG
WRC0418	TCAGTGGTCTTGTGGCCGGA	CAGTGGTCTTGTGGCCGGG	GAGCAGCAAGCAAGCAAAGCCAGT
WRC0436	CACTGCTTACTCTAATGAAAAGGAAAACT	CTGCTTACTCTAATGAAAAGGAAAACC	CTTGGGGTTGGCCTCATTGCTCTT
WRC0439	AGAGCATGTTTATCAATTCACTGCT	AGAGCATGTTTATCAATTCACTGCC	ACAATTCAGAGAAGCATATCATGAAGCAAA
WRC0461	CAGTGCAAGGCTACAAGCA	CTCAGTGCAAGGCTACAAGCG	CATGCCSCATCGACTCCCGAT
WRC0464	ACTGGGACTGGAATGCAACCAAA	CTGGGACTGGAATGCAACCAAC	GGTTCGTCACCTATCGTTGCCTGT
WRC0466	GTGTGCGGCTGATCTCTCCC	CGTGTGCGGCTGATCTCTCCT	CTGCTTCGGCTCATGCCCTCTA

WRC0474	GCCCCGGATCCTGATTTTGCC	GCCCCGGATCCTGATTTTGCA	CATCGTTAACTTTGCAGGTGGACGTT
WRC0476	GAAGTGATTAGATAACACTAGTTTGTG	GAAGTGATTAGATAACACTAGTTTGTC	TCGTGATGATGTTGCTAGTTGTACCTTTA
WRC1360	CGATTATCGCACCGTCACAT	CGATTATCGCACCGTCACAC	ATTATCGAGCCTCAGCGA
WRC0492	AACATCATGGTCCAAGAGTAGAACG	AACATCATGGTCCAAGAGTAGAACC	CCTGAAGTATTCCCTTCTCTGAGCAA
WRC0494	CATCCAACATGAATAGCTAGACAACAT	ATCCAACATGAATAGCTAGACAACAG	ATCAAAGATTTAGCAGGAAAGGAAGTTTAG
WRC1364	GGAAGTGAGATGGTGGTACT	GGAAGTGAGATGGTGGTACC	ATCAGTGCTACATACAGTAGG
WRC0518	CGGCACCAAGCCTGGC	CCTCGGCACCAAGCCTGGT	CCATACGGCAGATAGTGGCAGATT
WRC0521	GATTGTCAACCAATGGCAGGAGT	GTCAACCAATGGCAGGAGC	CCTACTATCAGTGTATGGATATGGCCA
WRC0523	GGTCCATCAGCAGCCCTGG	GGTCCATCAGCAGCCCTGA	GTTTCAGTTTTGAGAGACGAGGTTAACAA
WRC0531	GAAACTGATTACGCGACTGACTACAA	AAACTGATTACGCGACTGACTACAG	GTCATGTCTTGAACTCTRGCTAGATCTT
WRC0536	ACCAATTGAGTCCAATCTTCTGCTAT	CCAATTGAGTCCAATCTTCTGCTAG	ATCTTGAGAATGGATGCTCATCCTGTA
WRC0538	GTAGCACCAAAAACATCGTTAAAATCG	CGTAGCACCAAAAACATCGTTAAAATCA	GATGTTGTCCTTCCTGAAAATGTTGAGTT
WRC0542	GATTTACCGGACCAGCAAGCCA	TACCGGACCAGCAAGCCG	GTCCGCTGTCAGCGTAGATAGATTAAA
WRC0544	CATCCTTTTGTGAATAGATCTATGCGA	ATCCTTTTGTGAATAGATCTATGCGG	AAGAATTTGGGCGTGTATATGTCCTCATA
WRC0549	CAACSTCACCCATGACAGCGA	CAACSTCACCCATGACAGCGC	GCGGGGGTGTATTGACGGCRA
WRC0551	CCGTTGCCTGAGAAATCTGTC	CCGTTGCCTGAGAAATCTGTG	GTTATCGTACATCTTCAGAGGGTTCATT
WRC1302	CAGATTTCGACACCCATGGAGC	CAGATTTCGACACCCATGGAGT	GACAAAGAGTTCCCAGCTTCTGGTA
WRC1368	CCAACAGCTCAATAGTGCG	CCAACAGCTCAATAGTGCA	AGATTTTCAGGCTTCAGG

WRC1370	AAATTCTCAGTTCTGCCTGGT	AAATTCTCAGTTCTGCCTGGC	GGAAGAAACAGCACAGCGA
WRC0569	GAAGGTGTACACTTCTTGGTATCCT	AAGGTGTACACTTCTTGGTATCCC	CCAGGGATTGAAAAACTGGTCCTYTTT
WRC0571	ATGCAGAAACAGTGAGGCAGCTT	GCAGAAACAGTGAGGCAGCTG	CAGTGAAATAATCTGCCTGAAGATACCAT
WRC0575	TTGGCTGTCTCGTTCATAACCACA	GGCTGTCTCGTTCATAACCACG	GCCTTATCAGTATCAACTTGAGCTTCAAA
WRC0578	CTAATCCCTCACACGCGTCTG	TCTAATCCCTCACACGCGTCTA	GSCGTGGAACCTTTCGATCCG
WRC0596	GACATCACAAACAGCTGCA	GCTGACATCACAAACAGCTGCG	AAATTGCGAAYATCTGTCGGACTAAAATAC
WRC0597	TCGTACTGCAACCGGAATTCTCAT	CGTACTGCAACCGGAATTCTCAC	ACACCTTTGGGAGCTTGACATGGAA
WRC0599	ACGAGGACTCGATTGCATTTTAGC	GACGAGGACTCGATTGCATTTTAGT	CGGTCCACACACTTTGCAAACATGAT
WRC0601	TGCCACGAAAACTGCAGGTACTA	GCCACGAAAACTGCAGGTACTC	ACCTTCTTCCAATTCCAATCTTTCCCAT
WRC0607	GAGGCTACGGGAGGAAGCCA	AGGCTACGGGAGGAAGCCG	CTCTCCTCGAAGAATTGATTCTGTGC
WRC0608	CCTTGAAGAATTGACTCTGTGCCT	CTTGAAGAATTGACTCTGTGCCG	GCACAGAATCAACTCTTYGAGGAGA
WRC1137	CACCGTGTATAGCAACGAAACCC	ATCACCGTGTATAGCAACGAAACCT	GTAGGTGAAGGAGAAGCCAACGAT
WRC0609	ACATCTCCTTTGGCAACAACARCTA	ACATCTCCTTTGGCAACAACARCTT	GGAGGCGTRGGCATCATCTTTAGAA
WRC0614	CCCGTAAGTTTGTTTGTGGAAATGAC	CCCGTAAGTTTGTTTGTGGAAATGAT	TGAATGGCATTCCACGAAATTCAGAGAAT
WRC0617	GATAAAACAAGTAAGGCTCCAATATCCAT	AAAACAAGTAAGGCTCCAATATCCAG	CATTAGGGTGGTTGAGCAAGGTGG
WRC0620	CGTGAAGGTTAAGGCGATGGCA	GTGAAGGTTAAGGCGATGGCG	CGGAGGACAGTTGCAACTACCG
WRC0623	GATCTGAAATTGTTCATATATTGCTGCG	GATCTGAAATTGTTCATATATTGCTGCA	TGAGGTGTTGTCTAATGGTGAGAAAATTG
WRC0624	CAGGATGACACCAGGTTTCAG	CTCAGGATGACACCAGGTTTCAC	ACATATGGGCCAGGCAAATGGTCAT

WRC0633	AGGAGAAGGGATCTACGAGGAC	AAGGAGAAGGGATCTACGAGGAT	GCCCACAGGACGCCGACG
WRC0646	CCAAGATAAGGTGGGCAGATAACTTA	CAAGATAAGGTGGGCAGATAACTTG	GACATGGGAATGGAAATGAATGAAAGYAAT
WRC0655	GTGCATGATAGAGGCCAACTACA	GTGCATGATAGAGGCCAACTACT	CCTCCCCAGTCGTCATGGAGTT
WRC1229	GCTGCATGGGTTCACATGGA	GCTGCATGGGTTCACATGGG	AAGATCTCCTCCATGCTCAGCTCTA
WRC0665	TCGGCCGCTCCTGCGGA	CGGCCGCTCCTGCGGG	GCGGAGGAGGTGGTTCTCAACAA
WRC0668	GTTTACTTCCAGATTTGATTGGATTCG	AAAGTTTACTTCCAGATTTGATTGGATTCA	TCCGATACRAAGCAAAATAACATGCTGTTT
WRC0672	AAGGGATAGATGTTGTATACTTGCCAT	AAGGGATAGATGTTGTATACTTGCCAA	ATGTGATCAAGATAGATGGGTACTCAAG
WRC0681	GTGGTTGATTGGTGGCTCTGA	GTGGTTGATTGGTGGCTCTGG	CGCAAATTCGATGAAAAGAACCGCAG
WRC0684	CCCTCACTGGTACCATCCCAAA	CCTCACTGGTACCATCCCAAC	ACTGAGATGGTTGTGACTAAGATTGAGAA
WRC0685	ACCTCGAAAGAAGTAGCAATACG	ATCTACCTCGAAAGAAGTAGCAATACT	TACATCATAACATTCCATACAAGCGAAGTA
WRC0688	GTCGTTTACTTGGCTTATCAAACTA	GTCGTTTACTTGGCTTATCAAACTC	TGGGCCACGCGCCTCTTTCAT
WRC0690	CAYCGAGGCGGGCGTCG	CAYCGAGGCGGGCGTCC	TWTCGCATGGGGTCGAGCGTG
WRC0698	GGAATGGGCGCTCCACTGC	GGAATGGGCGCTCCACTGG	CTCTTGACAGTAACATTACACCAACTCAA
WRC0701	TTGTACTTTTAAGCTGCCACATCTCA	GTACTTTTAAGCTGCCACATCTCG	RACAGGTTTTGTACAATTTTACAAATCAA
WRC0706	ATGGCCATGTACATCTCGCTGG	CATGGCCATGTACATCTCGCTGT	TCSGCGGCCGACACGGCA
WRC0712	GTCTCAATAATAYTATTCTCTCTTTTGACTT	CTCAATAATAYTATTCTCTCTTTTGACTC	AAAATGTGGAAAGAGATRCATCGATCATC
WRC0725	CAAGGCCCAACCCTAGGCC	CCAAGGCCCAACCCTAGGCT	GTTCGTCTAAATGTTGCGATCACAC
WRC0731	TCCCTACCGTGTGCC	CATGCTTCCCTACCGTGTGCT	GCTAGGTAGTCAAGGACAAGGTGAT

