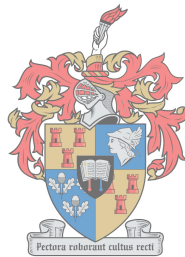# Improving wheat grain yield by employing an integrated biotechnology approach

by

Rayganah Rhoda

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in the Faculty of Genetics at Stellenbosch University.

Supervisor: Willem Botes

March 2018

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third-party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: March 2018

Abstract

Wheat is a widely cultivated crop as it forms a significant part of the world's diet, especially within developing countries. However, food insecurity is increasing at a rapid rate and to meet this demand, wheat yields need to increase by 50%, by the year 2050. To gain higher wheat yields, breeding efficiency needs to increase which can be done through employing biotechnological approaches that can aid in achieving increased yields. Yield, however, is quantitatively inherited and strongly influenced by the genotype x environment interaction. Therefore, yield-determining traits that have less genotype x environment interaction should be investigated to identify underlying inheritance of high yield, along with good husbandry practises that can result in increased wheat yield.

The aim of this study was to assess high-yielding genotypes through validating yield-determining traits using genotypic and phenotypic screening as well as the use of these high-yielding genotypes as male crossing parents within the male-sterility marker-assisted mediated recurrent selection breeding (MS-MARS) scheme for the improvement of grain yield. The yield-determining traits as well as molecular markers associated to some of the yield-determining traits were identified through literature. The molecular markers were validated through genotypic screening and each yield-determining trait was phenotypically screened for each genotype and statistically analysed. The validation of two mobile applications, SeedCounter and 1KK, that measures grain morphology was also executed.

All molecular markers were validated as reliable diagnostic markers to be used in marker-assisted selection (MAS) for identifying its specific yield-determining trait, except for one marker. The statistical analysis for the yield-determining traits displayed that three genotypes were better performing among this set of genotypes and therefore can be used as the male crossing parents within the next MS-MARS cycle. The association of the molecular marker with the yield-determining traits displayed negative correlations that suggests that the function of the high-yielding genes are different within this set of genotypes. Only the SeedCounter application was validated to be used as a future phenotyping tool for grain morphology and the MS-MARS cycles were successfully executed.

Future studies should include the validation of more mobile applications, the identification of the relationship between yield and these molecular markers identified and QTL mapping to contribute to the understanding of the underlying genetic control of the desired phenotypes that contribute to higher grain yield.

Uittreksel

Koring is 'n wyd verboude gewas aangesien dit 'n belangrike deel van die wêreld se dieët vorm, veral in ontwikkelende lande. Voedselonserkheid verhoog egter teen 'n vernuelde tempo en daarom moet die opbrengs van koring met 50%, teen die jaar 2050 verhoog. Om hoër koringopbrengste te behaal moet die effektiwiteit van die teling toeneem. Dit kan beruik word deur die gebruik van biotegnologiese tegnieke. Opbrengs is egter 'n kwantitatief oorgeërfde eienskap en word beduidend beïnvloed deur genotipe x omgewing-interaksie. Daarom moet opbrengsbepalende eienskappe wat minder beïnvloed genotipe x omgewing-interaksie, ondersoek word om eienskappe te identifiseer, tesame met goeie verbouingspraktyke wat tot verhoogde koringopbrengste kan lei.

Die doel van hierdie studie was om hoë opbrengs genotipes te evalueer deur opbrengsbepalende eienskappe te identifiseer deur gebruik te maak van genotipiese en fenotipiese sifting sowel as die gebruik van hierdie hoë opbrengs genotipes as manlike kruisingsouers binne die manlik steriliteits merker bemiddelde herhalende seleksie MS-MBHS-telingsskema vir die verbetering van graanopbrengs. Die opbrengsbepalende eienskappe sowel as molekulêre merkers wat geassosieer word met sommige van die opbrengsbepalende eienskappe is deur middel van literatuur geïdentifiseer. Die molekulêre merkers is deur genotipiese sifting bevestig en elke opbrengsbepalende eienskap is fenotipies gesif vir elke genotipe en is statisties ontleed. Die bevesting van twee mobiele toepassings, "SeedCounter" en "1KK" wat korrelmorfologie meet, is ook uitgevoer.

Alle molekulêre merkers is as betroubare diagnostiese merkers bevestig om in merkerbemiddelde seleksie (MBS) gebruik te word om spesifieke opbrengsbepalende eienskap te identifiseer, behalwe vir een merker. Die statistiese analise vir die opbrengsbepalende eienskappe het getoon dat drie genotipes beter presteer as die ander en dus as die manlike kruisingsouers binne die volgende MS-MBHS- telingskema gebruik kan word. Die assosiasie van die molekulêre merker met die opbrengsbepalende eienskappe het negatiewe korrelasies vertoon wat daarop dui dat die funksie van die hoë opbrengs gene verskil in hierdie stel genotipes. Daar is bevestig dat slegs die SeedCounter-toepassing geskik is om as 'n toekomstige fenotiperingstoepassing vir korrelmorfologie gebruik te kan word en die MS-MBHS siklusse was suksesvol uitgevoer.

Toekomstige studies moet die bevestiging van meer mobiele-toepassings insluit, tesame met die identifisering van die verhouding tussen opbrengs en die molekulêre merkers wat

geïdentifiseer was in kombinasie met QTL-kartering moet verder ondersoek word om die begrip van die onderliggende genetiese beheer van die gewenste fenotipes wat bydra tot hoër graanopbrengs beter te verstaan.

Acknowledgements

I would like to extend my gratitude to:

List of abbreviations

%	Percent
µl	microlitre
µM	Micromolar
µS/cm	Microsiemens per centimetre
ABC	ATP binding cassette
AFLP	Amplified fragment length polymorphism
ANOVA	Analysis of Variance
BAC	Bacterial artificial chromosome
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
Br	Brittle rachis
BSD	Berkley Software Distribution
CAPS	Cleaved amplified polymorphic sequence
CGIAR	Consultative Group on International Agriculture Research
Chl	Chlorophyll content
CIMMYT	The International Maize and Wheat Improvement Center
cm	centimetre
cM	CentiMorgan
CSS	Chromosome Survey Sequence
CT	Canopy temperature
CTAB	N-Cetyl-N, N, N-trimethyl Ammonium Bromide
CV	Coefficient of variation
CWR	Crop Wild Relative
DArT	Diversity arrays technology

dH$_2$O	Distilled water
DNA	Deoxyribonucleic Acid
dwb	dry weight basis
EDTA	Ethylenediaminetetraacetic acid
EFS	Expected Fragment Size
EST	Expressed transposable element
EtBr	Ethidium Bromide
F	Forward primer
f. sp.	Forma specialis
F$_1$	First generation
F$_2$	Second generation
FAO	Food and Agriculture Organization
Fig	Figure
g	Gram
g/L	Grams per litre
Gbp	Gigabase pairs
gDNA	Genomic Deoxyribonucleic Acid
GLM	General Linear Model
H^2	Heritability
H$_2$O	Water
Ha	Hardness locus
ha	hectares
Hap	Haplotype
HI	Harvest index

HTTP	High Throughput Plant Phenotyping
HYLD	High-yield
Hz	Hertz
i.e.	that is
Inc.	Incorporation
InDel	Insertion or Deletion
IWGSC	International Wheat Genome Sequencing Consortium
kbp	Kilobase pairs
kg	Kilograms
kg/ha	Kilograms per hectare
kg/hl	Kilograms per hectolitre
L	Litres
LED	Light Emitting Diode
LI	Light intercepted
Lr	Leaf rust resistance gene
LSD	Least significance difference
LTN	Leaf tip necrosis
M	Molar
MABC	Marker-Assisted Backcrossing
MAP	Mexican Agricultural Programme
MARS	Marker-Assisted Recurrent Selection
MAS	Marker-Assisted Selection
Mbp	Megabase pair
min	minute

ml	millilitre
mm	Millimetre
mM	Millimolar
mm^2	Squared millilitre
Ms	Male sterility
MS-MARS	Male Sterility Mediated Marker Assisted Recurrent Selection
MTP	minimum tiling path
NaCl	Sodium chloride
NDVI	Normalised difference in vegetation index
ng/µl	Nanogram per microlitre
NIR	Near Infrared
NNA	Nearest Neighbour Analysis
ºC	Degrees Celsius
ºN	Degrees North
ºS	Degrees South
P.	*Puccinia*
PBC	Pseudo black chaff
PCR	Polymerase chain reaction
Ph	Pairing homoeologous
pH	Percentage hydrogen
Pin	Puroindoline
Pty Ltd	Propriety Limited
QTL	Quantitative trait loci
R	Reverse primer

R^2	Coefficient of determination
RAPD	Random amplified polymorphic DNA
RCBD	Randomised complete block design
rcf	Relative centrifugation force
RefSeq	Reference Sequence
RFLP	Restriction fragment length polymorphism
RGB	Red, Green and Blue
Rht	Reduced Height
RSA	Republic of South Africa
RUE	Radiation use efficiency
SAGIS	South African Grain Information Service
SCAR	Sequence characterised amplified region
secs	Seconds
SNP	Single nucleotide polymorphism
SPAD	Soil-plant analysis development
Sr	Stem rust resistance gene
SSR	Simple sequence repeat
STARs	Sequence-tagged amplified regions
STS	Sequence tagged site
SU-PBL	Stellenbosch University Plant Breeding Lab
T.	*Triticum*
TBE	Tris/Borate/EDTA
TE	Transposable elements
Tg	Tenacious glume

TGAC   The Genome Analysis Centre

TKW   Thousand kernel weight

tin   Tiller inhibition

Tris-Cl   Tris-chloride

USA   United States of America

USD   United States Dollars

UV   Ultra Violet

v   version

V   Volt

VIR   Virulence

Vrn   Vernalisation

w/v   weight per volume

WES   Welgevallen Experimental Station

WGA   Whole Genome Assembly

YP   Yield potential

Yr   Stripe rust resistance gene

List of Figures

List of Tables

Table of Contents

Chapter 1:   Introduction

The wheat crop is of great importance to mankind as it serves as a staple to over a third of the global population. Among the cereal crops, it is one of the primary sources of proteins and calories to the world diet, where approximately 20% of food calories and 55% of carbohydrates are provided by this crop. Globally, several billion people rely on wheat as a significant proportion of their diet. Therefore, wheat proteins' nutritional value should not be misjudged as wheat products such as bread and noodles, for instance, may provide a considerable proportion of the diet within developing countries (Kumar *et al*., 2011).

The global population, however is at a point where food insecurity is at its greatest with over 800 million people suffering from chronic hunger and many more at risk of it. Even though progress has been made in some countries to eradicate hunger; other areas which include the Middle-East and Africa, the hungry population is escalating (Cheeseman 2016). In order to meet the growing population demand, the production of wheat needs to increase by 50% by the year 2050, but yield plateaus are currently being observed which results in a significant challenge to increase crop yields (Allen *et al*., 2017; Araus & Cairns, 2014).

Over the past 50 years, extensive breeding and agronomic efforts has been responsible for boosting cereal yields (Araus & Cairns, 2014). However, global climate changes such as the rise in temperature averages, severe droughts and extreme inconsistency in weather patterns present an ever-increasing challenge within an already-stressed agriculture ecosystem (Cheeseman 2016). Crop yields are largely limited by abiotic and biotic stresses; therefore, the main aim of breeding programmes is to increase yield, productivity, quality and adaptation of the crop along with optimising resource use (Lado *et al*, 2017). This is achieved through the incorporation of resistance genes and the development of more climate resilient cultivars (Singh *et al*., 2016).

For the future demand to be met, the efficiency of breeding needs to increase and this is only possible with the use of high-throughput genotyping which is a time-efficient and cost-effective way to gain genomic information (Araus & Cairns, 2014). High-throughput genotyping such as molecular marker technology allows the breeder to develop high-yielding disease resistant cultivars. This technology can identify the presence of important genes with accuracy and at a rapid speed which results in increased selection efficiency through indirectly selecting for desired traits using marker-assisted selection (MAS) (Goutam *et al*., 2015). Although, more focus is placed on selecting according to genotypic information, the phenotypic data is still

required as to confirm that the gene selected by MAS is indeed functional (Araus & Cairns, 2014; Lagudah *et al*., 2009).

Among the primary objectives in wheat breeding is achieving high yield (Cui *et al*., 2014). However, yield and yield-determining traits are quantitatively inherited and strongly influenced by the genotype x environment interaction; but some yield-determining traits are less influenced by the genotype x environment interaction than others and possess higher heritability values as compared to grain yield. Thus, examining yield-determining traits is useful when assessing yield in order to gather specific information with regards to the genetic control and the relationship that exists between yield and yield-determining traits which is essential for continued wheat improvement (Wu *et al.,* 2012).

The aim of this study was the assessment of high-yielding genotypes through the validation of yield-determining traits with the use of genotypic and phenotypic screening, as well as the incorporation of high-yielding traits into the MS-MARS facilitated pre-breeding programme to achieve grain yield improvement.

In order to achieve the aim stated in the study, the following objectives were identified.

a) Identification of yield-determining traits through reviewing literature, as well as identifying genes or QTL's associated with these yield-determining traits. Followed by MAS of the male population to validate and optimise molecular markers associated with genes related to yield-determining traits.

b) Phenotyping of the field trials of the high-yielding genotypes, in order to obtain phenotypic data concerning yield-determining traits identified. Along with identifying the relationship between the yield-determining trait and its respective molecular marker through phenotypic and genotypic data obtained.

c) Execution of a MS-MARS scheme, where the recurrent population are used as the female crossing parents and the high-yielding genotypes are used as the male crossing parents that are cross-pollinated in order to achieve the transfer of high-yielding traits into the recurrent population.

Chapter 2:   Literature Review

2.1.   Wheat

2.1.1.   The evolution and origins of wheat

Wheat was first cultivated approximately 10 000 years ago, forming part of the 'Neolithic Revolution' which was the shift from the lifestyle of a hunter-gatherer to settled agriculture (Shewry, 2009). The earliest cultivated forms of wheat included diploid and tetraploid genomes (Shewry, 2009). The tetraploid wheat that evolved around that time was *Triticum turgidum* which was the result of the hybridisation between *Triticum urartu* and *Aegilops speltoides* (Kamran *et al*., 2014). Through this hybridisation, the genetic make-up of the wild tetraploid wheat consisted of the A genome from *Triticum urartu* and the B genome from *Aegilops speltoides* (Shewry, 2009). Another hybridisation occurred between *Triticum turgidum* and *Aegilops tauschii* about 8000 years ago, which resulted in the origins of the hexaploid wheat *Triticum aestivum*, more commonly known as bread wheat (Figure 2.1) (Jia *et al*., 2013). This hexaploid wheat consisted of the A and B genome donated from the tetraploid wheat; and the D genome donated from *Aegilops tauschii* (Valkoun, 2001). When crops are domesticated, it usually leads to loss in genetic diversity, and therefore there is a substantial reduction in nucleotide diversity when comparing with ancestral populations. For instance, when domestication of the tetraploid emmer wheat occurred there was a reduction of 30% - 50% in nucleotide diversity of the A and B genome (Chao *et al.,* 2010, Brenchley *et al.,* 2012). However, as time passes, new mutations accumulate which results in an increase of diversity that is more uniformly distributed across the genome (Akhunov *et al*., 2010).

In the early days of when cultivation was first established; farmers selected from wild populations, essentially landraces, for higher yield and other desired characteristics which was undoubtedly a non-scientific form of plant breeding. Nevertheless, domestication was also linked to selecting for genetic traits that distinguished them from their crop wild relatives (Shewry, 2009). One of the first symbols of domestication was the selection for the transformation of brittle rachis (*Br*) to non-*Br* (Peng *et al.*, 2011). Spike-shattering at maturity resulted in seed loss during harvesting, thus the alteration of the brittle rachis trait was critical. It was then identified that in a non-*Br* spike, breakage zones are suppressed until the harvesting of mature spikes. Therefore, early farmers consciously selected the mutated plants that displayed non-brittle spikes and thus, their frequency had a constant increase in the cultivated fields (Peng *et al.*, 2011; Shewry, 2009). Later, the brittle rachis trait was mapped to the

chromosome 3 homeologous group in wheats. In an undomesticated form of *T. aestivum*, it was identified that on chromosome 3DS a single dominant gene, *Br1*, was responsible for fragile rachis and the non-brittle rachis characteristic in domesticated emmer wheat is under the control of two main genes, *Br2* and *Br3* localised on chromosome 3AS and 3BS, respectively (Peng *et al.*, 2011).
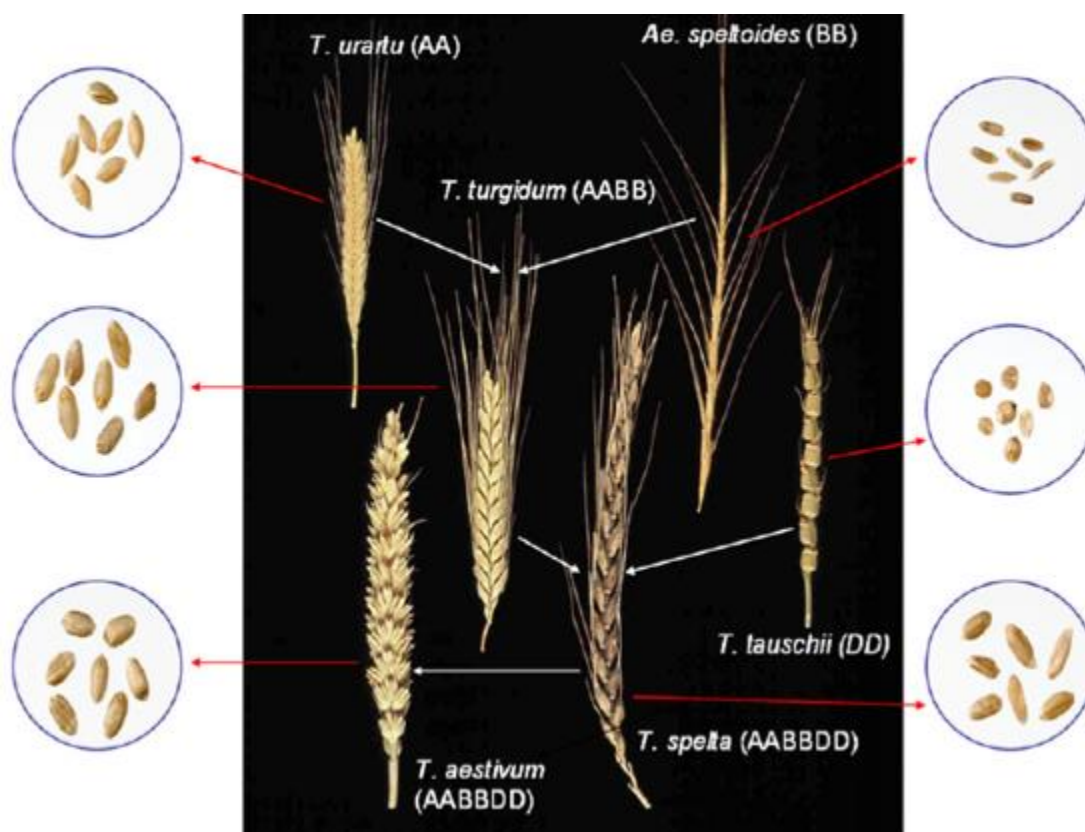
Figure 2.1: The evolutionary events that occurred during the domestication of the hexaploid bread wheat (Shewry, 2009).

Another key trait during the domestication process was glume tenacity which is closely related to the free-threshing trait. The floret of wild wheat is wrapped with tough glumes which makes the threshing of spikes difficult, and in contrast, the cultivated wheat florets possess soft glumes which allows for free-threshing. The detection of several QTLs that affect the free-threshing trait were found on chromosomes 2A, 5A, 6A, 2B, 7B, 2D and 6D. However, the free-threshing trait is largely affected by the partially recessive allele at the tenacious glume (*Tg*) loci on chromosome 2DS and the partially dominant allele at the *Q* loci on chromosome 5AL (Peng *et al.*, 2011, Matsuoka, 2011). The interaction of the *Tg* and *Q* loci largely affect the morphology

of a spike and since the expression of *Q* is suppressed by *Tg,* a *QQTgTg* genotype will display a non-free-threshing phenotype. Essentially, the change from a *qqTgTg* genotype to a *QQtgtg* genotype was necessary for the development of the free-threshing trait in the phenotype of wheat (Matsuoka, 2011).

Each ploidy level of wheat (i.e. diploid, tetraploid, hexaploid) possesses its own geographical centre of diversity and therefore, it can be concluded that the geographical place of origins for each ploidy level is different. The centre of diversity of the tetraploid wheat (*T. turgidum*) as well as other free-threshing tetraploid wheat was placed to Northeastern Africa and the Eastern Mediterranean. The diploid wheat, *Aegilops tauschii* was spread from the Caucasus region, across Eurasia through central Asia towards the east in China and the hexaploid wheat was placed to be across areas from Afghanistan and Turkmenistan to Transcaucasia (Dvorak *et al.*, 2011; Jones *et al.*, 2013).

The genome of wheat is among the largest genomes among field crop plants and due to this, whole genome shotgun sequencing of the wheat genome was delayed by computational power and the decision was made to sequence the chromosomes of wheat, individually. The assembly of a reference sequence genome of wheat was released in 2017 and with this reference sequence available, it enables the identification of genes and markers that are related to important agronomic traits and this will assist in accelerating the development of better adapted cultivars (Bierman & Botha, 2017; IWGSC, 2017).

2.1.2.   Sequencing the genome of bread wheat

The reality of sequencing the isolated wheat chromosomes and progenitor genomes first arose in 2003, when the international wheat genome sequencing consortium (IWGSC) first assembled at a workshop. Nine years later, Brenchley *et al*. (2012) reported on the first draft sequence for hexaploid wheat. This was achieved through the use of shotgun sequencing, where lower coverage (five-fold) with longer read lengths were obtained using Roche 454 pyrosequencing technology. With this draft sequence, it was then estimated that the total number of genes within the wheat genome was between 94000 and 96000 (Brenchley *et al*., 2012; Bierman & Botha, 2017).

In 2014, the wheat genome sequence by the IWGSC was published. Mayer *et al.* (2014) isolated the individual chromosomes of "Chinese Spring" by using double ditelosomic wheat

lines of the cultivar for sequencing, using Illumina technology platforms. The paper reported that the wheat crop is made up of a 17-gigabase-pair genome with an allopolyploid structure that consists of three homeologous sets, each consisting of 7 chromosomes in the A, B and D genome. This results in a total of 21 pairs of chromosomes that make up the whole hexaploid wheat genome. It does, however, genetically behave as a diploid due to the prevention of homeologous pairing by the action of pairing homoeologous (*Ph*) genes. Each genome is approximately 5.5 Gbp in size and carries more than 80% of highly repetitive transposable elements (TEs) (Mayer *et al.*, 2014; Bierman & Botha, 2017). When the TEs and the sequence repeats across the whole genome was assessed, it was discovered that 76.6% of assembly sequences and 81% of raw sequence reads contain repeats. On average, the duplication of genes for all chromosomes are 23.6%, and was stated by Mayer *et al.* (2014) to be an underestimation (Bierman & Botha, 2017; Mayer *et al.*, 2014).

Several whole genome assemblies were released by the IWGSC for wheat and its progenitors. These assemblies include the TGAC wheat reference genome assembly which was publicly released in April 2016 and represented 78% of the genome with N50 for 88.8 kbp scaffolds. Later within that year, the IWGSC WGA v0.4 assembly was also made available to the public for downloading or BLAST analysis. This assembly provided researchers with a chromosome-based draft sequence of "Chinese Spring". The scaffolds that were produced by Illumina short sequence reads added up to 14.5 Gbp (Bierman & Botha, 2017). In 2017, IWGSC generated the first version of the reference sequence for bread wheat. This reference sequence includes highly diverse community resources such as physical chromosomal maps, BAC (bacterial artificial chromosome)-based MTP chromosome sequences, CSS assemblies, Hi-C scaffolding, a high quality whole genome shotgun assembly and many genetic markers (Stein, 2017; IWGSC, 2017).

The expectations that IWGSC has for this reference sequence is to reduce time and to improve successful cloning of genes and QTLs; as well as providing unlimited access to high quality markers to be used in MAS and lastly, to facilitate in exploring diversity in genetic resources for pre-breeding programmes. The reference sequence of bread wheat is now available for BLAST and download, but the planned publication of this reference sequence has yet to be released (Stein, 2017).

2.1.3.   The wheat crop and its importance

Wheat is one of the most widely cultivated crops due to its ability to adapt to a varied range of climates and its improved grain quality to produce baker's flour (Jia *et al*., 2013). Wheat's adaptability to different environments enables it to be grown in countries that fall within the range between the equator and latitudes of 60°N and 44°S, and hence is cultivated in many countries which includes Scandinavia, Russia as well as countries such as Argentina (Shewry, 2009; Singh *et al*., 2011). The temperature that is optimum for wheat growth is 25°C, however, temperatures ranging between 3°C and 32°C can achieve satisfactory wheat yields as well (Kamran *et al*., 2014). Wheat's versatility is also evident from the fact that it can be grown in areas that receive low or high precipitation, ranging between 250mm and 1750mm (Monneveux *et al*., 2012).

Wheat can be divided into three types that can be classified as winter, spring and facultative types which are determined by vernalisation (Muterko *et al.*, 2015). Vernalisation is the requirement of a crop to be exposed to low temperature for a long period of time in order to flower. The allelic variants (*Vrn-A1, Vrn-B1, Vrn-D1*) at the *Vrn* genes are responsible for vernalisation and confers vernalisation insensitivity or sensitivity. Winter wheat contains all the recessive alleles and therefore, must undergo vernalisation (Iqbal *et al.*, 2007; Yan *et al*., 2004). Spring wheat, however, contains the dominant *Vrn-A1* allele and does not need to undergo vernalisation. The facultative type contains either the dominant *Vrn-B1* or *Vrn-D1* alleles alone which is associated with minimum vernalisation requirement (Muterko *et al.*, 2015). Since winter wheat becomes dormant during the winter season, it is planted in autumn and resumes growth in spring again. Spring wheat, however, continues growing from planting to harvest and is grown in areas where temperatures never reach cold enough conditions for vernalisation. Spring wheat are also grown in areas that experience such severe winters that the dormant wheat plant would die (Sacks *et al.*, 2010). In South Africa, wheat production is quite unique as three distinct areas exist in which wheat is produced. Large parts of the Western Cape province experiences a Mediterranean climate and therefore, dryland spring wheat are grown within those parts of the region; whereas the summer rainfall areas grow irrigated spring wheat types. However, the Free State province experiences dryland conditions where winter wheat are grown in soil with stored moisture that accumulated throughout the previous summer and autumn seasons (Smit *et al*., 2010).

The intimate knowledge of growth stages of wheat can provide valuable information on factors that positively and negatively influence forage, and grain yield potential to improve management decisions. The Feekes scale which can be seen in figure 2.2, is the most widely used scale to describe growth stages of wheat; other scales also commonly used include the Zadoks and Haun scales. The growth stages of the Feekes scale is divided into 11 stages and the four main sections are tillering (stages 1-5), stem extension (stages 6-10), heading (stages 10.1-10.5) and ripening (stage 11) (Miller, 1999).
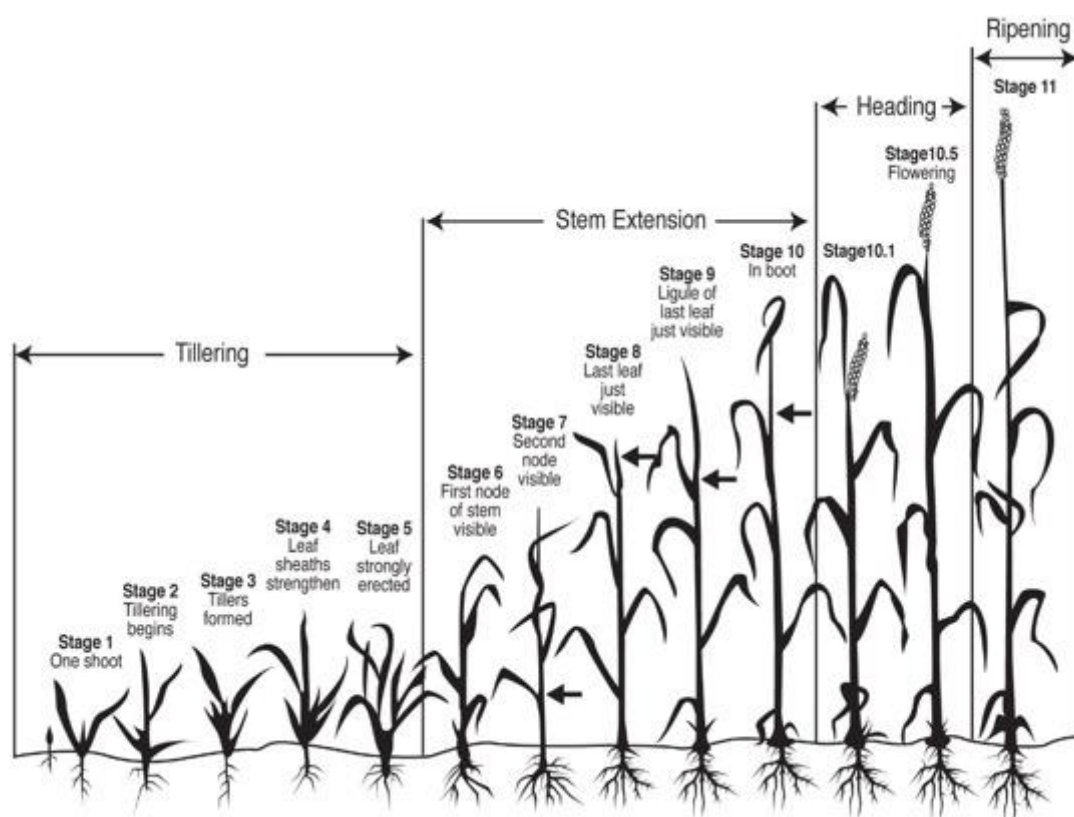
Figure 2.2: The growth stages of wheat according to the Feekes scale (Marsalis & Goldberg, 2016).

2.1.4.   Global and local production of wheat

The production of wheat is grown over approximately 225 million hectares of land worldwide as it is one of the major food crops that provides 20% of the global caloric intake. The total wheat produced annually approximates to 700 million tons, where nearly half is produced within developing countries (Kaur *et al*., 2017). Being one of the main food crops, wheat is an essential component to the dietary intake of 2.5 billion poor people whose standard of living is equivalent to less than 2 USD per day, with the majority being woman and children. High

dependency on wheat as a primary source of cereal calories and protein are found within the following countries; Central Asia, West Asia, Eastern Europe, North America, North Africa, Australia and Russia. Within these regions, wheat is solely responsible for more than three-quarters of the total cereal consumption (Shiferaw *et al*., 2013).

According to the Food and Agriculture Organisation (FAO) of the United Nations, the global wheat output is estimated at 754.8 million tons in 2017, which is a percentage lower than last year (FAO, 2017). This is concerning as wheat yields need to increase by at least 50% by 2050 as the world's population is estimated to increase to 9 billion people by then (Allen *et al*., 2017; Stratonovitch & Semenov, 2015). Therefore, wheat is currently a major global priority and needs to achieve increased yields in order to feed this growing population. However, to meet this demand would be challenging as yield plateaus are being exhibited and the possibility of extending crop-growing areas are limited (Allen *et al*., 2017; Stratonovitch & Semenov, 2015).

In South Africa, the total wheat produced in 2016 amounted to 1 910 000 tons that was planted on a total area of 508 365 ha. The area and production of wheat for 2017 was reportedly a final total area of 491 600 ha of wheat that was planted and a final total of 1 475 450 tons of wheat produced (Figure 2.3). The Western Cape, Northern Cape and Free State planted 326 000 ha, 38 000 ha and 80 000 ha of wheat, respectively. The reduction in wheat produced in 2017 is mainly due to the drought experienced in the Western Cape during the past season, where production declined by 37.6% from the previous year. Therefore, the final production of wheat in the Western Cape, Free State and Northern Cape was estimated at 586 800 tons, 296 000 tons and 304 000 tons, respectively. The negative effect of the drought experienced in the Western Cape can clearly be seen when comparing the production of wheat against the previous years (Figure 2.4).

The decline in wheat production forecasts from August to December can be seen on figure 2.3. Wheat production in South Africa during 2017 is 22.8% less than wheat produced in 2016 (SAGIS, 2017). When comparing the total South African production of wheat to the total production of wheat in the Western Cape, it is seen that the wheat produced in 2017 consists of 47% of the wheat produced in the Western Cape. In 2016, 42% of the total wheat produced was from the Western Cape and this shows that more than 40% of the total wheat produced in South Africa is dependent on the Western Cape wheat production. With the current drought, the production of wheat is suffering a huge loss. This, however, highlights the need to breed

9

for resistance against drought within South Africa in order to avoid losses incurring due to environmental conditions.
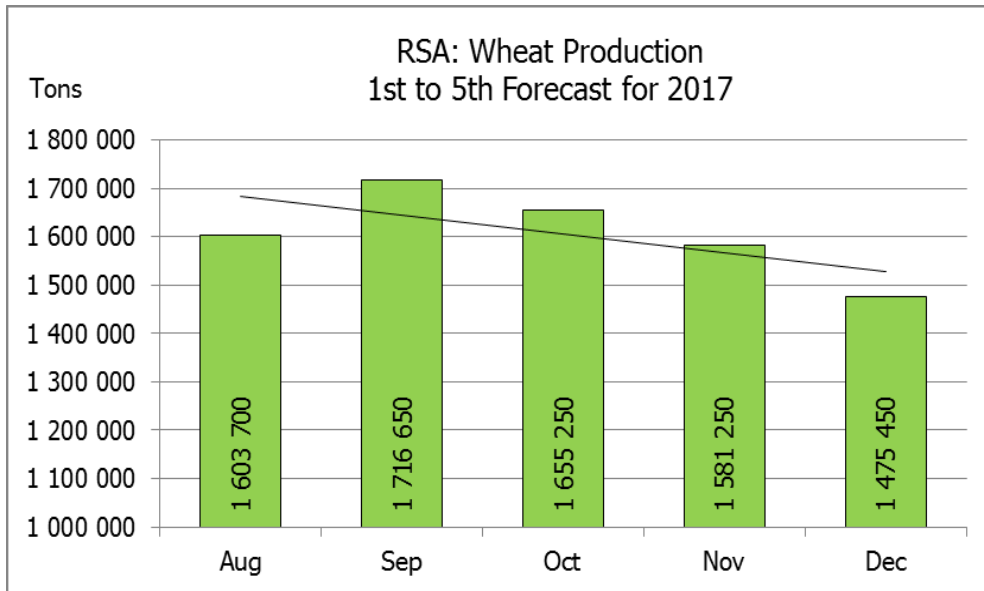
Figure 2.3: The 2017 production forecasts for wheat in South Africa from August to December (SAGIS, 2017).
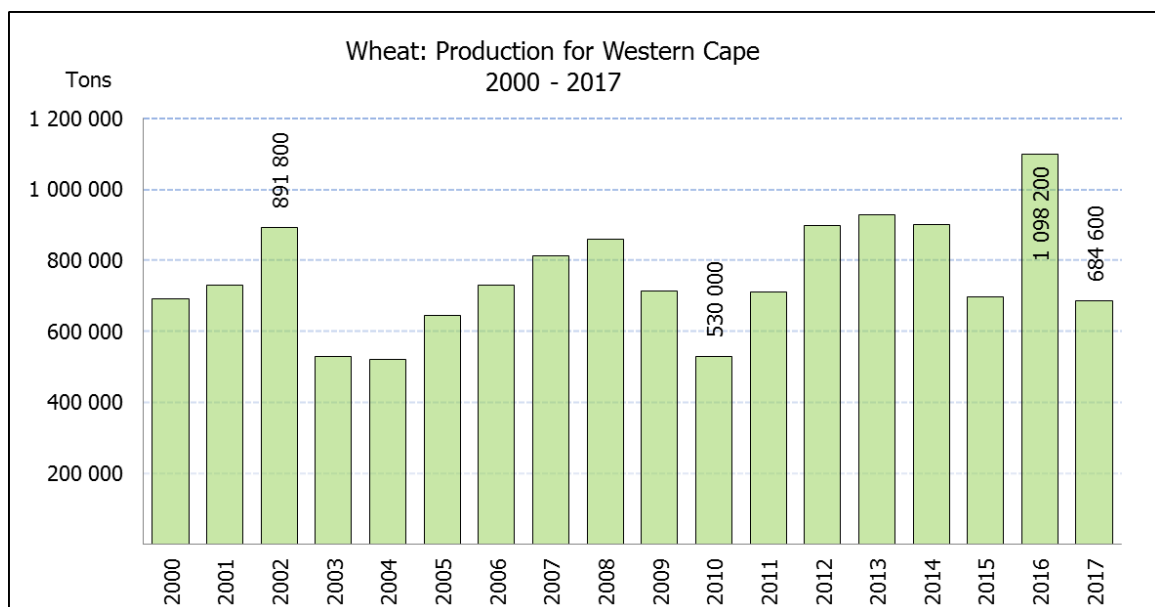
Figure 2.4: The total production of wheat in the Western Cape from 2000-2017 (SAGIS, 2017).

2.1.5.   Limitations of wheat production

The maintenance of global food security is a challenge as limitations are caused by abiotic and biotic stresses that negatively influence production, yield and the general survival of wheat; therefore, giving rise to the increase in food insecurity and poverty (Agarwal *et al*., 2006; Ortiz *et al*., 2008). These limitations cause large yield gaps, therefore to rectify this and maintain food security, the implementation of system management, mitigation and germplasm adaptation is needed. Germplasm adaptation can be achieved through introducing novel genes from new sources that confer resistance against abiotic and biotic stresses in wheat. This paired with efficient breeding and selection methods can increase the genetic yield potential significantly (Ortiz *et al*., 2008).

2.1.5.1. Biotic Stresses

Biotic stresses are living factors that have negative effects on wheat and its production (Agarwal *et al*., 2006). One of the major biotic factors are fungal pathogens, particularly rust pathogens that causes significant losses in wheat yield. *Puccinia graminis* Pers.f. sp. *tritici Eriks. & Henn.*, *Puccinia triticina* Eriks and *Puccinia striiformis* Westend are the rust fungi that causes stem, leaf and stripe rust, respectively (Mallick *et al.,* 2015). These pathogens pose as a huge threat to food security as rust epidemics can result in yield losses up to 50% (Kaur *et al.*, 2017). Practices in controlling pest and diseases are implemented, but annual losses within the developing countries still amounts to an estimated 13% which is solely due to pests, pathogens and viruses. The incidence and impact of the rust pathogens increase with monocropping, cropping intensity as well as uniformity in genes that confer resistance (Shiferaw *et al.,* 2013).

2.1.5.2. Abiotic Stresses

Abiotic stresses are non-living factors such as drought, temperature extremities, salinity and nutrient stress that negatively impact agriculture. Therefore, the average yields can be reduced between 50-100% simply because of the presence of abiotic stresses. Additionally, when a plant is stressed with biotic stresses, the effect of the abiotic stress is increased (Atkinson & Urwin, 2012; Barlow *et al*., 2015). Various physiological, metabolic and biochemical

approaches have been adapted by plants in order to encounter abiotic stresses (Younis *et al.*, 2014).

There are two main strategies implemented to counter abiotic stresses, namely stress tolerance and stress avoidance. Stress tolerance is the potential ability of a crop to adapt to stressful conditions and stress avoidance includes various protective mechanisms that prevent or alternatively delay the negative influence that the stress factor has on the crop (Krasensky & Jonak, 2012).

2.2. The development of wheat improvement

2.2.1. Breeding programmes

The need to improve the resistance of wheat is critical and therefore, the International Maize and Wheat Improvement Center (CIMMYT) are leading the Global Wheat Programme from the Consultative Group on International Agriculture Research (CGIAR), with the main aim of increasing production of wheat cropping systems to achieve food security within developing countries (Guzman *et al.*, 2016). CIMMYT prioritises the improvement of disease resistance, grain yield, tolerance to abiotic resistance, as well as desirable quality in wheat through their global wheat breeding programmes. Plant breeding programmes play a crucial role in the efforts to increase food production. Therefore, CIMMYT is constantly working to develop new wheat germplasm to be used by national partners for the improvement of their own germplasm, or to be released directly as a cultivar (Guzman *et al.,* 2016). Another collaborative network that is coordinated by CIMMYT is the International Wheat Yield Partnership (IWYP). The primary aim of the IWYP is improving wheat yield potential through an international strategy where agricultural experts and wheat scientists from public or private institutions unite as members of the IWYP to achieve increased wheat yields (Solis-Moya *et al.,* 2017).

Breeding programmes with the aim of releasing successful cultivars commercially tend to grow several thousand genotypes within a set of targeted environments so that phenotypic selection can be taken for grain yield and other major traits. Grain yield along with other traits are selected for in breeding programmes when selecting for new cultivars to be commercially released. However, requirements have been set for these cultivars by the marketers, processors, consumers and especially the farmers that need to be met. Ultimately, the farmer wants a

cultivar that possesses traits such as high grain yield, durable resistance against a wide range of pests, diseases as well as tolerance against abiotic stresses that includes temperature, drought, soil acidity and salinity. Therefore, the new cultivars need to have superior traits to the cultivars currently grown by farmers (Richards *et al*, 2010).

2.2.2. Pre-breeding programmes

Pre-breeding is all the activities involved in identifying desired genes and/or traits obtainable from exotic or semi-exotic (unadapted) material (Iqbal *et al.,* 2013). Plant material such as popular cultivars, wild type germplasms and landraces are used as donors that cannot directly be used in breeding populations, to transfer desired traits into recipients that have well-adapted genetic backgrounds. This results into the development of intermediate material that may be utilised by plant breeders in specific breeding programmes to aid in the development of new cultivars that possess a broad genetic background (Figure 2.5) (Sharma *et al*., 2013).
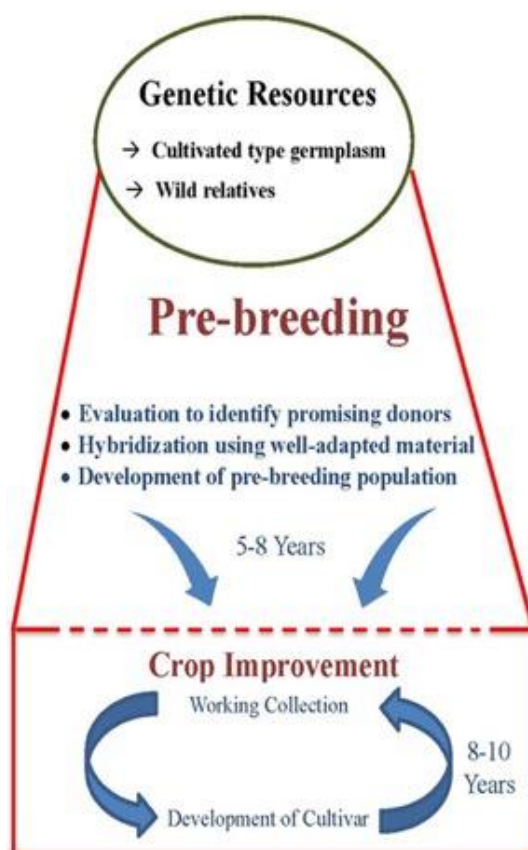
Figure 2.5: An illustration on the process of how pre-breeding forms part of the crop improvement process (Sharma *et al.,* 2013).

Pre-breeding programmes can also serve as a means to attempt to reset the genetic diversity of crops as genetic diversity, in some extent were lost during domestication and during the improvement process of crops. Wild species that are closely related to the domesticated species are known as crop wild relatives (CWRs) and these can be used as a rich source of genetic diversity that can reintroduce genetic variation into cultivated cultivars (Brozynska *et al.*, 2016; Dempewolf, 2017). Pre-breeding has even made use of genetic diversity that was not formerly accessible because of either non-overlapping geographic ranges or genetic incompatibilities. Therefore, pre-breeding is a key step for the linkage between the desired traits possessed by CWRs to the development process of a modern cultivar by providing breeders with a more immediate usable form of wild genetic diversity (Dempewolf, 2017).

Wheat can obtain these desirable traits from three gene pools that is defined according to the amount of effort needed to employ them. Each gene pool is classified according to the degree of phylogenetic relatedness to common wheat (Feldman & Sears, 1981; Reynolds *et al.,* 2009a). The three gene pools are known as the primary, secondary and tertiary gene pools. The primary gene pool is the easiest to use as it contains germplasm that have a shared common genome but were separated from the mainstream genome, such as landraces. The secondary gene pool contains closely related genomes that can be used by undergoing interspecific hybridisation. An example being the hybridisation that took place to form the hexaploid wheat. The tertiary gene pool contains related types of grasses, but to transfer genes would require special techniques (Reynolds *et al.,* 2009a).

For the identification of the desirable traits, plant breeders make use of genotypic and/or phenotypic approaches. The selection method is an important tool in breeding and with breeders recognising molecular marker technology within modern plant breeding; the selection approach has shifted from phenotypic selection to direct or indirect selection of genes (van Bueren *et al.,* 2010). The application of molecular markers has many advantages which include pyramiding of desirable gene combinations into a single genetic background and it can be applied during the early stages of the plant's growth which cannot be done for certain phenotyping methods. Molecular markers also accelerate the transfer of desired genes from unadapted plant material into the desired germplasm through cross-pollination (Randhawa *et al*., 2013). The use of molecular markers provides a vital alternative to phenotypic-based selection as molecular markers are cost-effective, efficient and reliable; especially when phenotyping becomes difficult and costly (Tester & Langridge, 2010).

2.2.2.1. Marker Assisted Selection

Marker-assisted selection (MAS) is frequently used in pre-breeding programmes. This approach employs molecular markers to identify genes of interest in a cost-effective way, and can also be used as a highly precise selection tool for desired wheat lines (Collard & Mackill, 2008; Prabhu *et al*., 2009; Wessels & Botes, 2014). The use of available molecular markers has given breeders the chance to combine desirable alleles at numerous loci within a short time and thus, MAS has successfully contributed in pyramiding targeted genes in major crops, including wheat. Many studies reported the successful pyramiding of target genes with the use of MAS to achieve leaf rust resistance (Cox *et al*., 1994; Gupta *et al*., 2005; Nocente *et al*., 2007; Singh *et al*., 2004), powdery mildew resistance (Liu *et al*., 2000; Wang *et al*., 2001) and fusarium head blight resistance (Badea *et al*., 2008; Shi *et al*., 2008; Tamburic-Ilincic *et al*., 2011) (Tyagi *et al*., 2014).

Marker-assisted selection brings about the opportunity for the selection of desired lines on the basis of genotypic screening rather than phenotypic screening (Prabhu *et al*., 2009). It is, therefore, especially beneficial to select for traits that are difficult to identify phenotypically, due to environmental error or the costly expenses needed to assess these traits (Inostroza-Blancheteau *et al.*, 2010). Additionally, MAS can be executed on DNA extracted from the leaf tissue of a plant and therefore provides a non-destructive substitute to selection based on phenotype (Kuchel *et al*., 2007). Although molecular markers are applied with high precision, MAS cannot completely replace conventional selection techniques because of the genetic complexity and high number of the traits selected (Miedaner & Korzun, 2012). Breeders has acknowledged that a certain gap between genotype and phenotype will always exist as molecular markers will never be fully informative, especially when more genotype x environment interaction exists. Thus, the development of easier phenotypic selection methods will remain a research priority. However, MAS has many advantages over phenotypic selection such as increased flexibility that enables breeders to work with smaller populations, leading to more effective use of field trial capacity. In conclusion, the use of molecular markers need not be an exclusive selection tool to be utilised, but rather considered as a complement to phenotypic selection (van Bueren *et al.*, 2010).

2.2.2.2. Recurrent Mass Selection

Recurrent mass selection is a breeding technique that is well-established and used for cross-pollinating species to achieve genetic improvement. Primarily this technique was developed to improve quantitatively inherited traits with the objective of increasing the frequency of desired genes within a breeding population. Thus, increasing the opportunity to exploit superior genotypes with desired traits. However, the use of recurrent mass selection for self-fertilising crops was discouraged as intercrossing was difficult for each cycle, with low seed production results. As a solution, Hallauer (1981) suggested to integrate recurrent selection procedures with other selection methods, but to not expect its products to be directly useful in developing commercial cultivars. Many pilot studies made use of recurrent selection and achieved positive outcomes, but these studies consisted of less than five cycles and therefore was short-term. Single traits were also only pursued and the number of possible intercrosses to be made were also restricted (Marais & Botes, 2009). Effectiveness of the recurrent selection was highlighted by various studies for genetic improvement of many traits with crops which include barley, oats, soybean and wheat (Diaz-Lago *et al.,* 2002; Liu *et al.*, 2007; Wiersma *et al.*, 2001). Even though, impressive selection progress was made in numerous studies, the usefulness of this strategy was still flawed and therefore a holistic approach became necessary for cultivar improvement. Strategies for improved approaches were reported by Wallace & Yan (1998), Jensen (1970), Falk (2002) and Huang & Deng (1998). These strategies included the use of recurrent mass selection in combination with conventional breeding strategies and male sterility (Marais & Botes, 2009).

2.2.2.3. Marker Assisted Recurrent Selection scheme

One of the first reported recurrent selection breeding populations in wheat were established by Huang and Deng (1988) that made use of a dominant male sterile gene, *Ms2*, with the population segregating into male fertile and male sterile (i.e. female) plants. In their scheme, a selection of female plants undergo natural (field) pollination by selected male fertile plants. Recurrent selection-based applications were then pursued by Chinese researchers, forming a small nation-wide network. Cox *et al.* (1991) was next to develop and register a germplasm source that segregates for the dominant *Ms3* gene that confers male sterility. A duration of several years were spent to develop a highly heterogeneous base population with the use of parents that are a source for diverse desired genes. Some ways in which cross pollination of

female spikes by male plants was enhanced, was with the use of fans and manual agitation (Marais & Botes, 2009).

At Stellenbosch University, Marais *et al.* (2000) used the dominant *Ms3* gene for the establishment of a recurrent selection base population that would segregate into a 1:1 ratio of male sterile and male fertile plants. This recurrent mass selection scheme enabled the establishment of a base population that is genetically diverse and therefore, rich with genes for quality, pest resistance, yield and adaptation through utilising the male sterility segregation in order to cross-pollinate male sterile plants with male donor plants that possess desired traits. The scheme consisted of breeding cycles with a duration of at least four years where males and females were differently handled (Marais & Botes, 2009).

To establish the recurrent base population, a male sterile winter wheat accession "KS87UP9" was cross-pollinated with the spring wheat "Inia 66". The produced F_1 plants that were sterile was then cross-pollinated with a selected spring-type wheat breeding line. This male sterile gene results in the recurrent selection base population obtaining 50% male sterile and male fertile plants for each recurrent cycle and therefore male sterile plants are easily obtained due to this gene. Marker-assisted selection (MAS) technology was also introduced into the recurrent mass selection scheme by 2005, for genotypic screening of rust resistance genes. There are many advantages to using molecular markers which includes ensuring economic efficiency and genetic effectiveness. The use of molecular and phenotypic markers during the different cycles provide important genetic information regarding the genotypes. This technique using molecular markers, phenotypic measurements and the male-sterility gene is known as male sterility marker-assisted recurrent selection (MS-MARS) scheme (Marais & Botes, 2009).