WRC0732	GTGGTAGAAATTTGTATGATCATCATGTAT	GGTAGAAATTTGTATGATCATCATGTAC	TGAGTAAGCGTGGTACTTTTCAGAATG
WRC0733	ACGAAATGCAACAGATGCTGCGA	CGAAATGCAACAGATGCTGCGG	ATGTTGGTAGATGCTCCACCCATG
WRC0737	CAAAAGTGGACTGAAGGGTGGC	TCAAAAGTGGACTGAAGGGTGGT	CAGCCTCCTCAGCCTTAACAATATGA
WRC0747	AGGTAGAATCGAATCTCTTTGCTTGTA	AGGTAGAATCGAATCTCTTTGCTTGTT	TCTGGAAGCCACGCGACGCAAT
WRC0751	ATCCTGGACCGAGTACTGTTCC	GATCCTGGACCGAGTACTGTTCT	TGCAGATCTCATTCCCCACCAGATT
WRC1161	AGTATGTTCCCGAGCTCAACCA	AGTATGTTCCCGAGCTCAACCT	GGTAGTCCTGATCAGAGACACCAAA
WRC0760	ARTGCCGTCCTCTTCTGCCAA	GCCGTCCTCTTCTGCCAG	CGCATTGAAGACCCTCCTCAAGAG
WRC0763	ACATCCAACAACTCATTTCTGTCTCT	CATCCAACAACTCATTTCTGTCTCC	TTGCTTCGCAGGTATCTAGAGCTATATTT
WRC0764	TTTCCRAATGTAAATTACATCTTTTTACC	CCTTTTCCRAATGTAAATTACATCTTTTTACT	CGGCAGTYGATGAGAGAGGGG
WRC0768	AGACGAACGTTAGCAGCTGTCG	AGACGAACGTTAGCAGCTGTCC	TATGCGAGACGTTTGACAGAAAATGGTTA
WRC0769	AACACCTGTAGTCCTAGATCACC	AACACCTGTAGTCCTAGATCACG	AGGTTTYAGGTCGAATATGAAGGAACAAT
WRC0771	GGATTGTGTACACCCTTCACCGA	GTGTACACCCTTCACCGG	TTTCTCGAGATATGGWGCGCGACAT
WRC0774	GACATCCATCGATAAAACTTGTTAACAAC	GACATCCATCGATAAAACTTGTTAACAAT	AGACCCCAAGTTTGATTTTGTGCAAGTTA
WRC0777	GCCATCTCGAAGTCCTCCTGT	GCCATCTCGAAGTCCTCCTGC	AGCCGGAATGTTTGCTCTCCGC
WRC0778	ACCATCAGTATGGGACACAACCT	CCATCAGTATGGGACACAACCC	GAAAAGCTTGCCCACTTACACTCTGT
WRC0793	ATAATATTTTGCCCCTTTTATATCGTTATTATC	ATAATATTTTGCCCCTTTTATATCGTTATTATA	ATGCCTGCATCAGGAAACCATGTTRTT
WRC0801	CCTCCTCCGACGGCTTG	GCTCCTCCTCCGACGGCTTA	ATCGCATTCTACATCTTTGGCCCGA
WRC0802	ATGGCTCAAGTTCACATGAATGCTG	ATGGCTCAAGTTCACATGAATGCTC	GACCTCATCAAACCCAAAACCTGCAAT

WRC0803	AAAGATGCTGAAACGAGTGGTCTC	AAAGATGCTGAAACGAGTGGTCTG	CGGCATCCTTTCCTGATGAACTCG
WRC0819	CATGTTGAATAAATGTACCTGAGTTGC	CATGTTGAATAAATGTACCTGAGTTGT	CRCCAATGTCAAGGGTCAAACAAAAGT
WRC0822	CGATTTTGAGTAGTTCGCCGGG	AATCGATTTTGAGTAGTTCGCCGGA	CTCACCCTAAACCCTAACATAGATAATGA
WRC0825	TTCGCAAAACAATCTTTCWGGTCAA	TCGCAAAACAATCTTTCWGGTCAC	AGTTTATTATAGTCTAGACGAAGGGTCATT
WRC0828	GGGTGGGAGAGGCGGGAG	GGGTGGGAGAGGCGGGAA	GCYGCCGGCATCCTCCGTT
WRC0847	GTGGCCTACTGGGCGCGC	GTGGCCTACTGGGCGCGG	ACCACCTTGAAGCTCATCCCGAAA
WRC0850	ACCTACATGCACCTCTGCGTC	ACCTACATGCACCTCTGCGTG	AGAAGAAYGACTTCCACCACCACTT
WRC0852	CAGACTGGATTCTTGGCCTGC	GCAGACTGGATTCTTGGCCTGT	TGAGGGTCTGAGRAAAGGCTCCAA
WRC1176	GAGAGTCGCCTCCTGTGCAGT	GAGTCGCCTCCTGTGCAGG	GGAGAGGAGAGGGGCCCGA
WRC0872	GGGCTTCTATCAGTTCTTTCCATGT	GGCTTCTATCAGTTCTTTCCATGC	CATAAAATAATCCAGCACATGYTGACAGTA
WRC0879	ATCCCTTTTGGCGCTCTACGTG	CATCCCTTTTGGCGCTCTACGTA	GAGGAAGTGCCTGTTGCTCAGAGA
WRC0885	GATGGCAAACACGAGACCAGTG	GATGGCAAACACGAGACCAGTA	CGGAGTCACCACCAAAGTTCACAAT
WRC0886	GATAAAATGTCCTGCCAAGCCGTTT	ATAAAATGTCCTGCCAAGCCGTTC	CGACTGGACTGATTTTGTGTATCCG
WRC0895	GAGTTGCGGCTGCCAGTCTAG	GAGTTGCGGCTGCCAGTCTAC	CACAACAATGGGTACAGTGTAGCAATAAAT
WRC0908	GAATTCTTGTAAATAGGTGCCAGC	CTGAATTCTTGTAAATAGGTGCCAGG	AACATTACATCATCTGGGAACATGTCAAC
WRC0910	GCCTGGTAASCAGACTACACTA	GCCTGGTAASCAGACTACACTG	CACAACCTGTAAGCTGAAAGGATTCC
WRC0913	CCTTCCCCCTCCTCCGCG	CCTTCCCCCTCCTCCGCC	GACRAAGAGGCAACTGGCACTGAT
WRC0923	TTGCCGATGTCGGCACCG	AACTTTGCCGATGTCGGCACCA	CACGTAGTAGGGTTTGTACACCACAT

WRC0924	CAGTGAACCTGGGCAAAGGAG	TCAGTGAACCTGGGCAAAGGAT	CCTGTGCACAATCTATGCCAATAATTGTA
WRC0926	GTTGATTCCGAATTTTCAATCAATCC	CCTGTTGATTCCGAATTTTCAATCAATCT	CGGATCRAGGAACCGCGTGCA
WRC0930	CAGTCTCCTGACATAACCAAGTGTA	CAGTCTCCTGACATAACCAAGTGTT	TGATTTACTGCTCTCTTCGTCTACGTATA
WRC0931	GCCGTGAAGGATAAATCAGCC	CTGCCGTGAAGGATAAATCAGCT	TTTGTTGGTATTGAGTGCCTCCGG
WRC0939	CACACAGCCATCATATCAGCG	CTCACACAGCCATCATATCAGCA	TCGTCGCTGCCAAGTAGATATTTTA
WRC0941	GACGTTGTCCTTTTATTTATGTCATC	CTGACGTTGTCCTTTTATTTATGTCATG	CAACAGCCTCCACAGATCATTAATTAGTT
WRC1291	AGTCGCTCACTCGCCTCGC	AGTCGCTCACTCGCCTCGG	CCTTTCTTCGCTTGATCGACACCAA
WRC0951	GGAACTGTGAAACCTTCATCT	CTGGAACTGTGAAACCTTCATCC	CGAGTCCAGAAGTCACTSTATGTTGA
WRC0962	GAGGGACATGTTGATGGTGGAG	GAGGGACATGTTGATGGTGGAC	CCTTCCACAACGGCAGATTCTACA
WRC0980	CACAAGGAAAGAAACTGTTCGTCC	CACAAGGAAAGAAACTGTTCGTCG	GTGCATTTGYCTTTTGACGATAAACGTTTT
WRC0984	CGGTTATAGGAACCTCTGATCTTG	CGGTTATAGGAACCTCTGATCTTC	CAGAGTAAAGATGGCAAGAAAGAGAACAT
WRC0998	TGGGACTTTGGTCTGCTC	GGCTTGGGACTTTGGTCTGCTA	CTCCTTCTTCCTCCTGGGCTTC

Table 2: BLASTN and BLASTX results for all chromosome-specific SNPs.