2.2.3.   Incorporation of genetic resistance and agronomic improvement

A crucial point in crop improvement was when dwarfing genes were introduced into the wheat and rice crop, this occurrence was later known as the 'green revolution'. This bought about higher yield due to the reduction in stem stature resulting in higher lodging tolerance and higher harvest index (Zhang *et al*., 2014a). The green revolution started when the Rockefeller Foundation had sent out a team to Mexico to survey their agriculture in 1941, resulting in the development of a programme called the Mexican Agricultural Programme (MAP). In 1944, a young biologist was hired by the name of Norman Borlaug and his ingenuity and dedication

resulted in the development of the "miracle wheat" in 1954 (Patel, 2013). A Japanese wheat cultivar, "Norin 10", was the original source for the two dwarfing alleles, named *Rht-B1b* and *Rht-D1b* that was located on chromosome 4BS and 4DS, respectively (Kang *et al.*, 2012; Zhang *et al.*, 2014a). The development of this "miracle wheat" caused the global food supply to triple during the last 30 years of the 20th century through exploiting the deployment of the dwarfing genes. The Rockefeller and Ford foundations contributed to spreading the "miracle wheat" during the 1950s and 60s, but even today, these genes are still globally deployed within wheat breeding programmes (Kang *et al.,* 2012; Langridge, 2014; Patel, 2013). These semi-dwarf genes complimented improved agricultural practices very well, but obtaining yield gains are becoming more difficult as compared to when the green revolution was first introduced (Shiferaw *et al.*, 2013).

Breeding for durable resistance is one strategy that could be implemented in breeding programmes to achieve higher yield gains. The two categories in which disease resistance can be divided into is pathotype-specific and pathotype non-specific. Pathotype-specific also referred to as vertical resistance is controlled by major genes that provide the crop with complete resistance at either seedling or adult plant stage (Figlan *et al.*, 2017; Lillemo *et al.*, 2008). The pathotype-specific resistance genes (R-genes) usually conform to a "gene-for-gene" model which confers resistance to pathogens that carry the corresponding avirulence (*Avr*) gene (Figueroa *et al.*, 2016). Therefore, the outcome of the infection is simply determined by these two genes. When the *Avr* gene is recognised by the plant that is carrying the corresponding R-gene, a defence response is activated which results in cell death during attempted infection. Alternatively, infection may occur if the plant is carrying a susceptible gene or the pathogen has a different virulence gene (Anderson *et al.*, 2016 Persoons *et al.*, 2017). However, R-genes are associated with short durability as it is easily overcome by pathogens which are continuously striving to gain virulence against resistant genes through mutation, genetic recombination and new introductions from different countries (Figlan *et al.*, 2014; Lillemo *et al.*, 2008). Through mutating to virulence (VIR), the recognition by the host plant is avoided and the defence response is not activated (Wu *et al.*, 2017). Additional genetic protection that can be employed against pathogens is the use of pathotype non-specific resistance genes (Figueroa *et al.*, 2016).

Pathotype non-specific also referred to as horizontal resistance which is controlled by minor genes that remain effective against all pathotypes at adult plant stage as a partial, slow rusting resistance. These resistance genes are associated with long term and durable resistance which

permits susceptibility of infection from the pathogen to the plant at seedling stage, but later effectively displays resistance towards numerous pathogens at adult stage (Figlan *et al*., 2017; Lan *et al*., 2015). However, pathotype non-specific genes deployed alone does not provide very effective resistance and therefore is usually deployed along with other minor genes and thus, is associated with quantitative inheritance (Lan *et al*., 2017). Thus far, many pathotype non-specific resistance genes against wheat rusts has been characterised, namely: *Sr2, Sr55, Sr56, Sr57, Sr58, Lr34, Lr46, Lr67. Lr68* and *Yr36* (Yu *et al*., 2017; Li *et al.,* 2017). There are three QTLs that have been cloned in wheat conferring resistance to wheat rust. One QTL confers resistance against stripe rust (Fu *et al.*, 2009) and the other two are more wide-spectrum QTLs as they confer resistance against leaf rust, stem rust and stripe rust, as well as powdery mildew (Moore *et al.*, 2015; Krattinger *et al*., 2009; Yeo *et al.*, 2017). Zou *et al.* (2017) has also identified three QTLs for both leaf and stripe rust; i.e. *QLr.dms-2D.1*, *QLr.dms-2D.2*, and *QLr.dms-3A* for leaf rust; and *QYr.dms-3A*, *QYr.dms-4A*, and *QYr.dms-5B* for stripe rust (Zou *et al*., 2017).

2.2.3.1. Rust resistance

Three wheat rusts are known for causing considerable damage to wheat worldwide, and therefore breeders have adopted the deployment of resistance genes to control losses that is caused by wheat rusts (Mallick *et al.,* 2015). However, numerous resistance genes that have been identified are pathotype-specific which makes these genes more prone to be overcome by pathogens that evolve into new virulent pathotypes. Rust resistance breeding therefore shifted its efforts to focusing on achieving durable resistance which is widespread, pathotype non-specific and expresses prolonged resistance. Exceptional durable resistance is known to be expressed by the slow-rusting, partial, adult plant resistance genes, *Lr34* and *Sr2* (Juliana *et al.,* 2015).

2.2.3.1.1. *Lr34* gene

The *Lr34* wheat resistance gene has retained its effectiveness for several decades as no virulence has evolved against this gene and therefore it is commonly deployed within breeding programmes (Krattinger *et al*., 2009). When this gene was first characterised by Dyck *et al.* (1966), it was also discovered that the presence of this gene was in many wheat cultivars as it

19

had form part of the improvement of wheat during the early twentieth century already (Krattinger *et al*., 2009; Muthe *et al*., 2016). The *Lr34/Yr18/Pm38* gene complex is valuable as it confers durable resistance against the two major wheat rust diseases; leaf rust (*P. triticina*) and stripe rust (*P. striiformis*), as well as the powdery mildew (*Blumeria graminis*) disease. The *Lr34* gene is also associated with tolerance towards stem rust (*P. graminis*), barley yellow dwarf virus (*Bdv1*) as well as spot blotch (*Bipolaris sorokiniana*) (Juliana *et al*., 2015; Krattinger *et al*., 2009).

Many wheat cultivars carrying the *Lr34* gene has flag leaves that develop a necrotic leaf tip which is described as leaf tip necrosis (*LTN*) as can be seen on figure 2.6c (Krattinger *et al*., 2009). *LTN* is considered as a phenotypic marker associated with *Lr34* as it shows complete linkage to this gene. However, under field conditions, the appearance of *LTN* will take time and predicting the presence of *Lr34* based on *LTN* is not always reliable as the expression of *LTN* may vary between different environments. Therefore, the *Lr34* gene should be identified through efficient screening methods such as the use of molecular markers (Muthe *et al*., 2016, Lagudah *et al*., 2009). The *Lr34* gene is found on chromosome 7DS, was cloned and sequenced; and reportedly encoded for an energy ATP-binding cassette (ABC) transporter, therefore the complete nucleotide sequence of the *Lr34* gene is known. It was also reported that the only difference between susceptible and resistant plants are three genetic polymorphisms (Krattinger *et al*., 2009). With this information available, Lagudah *et al.* (2009) was able to develop molecular markers that are gene-specific to identify the presence of the *Lr34/Yr18/Pm38* gene complex (Lagudah *et al*., 2009).

2.2.3.1.2. *Sr2* gene

In 1999, a new stem rust pathotype group was first reported in Uganda that quickly became a threat to the global production of wheat. This pathotype was termed Ug99, but later, Wanyera *et al.* (2006) designated the pathotype as TTKS which was named according to the North American nomenclature system. However, after the addition of a fifth set of differential lines, this pathotype was once again re-named as TTKSK (Figlan *et al*., 2014; Yu *et al.,* 2011). The Ug99 pathotype was virulent on many major genes including *Sr24* and *Sr31,* which were previously effective resistance genes. The resistance gene *Sr2* provides the wheat crop with partial resistance to adult plants from all identified stem rust pathogens as well as the Ug99

pathotype group members. Due to the partial resistance, protection of this gene alone is insufficient; but when combined with other stem rust resistance genes, protection in adult plants against stem rust is provided. Durable adult plant resistance against stem rust has been provided from the *Sr2* gene since it was selected by McFadden during the 1920s (Figlan *et al*., 2014; Yu *et al*., 2011).

The *Sr2* gene is closely linked to the pseudo-black chaff (*PBC*) phenotype as can be seen on figure 2.6d, which is the appearance of a dark pigmentation occurring on the glumes, stem internodes and peduncle of wheat. Therefore, this phenotype has served as a phenotypic marker for the *Sr2* gene in breeding programmes for many years (Yu *et al*., 2011). The expression of this phenotypic marker varies as it depends on genetic backgrounds as well as environments (Juliana *et al*., 2015). Therefore, a molecular marker that can predict *Sr2* with high accuracy within diverse wheat germplasm was developed. This molecular marker is a cleaved amplified polymorphic sequence (CAPS) marker which can detect the presence of the *Sr2* gene that is located on chromosome 3BS. The development of this molecular marker enables plant breeders with a useful tool to select one of the most vital wheat resistance genes (Mago *et al*., 2011).
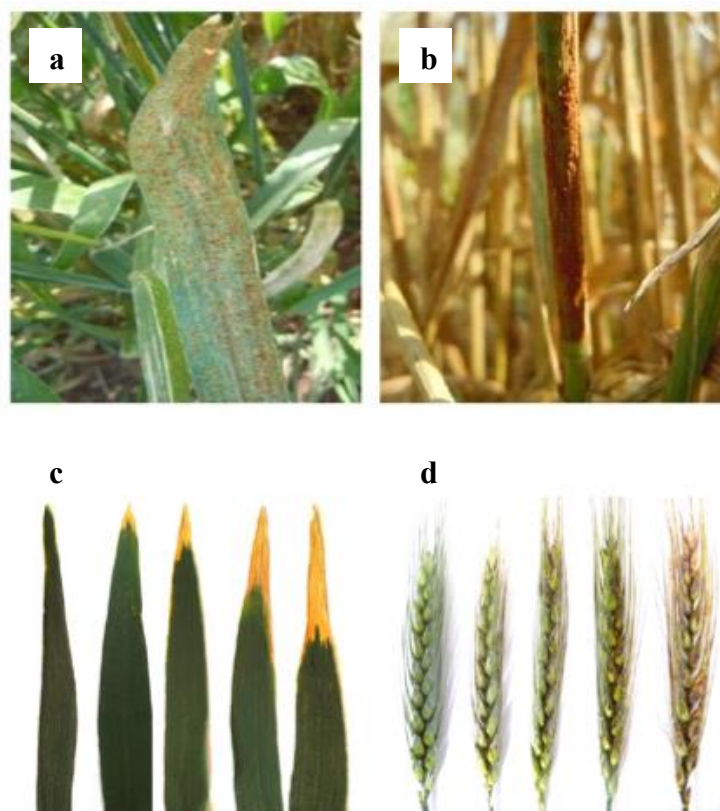
Figure 2.6: Pictures of an example of the rust diseases and associated phenotypes. a) Leaf rust, b) Stem rust (Photos by W.C. Botes), c) Leaf tip necrosis (*LTN*) and d) Pseudo black chaff (*PBC*) (Juliana *et al.*, 2015).

21

2.2.3.2. Agronomic improvement for yield progress

Since the challenge to breed at a similar rate as the growing demand for wheat is high, the increase in productivity through genetic gains need to be combined with good agronomy for the potential yield gap to be minimal (Lopes *et al*., 2014). Therefore, when better agronomic and physiological traits as well as disease resistance is combined into one cultivar, through genotyping and phenotyping, it will increase the genetic progress in wheat yield (Lopes *et al*., 2012). Systematic progress has been made with trait selection in breeding through the improvement of agronomic traits such as flowering time, height, resistance to various common diseases, end-use quality parameters and expression of yield at multiple locations (Lopes *et al*., 2014). However, in many countries yield gains are mainly achieved through further progress in genetic improvement, but in some countries agronomic improvement is still needed. Although, the necessity of genetic improvement could likely decrease if agronomic practices are able to provide sustainable control that is long-term (Hawkesford *et al*., 2013).

2.3.   Integrated biotechnological approach in breeding programmes

2.3.1.   Genetic markers in wheat breeding

The use of genomic tools benefits plant breeding through many ways which include more effective identification, quantification and characterisation of genetic variation from the available germplasm. As well as cloning, tagging, and the introgression of genes and/or QTLs that will be useful for enhancement of the targeted trait with the use of molecular marker technologies and genetic transformation. Another benefit is also the manipulation of genetic variation through pyramiding, selecting, differentiating and integrating within breeding populations (Xu & Crouch, 2008).

In the early 1980s, the development of molecular markers for applicational use became popularised in plant breeding when the use of isozyme markers sped up the process of introgressing monogenic traits obtained from exotic germplasm into the background of a cultivar. A few years later, the use of restriction fragment length polymorphism (RFLP) markers in crop improvement was described by Beckmann and Soller (1986) along with theoretical issues linked to marker-assisted backcrossing (MABC) for improving qualitative traits. Soon after, Lande and Thompson (1990) founded the theoretical studies of MAS for quantitative traits. This started the publication of a series of simulation studies throughout the

1990s. Later, further theoretical considerations with regards to the application of MAS was addressed as well as optimising the MABC systems and pyramiding desired alleles through the use of recurrent crossing schemes. The contribution of these theoretical studies has greatly added to our understanding of several fundamental genetical issues when considering MAS system development which included sample size, population type, genome size and marker number (Xu & Crouch, 2008).

2.3.2.   The use of genetic markers as a selection tool

The definition of genetic markers is simply the consideration of genetic markers as identifiers of specific aspects within the genome, for example nucleotide or short DNA sequences. It describes observed variation possibly due to a mutation or alterations that arose at the genomic loci. It is also possible to follow genetic markers through inheritance from generation to generation. The advantages of genetic markers include; the lack of influence from the environment, limitless in number, easily analysed and cost efficient as compared to other marker types (Dreisigacker, 2012).

Markers are used as an indirect selection tool as soon as it is identified through association or linkage mapping analysis, however it must first be validated by appropriate breeding populations. The value of a marker is dependent on its success in its ability to be integrated and easily applied in a large-scale breeding programme. Thus, the approach of MAS is widely utilised in wheat breeding programmes; but breeding strategies utilised is dependent on the objectives of the breeding programme, the available resources and the information provided from the genetic characterisation of the various traits (Dreisigacker, 2012).

For genetic markers to be used in MAS, five key considerations should be met, i.e. reliability, quality and quantity of required DNA, marker assay technical procedure, polymorphism level and cost. Reliable markers should be tightly linked to the loci of interest with a preferable genetic distance of less than 5cM. Flanking markers may be used as it increases the reliability of the markers for phenotype prediction. The quality and quantity of DNA is important as large quantities with high quality DNA are required for some marker techniques which may be disadvantageous as sometimes obtaining DNA may be difficult and add to procedure costs. The simplicity level and required time for technical procedures is a critical consideration as it is highly desirable for methods that are high-throughput, simple and quick. The level of

polymorphism is important as breeding material should be highly polymorphic because different genotypes should discriminate from each other, especially within core breeding material (Collard & Mackill, 2008).

2.3.3. Types of genetic markers used within breeding

Over the decades, the identification of chromosomal regions that carry important genes has been identified through molecular mapping with the use of SSR, AFLP, RAPD, RFLP, and DArT markers (Liu *et al*., 2012). However, types of molecular markers affect the simplicity and cost of implementation, for instance the implementation of microsatellites are easier as compared to diversity arrays technology (DArTs) and amplified fragment length polymorphism (AFLPs). However, one of the most responsive molecular markers with very high throughput are single nucleotide polymorphisms (SNPs) (Reynolds *et al*., 2012).

Molecular markers can also be classified as dominant markers such as microsatellites or co-dominant markers such as AFLPs or RAPDs. When comparing co-dominant markers to dominant markers, it was identified that dominant markers can easily be developed at a low cost. Thus, co-dominant markers may be represented by dominant markers as excellent alternative tools when addressing the questions that require the estimation of individual relatedness (Hardy, 2003). Molecular markers can also serve as a means to determine the genetic constitution of a set of individuals by specifying their genotypes. The genotype of an individual can be referred to as particular alleles that are present at all loci that affects the trait being investigated, where an allele is defined as the alternative forms of a gene (Hartl & Clark, 1998). Thus, genotypic frequencies and allelic frequencies can be determined when molecularly screening specific traits, which can be used for prediction of future expected genotypic and allelic frequencies.

The DArTs markers are used to detect single base changes which also includes insertions and deletions within a genome without reliance on the sequence information. The DArT markers has a high multiplexing level and the marker systems have an array based nature that ensures high-throughput and low analysis costs (Dreisigacker, 2012). Therefore, DArT markers are displaying potential as a high-throughput marker that is inexpensive and offers the opportunity to map the whole genome to be used by breeders (Reynolds *et al*., 2012).

Microsatellites or simple sequence repeats (SSRs) are tandem repeats that are dispersed throughout the genome. The monomer sequence lengths for tandem repeats range from 1 to 6bps that repeats several times. These microsatellites are amplified by PCR with the use of flanking primers of the regions. The use of microsatellite markers has previously set a good platform for the implementation of QTL mapping and then the use of MAS within breeding programmes (Dreisigacker, 2012). These markers are highly reliable, relatively cheap and simple to use, co-dominant in inheritance and normally highly polymorphic. A disadvantage of these markers, however, is the typical polyacrylamide gel electrophoresis requirement and the fact that a single assay gives information on only one locus. This has been overcome through selecting SSR markers with size differences that is large enough to be able to make use of multiplexing many markers within a single reaction (Collard & Mackill, 2008).

Markers such as single nucleotide polymorphism (SNP), sequence tagged site (STS) and sequence characterised amplified region (SCAR) are very useful markers for MAS as they are developed from specific DNA sequences of markers such as restriction fragment length polymorphism (RFLPs) which are linked to a QTL or gene (Collard & Mackill, 2008). The developed SCAR markers may be co-dominant or dominant markers. In literature SCAR markers has even been referred to as STS markers and in accordance to Rafalski and Tingey, 1993, the more appropriate acronym should be STARs, which is abbreviated for sequence-tagged amplified regions (Gupta *et al.*, 1999).

The occurrence of a single base change within the sequence is called a single nucleotide polymorphism (SNP) which is among the most abundant source of variation within animal and plant genomes (Xu & Crouch, 2008). The most applicable markers for high-throughput screening is SNPs but only once the associations between genotype and phenotype is determined (Liu *et al.*, 2012). SNP markers possess many advantages as compared to previous generation of markers which include the increased probability of identifying a marker within the desired gene because of the high density of SNPs that can be found across the genome. It should be noted that not all will display polymorphism in any given breeding population, but the probability of at least on SNP displaying polymorphism nearby and in the target gene increases with higher density of these markers. This offers a huge advantage within MAS programmes over markers previously used as they were closely linked but not within the desired loci which could result in the linkage being lost when applying the marker to various populations that have different recombination patterns (Xu & Crouch, 2008).

2.4.   Improvement of wheat yield

2.4.1.   Importance of improving yield in wheat

The main breeding objective in wheat is achieving and maintaining high yield as the human population is increasing with a decrease in cultivated land (Jiang *et al*., 2014). During the second half of the 20th century, the breeding of grain crops with the approach of selecting for yield has been successful. The annual genetic yield gain for cereals currently fall within the range of 0.5%–1%, almost entirely due to conventional breeding (Reynolds & Langridge, 2016). However, the rate of genetic gains has been decreasing over time which is undesirable as there is an urgency for the identification, understanding and incorporation of alleles that result in increased wheat yield potential across various environments to prevent global food insecurity (Pedro *et al*., 2012; Simmonds *et al*., 2016).

A trait that holds great importance for improving wheat is grain yield. Grain yield is the end result of various processes and is therefore determined directly through yield component traits, for instance spike number per plant, grain number per spike, density and other yield-determining traits such as the number of spikelets per spike, the number of fertile spikelets per spike, spike length, spikelet density per spike and plant height, which indirectly affect grain yield. Yield and yield-determining traits are complex quantitative traits that are controlled by various genes and are largely influenced by the genotype x environment interactions, resulting in low heritability especially for grain yield (McIntyre *et al*., 2010; Wu *et al.*, 2012). Grain yield could display heritability values that range between 0.40-0.70 (McIntyre *et al*., 2010) However, lower heritability values are possible as reported by Wu *et al.* (2012) where the heritability of yield was estimated at 0.272. Some yield-determining traits are less influenced by the environment and have higher heritability as compared to grain yield as such naley height and grain weight that display heritability values that are greater than 0.70 (McIntyre *et al*., 2010). Thus, examining yield-determining traits are useful for the evaluation of yield in order to gather specific information with regards to the genetic control and the relationship that exists between yield and yield-determining traits which could be vital for sustainable wheat improvement (Wu *et al*., 2012).

In order to gain better understanding of yield, many studies have used QTL mapping to dissect the genetic information of yield among different environments for hexaploid as well as tetraploid wheat (Mason *et al*., 2013). For example, a QTL located on chromosome 4A was identified by Kirigwi *et al.* (2007) and shown to be associated with an increase in grain yield

under drought conditions. Two major QTLs on chromosome 3B was also reported by Bennett *et al.* (2012) that is associated with yield and related physiological traits (Mason *et al*., 2013). Huang *et al.* (2006), Marza *et al.* (2006), Snape *et al.* (2007) and Rebetzke *et al.* (2008) are some of the studies that have also previously reported on QTLs for plant height, anthesis, grain number, grain weight and grain yield (McIntyre *et al*., 2010).

2.4.2.   Physiological breeding to increase wheat yield potential

Physiological trait breeding focuses its efforts on combining the traits associated with the three main genetic drives that determine yield, i.e. light interception, radiation use efficiency and total assimilate partitioning (Cossani & Reynolds, 2012). As the global demand for wheat increases, the current genetic gains need to meet this demand and therefore the use of physiological traits as an indirect selection criteria in wheat could cause the acceleration of breeding needed (García *et al*., 2013). Since the proposal of indirect selection, various efforts have been made in understanding the association between improvements in grain yield and the physiological trait changes (Garcia *et al*., 2014). Several speculations have also been made that a better understanding of the yield physiology of wheat would be able to contribute to regaining improved rates of genetic gains in yield (Pedro *et al*., 2012).

For the use of physiological criteria within a breeding programme, firstly the traits need to be identified that regulate grain yield followed by detecting the genetic factors that control them so that their manipulation can be made easier and/or prediction of their selection response. Physiological breeding is not only useful for improved understanding of relationships concerning grain yield and the traits determining it, but also improving the existing knowledge about the relationship between genotype and phenotype (Garcia *et al*., 2014). Therefore, the incorporation of physiological traits as additional criteria is necessary to accelerate genetic improvements in yield in the future.

There are various physiological traits which can be considered as genetic resources and selectively included in breeding programmes (Reynolds *et al*., 2012). Therefore, strategic trait-based crossing makes use of complementary physiological traits within a new generation of offspring and when comparing these offspring to advanced conventionally bred lines, it displays higher expression of yield and physiological traits (Lopes *et al.,* 2012). QTLs are rarely reported for physiological traits besides the few association studies for soil-plant analysis

development (SPAD) value of canopy temperature (CT), chlorophyll content (Chl) and normalised difference in vegetation index (NDVI) in spring wheat (Gao *et al.*, 2015). The study by Gao *et al.* (2015) identified QTLs related to physiological traits associated with yield, but only three stable QTLs were identified for chlorophyll content and NDVI that can be utilised in breeding. The three QTLs; *QNDVI-10.caas-5BL*, *QChl-10.caas-5AL* and *QChl-10.caas-5BL*, is located on chromosome 5A and 5B. The phenotypic variation explained by these QTLs ranged between 6.8 - 14.2% for their respective trait (Gao *et al.*, 2015).

2.4.3. The structure of the wheat grain

The structure of a wheat grain (figure 2.7) consists of 2-3% embryo, 80-85% endosperm and 13-17% outer layers that includes the pericarp layer. The embryo is found at the end of the grain and consists of 25% proteins and 8-13% lipids. The endosperm which is enclosed by the pericarp and the seed coat, differentiates into an outer aleurone layer and starchy endosperm cells in the inner columns. The endosperm is important as it contains the food reserves needed during seedling growth. The outer layers also known as the bran is comprised of many layers that protects the grain. However, during the milling of wheat grains, the aleurone, pericarp and embryo is removed, with the starchy endosperm only remaining as the primary contributor to the milled grain producing white flour (DuPont & Altenbach, 2003; Sramkova *et al.*, 2009). Wheat grains are mostly comprised of approximately 65-80% of starch and approximately 10-15% of proteins. Starch contains amylose and amylopectin that are synthesised to form distinct starch granules within the amyloplasts. These starch granules are the main storage organelle for carbohydrates. Quantitative trait loci (QTLs) has been previously identified, by Igrejas *et al.* (2002), that share close relation to the development of starch granule and is found on chromosome 1BL, 1BS and 7AS (Cao *et al.*, 2015).
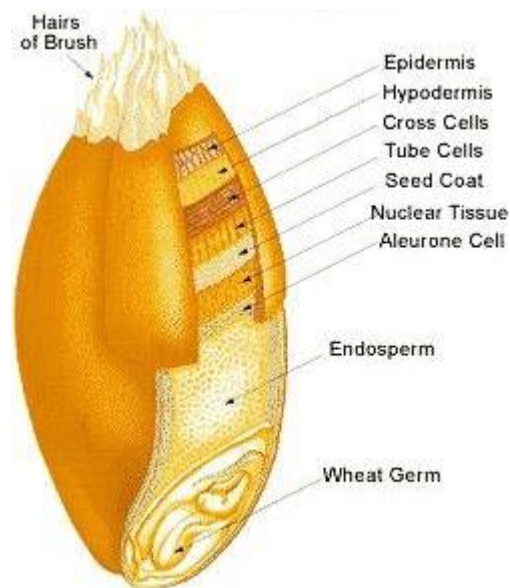
Figure 2.7: The structure of a wheat grain (NZFMA, 2017).

2.4.4. The wheat grain parameters influencing wheat quality

2.4.4.1. Grain hardness

Wheat grain quality is influenced by physical characteristics such as grain colour, weight, shape, vitreousness and hardness. Grain texture can be considered as the degree of softness and hardness of a grain and this also forms as an integral characteristic for the milling and baking quality of wheat. The hardness of a wheat grain can be classified into five levels, namely; soft, medium soft, medium hard, hard, and extra hard. This classification is essential for the growers, bakers and millers as to determine intended end-use. Soft wheat grains result in fine flour with a considerable number of intact starch granules which is good for making pastries, cakes and cookies. Hard wheat grains results in flour that has a coarser texture which is more fit to use for making bread-based products (Mikulikova, 2007; Pasha *et al.*, 2010). Grain hardness has been identified to be controlled by the *Puroindoline* (*Pin*) genes on chromosome 5DS found at the genetic locus (*Ha*) that controls hardness. The soft grain phenotype develops with the wild-type alleles, *Pina* (*Pina-D1a*) and *Pinb* (*Pinb-D1a*); however, a mutation or deletion occurring at either or both of the *Pin* genes will result in the development of a hard grain phenotype. The most frequently observed mutations found for the hard grain phenotype was the *Pina-D1b* and *Pinb-D1b* alleles (Chichti *et al.*, 2015; Nirmal *et al.*, 2016). Grain hardness can be measured indirectly as well by other parameters such as grain density and vitreousness as they influence grain hardness (Pasha *et al.*, 2010).

29

2.4.4.2. Specific weight

Milling yield, even agronomic yield is partially determined by the physical factors; grain weight, grain size and specific weight (Kunert *et al*., 2007). Specific weight, also termed as hectolitre or test weight, is the standard volume weight of grain and generally thought of as the bulk density measurement of a grain (Manley *et al.,* 2009). Since specific weight is a rough index for grain plumpness and soundness, it can serve as is an indication of the yield potential of flour. As flour yield and specific weight is associated; when low specific weight is observed, low flour yield is also observed. Additionally, low specific weight can be an indication of grains that are damaged such as shrunken, immature, frozen, diseased and sprouted grains, which is unfavourable for wheat quality (Hou, 2011). Many studies have reported specific weight as a wheat quality indicator which was identified by Marshall *et al.* (1986) to have strong positive correlations with milling yield (Kunert *et al*., 2007). Specific weight has been reported to be influenced by grain density, uniformity, shape and packing efficiency (Sun *et al*., 2010). Specific weight is also believed to be related to grain size or shape as the way a grain packs is determined by these parameters (Gegas *et al*., 2010). Identification of seven QTLs associated with specific weight were reported by Kumar *et al.* (2016), which were located on chromosomes 3A, 4A, 4B, 2D and 7D. These QTLs explained between 3.91 - 21.22% of the phenotypic variation for specific weight. Where two QTLs, *QTW.ndsu.4A* and *QTW.ndsu.4B.2*, were identified as major contributors to specific weight, with 10.80% and 21.22% phenotypic variation explained, respectively (Kumar *et al.*, 2016). Boehm *et al.* (2018) has also identified a major QTL associated with specific weight; *QTw.wql-5AL* is located on chromosome 5AL, with 15% of phenotypic variation explained (Boehm *et al.*, 2018).

2.4.4.3. The wheat grain morphology traits

The grain morphology traits grain weight (Nyiraneza *et al*., 2012), size (Gegas *et al*., 2010), length and width (Ramya *et al.*, 2010) have been reported to be highly important traits associated with the quality of wheat. During wheat domestication, grain shape was not considered as a main component to select for, as compared to other cereal crops. For instance, the domestication process of rice involved strong selection with regards to grain shape and size. However, due to market and industry requirements, grain shape and size in wheat became a target characteristic in breeding programmes. The grain market value of wheat is determined by the grain density, uniformity, shape and size as these are important attributes that influence

the flour quality and yield, i.e. milling performance (Gegas *et al*., 2010). Therefore, selection of larger grains resulted in significantly higher flour yield as compared to smaller grains (Ramya *et al.*, 2010). The optimum morphology of a grain is large in size and spherical in shape (Gegas *et al*., 2010). The size of a grain can be described by the seed length, width and area; as well as TKW (Guo *et al.*, 2017). Grain length and width have been tested to be cheap and non-invasive predictors of wheat milling quality (Ramya *et al.*, 2010). These traits are also strongly associated with grain yield and therefore the selection of these traits are beneficial (Maphosa *et al*., 2014). Grain weight is largely influenced by grain dimensions such as grain length and width. Therefore, it is important to gather information on the underlying genetic control of these two grain dimensions (Cui *et al.*, 2014). The genetic control for grain length and width, however, seems to differ from one another and is therefore under independent genetic control (Maphosa *et al*., 2014; Simmonds *et al*., 2014).

During the past decade, QTLs have been mapped for grain weight and length by many scientists. There have been many reports on numerous QTLs for grain length on chromosomes 1A, 1B, 2B, 2DL, 3A, 3B, 4A, 4B, 5A, 5B, 6A, 6BS, 6D, and 7BL (Zhang *et al*., 2014b). Kumar *et al.* (2016) mapped 10 QTLs located on eight different chromosomes. The four QTLs; *QKL.ndsu.3A.2, QKL.ndsu.4B, QKL.ndsu.5B.1* and *QKL.ndsu.7A,* were considered as major QTLs as more than 10% of the phenotypic variation was explained for grain length (Kumar *et al*., 2016). Some high-yielding genes are also associated with grain length, these include the *TaGW2-6A* gene, the *TaGS-D1* gene and the *TaGASR7-A1* gene (Hu *et al.,* 2016; Zhang *et al.*, 2014b). QTLs for grain width has been found to exist on chromosomes 2A, 2B, 4B, 4D and 6D (Williams *et al*., 2013). Nine QTLs located on eight different chromosomes were mapped by Kumar *et al.* (2016) as well. These QTLs explained between 3.09 – 19.30% of the phenotypic variation for grain width. Major QTLs; *QKW.ndsu.3B, QKW.ndsu.4A, QKW.ndsu.5A* and *QKW.ndsu.7A.1,* for grain width were identified and located on chromosomes 3B, 4A, 5A and 7A (Kumar *et al.*, 2016). Grain width has also been associated with high-yielding genes such as the *TaGW2-6A/B* gene, *TaTKW-7A* gene and *TaGS5-3A* gene (Ma *et al.,* 2016; Hu *et al.*, 2016; Yang *et al.,* 2012). Thousand kernel weight is another parameter that can be used as a wheat quality indicator as measuring the mass of a grain can provide an indication of flour extraction potential, but TKW is influenced by cultivar, year, site and grain content moisture (Nyiraneza *et al*., 2012). TKW can also be used to estimate the agronomic wheat yield as it is a yield-determining trait (Ramya *et al.*, 2010).

2.4.5.   Yield-determining traits

The ultimate expression of many different physiological processes is grain yield (Pietragalla & Pask, 2012). Grain yield is polygenically regulated and the environment has a strong effect on this trait, therefore overall grain yield can be divided into various components (Simmonds *et al*., 2016). The components that influence grain yield are the following traits: spike number (McIntyre *et al*., 2010), spike length (Wu *et al*., 2012), tiller number (Blum, 2005), plant height (Wu *et al*., 2012), grain number and weight (McIntyre *et al*., 2010), thousand kernel weight (Simmonds *et al*., 2016), grain size (McIntyre *et al*., 2010), density (Girma *et al*., 2006) as well as physical parameters such as grain width, length and area (Simmonds *et al*., 2016). Therefore, selecting these traits in breeding programmes are important to increase grain yield (Pietragalla & Pask, 2012).

2.4.5.1. Grain number and weight

Grain yield is commonly considered to be the product of grain number and final grain weight. The advantage of these two components is to some extent temporally separated and can easily be determined (Fischer, 2011). Identification of traits that could be used for the improvement of crop yield was previously based mainly on traits that were determined at the maturity growth stage of wheat (Pedro *et al*., 2012). However, it is now known that wheat yield is mainly determined at the anthesis growth stage; as grain number is mainly determined by the spike dry weight at the anthesis growth stage. Wheat yield has also been reported to be closely related to grain number, even more so than the average grain weight of wheat (Pedro *et al*., 2012). Many studies have indicated that an increase in grain yield has often been associated with an increase in grain number, hence there exists a high positive correlation between grain yield and grain number (Fischer, 2011; McIntyre *et al*., 2010; Pedro *et al*., 2012; Reynolds *et al*., 2009b).

The relationship between grain number and grain weight of wheat has been identified to have a significantly negative correlation (Garcia *et al*., 2013). This is due to an increased grain number that would be associated with additional grains per spikelet that have lower potential grain weight. It was identified that as grain weight increases, the grain number decreases thus the potential yield decreases as well (Fischer, 2011). Therefore, future breeding should consider rather maintaining grain weight than attempting to increase it. However, the CIMMYT-derived mapping population (Seri-Babax) seemed to have weakened the relationship that grain number and grain weight share as some recombinant inbred lines (RILs) were produced that combined

high grain number and grain weight; and still managed to result in higher yield as compared to the parents and best local checks (Fischer, 2011).

Thousand kernel weight is commonly used to express grain weight and is a vital agronomic trait as the improvement in wheat yield is also dependent on this trait, along with grain number (Hou *et al.*, 2014; Yue *et al.*, 2015). Due to this, one of the leading targets in wheat breeding is increasing TKW, as this trait is among the most stable components of grain yield. However, TKW is a trait that is quantitatively controlled and thus environmental conditions significantly influence this trait (Hu *et al.*, 2016). Kumar *et al.* (2016) identified 10 QTLs for TKW located on eight different chromosomes, with values ranging between 3.34 – 17.65% of phenotypic variation that explained the trait. The major QTLs identified was *QTKW.ndsu.5A.1* and *QKW.ndsu.5A.2* located on chromosome 5A (Kumar *et al.*, 2016).

Genes are also continuously being identified that contribute to this trait, for use in molecular selection throughout the wheat improvement process (Yue *et al.*, 2015). Ma *et al.* (2016) has cloned and characterised the *TaGS5* orthologs which is located on group 3 homoeologous chromosomes (refer to table 2.1). It was identified that *TaGS5-3A* contained two alleles, with the *TaGS5-3A-T* as the ideal allele which is associated with increased grain weight and size. To distinguish between the two alleles, a functional molecular marker was developed for the use of marker-assisted selection (MAS). A CAPS marker, *TaGS5-3A-CAPSF/R*, was developed which amplified an 863bp band associated with the desired *TaGS5-3A-T* allele and when cleaved, amplified a 718bp and 145bp band associated with the *TaGS5-3A-G* allele (Ma *et al.*, 2016). Qin *et al.* (2014) also cloned a gene associated with grain weight and width, located on chromosome 6B of wheat. The *TaGW2-6B* gene was identified to have four haplotypes formed from 11 SNPs in the upstream sequence of the gene. The four haplotypes were designated as *Hap-6B-1*, *Hap-6B-2*, *Hap-6B-3* and *Hap-6B-4* (refer to table 2.1). When the results were compared between the haplotypes, the favoured haplotype for increased grain weight was *Hap-6B-1*. This haplotype was distinguished from the rest with a developed CAPS marker, where there was no restriction enzyme recognition site within *Hap-6B-1* but existed within the rest of the haplotypes (Qin *et al.*, 2014). Another study by Zhang *et al.* (2014b) identified a gene, *TaGS-D1*, associated with lower and higher TKW (refer to table 2.1). No diversity was revealed in the coding sequence of exons within the *TaGS-D1* locus, however in the first intron a SNP was found and within the second intron, a 3-bp and 40-bp InDel. On the basis of the 40-bp InDel, a co-dominant STS marker was developed, to discriminate between two alleles, this marker was designated as *GS7D*. The two alleles were designated as *TaGS-*

D1a and *TaGS-D1b* which represented the 40-bp insertion and deletion, respectively. The *GS7D* marker amplified a 562-bp or 522-bp fragment which was generated for the *TaGS-D1a* or *TaGS-D1b* allele, respectively (Zhang *et al*, 2014b).

2.4.5.2. Tiller number

Since the green revolution introduced the semi-dwarf genes into wheat, these genes allowed more effective tiller development per plant, which increased biomass and led to a dramatic shift in yield potential (Kumar *et al.*, 2015). Tiller development is a yield component essential in determining plant architecture in wheat (Spielmeyer & Richards, 2004). Tillers can be defined as primary or secondary, where primary tillers form the main stem, and secondary tillers are formed from the main stem. Tiller numbers per plant is dependent on tiller appearance and tiller survival, therefore when tillers are fertile; it contributes significantly to grain yield (Kumar *et al*., 2015; Spielmeyer & Richards, 2004).

The identification of a molecular marker for tiller number was identified by Spielmeyer and Richards, (2004) (refer to table 2.1). It was identified that the tiller inhibition gene (*tin*) which reduces the tiller number of wheat was associated with an increased harvest index, a reduction in sterile tillers and an increase in grain size. Due to the agronomic potential of this gene a microsatellite marker, *Xgwm136*, was identified to be tightly linked with the *tin* gene and was mapped on the chromosome 1AS. Thus, this molecular marker was suggested to be used for marker-assisted selection (MAS) for the identification of low-tillering lines. To differentiate between high and low-tillering lines, the *Xgwm136* marker detects an approximate 300bp band which is associated to high tillering and an approximate 350bp band which is associated to low tillering (Kumar *et al*., 2015; Spielmeyer & Richards, 2004).

2.4.5.3. Plant height

Plant height is an important trait in wheat as the green revolution was based on reduced plant height in order to increase yield. Wheat plants that are tall did not have strong enough stems to support the heavy grains and therefore lodging (falling over) occurred. However, with the introduction of the semi-dwarf genes from the cultivar 'Norin 10', plant height was reduced and wheat did not lodge, preventing the reduction in yield losses that was previously incurred

34

(Hedden, 2003). 'Norin 10' carried the two dwarfing alleles, *Rht-B1b* and *Rht-D1b*, and was bred by the Japanese early in the 20th century, but little importance was given to this cultivar in Japan. Norman Borlaug was sent one cross from which he bred cultivars that was adapted to grow within tropical and sub-tropical climates. The progeny was distributed to various countries and wheat yields increased spectacularly (Hedden, 2003). The reduced plant height of the semi-dwarf cultivars resulted in an improved harvest index as well. The harvest index for cultivars were 0.3 due to significant yield losses that resulted from cultivars that grew excessively tall and fell over. However, the harvest index for semi-dwarf cultivars was approximately 0.5 due to their short stature conferring lodging resistance (Sakamoto & Matsuoka, 2004). Therefore, these genes that confer reduced plant height are still commonly deployed globally (Langridge, 2014).

2.4.5.4. Days to Heading

The days to heading trait is among the most important traits that is measured in cereal crops as cultivars that have suitable heading time according to the life cycle duration of the crop and the target environment, will contribute to maximising wheat yield potential within different environments (Kiseleva *et al*., 2016). There are three genetic systems that control the developmental and growth stages of wheat, namely; earliness *per se*, vernalisation and photoperiod response. Flowering time is mainly determined by photoperiod response, which can confer photoperiod sensitivity or insensitivity. A photoperiod sensitive wheat crop will continue in the vegetative phase until its photoperiod requirement is satisfied, i.e. a sufficient increase in the day length. However, a photoperiod insensitive wheat crop will immediately change from the vegetative to the reproductive phase as soon as there is a rise in temperature which allows the plant to develop completely and complete grain filling before the high summer temperatures are reached in the warmer environments (Beales *et al*., 2007; Iqbal *et al*., 2011).

Three major genes that have been identified to control photoperiod response are *Ppd-D1, Ppd-B1* and *Ppd-A1* (Seki *et al*., 2013). The *Ppd-D1* gene is a *pseudo-response regulato*r gene family member and is the major locus that confers photoperiod insensitivity as well as having minor effects on plant height in wheat and is located on chromosome 2D. A functional molecular marker for *Ppd-D1* (refer to table 2.1) has been developed that differentiates *Ppd-D1a* and *Ppd-D1b* (Beales *et al*., 2007, Kiseleva *et al*., 2016; Zhao *et al*., 2014). The molecular

marker amplifies a 288bp band associated with the *Ppd-D1a* allele and a 414bp band associated with the *Ppd-D1b* allele. The semi-dominant *Ppd-D1a* allele is the main source of photoperiod insensitivity and its presence will result in reduced number of days to heading contributing significantly to grain yield potential (Beales *et al*., 2007; Kiseleva *et al*., 2016).

2.4.5.5. Spike Length and spikelet number

A trait that holds considerable importance for increasing yield is spike length, as the larger a spike is, the more grains will be produced and thus, resulting in higher yield (Laghari *et al*., 2010). The spike structure possesses benefits such as the ability to stay green and functional for an extended amount of time which also includes awns (Sharma *et al*., 2003). As a result of these features, spike structure contributes between 20-30% on average to the accumulated dry matter in the kernel. (Sharma *et al*., 2003). Since larger spikes result in more florets and spikelets, another route for improved yield potential is breeding for large spikes that will produce an increased number of kernels (Wu *et al*., 2006).

Another component of grain yield potential is spikelet number per spike which is important as it is associated with increased number of grains produced per spike (Lewis *et al*., 2008). The spikelet number has been reported to be controlled by an earliness *per se* (*Eps*) gene, *Eps-A^{m}1*. This gene is located on the distal region of chromosome 1A^{m}L of *T. monococcum* and is also involved with the determination of grain number per spike. However, the relationship that spikelet number and grain yield share is complicated due to wheat being able to adjust the various yield components to resources that are available (Lewis *et al*., 2008; Zikhali & Griffiths, 2015).

2.4.5.6. Harvest Index

The simplest form to express yield potential is through the function of light intercepted (LI), radiation use efficiency (RUE), and harvest index (HI); i.e. YP = LI x RUE x HI. The improvement of wheat yield is greatly due to the increase of partitioning of the total above-ground biomass to grain, which is due the green revolution as well as continuing to select for yield during post-green revolution (Reynolds *et al*., 2009b).

The theoretical limit for harvest index was proposed by Foulkes *et al.* (2009) to be approximately 0.60, but harvest index values typically fall between 0.4 and 0.5 (Sadras & Lawson, 2011; Valluru *et al.,* 2015). However, this leaves scope for improvement and a 0.56 harvest index value for winter wheat has already been achieved by Foulkes *et al.* (2009) (Sadras & Lawson, 2011). Modern maize and rice cultivars are approaching a harvest index of 0.55 (Fischer, 2011; Jiang *et al.,* 2017). It has been shown in different studies that the harvest index of high-yielding cultivars has had little progress since the nineties and this suggests that this trait has been optimised in breeding, but under stress, the stability of the trait remains an issue (Garcia *et al.,* 2013). However, since harvest index has been exploited in breeding, the increase of improved grain yield potential might be more challenging to achieve (Rose *et al.*, 2017).

2.4.5.7. Flower fertility

Wheat improvement is dependent on crop wild relatives that possess desired agronomical traits such as broad adaptability, disease resistance, improved quality as well as superior floret number. As these traits are desirable for wheat improvement, transferring the useful genes from the crop wild relatives is an efficient method for cultivar development; for instance, the desired traits transferred from *A. cristatum* accession Z559. A Japanese cultivar was crossed with *A. cristatum* accession Z559 and several progeny lines displayed traits such as large spikes with multiple grains and florets (Wu *et al.*, 2006).

Wheat has an indeterminate nature that enables the possible formation of eight florets within a single spikelet. Final grain yield is influenced by the final number of grains per spikelet at the physiological maturity growth stage; but before the set final grain number, the floral structure has to undergo a sophisticated development and abortion process. This occurs after the maximum floret primordia number is reached, which is a representation of yield potential. The initiation of a floral degradation process determines the florets fertility at anthesis. During post-anthesis events, one to three florets are lost until reaching the final grain number at physiological maturity growth stage (Guo & Schnurbusch, 2015).

Table 2.1: Table of Molecular Markers associated with yield-determining traits

Trait	Gene/Locus	Chromosome	Marker	Enzyme	Fragment Size (bp)	Closely related Traits	Reference
Thousand Kernel Weight and Grain Width	TaGW2	6B	TaGW2-6B-CAPS-F	BstNI	Hap-6B-1 : 1361		Qin et al., 2014
			TaGW2-6B-CAPS-R		Hap-6B-3/4 : 464		
			TaGW2-6B-ACAS-1F	None		Earlier maturity (Hap-6B-1 & Hap-6B-2)	
			TaGW2-6B-ACAS-1R				
			TaGW2-6B-ACAS-2F		Hap-6B-1/2 : 626		
			TaGW2-6B-ACAS-2R				
			TaGW2-6B-Dcaps-F	Hpy166II	Hap-6B-3 : 263		
			TaGW2-6B-Dcaps-R		Hap-6B-4 : 240		
Thousand Kernel Weight and Larger Kernel	TaGS5	3A	TaGS5-3A-CAPS-F	Fnu4HI	TaGS5-3A-T : 863	Kernel width, length and thickness	Ma et al., 2016
			TaGS5-3A-CAPS-R		TaGS5-3A-G : 718, 145		
Tiller Number	tin1	1A	Xgwm 136-F	None	300 (tin1 -), 350 (tin1 +)	Larger grain size	Kumar et al., 2015
			Xgwm 136-R				
Days to Heading	Ppd-D1	2D	Ppd-D1-1F	None	Ppd-D1a - 288, Ppd-D1b - 414	Grain Filling	Wilhelm et al., 2013 Beales et al., 2007
			Ppd-D1-1R				
			Ppd-D1-2R				
Thousand Grain Weight	TaGS-D1	7D	GS7D-F	None	TaGS-D1a - 562	Grain Length	Zhang et al., 2014b
			GS7D-R		TaGS-D1b - 522		

2.5. High throughput plant phenotyping platforms

Plant phenotyping is defined as the set of methodologies and protocols that are performed to measure plant architecture, growth and composition with efficient accuracy and precision. The phenotype of a plant possesses complex plant traits that are evaluated through measuring root morphology, leaf and fruit characteristics, photosynthetic efficiency, biomass, yield-determining traits and the response to abiotic and biotic stresses (Golbach *et al*., 2016). One important condition with regards to phenotypic analysis is the absence of damage to the plant when measuring the phenotypic trait. However, some phenotyping methods require partial damage to plant organs, but this is acceptable as the plant is still able to survive and reproduce (Afonnikov *et al*., 2016).

The traditional way of phenotyping, even a simple trait such as height, is time-consuming and laborious especially with large populations (Tanger *et al*., 2017). Therefore, the use of high-throughput plant phenotyping (HTPP) can be implemented, as it is time-efficient and allows the quantification and acquisition of plant phenotypes from populations consisting between hundreds and thousands of plants (Acosta-Gamboa *et al.,* 2017). High-throughput plant phenotyping consist of various phenotyping platforms such as; field-based (Araus & Cairns, 2014; Tanger *et al*., 2017), aerial-based (Araus & Cairns, 2014; Haghighattalab *et al*., 2016), image-based (Acosta-Gamboa *et al*., 2017) and mobile-based (Afonnikov *et al*., 2016).

The ability to measure a plant's performance in a rapid and accurate way within the field has always been a limiting factor. However, the emergence of field-based high throughput phenotyping solves this through combining the advances being made in aeronautics, remote sensing and high-performance computing to develop the field-based HTPP platform (Araus & Cairns, 2014; Tanger *et al*., 2017). As field-based HTPP is developing, several new platforms have emerged such as ground-based and aerial-based HTPP. Ground-based high throughput phenotyping includes hand-held sensors as well as equipping modified vehicles with sensors and a GPS navigational system (Araus & Cairns, 2014). An example of ground-based HTPP includes tractor-based high throughput phenotyping, which is the use of a tractor that is mounted with multispectral reflectance and ultrasonic sensors that collect measurements on various physiological traits (Tanger *et al*., 2017). However, the ground-based platform is limited by its inability to measure all the plots simultaneously, in this instance aerial-based HTPP can be a viable alternative (Araus & Cairns, 2014). One of the most efficient aerial-based HTPP used are unmanned aerial system (UAS) platforms that are cost-effective and can

rapidly characterise field trials with a large number of plots. On the UAS platform, remote sensing of crops can make use of different imaging systems with the use of various cameras such as multispectral, thermal cameras, RGB and consumer grade cameras (which captures near infrared when modified) (Haghighattalab *et al*., 2016). High-resolution cameras are typically used in HTPP platforms to capture images for quantification of various plant traits. HTPP advantages are that the methods used are non-destructive, unbiased, fast and accurate (Acosta-Gamboa *et al.,* 2017).

Manual assessments of measurements can be laborious and time consuming, therefore a solution to this can be the use of mobile devices through image-based analysis that is effective and efficient with sufficient accuracy (Komyshev *et al.,* 2017). Modern mobile devices possess high resolution digital cameras and multicore processors with adequate computational power that allows for sufficient image processing and analysis. These features are needed for the capture and process images within an application. Many applications have been developed for mobile devices to use for the morphological measurements of plant organs. Some of the applications developed includes leafsnap (Kumar *et al.*, 2012) which is an application identifying a plant species simply based on the leaf images captured as well as Leaf Doctor (Pethybridge & Nelson, 2015) which is an application estimating the disease severity on the basis of the leaf images captured. These applications, evidently shows that the use of image-based analysis can serve as an effective tool for morphological measurements (Komyshev *et al.,* 2017).

Morphological parameters such as grain length, width and area can be measured in a cost-effective and efficient way through the use of a mobile application that was developed by Komyshev *et al.* (2017) called SeedCounter. This application is distributed under the Berkley Software Distribution (BSD) license and can be freely downloaded on Google play store. It runs on the Android platform of mobile devices to perform automated calculations of morphological parameters on the grains of wheat without the use of computer facilities. Wheat grains are scattered on a paper sheet (A3, A4, A5, B4, B5, B6, Letter or Legal) and the application estimates the grain count as well as morphological parameters which include grain length, width, area and the distance between the mass geometric centre of the grain and the intersection point of its principal axes. The SeedCounter application makes use of the OpenCV library to process images. Therefore, the grains are identified as contours through the application of the findContours() function. Grain ellipsoids approximate these contours and

allows the size estimates of the minor and major principal axes that correspond to grain length and width (Komyshev *et al.,* 2017).