Type	Chr	**contig_regions**
WRC0002	1A	chr1A:7294417-7294567
WRC0006	1A	chr1A:53900436-53900583
WRC0022	1A	chr1A:292077729-292077923
WRC0023	1A	chr1A:310254594-310254694|chr1B:341135658-341135758|chr1D:235673389-235673489
WRC0037	1A	chr1A:490411952-490412098|chr1D:394672602-394672748|chr1B:526748173-526748319
WRC0052	1B	chr1B:66937742-66937880
WRC0057	1B	chr1B:107509865-107510061
WRC0069	1B	chr1B:259026852-259027004
WRC0074	1B	chr1B:332750256-332750363
WRC0076	1B	chr1B:333628901-333629001
WRC0077	1B	chr1B:359198198-359198298|chr1D:253217431-253217531|chr1A:325049874-325049972
WRC0083	1B	chr1B:466938466-466938634|chr1D:346664793-346664889|chr1A:447007411-447007517
WRC0089	1B	chr1B:553822259-553822451
WRC0091	1B	chr1B:568568678-568568778|chr1A:516786447-516786547

WRC1323	1B	chr1B: 412895813
WRC0102	1D	chr1D:6399438-6399538
WRC0103	1D	chr1D:7350427-7350583
WRC0105	1D	chr1D:10382915-10383015
WRC0114	1D	chr1D:112770008-112770200
WRC0122	1D	chr1D:225709006-225709106|chr1B:325816731-325816831
WRC0125	1D	chr1D:261741435-261741624|chr1A:338583450-338583639
WRC0128	1D	chr1D:279233900-279234061
WRC0131	1D	1D:338942296-338942481|1A:438366873-438367058|1B:453514111-453514296
WRC0135	1D	chr1D:364384995-364385151|chr1A:463400946-463401102|chr1B:487680131-487680287
WRC0143	1D	chr1D:477907563-477907755
WRC0150	2A	chr2A:16523759-16523912
WRC0151	2A	chr2A:16525696-16525836
WRC0152	2A	chr2A:79205226-79205326|chr2D:78764406-78764495
WRC0155	2A	chr2A:126996126-126996226|chr2B:175459359-175459459
WRC0159	2A	chr2A:127000514-127000659
WRC0179	2A	chr2A:528109093-528109282
WRC0182	2A	chr2A:573368458-573368558|chr2D:416860333-416860433|chr2B:488867515-488867615

WRC0184	2A	chr2A:569584919-569585019|chr2D:420525583-420525683|chr2B:492958059-492958159
WRC0199	2B	chr2B:7436573-7436766|chr2A:2647315-2647438
WRC0208	2B	chr2B:188719278-188719378|chr2A:139963823-139963923|chr2D:133039485-133039582
WRC0217	2B	chr2B:339108554-339108654
WRC0218	2B	chr2B:346669582-346669682
WRC0219	2B	chr2B:359866839-359866939|chr2A:389219252-389219352|chr2D:308024854-308024954
WRC0225	2B	chr2B:445686268-445686441
WRC0226	2B	chr2B:454019108-454019208|chr2D:383560478-383560536|chr2A:518336853-518336911
WRC0230	2B	chr2B:539290274-539290431
WRC0236	2B	chr2B:625544261-625544406|chr2D:527906691-527906836|chr2A:672773183-672773324
WRC0239	2B	chr2B:636112065-636112254
WRC0244	2B	chr2B:743329095-743329240|chr2D:609243841-609243986
WRC0248	2B	chr2B:777830336-777830523
WRC0254	2D	chr2D:30748994-30749190|chr2A:33044750-33044888
WRC0255	2D	chr2D:36659900-36660073
WRC0258	2D	chr2D:115590216-115590316|chr2B:166047218-166047316|chr2A:117160829-117160927
WRC0260	2D	chr2D:132699246-132699346
WRC0267	2D	chr2D:257388324-257388424|chr2B:285054674-285054774|chr2A:293152348-293152448

WRC0275	2D	chr2D:382180561-382180736|chr2A:516432570-516432648
WRC0288	2D	chr2D:586839118-586839291|chr6D:260037979-260038115|chr6B:308515155-308515291
WRC0292	2D	chr2D:627167519-627167677
WRC0293	3A	chr3A:13029108-13029281
WRC0300	3A	chr3A:83722204-83722366
WRC0311	3A	chr3A:209406657-209406757|chr3B:250468146-250468243|chr3D:172024362-172024456
WRC0312	3A	chr3A:250041842-250041976|chr6A:273933723-273933857
WRC0316	3A	chr3A:365966470-365966570|chr3D:274728017-274728117|chr3B:367480531-367480631
WRC1339	3A	chr3A:441066
WRC0340	3B	chr3B:14162870-14163056
WRC0345	3B	chr3B:26650781-26650925
WRC0350	3B	chr3B:78398692-78398868|chr3D:48867249-48867425|chr3A:61349026-61349202
WRC0355	3B	chr3B:154224453-154224578|chr3D:103649207-103649297|chr3A:552611328-552611415
WRC0360	3B	chr3B:250803269-250803458|chr3D:172355227-172355416
WRC0363	3B	chr3B:353995489-353995589|chr3A:323437355-323437455|chr3D:236627104-236627204
WRC0367	3B	chr3B:406460917-406461069|chr3D:313312255-313312407
WRC0370	3B	chr3B:443805542-443805693|chr3D:344340151-344340302|chr3A:458680704-458680855
WRC0371	3B	chr3B:474317238-474317433|chr6D:459192843-459192965

WRC0377	3B	chr3B:602359193-602359381|chr3D:455874616-455874804|chr3A:598558667-598558855
WRC0379	3B	chr3B:622283236-622283362
WRC0384	3B	chr3B:738315469-738315664
WRC0389	3B	chr3B:808673669-808673776
WRC0391	3D	chr3D:14150114-14150271
WRC0392	3D	chr3D:14415316-14415462|chr3B:21345927-21346065
WRC0396	3D	chr3D:27883980-27884096
WRC0399	3D	chr3D:84676747-84676938
WRC0401	3D	chr3D:95248978-95249078
WRC0403	3D	chr3D:130503566-130503666|chr3A:142578013-142578113
WRC0405	3D	chr3D:163893549-163893744
WRC0412	3D	chr3D:325977158-325977350|chr3B:423364620-423364812|chr3A:439290385-439290577
WRC0417	3D	chr3D:365533680-365533780|chr3B:476041768-476041868|chr3A:487448743-487448843
WRC0418	3D	chr3D:384004424-384004524|chr3A:506670367-506670467
WRC0436	4A	chr4A:5465184-5465284|chr4D:464988190-464988290|chr4B:581308988-581309085
WRC0439	4A	chr4A:10636838-10636938|chr4B:572554840-572554940|chr4D:458385423-458385521
WRC0461	4A	chr4A:330085730-330085887
WRC0464	4A	chr4A:464695378-464695478|chr4D:112725498-112725568|chr4B:173175513-173175583

WRC0466	4A	chr4A:499953393-499953548
WRC0474	4A	chr4A:659125202-659125372
WRC0476	4A	chr4A:674336579-674336750
WRC1360	4A	chr4A:674336592
WRC0492	4B	chr4B:197305685-197305783
WRC0494	4B	chr4B:207112381-207112522
WRC1364	4B	chr4B:398559983
WRC0518	4D	chr4D:11665481-11665617
WRC0521	4D	chr4D:21151009-21151109|chr5B:86130168-86130265|chr4B:299459567-299459664
WRC0523	4D	chr4D:77126010-77126170
WRC0531	4D	chr4D:163743719-163743866
WRC0536	4D	chr4D:275759491-275759591|chr4B:273244598-273244698|chr4A:210165711-210165811
WRC0538	4D	chr4D:300783411-300783511|chr4B:373981820-373981920|chr4A:182219159-182219259
WRC0542	4D	chr4D:339077565-339077665|chr4B:418520176-418520276|chr4A:132992610-132992698
WRC0544	4D	chr4D:371793368-371793468|chr4A:92217351-92217451|chr4B:456597943-456598043
WRC0549	4D	chr4D:430838119-430838285
WRC0551	4D	chr4D:449729668-449729862|chr4A:17263499-17263622
WRC1302	4D	chr4D:455342727

WRC1368	4D	chr4D:297206435
WRC1370	4D	chr4D:480873934-480874034|5A:662559918-662560017|4B:608944986-608945085
WRC0569	5A	chr5A:113607411-113607592
WRC0571	5A	chr5A:113607669-113607855
WRC0575	5A	chr5A:194739972-194740159
WRC0578	5A	5A:267895085-267895233|4B:647658944-647659042
WRC0596	5A	chr5A:535137744-535137876|chr5D:421533220-421533352|chr5B:507873558-507873690
WRC0597	5A	chr5A:549800856-549800995
WRC0599	5A	chr5A:595537998-595538179
WRC0601	5A	chr5A:653690878-653691024
WRC0607	5A	chr5A:702000301-702000400|chr4B:661612804-661612903
WRC0608	5A	chr5A:702000387-702000511
WRC1137	5A	chr5A:671551538
WRC0609	5B	chr5B:10431739-10431956
WRC0614	5B	chr5B:41816696-41816902
WRC0617	5B	chr5B:87975481-87975699
WRC0620	5B	chr5B:129250641-129250741|chr5D:117904322-117904422
WRC0623	5B	chr5B:158131458-158131558|chr5D:179768893-179768993

WRC0624	5B	chr5B:209650604-209650704|chr5D:198952444-198952544|chr5A:227707889-227707989
WRC0633	5B	chr5B:356717027-356717174|chr5D:310694991-310695138|chr5A:404228641-404228788
WRC0646	5B	chr5B:496146204-496146347
WRC0655	5B	chr5B:638605615-638605749
WRC1229	5B	chr5B:378851432
WRC0665	5D	chr5D:3609717-3609871
WRC0668	5D	chr5D:11310358-11310519
WRC0672	5D	chr5D:43100367-43100525|chr5A:32992606-32992764
WRC0681	5D	chr5D:170317349-170317449|chr5B:162494106-162494206|chr5A:176446127-176446227
WRC0684	5D	chr5D:193131346-193131546|chr5B:198419753-198419953|chrUn:68206149-68206349
WRC0685	5D	chr5D:228553359-228553514
WRC0688	5D	chr5D:251425995-251426182
WRC0690	5D	chr5D:273533070-273533210
WRC0698	5D	chr5D:361650097-361650249|5B:427248166-427248318|5A:462155095-462155247
WRC0701	5D	chr5D:396451337-396451503
WRC0706	5D	chr5D:496548357-496548480|chr5A:620400814-620400937|chr5B:614983587-614983707
WRC0712	6A	chr6A:1880645-1880773
WRC0725	6A	chr6A:121975475-121975575|chr6D:101495310-101495410|chr6B:186595530-186595630

WRC0731	6A	chr6A:235073433-235073608
WRC0732	6A	chr6A:276310378-276310478|chr6D:206549998-206550098
WRC0733	6A	chr6D:239043875-239043975|chr6A:319201803-319201903|chr6B:376293061-376293161
WRC0737	6A	chr6A:425281977-425282067|6D:295602096-295602186|6B:459805514-459805604
WRC0747	6A	chr6A:534961971-534962071
WRC0751	6A	chr6A:587556586-587556642
WRC1161	6A	chr6A:554104391
WRC0760	6B	chr6B:36228888-36229057
WRC0763	6B	chr6B:61793880-61793995
WRC0764	6B	chr6B:90281225-90281355|chr6A:51409692-51409822|chr6D:38888156-38888216
WRC0768	6B	chr6B:140849826-140849926
WRC0769	6B	chr6B:188683440-188683604|chr6A:124302475-124302573
WRC0771	6B	chr6B:210595944-210596132
WRC0774	6B	chr6B:229724853-229724953
WRC0777	6B	chr6B:261530999-261531099|chr6D:148768504-148768604|chr6A:198364352-198364452
WRC0778	6B	chr6B:292803423-292803523|chr6D:181700662-181700733
WRC0793	6B	chr6B:526482404-526482590
WRC0801	6B	chr6B:640900204-640900326

WRC0802	6B	chr6B:641396480-641396579|chr6D:426101493-426101592|chr6A:571021673-571021772
WRC0803	6B	chr6B:671159692-671159792|chr6D:445497124-445497224|chr6A:592034076-592034176
WRC0819	6D	chr6D:112429539-112429718|chr6A:137375581-137375702
WRC0822	6D	chr6D:142146392-142146586|chr3B:749494962-749495067
WRC0825	6D	chr6D:157442117-157442290|chr6B:280367050-280367223
WRC0828	6D	chr6D:187500436-187500544|chr6B:389242818-389242926|chr6A:259771513-259771621
WRC0847	6D	chr6D:436223903-436224086|6B:659235108-659235291|6A:583271039-583271222
WRC0850	6D	chr6D:444136522-444136715|chr6B:668776813-668777006|chr6A:591732120-591732313
WRC0852	6D	chr6D:451908916-451909109|6B:686417036-686417229|6A:597832388-597832581
WRC1176	6D	chr6D:455736155-455736225|4D:378265925-378265996
WRC0872	7A	chr7A:172979136-172979278
WRC0879	7A	chr7A:247063883-247063983|chr7D:234127845-234127930
WRC0885	7A	chr7D:348130082-348130182|chr7B:289162321-289162421|chr7A:340975488-340975588
WRC0886	7A	chr7A:427819274-427819374
WRC0895	7A	chr7A:635499300-635499489
WRC0908	7B	chr7B:104771883-104771983|chr7D:139847653-139847753|chr7A:139545572-139545672
WRC0910	7B	chr7B:115404311-115404454
WRC0913	7B	chr7B:152782747-152782920

WRC0923	7B	chr7B:387645929-387646029
WRC0924	7B	chr7B:421699376-421699546
WRC0926	7B	chr7B:465118574-465118718
WRC0930	7B	chr7B:479781875-479782005
WRC0931	7B	chr7B:481070104-481070204|chr7A:540951962-540952062|chr7D:459080322-459080422
WRC0939	7B	chr7B:629782400-629782500|chr7A:661581898-661581998
WRC0941	7B	chr7B:693169017-693169167
WRC1291	7B	chr7B:730115080
WRC0951	7D	chr7D:5445486-5445662
WRC0962	7D	chr7D:148164970-148165070|chr7A:149317393-149317493
WRC0980	7D	chr7D:388636824-388636968
WRC0984	7D	chr7D:459079654-459079754|chr7B:481069438-481069538|chr7A:540952627-540952727
WRC0998	7D	chr7D:625207448-625207548|chr7A:720912915-720913015

Table 3: PCR layout for the first genotyping analysis in September 2018, and 84 KASP markers were used for this set.

	1	2	3	4	5	6	7	8	9	10	11	12
A	BC3F1 265-2	BC4 269-4	BC4F2 26-3	BC5F1 8-3	BC5F1 11-2	BC5F1 13-4	BC5F2 1-1	55/3-1	59/5	62/2/4-1	BC4F1-108D	BC4-118E
B	BC3F1 221-1	BC3F1 224-4	BC4F1 110-6	BC4F1 96-4	BC4F1 112-2	BC4F1 103-4	BC4F1 93-6	BC4F1 101-2	BC4-245B	BC5-76A	BC3-677A	Paragon 3
C	BC3F1 265-3	BC4 269-5	BC4F2 26-4	BC5F1 9-1	BC5F1 12-1	BC5F1 14-1	BC5F2 1-2	55/3-2	59/8-1	62/2/4-2	BC5-73C	BC4-179C
D	BC3F1 221-2	BC4 266-1	BC5 98-1	BC4F1 88-1	BC4F1 112-3	BC4F1 103-5	BC4F1 95-1	BC4F1 101-3	BC4-250B	BC5-93A	BC3-677B	Th. elong 007 1
E	BC3F1 233-1	BC4F1 121-1	BC4F2 27-1	BC5F1 9-2	BC5F1 12-2	BC5F1 14-2	BC5F2 1-3	55/3-3	59/8-2	64/1	BC5-74C	BC4-179D
F	BC3F1 221-3	BC3F1 251-1	BC5 98-2	BC4F1 88-2	BC4F1 112-4	BC4F1 103-6	BC4F1 95-2	BC4F1 101-4	BC4-250C	BC5-93C	BC3-677C	Th. elong 007 2
G	BC3F1 233-2	BC4F1 121-2	BC4F2 27-1	BC5F1 10-1	BC5F1 12-3	BC5F1 14-3	BC5F2 1-4	55/3-4	59/8-3	67/2	BC5-79B	BC4-181B
H	BC3F1 221-4	BC4F1 110-1	BC5 98-3	BC4F1 88-3	BC4F1 112-5	BC4F1 93-1	BC4F1 95-3	BC4F1 101-5	BC4F1-112A	BC3-388E	BC3D1-230D	Th. elong 007 3
I	BC3F1 233-3	BC4F1 121-3	BC4F2 27-3	BC5F1 10-2	BC5F1 12-4	BC5F1 15-1	48/1/6	55/3-5	59/8-4	48/1/14-1	BC5-83A	BC4-183A
J	BC3F1 234-1	BC4F1 110-2	BC5 98-4	BC4F1 88-4	BC4F1 112-6	BC4F1 93-2	BC4F1 95-4	BC4F1-103B	BC4F1-112B	BC3-667B	BC3F1-234B	Th. elong 008 1
K	BC4 269-1	BC4F1 121-4	BC4F2 27-4	BC5F1 10-3	BC5F1 13-1	BC5F1 15-2	48/1/10	55/16-1	59/8-5	BC4-249B	BC5-83B	BC4F1-101A
L	BC3F1 234-2	BC4F1 110-3	BC4F1 96-1	BC4F1 88-5	BC4F1 103-1	BC4F1 93-3	BC4F1 95-5	BC4F1-110A	BC4F1-92D	BC3-667E	BC3F1-234C	Th. elong 008 2