Another application designed to analyse seeds is the 1KK application which forms part of a new initiative by Kansas State University. This initiative is called BreadPheno, which aims at converging new advances in image processing and machine vision in order to develop transformative mobile applications that can be distributed worldwide by established breeder networks (Neilsen *et al.,* 2016). It provides cost-effective, easy-to-use phenotyping tools to more traditional farmers that are not able to adopt HTPP technologies due to greater costs and more learning curves and barriers. Therefore, several mobile applications were developed that promotes proper data management, collection and ontology integration. These phenotyping applications include the Field Book, Inventory, 1KK and Coordinate (Rife, 2016).

The 1KK application is an open source software that is released under the GNU/GPL2 license, runs on the Android platform and can be used to measure morphological parameters based on imaging. The name of the application originates from one thousand (1K) kernel weight which is one of the traits selected for in breeding programmes. The application identifies length, width and area of an object and uses the algorithm also implemented in SmartGrain (Rife, 2016). SmartGrain is high-throughput phenotyping tool developed in 2012 that makes use of image-based analysis for the determination of seed shape. Within a digital image, SmartGrain can automatically recognise seeds and detect outlines to calculate seed length, width and area as well as other parameters (Tanabata *et al*., 2012).

The values obtained for these parameters are calculated by accumulating sequential points along the perimeter of the seed and minimising or maximising the values; this is done in the order displayed by figure 2.8b. SmartGrain makes use of the OpenCV functions, where the "cvFindContour" function detects the objects perimeter which is used to calculate length and width. The length is measured by selecting the maximum distance between the perimeter points and width is measured as the longest segment perpendicular to the measured length. Area is determined by the OpenCV function "cvContourArea" which calculates the area of the seed within the perimeter (Tanabata *et al.*, 2012).
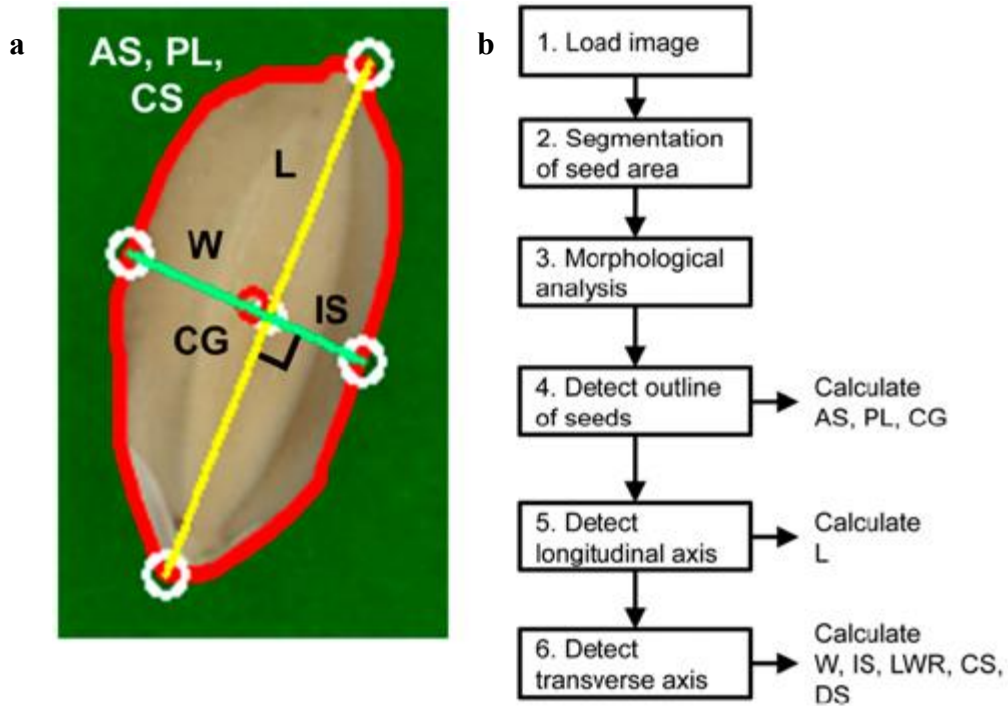
Figure 2.8: The method in which SmartGrain measures grain morphology. (a) The process in which the grain is identified on the image and (b) the process of how the measurements are calculated (Tanabata *et al.*, 2012).

Chapter 3:   Methods and Materials

The aim of this study was to assess high-yielding genotypes through validating yield-determining traits identified in literature, together with screening high yielding types genotypically and phenotypically screening the high-yielding genotypes for identified yield-determining genes and traits, respectively. The high-yielding genotypes was also used as male crossing parents in the MS-MARS facilitated pre-breeding programme to achieve grain yield improvement through the transfer of high-yielding traits into the recurrent population. The work flow of this study is illustrated in figure 3.1.

DNA extractions were performed on all the male and female plants. The female population was obtained from the existing MS-MARS pre-breeding programme in SU-PBL. Both populations were screened with rust markers that form part of the SU-PBL's standard panel of markers that was developed for genotypic screening (Wessels & Botes, 2014). Additionally, the male population was genotypically screened with molecular markers associated with yield-determining traits. These markers were identified and validated for their use in MAS in order to be added into the lab's standard panel of markers.

The recurrent population segregated into male sterile and male fertile plants from which the male sterile (female) plants was carefully selected and cross pollinated with the male plants that possess high-yielding traits, to produce F_1 seeds. The male fertile plants were allowed to self and then harvested to be planted in the field, the following season. The F_1 seeds produced in the first cycle of the MS-MARS was planted to form a new population that segregated into male sterile and male fertile plants once again. This population was also screened with the rust markers. For the second cycle of the MS-MARS, the SU-PBL 2017 nursery and a selected few wheat lines (according to phenotyping results) from the high-yielding genotypes were selected as male-crossing parents for cross pollination with the F_1 population.

Phenotypic screening was executed on the high-yielding genotypes, where the following traits were measured: tiller number, plant height, spike length, spikelet number, grain number/spike, grain weight, harvest index, thousand kernel weight, days to heading, flower fertility, grain length, grain width and area. These traits have been identified to be associated with high yield in wheat. Other measurements such as wet gluten, specific weight, moisture, protein fixed and protein dry weight basis were analysed on the NIR grain analyzer. Validation of image-based applications were also executed for future use as a phenotyping tool.

	Objective 1	Objective 2	Objective 3

2016

Identification of yield-determining traits associated with yield improvement in wheat.

Yield measurements taken by hand and on the Inframatic 9500 NIR grain analyzer on the 65 HYLD wheat genotypes.

Molecular Screening of male (donors) and recurrent (female) populations with rust resistance markers.

MS-MARS cycle 1.

2017

Identification of molecular markers associated with yield-determining traits.

Use of MAS to screen 65 HYLD wheat genotypes.

Selection of a few wheat lines from the HYLD wheat genotypes according to the statistical analysis from measurements taken.

Selected wheat lines used as male crossing parents along with SU-PBL's 2017 nursery.

Validation and execution of image-based applications.

Identification of relationship between identified molecular marker and specific yield-determining trait with regard to the 65 HYLD wheat genotypes.

Molecular Screening of F_1 recurrent (female) population.

MS-MARS cycle 2.

Figure 3.1: Work Flow diagram of this study

44

3.1.   Plant Material

The wheat population that was used in the study as the female population was derived from a vastly diverse MS-MARS pre-breeding wheat population established in 1999. This population segregates for various rust resistance genes (Marais & Botes, 2009). Seeds obtained from the MS-MARS base population carry the dominant *Ms3* gene located on chromosome 5AS (Maan *et al.,* 1987), which confers male sterility and results into a 1:1 ratio of male fertile and male sterile plants. The seeds were planted within a greenhouse located at Welgevallen Experimental Station (WES), Stellenbosch.

The wheat population that was used in the study as the male population was the 65 high-yielding genotypes whose pedigrees are withheld due to commercial purposes. These genotypes form part of a multi-year assessment being executed at SU-PBL. The male population was first planted at WES in 2015, forming the first year of phenotypic assessment and introduction into the MS-MARS scheme. In 2016, the high-yielding wheat genotypes were phenotypically screened within this study and was used as male donors for the two MS-MARS cycles executed in this study (Figure 3.1).

The male and female populations were planted within the greenhouse for both MS-MARS cycles. For the first cycle, the seeds were planted over four benches per population which resulted in 465 pots planted for the female population and 520 pots planted for the male population. For both populations, five seeds were planted in each pot, clockwise. For the second cycle, the seeds were planted over four benches per population which resulted in 420 pots planted for the female population and 480 pots planted for the male population. For both populations, six seeds were planted in each pot, clockwise.

The pots in which the seeds were planted were 3L pots consisting of a coarse sand mixture and the temperatures at planting time ranged between $10\text{-}25^\circ\text{C}$. The plants were irrigated twice daily with nutrient solution. The nutrient solution consisted of 12.5 kg Sol-u-fert T3T (by Kynoch Fertilizers Pty Ltd, Milnerton, South Africa), 150g Microplex (by Ocean Agriculture Pty Ltd, Muldersdrift, South Africa) in 100L of H_2O and 2: 6.8 kg of Calcinit was mixed in the 100L H_2O.

3.2.   DNA extraction of plant material

3.2.1.   Protocol for CTAB extraction

DNA extraction was executed according to a modified version of the CTAB-based method as described by Doyle and Doyle (1990). Leaves were cut from the plant material, then further cut up into smaller pieces and placed into a 2ml microcentrifuge tube along with three 3mm stainless steel balls. Five hundred microlitres of 2% (w/v) cetyltrimethylammoniumbromide (CTAB) [20 mM EDTA (pH 8), 1.4 M NaCl, 100mM Tris-Cl (pH 8)] extraction buffer was added to the microcentrifuge tube, which was then placed into the Qiagen® Tissue Lyser (Qiagen, Southern Cross Biotech, Claremont, RSA) for three rounds of grinding the plant material at 30 Hz per 30 seconds. The plant mixture was then incubated in the water bath at 60ºC for 15-20 minutes. When the plant mixture was removed from the water bath, five hundred microliters of chloroform:isoamylalcohol (24:1) was added into the microcentrifuge tube and inverted twice before being placed into the centrifuge to be spin down for 10 minutes at 14 000 rcf. The resulting supernatant was then transferred into a clean 1.5ml microcentrifuge tube, to which four hundred microlitres of chloroform:isoamylalcohol (24:1) was added and again inverted twice before placing into the centrifuge to be spun down for 5 minutes at 14 000 rcf (xg). The resulting supernatant was once again transferred into a clean 1.5ml microcentrifuge tube to which fifty microlitres of 3M Sodium Acetate was added along with five hundred microlitres of ice cold 100% ethanol. The microcentrifuge tube was then slowly inverted thrice for DNA precipitation to take place. The tube was then placed into the centrifuge and spun down for 2 minutes at 14 000 rcf (xg) to isolate precipitate. The supernatant was cast-off with only the pellet remaining. The pellet was then washed by adding four hundred microlitres of 70% ethanol to the microcentrifuge tube which was placed into the centrifuge to be spun down for 2 minutes at 14 000 rcf (xg), the ethanol was discarded and the pellet was washed again. After the washing step, the pellet was then air-dried and re-suspended with thirty microlitres of distilled water (dH$_2$O) and stored at -20ºC.

The quantity and quality of the extracted DNA was measured using a Nanodrop® ND-1000 spectrophotometer (Thermo Fisher Scientific Inc., Kempton Park, RSA) by pipetting 1.5 µl of DNA. The Nanodrop results for the DNA were then used to calculate the amount of DNA needed to make a dilution with a final concentration of 100ng/µl. DNA was then diluted with distilled water (dH$_2$O) and dilutions were stored at 4ºC.

46

3.3. Genotyping of plant material

All PCR reactions were performed using a 2720 Thermal Cycler (Applied Biosystems, Fairlands, RSA), or a MyCycler[TM] Thermal Cycler (Bio-rad, USA) or a TECHNE TC-5000 (Lasec, Cape Town, RSA). The primers used in this study had an initial concentration of 10μM and was obtained from Integrated DNA Technologies (Whitehead Scientific Inc, Stikland, RSA) and the KAPA Green Readymix was obtained by KapaBiosystems (distributed by Lasec SA (Pty) Ltd, Cape Town, RSA) as well as One Taq Quick-Load 2x Master Mix with Standard buffer from New England BioLabs (distributed by Inqaba Biotechnical Industries (Pty) Ltd).

All the molecular markers used within this study was dominant markers and therefore, was recorded as follows. When the band of the desired allele was present, it was recorded as 1 to represent the presence of the desired allele; and when the undesired allele was present, it was recorded as 0 to represent the absence of the desired allele. Since the markers were dominant, the allele frequency was calculated by adding all the plants that contained the desired allele and dividing it by the total number of plants.

3.3.1. Screening of the rust resistance genes

The molecular markers for rust genes *Lr34* and *Sr2* were used for screening the male (HYLD wheat lines with desired traits as male crossing parent) and female (recurrent population that segregates into 1:1 ratio male sterile and male fertile lines) populations of the MS-MARS cycles. These two molecular markers form part of the routinely used standard panel of markers within the lab (Wessels & Botes, 2014).

The PCR reaction for the *Lr34* molecular marker had a final volume of 17.8μl which contained 10μl of KAPA Green Readymix with final concentrations of: 0.34μM of *L34DINT9* forward primer and *L34PLUS* reverse primer, 0.14μM of *CSLV34* forward and reverse primer, 100ng/μl of DNA and brought to volume with dH_2O. The conditions for the PCR reaction were the following: denaturation was at 94ºC for 5 min, a duration of 35 cycles with 94ºC for 1 min, 57ºC for 1 min and 72ºC for 1 min and a final extension of 72ºC for 7 min. The PCR products were electrophoresed on a 1.5% (w/v) agarose gel stained with 4% (v/v) of Ethidium Bromide (EtBr) within a 1X TBE buffer at 100V. The positive control used for this molecular marker was the 'Chinese Spring' cultivar and the negative control used was dH_2O.

The PCR reaction for the *Sr2* molecular marker had a final volume of 12.9µl which contained 7.5µl of KAPA Green Readymix with final concentrations of: 0.35µM of *csSr2* forward and reverse primer, 100ng/µl of DNA and brought to volume with dH_2O. The conditions for the PCR reaction were the following: denaturation was at 94ºC for 5 min, a duration of 40 cycles with 92ºC for 30 secs, 60ºC for 40 secs and 72ºC for 50 secs and a final extension of 72ºC for 5 min. Five microlitres of each PCR product was firstly electrophoresed on a 1.0% (w/v) agarose gel stained with 4% (v/v) of EtBr within a 1X TBE buffer at 120V. This initial visualisation is to confirm the number of samples needed to be digested with the enzyme. All samples containing a 337bp band had to be digested. The enzyme digestion had a final volume of 2.5µl and contained 1X Buffer O, 1U/µl of *PagI* enzyme (Thermo Scientific) and brought to volume with nuclease-free water. The enzyme digestion was added to the 7.9µl PCR product remaining and incubated in the oven at 37ºC for 15 mins. The digested products were electrophoresed on a 2% (w/v) agarose gel stained with 4% (v/v) EtBr within a 1X TBE buffer at 100V. The positive control used for this molecular marker was the 'Steenbras' cultivar and the negative control used was dH_2O.

3.3.2.   Screening of the molecular markers associated with yield-determining traits

The PCR reactions for the *TaGS5-3A, TaGW2-6B* and *TaGS-D1* molecular markers each had a final volume of 10µl which contained 5µl of One Taq Quick-Load MasterMix and final concentrations of 0.25µM of forward and reverse primer, 100ng/µl of DNA and brought to volume with dH_2O. The conditions for the PCR reactions for *TaGS5-3A* and *TaGW2-6B* was performed using the touchdown method (Table 3.1) and the conditions for the PCR reaction for *TaGS-D1* was the following: denaturation was at 94ºC for 30 secs, a duration of 30 cycles with 94ºC for 30 secs, 52ºC for 50 secs and 68ºC for 1 min and a final extension of 68ºC for 5 min. The PCR products were electrophoresed on a 3% (w/v) agarose gel stained with 4% (v/v) of SYBR® Safe DNA gel stain within a 1X Lithium acetate buffer at 125V. The positive controls used for this molecular marker was 'Pavon F76', 'Thatcher' and 'Opata85' cultivars; and the negative controls used were 'Inia F 66' and 'Kite' cultivars.

Table 3.1: PCR conditions for the *TaGS5-3A-CAPS* and *TaGW2-6B-CAPS* molecular markers.

Temperature (°C)	TaGS5-3A		TaGW2-6B	
94	30 secs		30 secs	
94	30 secs	13 Cycles	30 secs	13 Cycles
62	50 secs		50 secs	
68	1 min		1 min	
94	30 secs	17 Cycles	30 secs	17 Cycles
55	50 secs		50 secs	
68	1 min		1 min	
68	5 min		5 min	

The PCR products of both *TaGS5-3A* and *TaGW2-6B* were digested with an enzyme. The enzyme digestion for *TaGS5-3A* had a final volume of 5µl and contained 1X NE Buffer, 0.2U/µl of *Fnu4HI* enzyme (Thermo Scientific) and brought to volume with nuclease-free water. The enzyme digestion was added to 5µl of the PCR product and incubated at 37°C for 1 hour. These products were electrophoresed on a 3% (w/v) agarose gel stained with 4% (v/v) of SYBR® Safe DNA gel stain solution within a 1X Lithium acetate buffer at 125V. The positive controls used for this molecular marker was 'Inia F 66', 'Kite' and 'Opata85' cultivars; and the negative controls used were 'Pavon F76' and 'Chinese Spring' cultivars.

The enzyme digestion for *TaGW2-6B* had a final volume of 10µl and contained 1X NE Buffer, 0.5U/µl of *BstNI* enzyme (Thermo Scientific) and brought to volume with nuclease-free water. The enzyme digestion was added to 5µl of the PCR product and incubated at 60°C for 15 mins, with an inactivation step of 80°C for 2 mins. These products were electrophoresed on a 3% (w/v) agarose gel stained with 4% (v/v) of SYBR® Safe DNA gel stain solution within a 1X Lithium acetate buffer at 125V. The positive controls used for this molecular marker was 'Pavon F76' and 'Kite' cultivars; and the negative controls used were 'Inia F 66', 'Chinese Spring' and 'Opata85' cultivars.

The PCR reaction for the *tin1* molecular marker had a final volume of 14µl which contained 6.25µl of KAPA Green Readymix and the final concentrations of 0.36µM of *Xgwm136* forward

49

and reverse primer, 100ng/µl of DNA and brought to volume with dH$_2$O. The conditions for the PCR reaction were the following: denaturation was at 94ºC for 3 min, a duration of 35 cycles with 94ºC for 30 secs, 60ºC for 1 min and 72ºC for 1 min and a final extension of 72ºC for 10 min. The PCR products were electrophoresed on a 2.5% (w/v) agarose gel stained with 4% (v/v) of EtBr within a 1X TBE buffer at 100V. The positive control used for this molecular marker was 'Kite' cultivar and the negative control used was 'Chinese Spring' (CS) cultivar.

The PCR reaction for the *Ppd-D1* molecular marker had a final volume of 15µl which contained 8.5µl of KAPA Green Readymix and the final concentrations of 0.33µM of *Ppd-D1* forward and both reverse primers, 100ng/µl of DNA and brought to volume with dH$_2$O. The conditions for the PCR reaction were the following: denaturation was at 94ºC for 3 min, a duration of 35 cycles with 94ºC for 30 secs, 54ºC for 1 min and 72ºC for 1 min and a final extension of 72ºC for 10 min. The PCR products were electrophoresed on a 2.5% (w/v) agarose gel stained with 4% (v/v) of EtBr within a 1X TBE buffer at 100V. The positive control used for this molecular marker was the 'Opata85' cultivar and the negative control was the 'Chinese Spring' cultivar.

3.4.   MS-MARS scheme

The MS-MARS scheme is a strategy used for pollination between large numbers of selected male sterile (female) and male (donor) plants. It makes use of a hydroponic system which maximises cross pollination and maintains female tillers in nutrient solution until seeds are ripe (Marais & Botes, 2009). Within the scheme, MAS is applied to screen male and female plants for desired genes in order to incorporate wheat lines with desired genes into the successive cycles to achieve allele enrichment of desired traits (Figure 3.2).

3.4.1.   Validation of the MS-MARS scheme

The male sterile (female) tillers were carefully selected at the flowering growth stage, cut and placed into water prior to cross pollination. All the leaves on the female tiller was removed except for the flag leaf. The florets of the female tillers were cut open for maximisation of cross pollination to take place (Figure 3.3a). The male tillers were cut at the flowering growth stage where they were about to shed pollen and tillers were then placed into water prior to cross pollination (Figure 3.3b). All the leaves on the male tiller was removed and both male and female tillers were cut into equal lengths.

The female tillers were arranged into the galvanized iron trays that are 600mm x 450mm x 160mm in dimension, painted with black anti-fungal paint on the inside and can hold up to 230 tillers per tray (Figure 3.3c). Nutrient solution, which is the same as previously mentioned, was used to fill the galvanized iron trays. The male tillers were arranged into a canopy-like order in two narrow galvanized iron trays which was placed on both sides of the bottom tray, about 600mm above the female tillers and filled with nutrient solution as well (Figure 3.3c). Pollination of the female tillers was allowed for 5-6 days, thereafter the male tillers were removed and discarded. The galvanized iron trays with the female tillers were kept under LED lighting for the duration of pollination and grain-filling, at optimum conditions within the growth chamber (Figure 3.3d).

Fresh nutrient solution was changed every two weeks in the galvanized iron trays containing the fertile tillers until the seeds were ripe. The seeds tend to be fully ripened after 5-7 weeks, after which it was placed into brown paper bags and incubated in an oven at 18°C for the seeds to properly dry before threshing them by hand (Figure 3.3f).

3.5.   Phenotyping of high yielding genotypes

3.5.1.   Phenotypic data of the high-yielding genotypes from the field

The high-yielding genotypes were planted with a Wintersteiger Rowseed XL planter at WES in plots. The plots consisted of 6 rows with 30cm spacing between rows. Each plot was 1,6m in width, 5m in length and plots were separated by 1m spacing. Various measurements were taken on the high-yielding genotypes for traits that are yield-determining (Figure 3.4). In total, three randomly selected plants from the 65 HYLD entries in three replicates were pulled from the field and measured, where each tiller of each plant was measured separately. Three random plants per plot was selected in order to obtain an average per plot to give a better indication of the plot's overall performance for any given measurement. The measurements were done by hand or a mobile application, with the use of a ruler and scale when needed (Table 3.2). It should also be noted that some of the genotypes did not have these measurements taken as these genotypes were not ready to be pulled from the field due to stay green characteristics that was displayed by entries 34, 38 and 43 as well as entries 44, 45 and 46 as these genotypes were winter type wheat lines and therefore, flowered much later. Other measurements for all 65 entries were also done on the Inframatic 9500 NIR grain analyzer (by Perten); such as wet gluten, specific weight, moisture, protein fixed and protein dry weight basis.

51
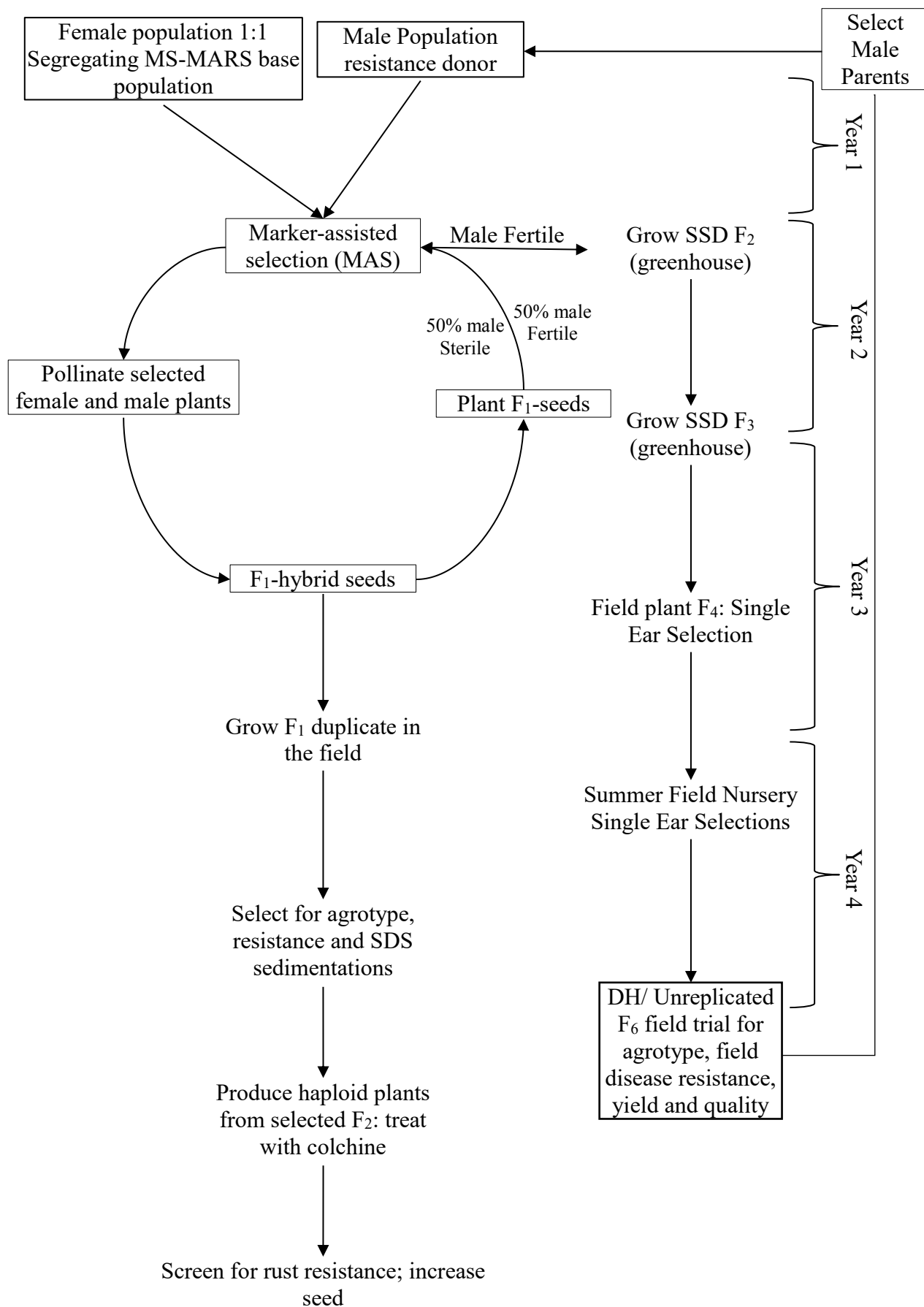
Figure 3.2: The MS-MARS Scheme executed within our lab (Springfield, 2014).
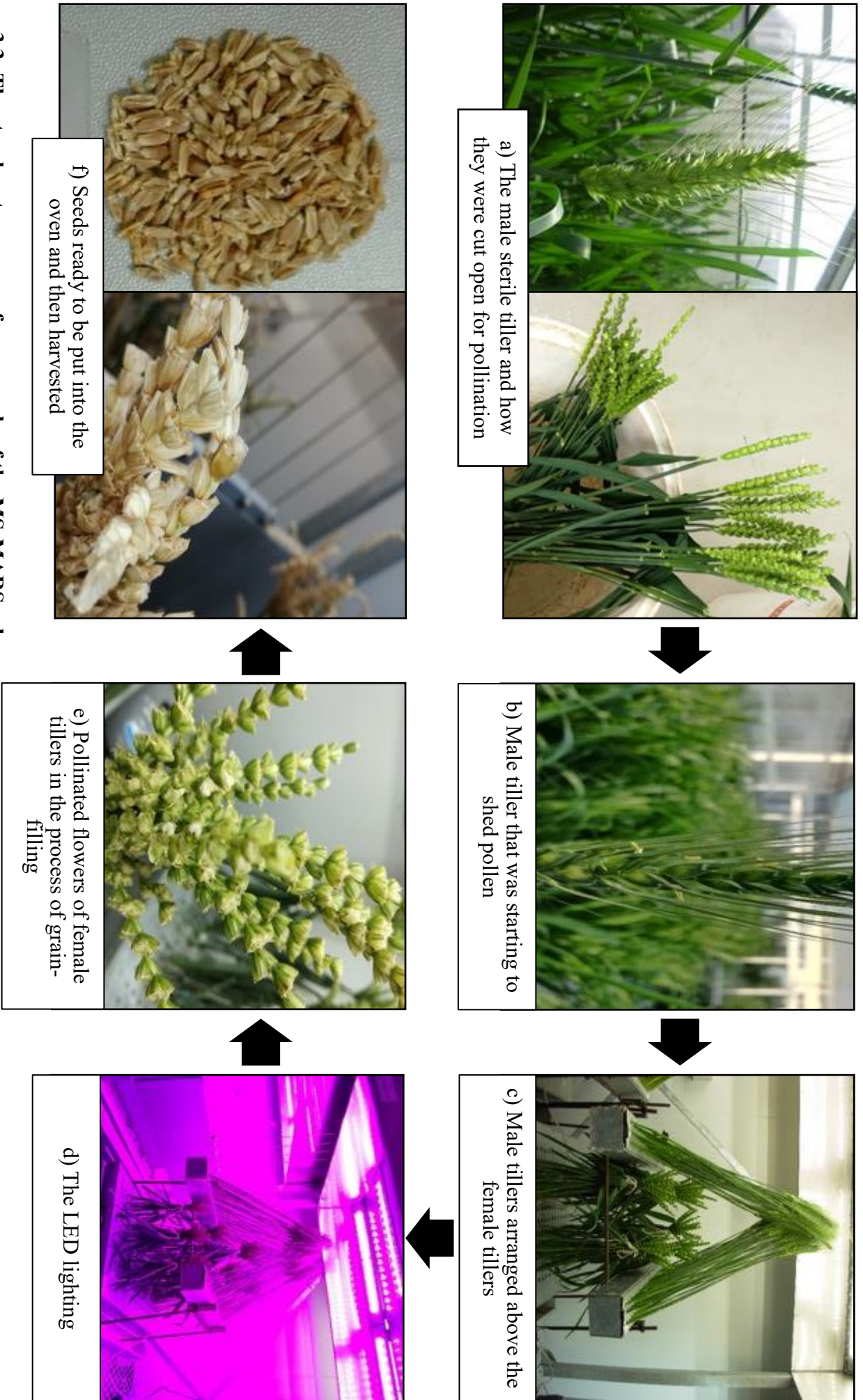
Figure 3.3: The step-by-step process for one cycle of the MS-MARS scheme

a) The male sterile tiller and how they were cut open for pollination

b) Male tiller that was starting to shed pollen

c) Male tillers arranged above the female tillers

d) The LED lighting

e) Pollinated flowers of female tillers in the process of grain-filling

f) Seeds ready to be put into the oven and then harvested

53

The phenotypic data collected (Figure 3.4) for these high-yielding genotypes were compiled into one input file to be statistically analysed in Agrobase Generation II version 34.4.18 (Agronomix Software, Winnipeg, Canada). The test for normality and homoscedasticity was done by the Agrobase software. As the measurements were taken on each tiller per plant, these measurements were averaged per plant and then averaged per genotype for each replicate in order to compile the input file. Thus, the input file contained the averaged measurements of all traits per genotype per replicate. The input file was used on Agrobase to perform a randomised complete block design (RCBD) ANOVA analysis which is a common statistical analysis used in field research to evaluate the high-yielding genotypes for all the traits measured. In addition to the RCBD analysis, the nearest neighbour analysis (NNA) analysis was also analysed on Agrobase, as this analysis accounts for spatial variability (Campbell & Bauer, 2007) and thus, adjusts outputs accordingly. The significant alpha value selected for these statistical analyses was 0.05 (5%).

The selections for the male donors for the recurrent cycle 2 were made according to the output of Agrobase for the dataset. The Agrobase output ranked the genotypes from the best to least best performing genotype according to each specific trait. The data obtained for the yield trait that was measured (on the NIR grain analyzer) for each genotype was used to identify the best possible lines to be selected for male crossing parents. With the use of least significance difference (LSD), it was possible to identify the genotypes that did not significantly differ from the best performing genotype. These genotypes identified were then compared to a data set analysed from a different study using the same set of genotypes. From this comparison, it was concluded that the best performing genotypes for the yield trait was the four entries; 4, 9, 26 and 64. These four genotypes were selected as the male crossing parents along with SU-PBL 2017 nursery for the second cycle of the MS-MARS.

Statistical analyses were also performed for the frequencies of the *Sr2* gene obtained for the MS-MARS cycle 1 and 2. Microsoft Excel (2016) was used to calculate whether the difference between the observed and expected frequencies for *Sr2* gene was statistically significant with a significant threshold of 0.05 (5%) (Microsoft, 2016). Contingency tables was also used to perform an exact test to identify whether the observed frequencies for the *Sr2* gene between the two cycles were statistically significant with a significant threshold of 0.05 (5%). The contingency table was generated on the GraphPad software website (GraphPad software, 2018).

Table 3.2: Table on how the yield-determining traits were measured for each tiller/plant

Trait	How traits were measured	Unit
Tiller Number	The number of tillers were counted for each plant.	numeric
Plant Height	The plant height of each tiller was measured from the base of the tiller to the spike (excluding awns).	mm
Spike Length	The spike length of each tiller was measured with a ruler.	mm
Spikelet Number	The spikelet number was counted for each spike, i.e. each ear per tiller.	numeric
Grain Number/Spike	The grains per spike was measured by counting the number of grains per spike of each tiller after hand-threshing.	numeric
Grain Weight	Grain weight of grains per spike was weighed on a scale.	g
Harvest Index	Harvest index was calculated by dividing the weight of the grains/plant by the weight of the complete plant including grains/plant.	%
TKW	TKW was calculated by dividing the grain weight/plant by the grain number/plant, multiplied by 1000.	g
Days to Heading	The heading date of each plot was determined when approximately 50% of the spikes of the tillers has completely emerged from the flag leaf sheath.	days
Flower fertility	Flower fertility was calculated by dividing grains/spike with the number of florets per spike and converted into percentage.	%
Grain Length, Width & Area	Grain length, width and area were measured on the SeedCounter and 1KK mobile applications.	mm, mm^2

Figure 3.4: How measurements were taken for the yield-determining traits. a) The plants pulled from the field. b) A pulled plant selected and tillers counted per plant. c) Measuring the height of the tillers. d) Measuring the spike length. e) Counting the number of spikelets per spike. f) Counting the number of florets for floret fertility. g) After the initial measurements, tillers kept separately per plant. h) Each tiller hand-threshed separately. i) The waste of the tillers per plant while threshing. j) All the waste of the tillers per plant was stored in brown bags. k) Counting the number of grains per spike using a seed counter. l) Weighing the waste in the brown bag to calculate HI. m) Weighing the grains per spike for grain weight as well as to calculate HI and TKW. n) The Inframatic 9500 NIR grain analyser.

3.5.2.   Phenotyping with the use of image-based analysis

To measure grain length, width and area with the SeedCounter; seeds were placed on an A4 white paper sheet which was placed on a dark background for the paper sheet to be recognised by the application (Figure 3.5a). The seeds were spaced from one another to avoid contact in order to reduce errors. Errors can occur from bright side lighting as well and thus, that should be avoided (Komyshev *et al.,* 2017). This application is compatible with most android devices (Huawei P9, Huawei P9 lite, and Vodacom Tablet) that was used to test the application within this study.

Each entry in three replicates of the high-yielding genotypes were measured for grain length, width and area as well as validation of seed count. A hundred seeds were counted for each sample and placed on the white paper sheet for analysis. The images were taken approximately 300mm above the paper using a Huawei P9 lite (OS version (Android) 6.0) and analysed with the SeedCounter (Figure 3.5b). The data was saved and exported to the mobile device's internal storage as a tab separated values (tsv) file. The tsv file was opened in notepad on Windows and copied into Microsoft Excel (2016) in which the data was properly organised (Microsoft, 2016). The 100 seeds were also used to calculate the TKW once again, by weighing the 100 seeds on the scale and multiplying that weight by 10.
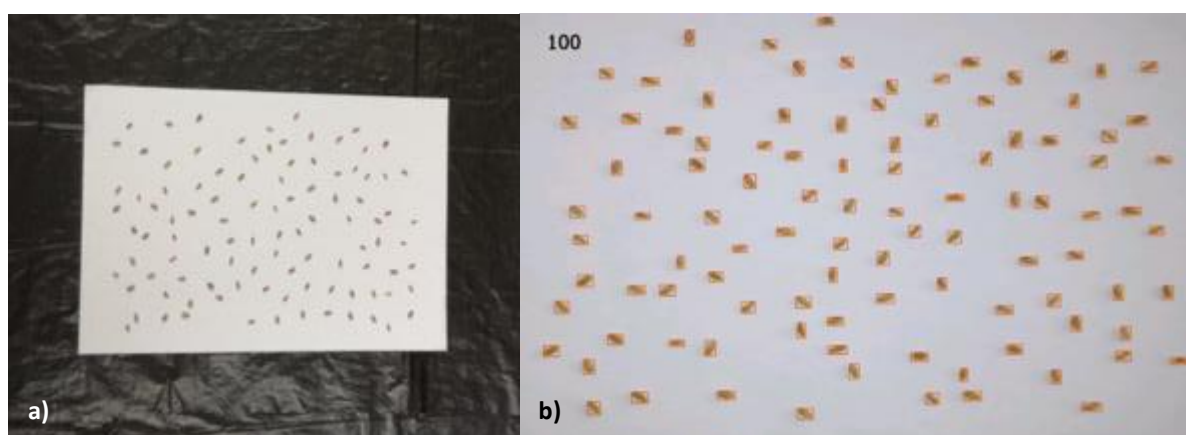
Figure 3.5: How the SeedCounter application measures grain morphology. a) The seeds placed on the white paper sheet with the dark background. b) The preview of the analysed image taken by SeedCounter, the seed count is given at the top left corner of the image.

To measure grain length, width and area with the 1KK application; seeds were placed on a coloured background (provided by the 1KK manual) which is green with blue reference circles of known size in order to translate the pixel measurements of the seeds to actual sizes (Figure 3.6b). The seeds were spaced from one another to avoid contact as seeds in contact are omitted from being analysed. The images were taken using an Android Vodacom Tablet (OS version 5.1) as this device was the most compatible with the 1KK application.

The application is also compatible with the Elane USB Plus 5kg scale and therefore this scale was used simultaneously with the 1KK appliaction to weigh the seeds while measuring grain length, width and area. The green background with blue circles was placed on the scale (Figure 3.7) and the following settings were inserted into the 1KK application. The display analysis preview and attempt to crop photos were selected, the reference circle size was set at 5mm and the minimum and maximum seed size remained 0 (Figure 3.6a). The tablet was placed approximately 230mm above the scale with direct light approximately 800mm above the scale.
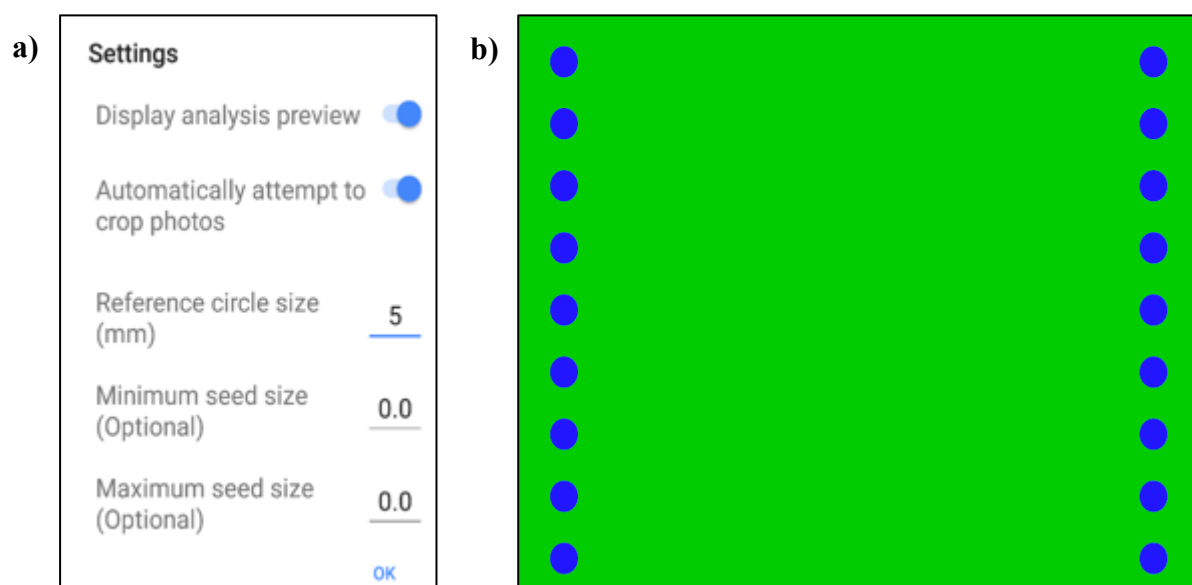
Figure 3.6: The procedure of the 1KK application for measuring grain morphology. a) The settings of the 1KK application. b) The coloured background on which the seeds were placed.

58

Figure 3.7: The tablet and scale that was used when analysing the seeds with the 1KK application.

Chapter 4:   Results and Discussion

4.1.   Genotyping of plant material

4.1.1.   Screening of rust resistance genes for MS-MARS crossing parents

The male and female (male sterile) populations were genotyped with the *Lr34* and *Sr2* rust markers forming part of the lab's standard panel of markers. This screening was done prior to cross-pollination. The frequencies for the male population of 65 high-yielding genotypes for *Lr34* and *Sr2* was 18% and 9%, respectively (Figure 4.2). The female population was screened for *Lr34* and *Sr2* for both MS-MARS cycles 1 and 2. For the first cycle, 465 plants of the female population were screened for the rust markers with frequencies of 46% for *Lr34* and 65% for *Sr2*. For the second cycle, 120 plants of the female F_1 population were screened for the rust markers with frequencies of 48% for *Lr34* and 48% for *Sr2* (Figure 4.1).
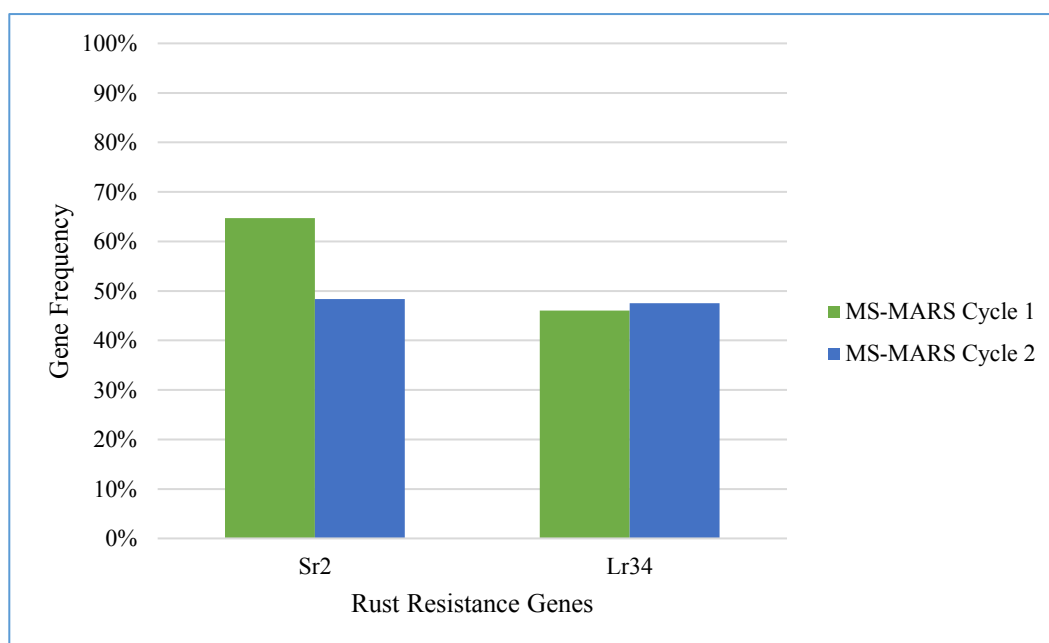
Figure 4.1: The allele frequencies for the rust markers for the female population of the MS-MARS cycle 1 and 2.

The expected allele frequencies for the second cycle was calculated with the use of the allele frequencies of both the male and female population. As per population genetics theory, the expected allele frequencies were calculated at 22.7% for the presence of the *Sr2* gene, which is lower than what was observed. The difference between the expected and observed allele frequencies was also observed to be statistically significant with a p-value of less than 0.05 (5%), were the significant threshold was set at 0.05 (5%). This result suggests that the

60

conditions of random mating was violated which could be due to an error in the sampling size as all cutting sessions did not consist of the same sample sizes, where some cutting sessions yielded more tillers than others, therefore causing somewhat biasness when cuttings were performed. Contingency tables were also used to perform an exact test for the allele frequencies of the *Sr2* gene between the two cycles, with a significant threshold of 0.05 (5%). It was then calculated that there was a statistical significant difference (p-value: 0.0015) in the frequencies for *Sr2* of the female populations between the first and second cycle. It was stated that when an allele frequency is at 60%, two additional seasons are required to raise the allele frequency to exceed 70%, but for a sharp increase in frequency, the male parents are also required to carry out the selection effect (Marais & Botes, 2009).

When comparing these results to the results obtained by previous MS-MARS cycles; it was displayed that for the MS-MARS cycles (2013 & 2014), the allele frequencies for the *Lr34* gene increased by 8% and the *Sr2* gene increased by 11% between cycles (Springfield, 2014). It was also displayed that for the MS-MARS cycles (2015 & 2016), the allele frequencies for the *Lr34* gene decreased by 1% and the *Sr2* gene increased by 2% between cycles (Meintjes, 2017). In both those studies, different male donors were used and therefore allele frequencies are dependent of the selection effect that the male donor carries. However, in this study, the allele frequencies for *Sr2* decreased between cycles and the reason for this is that the male population has a low allele frequency of 9% for *Sr2* as this population was not initially selected for its rust resistance but rather for its high yielding qualities. The low frequency for the *Sr2* gene in the male population could be due to the fact that this gene is associated with a phenotype undesired by breeders. The *Sr2* gene is associated with the pseudo-black chaff (*PBC*) phenotype and high expression of this phenotype may lead to grains that are shrunken and thus, reduces yield (Juliana *et al.,* 2015).

The frequencies for *Lr34* of the female population increased slightly between the first and second cycle. This is expected as a allele frequency that is continuously positively selected is supposed to increase over successive cycles and allele frequencies at 40% requires three additional seasons to raise the allele frequency to exceed 70% (Marais & Botes, 2009). However, if a male population with high frequencies of *Lr34* is used for the following cycles then a sharp increase in the allele frequency for *Lr34* will be experienced (Marais & Botes, 2009). It can also be noted that the allele frequencies for the *Lr34* gene in this male population is slightly higher than for the *Sr2* gene and therefore, there was an increase in the *Lr34* allele

61

frequency of the female population between the cycles instead of a decrease as displayed by the *Sr2* gene.
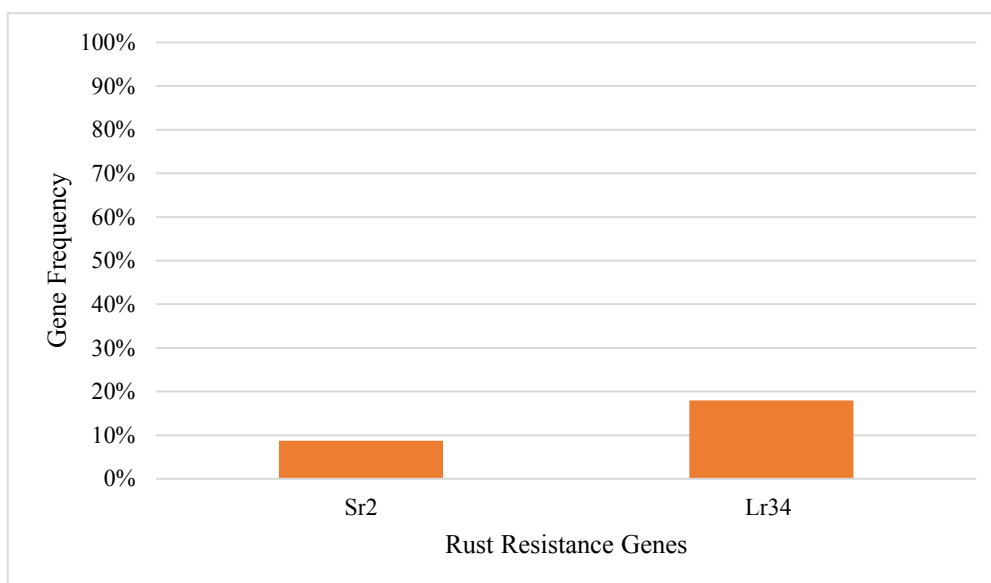
Figure 4.2: The allele frequencies for the rust markers of the male population.

4.1.2. Screening of the molecular markers associated with yield-determining traits

Molecular markers were identified through literature for use as a molecular tool to identify yield-determining traits and therefore, these molecular markers were validated for future use as a diagnostic marker. The optimisation and validation of these markers were executed through PCR tests. The markers for *TaGS5-3A*, *TaGW2-6B* and *TaGS-D1* genes were all CAPS markers. Only one marker was used for *tin1* and one marker with an additional reverse primer was used for *Ppd-D1*. The *TaGS5-3A*, *TaGW2-6B* and *TaGS-D1* genes are all associated with the TKW trait, however both *TaGS5-3A* and *TaGW2-6B* are additionally also associated with grain width and only *TaGS-D1* is also associated with grain length. The *tin1* and *Ppd-D1* genes are associated with the tiller number and days to heading traits, respectively. The 65 high-yielding genotypes was the only population tested with these markers, which resulted in the following frequencies for the genes; 77% for *TaGS5*, 49% for *TaGS-D1*, 32% for *TaGW2-6B*, 98% for *Ppd-D1* and 32% for *tin1* (Figure 4.3).

62
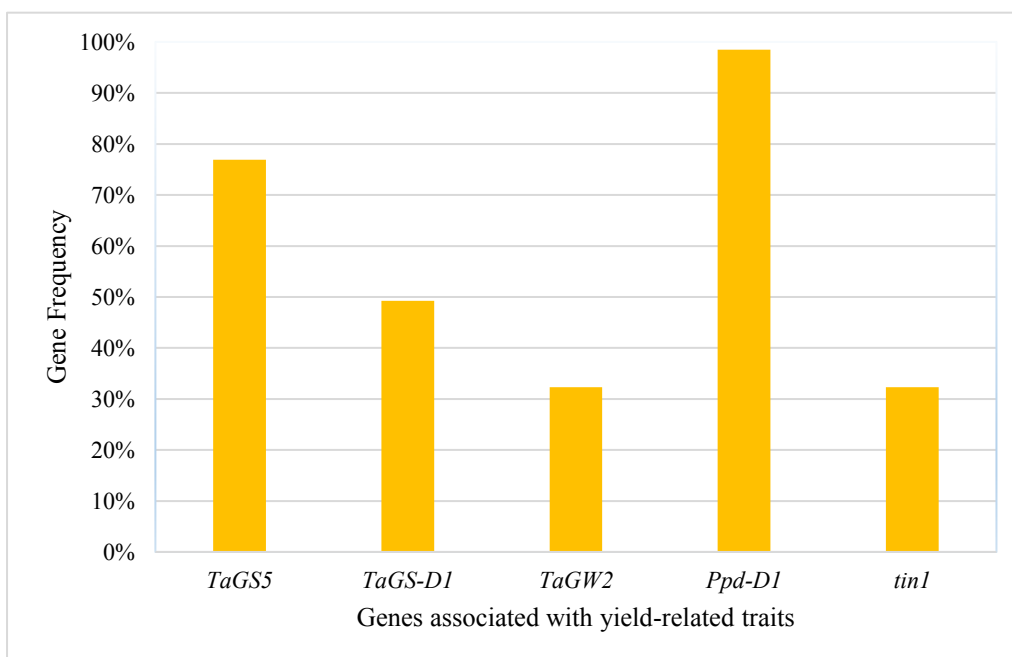
Figure 4.3: The allele frequencies for high-yielding genotypes for the markers associated with yield-determining traits.