M	BC4 269-2	BC4F2 26-1	BC5F1 8-1	BC5F1 10-4	BC5F1 13-2	BC5F1 15-3	55/1-2	55/20-1	59/11	BC4F1-108B	BC5F1-4A	BC4F1-101B
N	BC3F1 234-3	BC4F1 110-4	BC4F1 96-2	BC4F1 88-6	BC4F1 103-2	BC4F1 93-4	BC4F1 95-6	BC4-239D	BC4F1-95D	BC3-669A	Paragon 1	Th. elong 008 3
O	BC4 269-3	BC4F2 26-2	BC5F1 8-2	BC5F1 11-1	BC5F1 13-3	BC5F1 15-4	55/1-3	57/1/8	62/2-1	BC4F1-108C	BC4-89D	BC4F1-101D
P	BC3F1 224-1	BC4F1 110-5	BC4F1 96-3	BC4F1 112-1	BC4F1 103-3	BC4F1 93-5	BC4F1 101-1	BC4-245A	BC5-75A	BC3-669B	Paragon 2	NTC

Table 4: PCR layout for the second genotyping analysis in February 2019, and 203 KASP markers were used.

	1	2	3	4	5	6	7	8	9	10	11	12
A	BC3F1-221-1	BC3F2-133-3	BC4-266-1	BC4F2-43-4	BC4F2-45-3	BC4F1-93-1	BC4F2-46-8	BC4F2-47-5	BC4F2-48-3	BC4F2-49-2	BC4F1-101-3	elongatum3
B	BC3F2-132-1	BC3F2-133-4	BC4F1-144-3	BC4F2-43-5	BC4F2-45-4	BC4F2-46-1	BC4F2-46-9	BC4F2-47-6	BC4F2-48-4	BC4F2-49-3	BC4F2-51-1	elongatum4
C	BC3F2-132-2	BC3F2-133-5	BC4F1-144-4	BC4F2-43-6	BC4F2-45-5	BC4F2-46-2	BC4F2-46-10	BC4F2-47-7	BC4F2-48-5	BC4F1-101-1	BC4F2-51-2	elongatum5
D	BC3F2-132-3	BC3F2-133-6	BC4F1-144-6	BC4F2-43-7	BC4F2-45-6	BC4F2-46-3	BC4F1-93-2	BC4F2-47-8	BC4F2-48-7	BC4F2-50-1	BC4F2-51-3	paragon1
E	BC3F2-132-5	BC3F2-133-7	BC4F1-88-1	BC4F2-43-8	BC4F2-45-7	BC4F2-46-4	BC4F2-47-1	BC4F2-47-9	BC4F2-48-8	BC4F2-50-2	BC4F2-51-4	paragon2
F	BC3F1-234-2	BC3F2-133-8	BC4F2-43-1	BC4F1-88-6	BC4F2-45-8	BC4F2-46-5	BC4F2-47-2	BC4F1-93-5	BC4F2-48-10	BC4F2-50-3	BC4F2-51-5	paragon3
G	BC3F2-133-1	BC3F2-133-9	BC4F2-43-2	BC4F2-45-1	BC4F2-45-9	BC4F2-46-6	BC4F2-47-3	BC4F2-48-1	BC4F1-95-3	BC4F2-50-4	elongatum1	paragon4
H	BC3F2-133-2	BC3F2-133-10	BC4F2-43-3	BC4F2-45-2	BC4F2-45-10	BC4F2-46-7	BC4F2-47-4	BC4F2-48-2	BC4F2-49-1	BC4F2-50-5	elongatum2	X

Table 5: PCR layout for the Set 1 genotyping analysis in May 2019, and 235 KASP markers were used for Set 1.

	1	2	3	4	5	6	7	8	9	10	11	12
A	BC4F2 61-1	BC4F2 61-9	BC5F1 17-2	BC5F1 17-10	BC5F2 3-2	BC5F2 3-10	BC5F2 5-7	BC5F2 6-6	BC5F2 7-5	BC5F2 8-4	BC5F2 9-2	BC5F3 1-4
B	BC4F2 61-2	BC5F1 9-5	BC5F1 17-3	BC5F2 2-1	BC5F2 3-3	BC5F2 4-1	BC5F2 5-8	BC5F2 6-7	BC5F2 7-6	BC5F2 8-5	BC5F2 9-3	paragon1
C	BC4F2 61-3	BC5F1 9-6	BC5F1 17-4	BC5F2 2-2	BC5F2 3-4	BC5F2 5-1	BC5F2 5-9	BC5F2 6-8	BC5F2 7-7	BC5F2 8-6	BC5F2 9-4	paragon2
D	BC4F2 61-4	BC5F1 9-7	BC5F1 17-5	BC5F2 2-3	BC5F2 3-5	BC5F2 5-2	BC5F2 6-1	BC5F2 6-9	BC5F2 7-8	BC5F2 8-7	BC5F2 9-5	paragon3
E	BC4F2 61-5	BC5F1 9-8	BC5F1 17-6	BC5F2 2-4	BC5F2 3-6	BC5F2 5-3	BC5F2 6-2	BC5F2 7-1	BC5F2 7-9	BC5F2 8-8	BC5F2 9-6	elon1
F	BC4F2 61-6	BC5F1 9-9	BC5F1 17-7	BC5F2 2-5	BC5F2 3-7	BC5F2 5-4	BC5F2 6-3	BC5F2 7-2	BC5F2 8-1	BC5F2 8-9	BC5F3 1-1	elon2
G	BC4F2 61-7	BC5F1 9-10	BC5F1 17-8	BC5F2 2-6	BC5F2 3-8	BC5F2 5-5	BC5F2 6-4	BC5F2 7-3	BC5F2 8-2	BC5F2 8-10	BC5F3 1-2	elon3
H	BC4F2 61-8	BC5F1 17-1	BC5F1 17-9	BC5F2 3-1	BC5F2 3-9	BC5F2 5-6	BC5F2 6-5	BC5F2 7-4	BC5F2 8-3	BC5F2 9-1	BC5F3 1-3	X

Table 6: PCR layout for the Set 2 genotyping analysis in June 2019, and 235 KASP markers were used for Set 2.

	1	2	3	4	5	6	7	8	9	10	11	12
A	BC4F2-63-1	BC4F2-64-5	BC4F2-65-4	BC4F2-66-7	BC5F1-16-4	BC5F1-18-7	BC5F1-20-4	BC5F1-21-3	BC5F2-10-7	BC6-7-5	BC5F3-1-9	BC4F2 73-8
B	BC4F2-63-2	BC4F2-64-6	BC4F2-65-5	BC4F2-66-8	BC5F1-16-5	BC5F1-18-8	BC5F1-20-5	BC5F1-21-4	BC5F2-10-8	BC6-7-6	BC4F2 73-1	paragon1
C	BC4F2-63-3	BC4F2-64-7	BC4F2-66-1	BC4F2-66-9	BC5F1-18-1	BC5F1-18-9	BC5F1-20-6	BC5F2-10-1	BC5F2-10-9	BC6-7-7	BC4F2 73-2	paragon2
D	BC4F2-63-4	BC4F2-64-8	BC4F2-66-2	BC4F2-66-10	BC5F1-18-2	BC5F1-18-10	BC5F1-20-7	BC5F2-10-2	BC5F2-10-10	BC6-7-8	BC4F2 73-3	paragon3
E	BC4F2-64-1	BC4F2-64-9	BC4F2-66-3	BC4F2-67-1	BC5F1-18-3	BC5F1-19-1	BC5F1-20-8	BC5F2-10-3	BC6-7-1	BC5F3-1-5	BC4F2 73-4	elong1
F	BC4F2-64-2	BC4F2-65-1	BC4F2-66-4	BC5F1-16-1	BC5F1-18-4	BC5F1-20-1	BC5F1-20-9	BC5F2-10-4	BC6-7-2	BC5F3-1-6	BC4F2 73-5	elong2
G	BC4F2-64-3	BC4F2-65-2	BC4F2-66-5	BC5F1-16-2	BC5F1-18-5	BC5F1-20-2	BC5F1-21-1	BC5F2-10-5	BC6-7-3	BC5F3-1-7	BC4F2 73-6	elong3
H	BC4F2-64-4	BC4F2-65-3	BC4F2-66-6	BC5F1-16-3	BC5F1-18-6	BC5F1-20-3	BC5F1-21-2	BC5F2-10-6	BC6-7-4	BC5F3-1-8	BC4F2 73-7	X

Table 7: PCR layout for the Set 3 genotyping analysis in July 2019, and 235 KASP markers were used for Set 3.