4.1.2.1. *TaGS5-3A-CAPS* marker

To select for increased TKW as well as grain width, the *TaGS5* gene was identified by literature. This gene is associated with a *TaGS5-3A-CAPS* marker that detects the *TaGS5-3A-T* allele and when cleaved, detects the *TaGS5-3A-G* allele. The presence of the *TaGS5-3A-T* allele amplifies an 863bp product and the presence of the *TaGS5-3A-G* allele, after enzyme digestion, amplifies a 718bp and 145bp product (refer to figure 4.4). Within this reaction, the positive controls used were 'Inia F 66' (I66), 'Kite' (K) and 'Opata85' (OP); and the negative controls used were 'Pavon F76' (P76) and 'Chinese Spring' (CS).
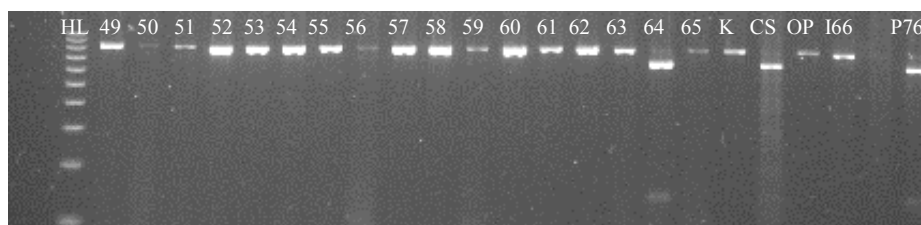
Figure 4.4: Marker analyses of an agarose gel stained with SYBR® Safe DNA gel stain solution, which displays the DNA bands produced by PCR amplification of the *TaGS5-3A-CAPS* marker for the *TaGS5* gene. Lane 1: 100bp DNA Ladder, Lane 2-17: High-yielding genotypes tested for the presence of the *TaGS5* gene, Lanes 18-24: positive and negative controls.

63

The bands observed for the genotypes displayed that majority of the genotypes amplified for the 863bp product. The positive and negative bands were observed as well as no contamination was observed. Within this set of genotypes, the favourable allele (*TaGS5-3A-T*) was found in 50 genotypes, therefore resulting in a frequency of 77%. This marker was successful in distinguishing between genotypes with the *TaGS5-3A-T* or *TaGS5-3A-G* allele which indicates the genotypes associated with higher or lower TKW, respectively. Therefore, this marker can be considered to be utilised as a diagnostic marker which detects the presence of either allele and can be used in MAS to detect increased grain weight and width.

4.1.2.2. *TaGW2-6B-CAPS* marker

Another gene associated with increased TKW as well as grain width was identified by literature. The *TaGW2* gene is associated with a *TaGW2-6B-CAPS* marker that detects the presence of the *Hap-6B-1* through an enzyme digestion which distinguishes *Hap-6B-1* from *Hap-6B-2, Hap-6B-3* and *Hap-6B-4*. *Hap-6B-1* has no restriction enzyme recognition site and amplifies a 1361bp product (refer to figure 4.5). Within this reaction, the positive controls used were 'Pavon F76' (P76) and 'Kite' (K); and the negative controls used were 'Inia F 66' (I66), 'Chinese Spring' (CS) and 'Opata85' (Op).

Figure 4.5: Marker analyses of an agarose gel stained with SYBR® Safe DNA gel stain solution, which displays the DNA bands produced by PCR amplification of the *TaGW2-6B* marker for the *TaGW2* gene. Lane 1-8: High-yielding genotypes tested for the presence of the *TaGW2* gene, Lanes 9-15: positive and negative controls, Lane 16: Quick-Load 50bp DNA Ladder.

The bands observed for the genotypes displayed that minority of the genotypes amplified for the 1361bp product. The positive and negative bands were observed and no contamination was observed. Within this set of genotypes, the favourable haplotype (*Hap-6B-1*) was found in 21 genotypes, therefore resulting in a frequency of 32%. This haplotype is associated with increased TKW as well as association with grain width. This marker was successful in distinguishing between genotypes with *Hap-6B-1* from the rest of the haplotypes. Therefore, this marker can be considered to be utilised as a diagnostic marker which detects the presence of *Hap-6B-1* and can be used in MAS. However, the *Hap-6B-1* displaying low frequencies is concerning but genotypic screening of *Hap-6B-2* can further be tested as this haplotype was reported not to be significantly different from *Hap-6B-1* and is also associated with increased grain weight (Qin *et al*., 2014).

4.1.2.3. *GS7D* marker

To select for increased TKW as well as grain length, the *TaGS-D1* gene was identified by literature. This gene is associated with the *GS7D* marker that is used to detect the *TaGS-D1a* or *TaGS-D1b* allele. The presence of the *TaGS-D1a* allele amplifies a 562bp product and the presence of the *TaGS-D1b* allele amplifies a 522bp product (refer to figure 4.6). The positive controls used were 'Pavon F76' (P76), 'Thatcher' (Th) and 'Opata85' (Op) and the negative controls used were 'Inia F 66' (I66) and 'Kite' (K).
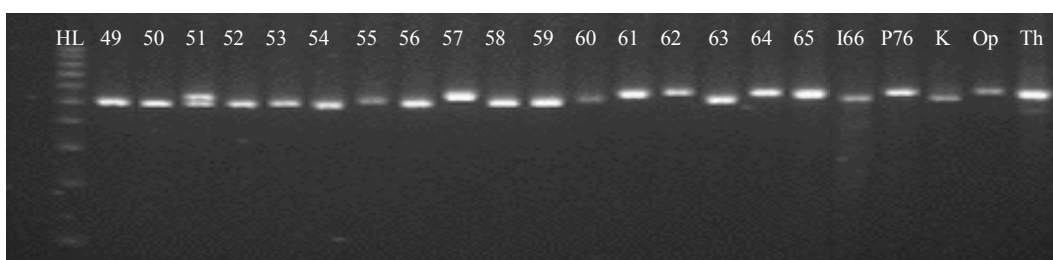
Figure 4.6: Marker analyses of an agarose gel stained with SYBR® Safe DNA gel stain solution, which displays the DNA bands produced by PCR amplification of the *GS7D* marker for the *TaGS-D1* gene. Lane 1: 100bp DNA Ladder, Lane 1-18: High-yielding genotypes tested for the presence of the *TaGS-D1* gene, Lanes 19-23: positive and negative controls.

The bands observed for the genotypes displayed an approximate 50% of the genotypes amplified for the 562bp product. The positive and negative band was observed and no contamination was observed. Within this set of genotypes, the favourable allele (*TaGS-D1a*) is found in 32 genotypes, therefore resulting in a frequency of 49%. This marker was successful in distinguishing between genotypes with the *TaGS-D1a* or *TaGS-D1b* allele. Therefore, this marker can be considered to be utilised as a diagnostic marker for which detects the presence of either allele and can be used in MAS to detect increased grain weight and length.

4.1.2.4. *Xgwm136* marker

To select for low-tillering number, the *tin1* gene was identified by literature. The *tin1* is associated with the *Xgwm136* marker that detects either a 300bp product associated with high tillering, or a 350bp product associated with low tillering. The positive control used was 'Kite' (K) and the negative control used was 'Chinese Spring' (CS).
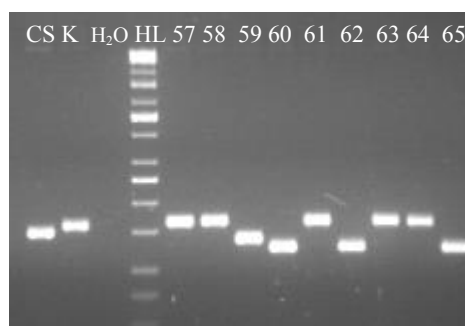
Figure 4.7: Marker analyses of an agarose gel stained with EtBr solution, which displays the DNA bands produced by PCR amplification of the *Xgwm136* marker for the *tin1* gene. Lanes 1-3: positive and negative controls, Lane 4: Universal Ladder, Lanes 5-13: High-yielding genotypes tested for the presence of the *tin1* gene.

The bands observed for the genotypes were inconsistent as shown on figure 4.7. The positive and negative band was observed and no contamination was observed. The genotypes were scored as any genotypes in line with or below the negative control were scored as high tillering genotypes and any genotypes in line with or above the positive control were scored as low tillering genotypes. Even though, Kumar *et al.* (2015) reported nearly perfect co-segregation of the tiller number and this molecular marker, a study by Spielmeyer and Richards (2004)

66

explained that the *Xgwm136* marker detected **approximately** the 300bp or 350bp. Due to the inconsistency of the marker, this marker cannot be considered as a reliable diagnostic marker for tiller number within this set of genotypes.

4.1.2.5. *Ppd-D1* marker

To select for earlier flowering resulting in less days to heading, the *Ppd-D1* gene was identified by literature. The *Ppd-D1* gene is associated with the *Ppd-D1* marker that detects either the *Ppd-D1a* or *Ppd-D1b* allele which is associated with photoperiod insensitivity or sensitivity, respectively. For the presence of *Ppd-D1a* allele, a 288bp product is amplified (refer to figure 4.8). For the presence of the *Ppd-D1b* allele, a 414bp product is amplified. The positive control used was 'Opata85' (Op) and the negative control used was 'Chinese Spring' (CS).
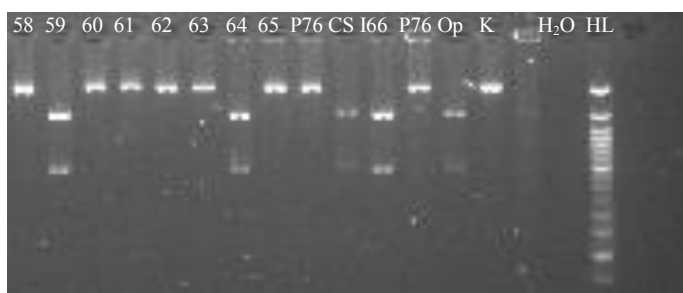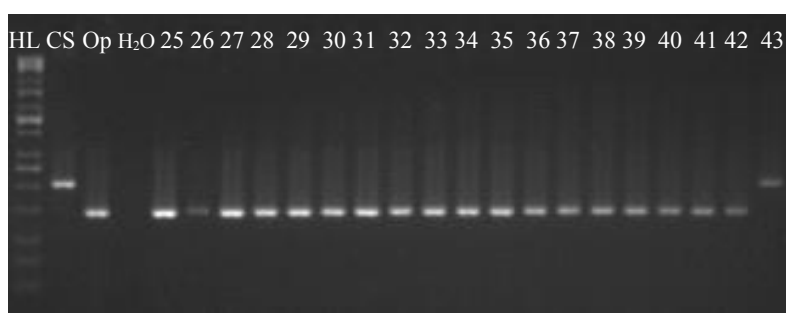

Figure 4.8: Marker analyses of an agarose gel stained with EtBr solution, which displays the DNA bands produced by PCR amplification of the *Ppd-D1* marker for the *Ppd-D1* gene. Lane 1: Universal Ladder, Lanes 2-4: positive and negative controls, Lanes 5-23: High-yielding genotypes tested for the presence of the *Ppd-D1* gene.

The bands observed for the genotypes displayed that all genotypes amplified for the 288bp product except for one genotype, i.e. entry 43. The positive and negative band was observed and no contamination was observed. Within this set of genotypes, the favourable allele (*Ppd-D1a*) is found in all except the one genotype. As the *Ppd-D1a* allele is associated with earlier flowering, it allows for wheat to be ready for harvesting before the warmer temperatures of the summer is reached (Beales *et al*., 2007). This marker was successful in distinguishing between genotypes with the *Ppd-D1a* or *Ppd-D1b* allele. Therefore, this marker can be considered to be utilised as a diagnostic marker which detects the presence of either allele and can be used in MAS as an indicator for early flowering resulting in less days to heading for wheat lines.

4.2.   MS-MARS scheme

4.2.1.   Validation of the MS-MARS scheme

The validation of male sterile plants was recognised by assessing the phenotype of the tillers. The spike of the tiller displayed a lighter in colour spike with sprawling awns and most importantly, the anthers within the floret was small, shrivelled and not powdery, i.e. absence of pollen as male sterile plants are unable to produce fertile pollen. Therefore, the selections of male sterile plants were carefully made in order to avoid selecting male fertile plants for cross-pollination. A successful cross-pollination was confirmed by the resulting hybrid seed that displayed a small, shrivelled phenotype (Marais & Botes, 2009). Although the hybrid seeds produced are small, the hybrid seeds that were planted within the greenhouses displayed quite high germinating percentage.

4.2.1.1. Recurrent cycle 1 (2016)

The cross-pollination for cycle 1 used the segregating base population (from the previously executed MS-MARS cycle) as the female (male sterile) population and the high-yielding genotypes as the male (donor) population. Following the genotypic screening of both populations for desired genes, the tillers for both populations were cut twice weekly for a duration of seven weeks. A total of 14 successful sessions of cuttings were completed with the number of tillers for the male and female populations varying between 60-110 and 60-150, respectively (Table 4.3). The number of tillers available for both male and female population per cutting were dependent on the number of tillers that satisfied the requirements needed for a male and female tiller to result in successful cross-pollination. A total number of 1263 male plants were successfully used to cross-pollinate with a total number of 1520 female tillers. The cross-pollination produced a total number of 10510 seeds over the duration of four harvesting periods (Table 4.3). These seeds were produced from 96% of the female plants that were sourced for cross-pollination. The overall average for successful cross-pollination rate was 23.88%. The average grain mass for the cross-pollinated seeds for this cycle was 243,5 g (Table 4.3).

4.2.1.2. Recurrent cycle 2 (2017)

The cross-pollination for cycle 2 used the F_1 seeds (that was harvested from the first recurrent cycle) as the female population and a selection of four high-yielding genotypes (based on phenotypic data) as well as SU-PBL 2017 nursery were used as the male population. A total of 18 successful sessions of cuttings were completed with the number of tillers for the male and female populations varying between 50-140 and 40-140, respectively. A total number of 1577 male plants were successfully used to cross-pollinate with a total number of 1811 female tillers. The cross-pollination produced a total number of 12268 seeds over the duration of four harvesting periods (Table 4.4). These seeds were produced from 99% of the female plants that were sourced for cross-pollination. The overall average for successful cross-pollination rate was 23.95%. The average grain mass for the cross-pollinated seeds for this cycle was 238.5 g (Table 4.4).

4.2.1.3. Cross-pollination of recurrent cycle 1 and 2

The percentage of successful cross-pollination for cycle 1 and cycle 2 resulted in an average of 23.88% and 23.95%, respectively. These averages are relatively low as the goal is to be able to achieve higher percentage in cross-pollination. This could be explained through selecting male and female tillers at the wrong reproductive time as optimal seed production can only occur if the females' flowers are opened simultaneously to the release of viable pollen from the male donors. This timing is important as wheat pollen is viable for a brief period of time (Kempe *et al*., 2014).

Another reason that could explain low cross-pollination is the structure of the floret, as it was observed that many sterile tillers possessed narrow glumes and others possessed wide glumes (Figure 4.9). The tillers with wider glumes had higher average percentage in cross pollination as compared to the tillers with narrow glumes. Therefore, the structure of a sterile tiller is important for maximising pollen reception (Whitford *et al*., 2013). It should also be noted that the male tillers should not be arranged too high above the female tillers as to ensure a higher chance of pollination for each flower of the female tiller.
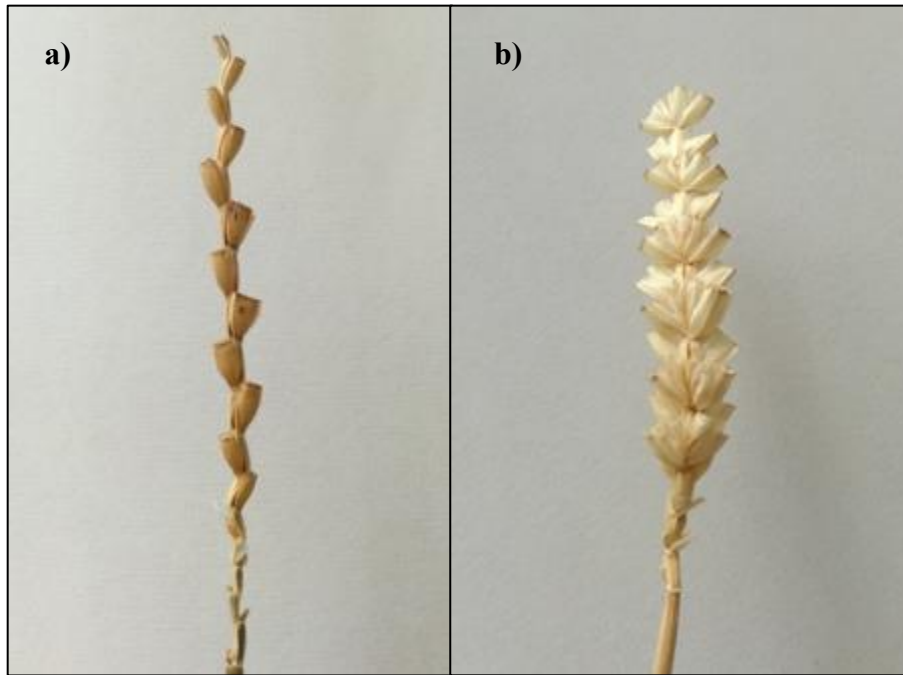
Figure 4.9: The two types of floret structures that a wheat tiller may possess. a) Male sterile tiller with florets that have narrow glumes. b) Male sterile tiller with florets that have wide open glumes.

4.2.1.4. The male sterility inheritance for recurrent cycle 1 and 2

When the segregating population was harvested, the male sterile to male fertile ratio for each pot was recorded in order to perform chi-squared analyses. The chi-squared analyses were done to determine the male sterility gene inheritance, i.e. whether the dominant male sterility gene displays a good fit to the 1:1 ratio of male sterile and male fertile plants. This was done for each of the female populations used in both recurrent cycles. The segregating population for both cycles were planted over four benches; therefore, chi-squared analyses were performed for each bench as well as an overall chi-squared analysis.

The results for the chi-square tests for the first cycle displayed bench 2, 3 and 4 had a p-value greater than 0.05 and therefore, those three benches showed a good fit to the 1:1 ratio (Table 4.1). However, bench 1 had a p-value of less than 0.05 and therefore, does not show a good fit to the 1:1 ratio. Due to the deviation of this bench, the overall p-value of all the benches are less than 0.05 and thus, deviates from the 1:1 ratio (Table 4.1). For the second cycle, benches 2 and 4 had a p-value greater than 0.05 and therefore, those two benches showed a good fit to the 1:1 ratio (Table 4.2). However, bench 1 and 3 had a p-value of less than 0.05 and therefore, does not show a good fit to the 1:1 ratio. Due to the deviation of these two benches, the overall

70

p-value of all the benches are less than 0.05 and therefore, deviates from the 1:1 ratio (Table 4.2). Thus, the single dominant gene within a heterozygous state that controls male sterility within the population could not be confirmed.

The deviation of the benches from the goodness fit to ratio could be an error that occurred during the recording of male sterile and male fertile plants. Temperatures may affect the sterility of a plant and may cause fertile anthers to form within a male sterile plant causing partial sterility (Meng *et al*., 2016). High temperatures during the anthesis growth stage may also cause fertile plants to become sterile and thus, largely reduce the grain yield (Farooq *et al*., 2011). The temperatures during the anthesis growth stage for both cycles were taken from the weather underground website from which the graphs in figure 4.10 and 4.11 was developed.

For cycle one and two, the maximum temperatures ranged between 15°C - 30°C and 12°C - 28°C, respectively. However, temperatures within the greenhouse tends to be 5°C higher as compared to the outside environment. Therefore, it can be noted that for certain days, the temperatures were higher than the threshold which could have been the cause of the partial sterility of plants as well as the reduction in grain set of the plants. Therefore, the temperatures experienced within the greenhouse could have caused the deviation of both populations from the 1:1 ratio.

Table 4.1: The male sterility inheritance within the recurrent population for cycle 1

Bench	Pots	Sterile Plants	Fertile Plants	X_2	Probability of fit to a 1:1 ratio
1	118	212	269	6,755	0,009
2	116	233	245	0,301	0,583
3	112	237	260	1,064	0,302
4	118	257	263	0,069	0,792
Overall	464	939	1037	4,86	0,027

Table 4.2: The male sterility inheritance within the F_1 recurrent population for cycle 2

Bench	Pots	Sterile Plants	Fertile Plants	X_2	Probability of fit to a 1:1 ratio
1	120	273	160	29,49	<0,001
2	120	254	231	1,091	0,296
3	121	307	222	13,658	<0,001
4	63	117	117	0	1
Overall	424	951	730	29,055	<0,001
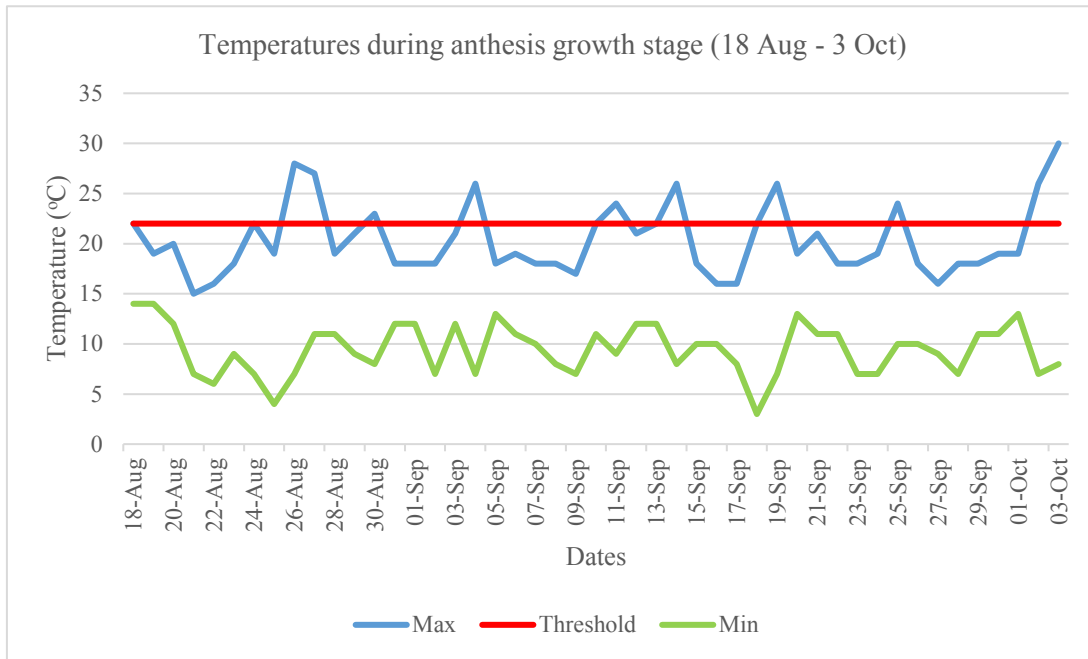
Figure 4.10: The temperatures during the anthesis growth stage of the first cycle from 18 Aug to 3 Oct 2016 (Weather underground, 2018).
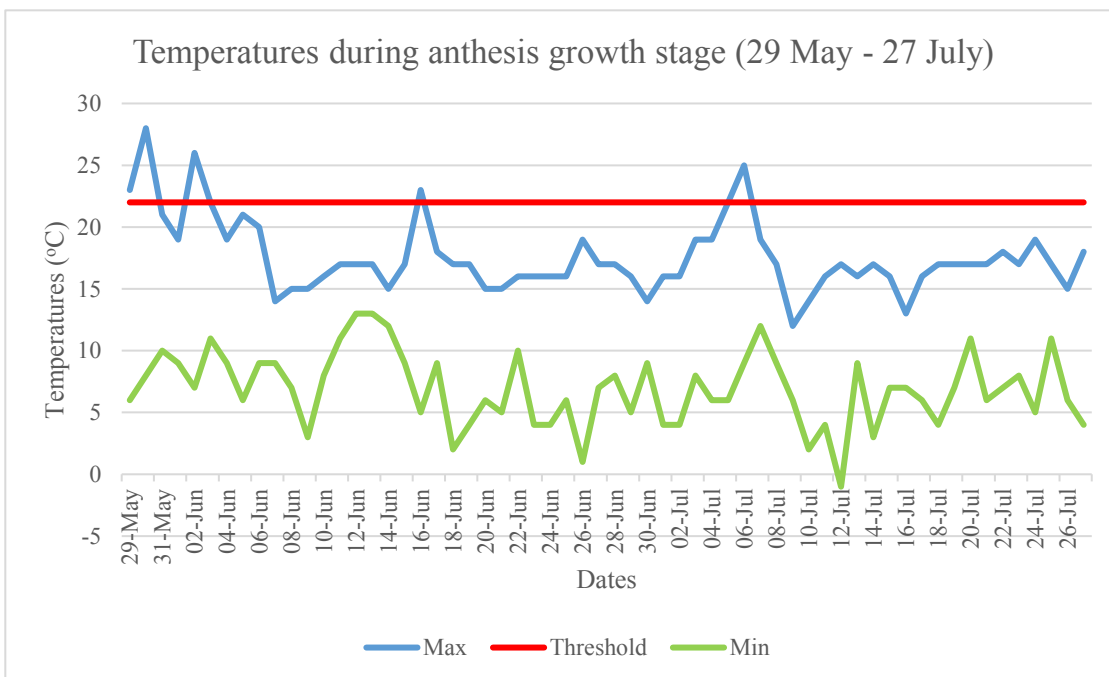
Figure 4.11: The temperatures during the anthesis growth stage of the second cycle from 29 May - 27 July (Weather underground, 2018).