	1	2	3	4	5	6	7	8	9	10	11	12
A	BC3F1-234-5	BC4F1-152-6	BC4F2-70-3	BC4F2-72-4	BC5F1-26-2	BC5F1-27-4	BC5F1-28-8	BC5F1-30-2	BC5F1-31-4	BC5F1-33-2	BC5F1-34-6	BC4F2-74-6
B	BC3F1-234-6	BC4F1-152-7	BC4F2-70-4	BC5F1-25-1	BC5F1-26-3	BC5F1-28-1	BC5F1-29-1	BC5F1-30-3	BC5F1-31-5	BC5F1-33-3	BC4F2-73-9	paragon1
C	BC3F1-234-7	BC4F2-69-1	BC4F2-71-1	BC5F1-25-2	BC5F1-26-4	BC5F1-28-2	BC5F1-29-2	BC5F1-30-4	BC5F1-31-6	BC5F1-33-4	BC4F2-73-10	paragon2
D	BC4F1-152-1	BC4F2-69-2	BC4F2-71-3	BC5F1-25-3	BC5F1-26-5	BC5F1-28-3	BC5F1-29-3	BC5F1-30-5	BC5F1-32-1	BC5F1-33-6	BC4F2-74-1	paragon3
E	BC4F1-152-2	BC4F2-69-3	BC4F2-71-5	BC5F1-25-4	BC5F1-26-6	BC5F1-28-4	BC5F1-29-4	BC5F1-30-6	BC5F1-32-2	BC5F1-34-1	BC4F2-74-2	elong1
F	BC4F1-152-3	BC4F2-69-4	BC4F2-71-6	BC5F1-25-5	BC5F1-27-1	BC5F1-28-5	BC5F1-29-5	BC5F1-31-1	BC5F1-32-3	BC5F1-34-3	BC4F2-74-3	elong2
G	BC4F1-152-4	BC4F2-70-1	BC4F2-72-2	BC5F1-25-6	BC5F1-27-2	BC5F1-28-6	BC5F1-29-6	BC5F1-31-2	BC5F1-32-6	BC5F1-34-4	BC4F2-74-4	elong3
H	BC4F1-152-5	BC4F2-70-2	BC4F2-72-3	BC5F1-26-1	BC5F1-27-3	BC5F1-28-7	BC5F1-30-1	BC5F1-31-3	BC5F1-33-1	BC5F1-34-5	BC4F2-74-5	NTC

Table 8: PCR layout for the Set 4 genotyping analysis in October 2019, and 198 KASP markers were used for Set 4.

	1	2	3	4	5	6	7	8	9	10	11	12
A	BC3F1-309-1	BC3F1-309-9	BC2F3-92-2	BC4F2-85-1	BC5F1-35-3	BC4F2-84-5	BC4F3-26-5	BC2F3-91-1	BC4F2-75-1	BC4F2-75-9	BC4F2-78-3	paragon 7
B	BC3F1-309-2	BC3F1-309-10	BC3F1-400-1	BC4F2-85-2	BC5F1-35-4	BC4F2-84-6	BC4F3-26-6	BC2F3-91-2	BC4F2-75-2	BC4F2-75-10	BC4F2-78-4	elong1
C	BC3F1-309-3	BC2F2-606-1	BC3F3-25-1	BC4F2-85-3	BC5F1-35-5	BC4F2-84-7	BC3F2-155-1	BC2F3-91-3	BC4F2-75-3	BC4F2-76-1	paragon1	elong2
D	BC3F1-309-4	BC2F2-606-2	BC3F3-25-2	BC4F2-85-4	BC5F1-35-6	BC4F2-84-8	BC3F2-155-2	BC2F3-91-4	BC4F2-75-4	BC4F2-76-2	paragon2	elong3
E	BC3F1-309-5	BC2F2-606-3	BC3F3-25-3	BC4F2-85-5	BC4F2-84-1	BC4F3-26-1	BC3F2-155-3	BC4F2-74-7	BC4F2-75-5	BC4F2-76-3	paragon3	elong4
F	BC3F1-309-6	BC2F2-606-4	BC3F3-25-4	BC4F2-85-6	BC4F2-84-2	BC4F3-26-2	BC2F3-93-1	BC4F2-74-8	BC4F2-75-6	BC4F2-76-4	paragon4	elong5
G	BC3F1-309-7	BC2F2-606-5	BC3F3-25-5	BC5F1-35-1	BC4F2-84-3	BC4F3-26-3	BC2F3-93-2	BC4F2-74-9	BC4F2-75-7	BC4F2-78-1	paragon5	elong6
H	BC3F1-309-8	BC2F3-92-1	BC3F3-25-6	BC5F1-35-2	BC4F2-84-4	BC4F3-26-4	BC2F3-93-3	BC4F2-74-10	BC4F2-75-8	BC4F2-78-2	paragon6	NTC

Table 9: PCR layout for the Set 5 genotyping analysis in November 2019, and 198 KASP markers were used for Set 5.

	1	2	3	4	5	6	7	8	9	10	11	12
A	BC3F2-156-1	BC3F2-157-3	BC3F2-158-8	BC4F1-154-5	BC4F2-55-11	BC4F2-86-3	BC4F2-88-1	BC4F2-89-4	BC4F2-91-2	BC4F2-92-4	BC4F2-94-4	BC4F3-28-5
B	BC3F2-156-2	BC3F2-158-1	BC4F1-153-1	BC4F1-154-6	BC4F2-55-12	BC4F2-86-4	BC4F2-88-2	BC4F2-89-5	BC4F2-91-3	BC4F2-93-1	BC4F2-94-5	pargn1
C	BC3F2-156-3	BC3F2-158-2	BC4F1-153-2	BC4F1-155-1	BC4F2-55-13	BC4F2-86-5	BC4F2-88-3	BC4F2-89-6	BC4F2-91-4	BC4F2-93-2	BC4F2-95-1	pargn2
D	BC3F2-156-4	BC3F2-158-3	BC4F1-153-3	BC4F1-155-2	BC4F2-55-14	BC4F2-86-6	BC4F2-88-4	BC4F2-90-1	BC4F2-91-5	BC4F2-93-3	BC4F2-95-2	pargn3
E	BC3F2-156-5	BC3F2-158-4	BC4F1-154-1	BC4F1-155-3	BC4F2-55-15	BC4F2-86-7	BC4F2-88-5	BC4F2-90-2	BC4F2-91-6	BC4F2-93-4	BC4F3-28-1	elong1
F	BC3F2-156-6	BC3F2-158-5	BC4F1-154-2	BC4F1-155-4	BC4F2-55-16	BC4F2-86-8	BC4F2-89-1	BC4F2-90-3	BC4F2-92-1	BC4F2-93-5	BC4F3-28-2	elong2
G	BC3F2-157-1	BC3F2-158-6	BC4F1-154-3	BC4F1-155-5	BC4F2-86-1	BC4F2-87-1	BC4F2-89-2	BC4F2-90-4	BC4F2-92-2	BC4F2-94-1	BC4F3-28-3	elong3
H	BC3F2-157-2	BC3F2-158-7	BC4F1-154-4	BC4F1-155-6	BC4F2-86-2	BC4F2-87-2	BC4F2-89-3	BC4F2-91-1	BC4F2-92-3	BC4F2-94-2	BC4F3-28-4	X

Table 10: PCR layout for the Set 6 (homozygous lines) genotyping analysis in November 2020, and 198 KASP markers were used for Set 6.

	1	2	3	4	5	6	7	8	9	10	11	12
A	BC2F3-111-1	BC2F3-111-2	BC2F3-111-3	BC2F3-111-4	BC3F3-56-1	BC3F3-56-2	BC3F3-56-3	BC4F2-121-1	BC4F2-121-2	BC4F2-121-4	BC4F3-43-1	BC4F3-43-2
B	BC4F3-43-3	BC4F3-43-4	BC4F3-44-1	BC4F3-44-3	BC4F3-45-3	BC4F3-46-1	BC4F3-46-2	BC4F3-46-3	BC4F3-46-4	BC4F3-47-1	BC4F3-47-2	BC4F3-47-3
C	BC4F3-48-3	BC4F3-49-1	BC4F3-49-2	BC4F3-49-3	BC4F3-49-4	BC4F3-50-1	BC4F3-50-2	BC4F3-50-3	BC4F3-51-1	BC4F3-51-2	BC4F3-51-3	BC4F3-51-4
D	BC4F4-7-1	BC4F4-7-2	BC4F4-7-3	BC4F4-7-4	BC5F2-17-1	BC5F2-17-2	BC5F2-17-3	BC5F2-18-1	BC5F2-18-2	BC5F2-18-3	BC5F2-18-4	BC5F2-19-1
E	BC5F2-19-2	BC5F2-19-3	BC5F2-19-4	BC5F3-6-2	BC5F3-7-1	BC5F3-7-2	BC5F3-7-4	BC5F4-1-1	BC5F4-1-2	BC5F4-1-3	BC5F4-1-4	BC4F3-47-6
F	BC4F3-43-7	BC4F3-43-8	BC4F3-44-5	BC4F3-44-7	BC4F3-46-5	BC4F3-46-6	BC4F3-46-7	BC4F3-46-8	BC4F3-47-5	BC4F3-47-7	BC4F3-47-8	BC4F3-48-5
G	BC4F3-48-6	BC4F3-49-5	BC4F3-49-6	BC4F3-49-7	BC4F3-49-8	BC4F3-50-5	BC4F3-50-6	BC4F3-50-7	BC4F3-50-8	BC4F3-51-5	BC4F3-51-6	BC4F3-51-7
H	BC4F3-51-8	BC4F4-7-6	BC4F4-7-7	BC5F2-17-5	BC5F2-17-6	paragon1	paragon2	paragon3	elong 1	elong 2	elong 3	x

Table 11: The table shows the list of markers used for the first genotyping analysis in September 2018.

Marker ID	LG 1	Marker ID	LG 2	Marker ID	LG 3	Marker ID	LG 4	Marker ID	LG 5	Marker ID	LG 6	Marker ID	LG 7
WRC0002	1A	WRC0150	2A	WRC0293	3A	WRC0439	4A	WRC0569	5A	WRC0712	6A	WRC0866	7A
WRC0023	1A	WRC0155	2A	WRC0311	3A	WRC0453	4A	WRC0578	5A	WRC0731	6A	WRC0879	7A
WRC0041	1A	WRC0179	2A	WRC0316	3A	WRC0467	4A	WRC0597	5A	WRC0747	6A	WRC0893	7A
WRC0043	1A	WRC0190	2A	WRC0336	3A	WRC0478	4A	WRC0608	5A	WRC0753	6A	WRC0898	7A
WRC0052	1B	WRC0198	2B	WRC0345	3B	WRC0482	4B	WRC0614	5B	WRC0763	6B	WRC0908	7B
WRC0057	1B	WRC0207	2B	WRC0355	3B	WRC0492	4B	WRC0623	5B	WRC0768	6B	WRC0919	7B
WRC0074	1B	WRC0220	2B	WRC0380	3B	WRC0504	4B	WRC0633	5B	WRC0774	6B	WRC0924	7B
WRC0091	1B	WRC0239	2B	WRC0384	3B	WRC0508	4B	WRC0660	5B	WRC0803	6B	WRC0931	7B
WRC0103	1D	WRC0255	2D	WRC0396	3D	WRC0521	4D	WRC0665	5D	WRC0816	6D	WRC0952	7D
WRC0114	1D	WRC0260	2D	WRC0403	3D	WRC0527	4D	WRC0681	5D	WRC0819	6D	WRC0962	7D
WRC0135	1D	WRC0270	2D	WRC0418	3D	WRC0542	4D	WRC0697	5D	WRC0829	6D	WRC0984	7D
WRC0144	1D	WRC0288	2D	WRC0425	3D	WRC0551	4D	WRC0706	5D	WRC0845	6D	WRC0994	7D

Table 12: The table shows the list of markers used for the second genotyping analysis in February 2019.

Marker ID	LG 1	Marker ID	LG 2	Marker ID	LG 3	Marker ID	LG 4	Marker ID	LG 5	Marker ID	LG 6	Marker ID	LG 7
WRC0002	1A	WRC0244	2B	WRC0293	3A	WRC0436	4A	WRC0571	5A	WRC0712	6A	WRC0866	7A
WRC0006	1A	WRC0248	2B	WRC0300	3A	WRC0438	4A	WRC0575	5A	WRC0725	6A	WRC0872	7A
WRC0017	1A	WRC0254	2D	WRC0311	3A	WRC0439	4A	WRC0578	5A	WRC0731	6A	WRC0879	7A
WRC0022	1A	WRC0255	2D	WRC0312	3A	WRC0444	4A	WRC0596	5A	WRC0732	6A	WRC0885	7A
WRC0023	1A	WRC0258	2D	WRC0316	3A	WRC0461	4A	WRC0597	5A	WRC0733	6A	WRC0886	7A
WRC0033	1A	WRC0260	2D	WRC0336	3A	WRC0464	4A	WRC0599	5A	WRC0737	6A	WRC0895	7A
WRC0037	1A	WRC0267	2D	WRC0340	3B	WRC0466	4A	WRC0601	5A	WRC0747	6A	WRC0907	7B
WRC0043	1A	WRC0275	2D	WRC0345	3B	WRC0472	4A	WRC0607	5A	WRC0749	6A	WRC0908	7B
WRC0047	1B	WRC0292	2D	WRC0350	3B	WRC0474	4A	WRC0608	5A	WRC0751	6A	WRC0910	7B
WRC0059	1B	WRC0151	2A	WRC0353	3B	WRC0476	4A	WRC0609	5B	WRC0760	6B	WRC0913	7B
WRC0069	1B	WRC0152	2A	WRC0355	3B	WRC0478	4A	WRC0614	5B	WRC0763	6B	WRC0923	7B
WRC0076	1B	WRC0155	2A	WRC0360	3B	WRC0489	4B	WRC0617	5B	WRC0764	6B	WRC0924	7B
WRC0077	1B	WRC0159	2A	WRC0363	3B	WRC0492	4B	WRC0620	5B	WRC0769	6B	WRC0926	7B
WRC0081	1B	WRC0179	2A	WRC0367	3B	WRC0494	4B	WRC0623	5B	WRC0771	6B	WRC0930	7B
WRC0083	1B	WRC0182	2A	WRC0370	3B	WRC0508	4B	WRC0624	5B	WRC0777	6B	WRC0931	7B
WRC0089	1B	WRC0183	2A	WRC0371	3B	WRC0518	4D	WRC0633	5B	WRC0778	6B	WRC0939	7B

WRC0091	1B	WRC0184	2A	WRC0373	3B	WRC0521	4D	WRC0646	5B	WRC0793	6B	WRC0941	7B
WRC0102	1D	WRC0199	2B	WRC0377	3B	WRC0523	4D	WRC0655	5B	WRC0801	6B	WRC0951	7D
WRC0103	1D	WRC0202	2B	WRC0379	3B	WRC0531	4D	WRC0658	5B	WRC0802	6B	WRC0954	7D
WRC0105	1D	WRC0208	2B	WRC0384	3B	WRC0536	4D	WRC0665	5D	WRC0803	6B	WRC0962	7D
WRC0114	1D	WRC0217	2B	WRC0389	3B	WRC0538	4D	WRC0668	5D	WRC0811	6D	WRC0980	7D
WRC0122	1D	WRC0218	2B	WRC0391	3D	WRC0542	4D	WRC0669	5D	WRC0818	6D	WRC0984	7D
WRC0125	1D	WRC0219	2B	WRC0392	3D	WRC0544	4D	WRC0670	5D	WRC0819	6D	WRC0998	7D
WRC0128	1D	WRC0225	2B	WRC0396	3D	WRC0549	4D	WRC0671	5D	WRC0822	6D		
WRC0130	1D	WRC0226	2B	WRC0399	3D	WRC0551	4D	WRC0672	5D	WRC0825	6D		
WRC0131	1D	WRC0230	2B	WRC0401	3D			WRC0681	5D	WRC0828	6D		
WRC0135	1D	WRC0236	2B	WRC0403	3D			WRC0684	5D	WRC0847	6D		
WRC0141	1D	WRC0239	2B	WRC0405	3D			WRC0685	5D	WRC0850	6D		
WRC0143	1D	WRC0242	2B	WRC0412	3D			WRC0688	5D	WRC0852	6D		
				WRC0417	3D			WRC0690	5D				
				WRC0418	3D			WRC0698	5D				
				WRC0421	3D			WRC0701	5D				
				WRC0426	3D			WRC0704	5D				
				WRC0435	3D			WRC0706	5D				

Table 13: The table shows the list of markers used for the Set 1, Set 2, and Set 3 genotyping analyses.

Marker ID	LG 1	Marker ID	LG 2	Marker ID	LG 3	Marker ID	LG 4	Marker ID	LG 5	Marker ID	LG 6	Marker ID	LG 7
WRC0002	1A	WRC0150	2A	WRC1339	3A	WRC0436	4A	WRC0569	5A	WRC0712	6A	WRC0872	7A
WRC0006	1A	WRC0151	2A	WRC0293	3A	WRC0438	4A	WRC0571	5A	WRC0725	6A	WRC1380	7A
WRC0017	1A	WRC0152	2A	WRC0300	3A	WRC0439	4A	WRC0575	5A	WRC0731	6A	WRC0879	7A
WRC0022	1A	WRC0155	2A	WRC0311	3A	WRC1354	4A	WRC0578	5A	WRC0732	6A	WRC0885	7A
WRC0023	1A	WRC0159	2A	WRC0312	3A	WRC1356	4A	WRC0596	5A	WRC0733	6A	WRC0886	7A
WRC0033	1A	WRC1329	2A	WRC0316	3A	WRC1357	4A	WRC0597	5A	WRC0737	6A	WRC1382	7A
WRC0037	1A	WRC1330	2A	WRC0336	3A	WRC0461	4A	WRC0599	5A	WRC0747	6A	WRC0895	7A
WRC0043	1A	WRC0179	2A	WRC1343	3A	WRC0464	4A	WRC0601	5A	WRC1161	6A	WRC0908	7B
WRC0047	1B	WRC0182	2A	WRC0340	3B	WRC0466	4A	WRC1137	5A	WRC0749	6A	WRC0910	7B
WRC1319	1B	WRC0183	2A	WRC0345	3B	WRC0472	4A	WRC0607	5A	WRC0751	6A	WRC0913	7B
WRC1320	1B	WRC0184	2A	WRC0350	3B	WRC0474	4A	WRC0608	5A	WRC0760	6B	WRC1385	7B
WRC0052	1B	WRC0199	2B	WRC0353	3B	WRC0476	4A	WRC0609	5B	WRC0763	6B	WRC0923	7B
WRC1321	1B	WRC0202	2B	WRC0355	3B	WRC1360	4A	WRC0614	5B	WRC0764	6B	WRC0924	7B
WRC0059	1B	WRC0208	2B	WRC0360	3B	WRC0478	4A	WRC0617	5B	WRC0768	6B	WRC0926	7B
WRC0069	1B	WRC0217	2B	WRC0363	3B	WRC0489	4B	WRC0620	5B	WRC0769	6B	WRC0930	7B

WRC0074	1B	WRC0218	2B	WRC0367	3B	WRC0492	4B	WRC0623	5B	WRC0771	6B	WRC0931	7B
WRC0076	1B	WRC0219	2B	WRC0370	3B	WRC0494	4B	WRC0624	5B	WRC0774	6B	WRC0939	7B
WRC0077	1B	WRC0220	2B	WRC0371	3B	WRC1364	4B	WRC0633	5B	WRC0777	6B	WRC0941	7B
WRC0081	1B	WRC0225	2B	WRC0373	3B	WRC0508	4B	WRC1229	5B	WRC0778	6B	WRC1291	7B
WRC1323	1B	WRC0226	2B	WRC0377	3B	WRC0518	4D	WRC0646	5B	WRC0793	6B	WRC0951	7D
WRC0083	1B	WRC0230	2B	WRC0379	3B	WRC0521	4D	WRC0655	5B	WRC0801	6B	WRC0962	7D
WRC0089	1B	WRC0236	2B	WRC0384	3B	WRC1314	4D	WRC0658	5B	WRC0802	6B	WRC0980	7D
WRC0091	1B	WRC0239	2B	WRC0389	3B	WRC1315	4D	WRC0549	5D	WRC0803	6B	WRC0984	7D
WRC1327	1B	WRC1099	2B	WRC0391	3D	WRC0523	4D	WRC0665	5D	WRC0811	6D	WRC0998	7D
WRC0102	1D	WRC0242	2B	WRC0392	3D	WRC0531	4D	WRC0668	5D	WRC0818	6D		
WRC0103	1D	WRC0244	2B	WRC0396	3D	WRC0536	4D	WRC0669	5D	WRC0819	6D		
WRC0105	1D	WRC0248	2B	WRC0399	3D	WRC1368	4D	WRC0672	5D	WRC0822	6D		
WRC0114	1D	WRC1337	2B	WRC0401	3D	WRC0538	4D	WRC0681	5D	WRC0825	6D		
WRC0122	1D	WRC1231	2D	WRC0403	3D	WRC0542	4D	WRC0684	5D	WRC0828	6D		
WRC0125	1D	WRC0254	2D	WRC0405	3D	WRC0544	4D	WRC0685	5D	WRC0847	6D		
WRC0128	1D	WRC0255	2D	WRC0412	3D	WRC0551	4D	WRC0688	5D	WRC0850	6D		
WRC0130	1D	WRC0258	2D	WRC0417	3D	WRC1302	4D	WRC0690	5D	WRC0852	6D		
WRC0131	1D	WRC0260	2D	WRC0418	3D	WRC1370	4D	WRC0698	5D	WRC1176	6D		

WRC0135	1D	WRC0267	2D	WRC0426	3D	WRC0701	5D
WRC0141	1D	WRC0275	2D	WRC0435	3D	WRC0704	5D
WRC0143	1D	WRC0288	2D			WRC0706	5D
		WRC0292	2D			WRC1150	5D

Table 14: The table shows the list of markers used for the Set 4, Set 5, and Set 6 genotyping analyses.

Marker ID	LG 1	Marker ID	LG 2	Marker ID	LG 3	Marker ID	LG 4	Marker ID	LG 5	Marker ID	LG 6	Marker ID	LG 7
WRC0002	1A	WRC0150	2A	WRC1339	3A	WRC0436	4A	WRC0569	5A	WRC0712	6A	WRC0872	7A
WRC0006	1A	WRC0151	2A	WRC0293	3A	WRC0438	4A	WRC0571	5A	WRC0725	6A	WRC0879	7A
WRC0022	1A	WRC0152	2A	WRC0300	3A	WRC0439	4A	WRC0575	5A	WRC0731	6A	WRC0885	7A
WRC0023	1A	WRC0155	2A	WRC0311	3A	WRC1357	4A	WRC0578	5A	WRC0732	6A	WRC0886	7A
WRC0037	1A	WRC0159	2A	WRC0312	3A	WRC0461	4A	WRC0596	5A	WRC0733	6A	WRC1382	7A
WRC0052	1B	WRC0179	2A	WRC0316	3A	WRC0464	4A	WRC0597	5A	WRC0737	6A	WRC0895	7A
WRC0069	1B	WRC0182	2A	WRC0340	3B	WRC0466	4A	WRC0599	5A	WRC0747	6A	WRC0908	7B
WRC0074	1B	WRC0184	2A	WRC0345	3B	WRC0472	4A	WRC0601	5A	WRC1161	6A	WRC0910	7B
WRC0076	1B	WRC0199	2B	WRC0350	3B	WRC0474	4A	WRC1137	5A	WRC0751	6A	WRC0913	7B
WRC0077	1B	WRC0202	2B	WRC0355	3B	WRC0476	4A	WRC0607	5A	WRC0760	6B	WRC0923	7B
WRC1323	1B	WRC0208	2B	WRC0360	3B	WRC1360	4A	WRC0608	5A	WRC0763	6B	WRC0924	7B
WRC0083	1B	WRC0217	2B	WRC0363	3B	WRC0478	4A	WRC0609	5B	WRC0764	6B	WRC0926	7B
WRC0089	1B	WRC0218	2B	WRC0367	3B	WRC0489	4B	WRC0614	5B	WRC0768	6B	WRC0930	7B
WRC0091	1B	WRC0219	2B	WRC0370	3B	WRC0492	4B	WRC0617	5B	WRC0769	6B	WRC0931	7B
WRC0102	1D	WRC0220	2B	WRC0371	3B	WRC0494	4B	WRC0620	5B	WRC0771	6B	WRC0939	7B

WRC0103	1D	WRC0225	2B	WRC0377	3B	WRC1364	4B	WRC0623	5B	WRC0774	6B	WRC0941	7B
WRC0105	1D	WRC0226	2B	WRC0379	3B	WRC0518	4D	WRC0624	5B	WRC0777	6B	WRC1291	7B
WRC0114	1D	WRC0230	2B	WRC0384	3B	WRC0521	4D	WRC0633	5B	WRC0778	6B	WRC0951	7D
WRC0122	1D	WRC0236	2B	WRC0389	3B	WRC0523	4D	WRC1229	5B	WRC0793	6B	WRC0962	7D
WRC0125	1D	WRC0239	2B	WRC0391	3D	WRC0531	4D	WRC0646	5B	WRC0801	6B	WRC0980	7D
WRC0128	1D	WRC0244	2B	WRC0392	3D	WRC0536	4D	WRC0655	5B	WRC0802	6B	WRC0984	7D
WRC0131	1D	WRC0248	2B	WRC0396	3D	WRC1368	4D	WRC0549	5D	WRC0803	6B	WRC0998	7D
WRC0135	1D	WRC1231	2D	WRC0399	3D	WRC0538	4D	WRC0665	5D	WRC0818	6D		
WRC0143	1D	WRC0254	2D	WRC0401	3D	WRC0542	4D	WRC0668	5D	WRC0819	6D		
		WRC0255	2D	WRC0403	3D	WRC0544	4D	WRC0672	5D	WRC0822	6D		
		WRC0258	2D	WRC0405	3D	WRC0551	4D	WRC0681	5D	WRC0825	6D		
		WRC0260	2D	WRC0412	3D	WRC1302	4D	WRC0684	5D	WRC0828	6D		
		WRC0267	2D	WRC0417	3D	WRC1370	4D	WRC0685	5D	WRC0847	6D		
		WRC0275	2D	WRC0418	3D			WRC0688	5D	WRC0850	6D		
		WRC0288	2D					WRC0690	5D	WRC0852	6D		
		WRC0292	2D					WRC0698	5D	WRC1176	6D		
								WRC0701	5D				
								WRC0706	5D				