72

Table 4.3: The results of the cross-pollination for the recurrent cycle 1 (2016)

Threshing Dates	Week	Male Fertile	Male Sterile	Possible Combinations	Recurrent Cycle 1				
					Harvest	Amount of sterile plants sourced from	Seeds per Harvest	Average cross pollination (%)	Total mass (g)
11 Oct	1	193	179	34547	1	428	4536	26,22	105,1
	2	210	250	52500					
3 Nov	3	200	265	53000	2	242	713	10,72	16,5
7 Nov	4	189	243	45927	3	234	1450	28,57	33,6
	5	192	290	55680					
15 Dec	6	187	217	40579	4	558	3811	30,01	88,3
	7	92	76	6992					
Total		1263	1520	289225	4	1462	10510	23,88	243,5

Table 4.4: The results of the cross-pollination for the recurrent cycle 2 (2017)

Threshing Dates	Week	Male Fertile	Male Sterile	Possible Combinations	Recurrent Cycle 2				
					Harvest	Amount of sterile plants sourced from	Seeds per Harvest	Average cross pollination (%)	Total mass (g)
26 July	1	144	251	36144	1	432	938	8,36	32,8
	2	174	191	33234					
	3	158	231	36498					
10 Aug	4	187	226	42262	2	431	4470	34,80	91,8
	5	176	199	35024					
31 Aug	6	192	226	43392	3	427	2515	22,06	49,2
	7	267	186	49662					
6 Nov	8	189	252	47628	4	500	4345	30,58	64,7
	9	90	49	4410					
Total		1577	1811	328254	4	1790	12268	23,95	238,5

4.3.   Phenotyping of high yielding genotypes

4.3.1.   Phenotypic data of the high-yielding genotypes from the field

The phenotypic data collected for the high-yielding genotypes were analysed with the Randomised Complete Block Design (RCBD) ANOVA analysis and Nearest Neighbour Analysis (NNA) in Agrobase Generation II version 34.4.18 (Agronomix Software, Winnipeg, Canada). The use of NNA was to determine whether data will be adjusted or remain unadjusted (as was determined by the RCBD analysis). The difference between these two analyses is that NNA takes into account the field trends of the dataset. With the use of the NNA's adjacent residual method that is available in Agrobase, the field trends were accounted for and adjusted accordingly (Campbell & Bauer, 2007; Stroup & Mulitze, 1991). The data generated with RCBD analysis for all the yield-determining traits is displayed in table 4.21.

Agrobase generated various results, but the most relevant results to observe for each trait is the coefficient of variation (CV), coefficient of determination (R^2) and broad-sense heritability (H^2) values. These values are important and it serves as an indication of the variability of the trait, the goodness-to-fit of the model and the transmissibility of the trait into future generations. The CV is described as a statistical measurement that computes phenotypic variation and thus, represents variation that is induced by the environment and can also be used as an assessment for stability (Fasoula & Tokatlidis, 2012). Within this study, the CV-values ranged between 0.67% and 26.68% for the various yield-determining traits (Table 4.21). The low CV-values generally indicate that low variability exist for that trait and that the trait has little genotype x environment interaction which makes the trait more reliable to select for in breeding programmes (Bello *et al.,* 2012; Racz, 2015). The R^2 is used to gain information on the goodness-of-fit of a model and describes the proportion of the variance total within the observed data which can be explained by the model. R^2-values range between 0 and 1, where the high values are an indication of better agreement to the model (Legates, 1999). The R^2-values within this study range between 0.4845 and 0.9521 (Table 4.21). Low R^2-values within this study indicates that the variance in the experiment was not explained very well by the RCBD analysis and high R^2-values indicate that the RCBD analysis can explain the majority of the variation. Heritability plays an important role when planning for efficient breeding programmes as the genetic improvement of plants for quantitative traits do require a reliable heritability estimate. It is essential to have reliable knowledge on heritability for selection improvement as it is an indication of the degree of transmissibility of a trait into the future

generations (Akinwale *et al*., 2011). Agrobase estimated the heritability for this dataset using the broad-sense heritability equation, V_g/V_p (Falconer & Mackay, 1996). The genetic variation was based on the variation of the entries and the environmental variation was based on the variation of the blocks, i.e. the variation of the entries over the different plots. The H^2-values within this study ranged between 0.110 and 0.928 (Table 4.21). The low H^2-values observed for some traits within this study indicates that directly selecting for those traits will be ineffective (Akinwale *et al*., 2011).

It should also be noted that when no data is mentioned for the NNA analysis, no data was generated for that analysis. This is due to the NNA analysis generating the same data as the RCBD analysis and therefore no data was reported by NNA, thus, no trends were observed and therefore the data remained the same. Hence, the results mentioned was generated from the RCBD analysis only.

The results generated for the TKW trait showed desirable results with a CV-value of 2.85%, R^2-value of 0.9521 and H^2-value of 0.928 (Table 4.21). These values suggest that this trait is a reliable trait to select for in breeding programmes as one of the yield-determining traits to achieve higher yield gains. The best performing genotypes for the TKW trait was identified as entries 64 (53.00g), 36 (52.00g), 37 (52.00g), 28 (48.67g) and 58 (48.00g) (Table 4.5). The average of this trait was 42.636g and these five genotypes were identified to be statistically significantly better than the average. Therefore, these genotypes are good candidates for future use as male crossing parents that possess the TKW trait.

Table 4.5: Best performing high-yielding genotypes for TKW

TKW			
Rank	Entry Number	Average	CV (%)
1	64	53.00	1.9
2	36	52.00	1.9
3	37	52.00	1.9
4	28	48.67	3.1
5	58	48.00	3.6

The grain length displayed desirable results as the CV, R^2 and H^2 for this trait was 1.47%, 0.9179 and 0.876, respectively (Table 4.21). Due to these values, it is evident that this trait can be selected for in breeding programmes as a reliable yield-determining trait that can contribute

to achieving increased yield. The best performing genotypes for grain length was identified to be entries 37 (5.29mm), 64 (5.19mm), 36 (5.18mm), 48 (5.17mm) and 49 (5.06mm) (Table 4.6). The average generated for this trait was 4.728mm and from this, it was identified that all the five entries were statistically significantly better than the average. These five genotypes that display low CV-values are therefore good candidates for the selection as male crossing parents that possess the grain length trait.

Table 4.6: Best performing high-yielding genotypes for grain length

Grain Length			
Rank	Entry Number	Average	CV (%)
1	37	5.29	1.8
2	64	5.19	2.2
3	36	5.18	1.3
4	48	5.17	0.6
5	49	5.06	0.2

The specific weight trait was analysed for both the RCBD and NNA analyses. The results obtained from the RCBD analysis was favourable as the CV-value generated was 0.67%, the R^2-value generated was 0.9138 and the H^2-value generated was 0.868 for this trait (Table 4.21). The NNA analysis adjusted the results obtained by the RCBD analysis. The CV-value adjusted to 0.654% and the H^2-value adjusted to 0.958; and thus, making specific weight the most heritable trait that has very little to no genotype x environment interaction. These values indicate that this trait should be selected for in breeding programmes as it is a very reliable and stable yield-determining trait that can contribute to achieving higher yield. When considering the five best performing genotypes for specific weight, it was identified to be entries 13 (85.93kg/hl), 32 (85.87 kg/hl), 5 (85.80 kg/hl), 21 (85.80 kg/hl) and 9 (85.60 kg/hl) (Table 4.7). The average for specific weight was identified to be 84.033kg/hl and the five entries displayed averages that were statistically significantly better than the average. Therefore, good candidates for future use as male crossing parents that possess the specific weight trait are the five entries in table 4.7.

Table 4.7: Best performing high-yielding genotypes for specific weight

Specific Weight			
Rank	Entry Number	Averages	CV (%)
1	13	85.93	0.2
2	32	85.87	0.4
3	5	85.80	0.4
4	21	85.80	0.8
5	9	85.60	0.4

The grain area trait showed desirable results as the CV, R^2 and H^2 for this trait was 2.77%, 0.9065 and 0.859, respectively (Table 4.21). Due to these results, this trait can be a reliable trait to select for in breeding programmes. The best performing genotypes for grain area was identified to be entries 64 (10.17mm^2), 37 (10.07 mm^2), 36 (9.70 mm^2), 16 (9.09 mm^2) and 63 (9.01 mm^2) (Table 4.8). The average obtained for this trait was 8.424mm^2 and it was then identified that the five entries in table 4.8 are statistically significantly better than the average. These genotypes can therefore be selected for future use as male crossing parents that possess the grain area trait.

Table 4.8: Best performing high-yielding genotypes for grain area

Grain Area			
Rank	Entry Number	Average	CV (%)
1	64	10.17	2.0
2	37	10.07	1.2
3	36	9.70	1.3
4	16	9.09	1.0
5	63	9.07	1.6

The results displayed for the days to heading trait was 2.48% for the CV-value, 0.8977 for the R^2-value and 0.846 for the H^2-value (Table 4.21). These values obtained is an indication that this trait can be selected for in breeding programmes as it displays desirable results and therefore, this trait can contribute to increased yield. The five best performing genotypes for days to heading was identified to be entries 57 (99.33 days), 61 (100.33 days), 59 (100.33 days), 48 (100.67 days) and 39 (101.00 days) (Table 4.9). The days to heading average obtained

by the analysis was 105.104 days, and it was identified that of the best performing genotypes, only the first four entries are statistically significantly better than the average. Therefore, when selecting male crossing parents, only entries 57, 61, 59 and 48 can be considered as good candidates that possess earlier flowering which results in less days to heading.

Table 4.9: Best performing high-yielding genotypes for days to heading

Days to Heading			
Rank	Entry Number	Average	CV (%)
1	57	99.33	1.5
2	61	100.33	1.5
3	59	100.33	1.5
4	48	100.67	2.1
5	39	101.00	2.6

The grain width trait also formed among the traits that displayed satisfactory results such as a CV-value of 1.94%, a R^2-value of 0.8835 and a H^2- value of 0.825 (Table 4.21). These results can conclude that this trait can be selected for as a reliable yield-determining trait which can be exploited in breeding programmes in order to contribute to increased yield. The best five performing genotypes for grain width was identified to be entry 64 (2.54mm), 37 (2.47mm), 17 (2.45mm), 63 (2.44mm) and 36 (2.44mm) (Table 4.10). The average generated for grain width was 2.314mm and the five best performing genotypes form part of the top 12 entries that are statistically significantly better than the average. Therefore, the selection of entries 64, 37, 17, 63 and 36 can be selected as male crossing parents that possess the grain width trait.

Table 4.10: Best performing high-yielding genotypes for grain width

Grain Width			
Rank	Entry Number	Average	CV
1	64	2.54	0.7
2	37	2.47	0.9
3	17	2.45	0.9
4	63	2.44	1.1
5	36	2.44	0.2

The plant height trait showed moderate results for the values of the CV (6.63%), R^2 (0.645) and H^2 (0.438) (Table 4.21). These results are an indication that a certain degree of genotype x environment interaction occurred and that the transmissibility of the trait into future generations are less than 50%, which is not a good heritability estimate. The R^2-value also indicates that the RCBD analysis can to some extent explain the variance in the experiment but not very well. Therefore, based on these values, it could be suggested that this trait is not very reliable to select for, but since molecular markers for the *Rht* genes have been developed (Ellis *et al*., 2002), genotypic screening can assist in the selection for this trait as plant height is an important trait when breeding for increased yield, since the green revolution was based on reduced plant height. The five best performing genotypes for plant height was identified to be entries 48 (846.85mm), 64 (842.59mm), 50 (832.95mm), 43 (829.69mm) and 49 (826.32mm) (Table 4.11). The plant height average was displayed to be 726.124mm and the five best performing entries were identified as statistically significantly better than the average. Thus, when selecting for male crossing parents that possess the plant height trait, the entries 64, 43 and 49 can be considered as these genotypes have the lowest CV-values (Table 4.11).

Table 4.11: Best performing high-yielding genotypes for plant height

Plant Height			
Rank	Entry Number	Average	CV (%)
1	48	846.85	4.9
2	64	842.59	3.3
3	50	832.95	6.8
4	43	829.69	0.0
5	49	826.32	3.5

The yield trait was analysed using both the RCBD and NNA analyses. The yield trait displayed unsatisfactory results as the CV-value generated was 25.16%, the R^2-value generated was 0.6568 and the H^2-value generated was 0.345 (Table 4.21). Evidently, these results display that the yield trait is largely influenced by the genotype x environment interaction and generates a low heritability value which is in accordance with other studies (McIntyre *et al*., 2010; Wu *et al.*, 2012). Therefore, the selection of yield-determining traits as an indirect selection for yield are implemented in breeding programmes. However, the NNA analysis adjusted the results obtained by the RCBD analysis. The CV-value was adjusted to 25.091% and the H^2-value was adjusted to 0.736; and thus, the yield trait obtained a higher heritability value due to observed

field trends analysed by NNA. This H^2-value implies that this trait can be transmissible into future generations, however, the CV-value adjusted was still high and therefore confirms that strong genotype x environment interaction occurred. Thus, this trait can be selected for, but along with additional yield-determining traits in breeding programmes for the aim of achieving increased yield. The five best performing genotypes for yield was identified to be entries 26 (3.02kg/ha), 50 (2.91 kg/ha), 41 (2.87 kg/ha), 48 (2.81 kg/ha) and 64 (2.73 kg/ha) (Table 4.12). The average for this trait was 1.913kg/ha and all five entries were statistically significantly better than the average. However, among the five best performing genotypes, only entry 48 should be considered as a good candidate for future use as a male crossing parent for the yield trait, as this entry possess the lowest CV-value (10.9%) among the five entries.

Table 4.12: Best performing high-yielding genotypes for yield

Yield			
Rank	Entry Number	Average	CV (%)
1	26	3.02	26.5
2	50	2.91	22.4
3	41	2.87	17.7
4	48	2.81	10.9
5	64	2.73	23.2

The results obtained for the spike length trait suggested that this trait was unreliable to consider for selection in a breeding programme. This is due to the CV-value of 9.36%, the R^2- value of 0.5815 and the H^2-value of 0.313 as these values are an indication of whether a trait is stable enough for successful inheritance and are not strongly influenced by the genotype x environment interaction (Table 4.21). The best performing genotypes for spike length was identified to be entries 19 (96.33mm), 27 (95.67mm), 16 (94.33mm), 47 (92.33mm) and 48 (91.67mm) (Table 4.13). The trait average generated was 80.594mm and all five genotypes was shown to be statistically significantly better than the average. Good candidates to be considered for future use as male crossing parents for spike length are entries 19 and 48, as they possess low CV-values of 3.3% and 2.5%, respectively.

Table 4.13: Best performing high-yielding genotypes for spike length

Spike Length			
Rank	Entry Number	Average	CV (%)
1	19	96.33	3.3
2	27	95.67	11.9
3	16	94.33	5.8
4	47	92.33	5.1
5	48	91.67	2.5

The grain number trait has previously been identified by many studies to have a high positive correlation with grain yield, as an increase in grain number is associated with increased grain yield (Fischer, 2011; McIntyre *et al*., 2010; Pedro *et al*., 2012; Reynolds *et al*., 2009b). Due to this, grain number is selected as a yield-determining trait, but the results obtained for grain number was undesirable as the CV obtained was 15.40%, the R^2-value obtained was 0.5631 and the H^2-value obtained was 0.286 (Table 4.21). Based on these values, this trait is largely influenced by the genotype x environment interactions and transmissibility of the trait into future generations are quite low, thus no selection of this trait should occur, but this trait is positively correlated to grain yield and therefore will continue to be selected in breeding programmes for the achievement of increased yield. Although, the results obtained could possibly be explained by the fact that the measurement of the grain number trait was performed on hand threshed tillers and due to human error caused this trait to statistically deviate. The five best performing genotypes for grain number was identified to be entries 23 (62.31), 57 (61.12), 13 (59.18), 15 (58.49), 14 (58.38) (Table 4.14). The average generated for this trait was 48.263 and it was identified that only entries 23 and 57 were statistically significantly better than the average. Therefore, these two genotypes can be considered to be selected as male crossing parents that possess the grain number trait.

Table 4.14: Best performing high-yielding genotypes for grain number

Grain Number			
Rank	Entry Number	Average	CV (%)
1	23	62.31	11.2
2	57	61.12	13.4
3	13	59.18	7.3
4	15	58.49	17.6
5	14	58.38	21.7

81

The floret fertility trait displayed undesirable results as the values of CV, R^2 and H^2 was 7.63%, 0.5363 and 0.284, respectively (Table 4.21). Therefore, this trait statistically displayed that it is not a reliable yield-determining trait to select for when attempting to increase yield. However, one of the aspects to consider regarding this trait is that the determination of floret fertility might have been done too late as the ideal stage for the measurement of this trait is when the primordia has formed for florets and therefore this could have been the reason for the poor statistical results obtained for this trait. The best performing genotypes for floret fertility was identified to be entries 13 (0.87), 40 (0.83), 57 (0.80), 28 (0.79), 21 (0.79) (Table 4.15). The floret fertility average was displayed to be 0.725 and within the five entries, only two entries were statistically significantly better than the average which were entries 13 and 40, which also happen to have moderate CV-values (Table 4.15). These two genotypes can therefore statistically be considered as good candidates to select for as male crossing parents for the floret fertility trait.

Table 4.15: Best performing high-yielding genotypes for floret fertility

Floret Fertility			
Rank	Entry Number	Average	CV (%)
1	13	0.87	5.2
2	40	0.83	5.0
3	57	0.80	8.8
4	28	0.79	6.3
5	21	0.79	6.7

The values of the CV, R^2 and H^2 for the protein (dwb) trait was 6.72%, 0.5235 and 0.270, respectively (Table 4.21). These values assist in determining whether a trait is worth selecting for and for the protein trait, due to its low H^2-value, the selection of this trait would be ineffective and result in no contribution to increasing yield, therefore this trait should not be selected in breeding programmes. The best performing genotypes for protein was identified to be entries 65 (15.96%), 45 (15.96%), 38 (15.67%), 6 (15.30%) and 60 (15.30%) (Table 4.16). The protein average was identified to be 14.103% and it was then identified that only entries 65, 45 and 38 is statistically significantly better than the average. However, only entry 45 can statistically be considered as a good candidate for selection as a male crossing parent for protein, due to its low CV-value (Table 4.16).

Table 4.16: Best performing high-yielding genotypes for protein

Protein (dry weight basis)			
Rank	Entry Number	Average	CV
1	65	15.96	10.1
2	45	15.96	4.9
3	38	15.67	15.2
4	6	15.30	3.3
5	60	15.30	1.1

The harvest index trait displayed a moderate CV-value (6.80%) and R^2-value (0.5372), but the H^2-value (0.266) was quite low (Table 4.21). Thus, the selection for harvest index should not be considered as this trait is not a reliable yield-determining trait to select for when attempting to increase yield. The five best performing genotypes for harvest index was identified to be entries 39 (0.54), 23 (0.54), 29 (0.54), 33 (0.53) and 12 (0.52) (Table 4.17). The average of the trait was identified to be 0.485 and when selecting for the entries that was statistically significantly better than the average, entries 39, 23, 29 were identified. However, only the two entries, 39 and 23, can be statistically considered as male crossing parents for harvest index. These entries 39 and 23 have CV-values that are much lower than entry 29 and thus, displayed to have less genotype x environment interaction.

Table 4.17: Best performing high-yielding genotypes for harvest index

Harvest Index			
Rank	Entry Number	Average	CV (%)
1	39	0.54	4.9
2	23	0.54	3.1
3	29	0.54	13.6
4	33	0.53	0.6
5	12	0.52	3.7

The spikelet number trait generated resulted in a CV-value of 7.47%, a R^2-value of 0.5926 and a H^2-value of 0.262 (Table 4.21). These results, evidently displays that the spikelet number trait cannot be considered as a reliable yield-determining trait that will contribute to increasing yield within breeding programmes. The five best performing genotypes for spikelet number was identified to be entries 16 (17.81), 57 (17.35), 6 (17.24), 53 (17.16) and 47 (17.10) (Table

4.18). The average for spikelet number was generated to be 15.805 and only entry 16 was identified to be statistically significantly better than the average. This genotype also possesses a low CV-value and can thus be statistically considered for selection as a male crossing parent that possesses the spikelet number trait.

Table 4.18: Best performing high-yielding genotypes for spike number

Spike Number			
Rank	Entry Number	Average	CV (%)
1	16	17.81	3.7
2	57	17.35	4.8
3	6	17.24	6.9
4	53	17.16	3.9
5	47	17.10	3.8

The results for the tiller number trait displayed a high CV-value (19.44%), a moderate R^2-value (0.4845) and a low H^2-value (0.158) (Table 4.21). Due to these values, this trait displays unreliability as this trait is strongly influenced by genotype x environment interactions and displays a low H^2-value. Thus, selection of the tiller number trait would be ineffective in contributing to the aim of increasing yield. The five best performing genotypes for tiller number was identified to be entries 9 (5.00), 62 (4.78), 28 (4.78), 40 and 7 (4.56) (Table 4.19). The trait average was displayed to be 3.614 and when selecting genotypes that are statistically significantly better than the average, the three entries 9, 62 and 28 was identified. However, none of these genotypes are good candidates to select for as male crossing parents for tiller number as the CV-values for the genotypes indicate large genotype x environment interaction (Table 4.19). Even though, a molecular marker was identified for this trait, it is evident that selecting for this trait will be ineffective and have little to no contribution to yield.

Table 4.19: Best performing high-yielding genotypes for tiller number

Tiller Number			
Rank	Entry Number	Average	CV (%)
1	9	5.00	17.6
2	62	4.78	17.6
3	28	4.78	28.2
4	40	4.56	11.2
5	7	4.44	8.7

The results generated for the grain weight trait was 26.68% for the CV-value, 0.5019 for the R^2-value and 0.110 for the H^2-value (Table 4.21). These results suggest that this trait should not be selected in breeding programmes as it is largely influenced by genotype x environment interaction and the low heritability suggests that the selection of this trait will be ineffective. However, grain weight which can also be expressed as TKW displayed contradicting results as TKW is highly heritable with little genotype x environment interaction. This contradiction could possibly be explained by the fact that the measurement of this grain weight trait was performed on hand threshed tillers and due to human error caused this trait to statistically deviate. The five best performing genotypes for grain weight was identified to be 23 (9.87g), 22 (9.67g), 9 (9.36g), 35 (9.24g) and 16 (9.17g) (Table 4.20). The average for grain weight was 6.981g and it was identified that no entry was significantly better than the average. These genotypes are therefore not good candidates to select as male crossing parents since no genotype is significantly better than the average.

Table 4.20: Best performing high-yielding genotypes for grain weight

Grain Weight			
Rank	Entry Number	Average	CV (%)
1	23	9.87	2.9
2	22	9.67	51.9
3	9	9.36	17.3
4	35	9.24	36.5
5	16	9.17	20.5

The results obtained from Agrobase for the yield-determining traits could be used to determine a new selection of high-yielding genotypes that possess traits that has little genotype x environment interactions and have high heritability values to be selected as future male crossing parents. The traits that was identified to have high heritability and little genotype x environment interaction was TKW, grain length, specific weight, grain area, days to heading and grain width. Comparing the best performing genotypes for each of these traits, it was discovered that entry 36, 37 and 64 are found to be part of the best performing genotypes for TKW, grain length, grain area and grain width. Additionally, entry 64 was also found to form part of the best performing genotypes for grain width, plant height and yield. Interestingly, the specific weight trait which was identified as the most stable trait had no genotypes in common with the other stable traits that also possess high heritability and therefore, no selections were

made from the best performing genotypes for specific weight. It can thus be concluded that entries 36, 37 and 64 possess desirable yield-determining traits and should be selected as male crossing parents within the pre-breeding programme in order to contribute to achieving higher grain yield.

The Agrobase results obtained for the four wheat lines (entries 4, 9, 26 and 64) selected as male crossing parents for the MS-MARS cycle 2 was also observed. Entry 4 was found to be missing from all the best performing genotypes of all the traits and thus, does not possess strong association with any yield-determining traits and therefore, is not a good candidate to be selected as a male crossing parent. Entry 9 was found to be one of the best performing genotypes for specific weight and tiller number. It was identified that specific weight is a very reliable trait and thus, entry 9 could be considered to be selected as a male crossing parent that possesses the specific weight trait. Entry 26 was found to be one of the best performing genotypes for yield only, but since it has a CV of 26.5% which suggests strong genotype x environment interaction, this entry is not a good candidate to be selected as a male crossing parent. Entry 64, by far out-performed these three entries as entry 64 was found to form part of the best performing genotypes for TKW, grain length, grain area, grain width, plant height and yield. In conclusion, when reviewing these entries once again, selection for only entry 64 as the only male crossing parent would be made as it is associated with many traits that are highly heritable and able to contribute in increasing yield.

The marker data of the three best performing entries, 36, 37 and 64 displayed that entry 36 did not contain any favourable alleles for all three molecular markers associated with TKW. Entry 37, contained two favourable alleles for the genes *TaGS5* and *TaGW2* that are associated with TKW and grain width. However, entry 64 only contained one favourable allele for the *TaGS-D1* gene that is associated with TKW and grain length. This could possibly indicate that these genes that are associated with TKW do not necessarily have the same function within these genotypes as within the study they were identified from. Therefore, these three entries are selected as the male crossing parents to be used within the next MS-MARS cycle for further incorporation of yield-determining traits to achieve higher grain yield.

The results obtained from the RCBD analysis (Table 4.21) was used to identify traits that have relatively high heritability and associated with a specific molecular marker identified through literature. These traits are then correlated with the molecular marker in order to identify the relationship that exists between trait and marker within these high-yielding genotypes.

86

Table 4.21: Summary of the RCBD results obtained for the yield-determining traits that was measured.

Trait	Grand Mean	R-squared	C.V	LSD for ENTRY	S.E.D	p-value	Heritability	t (2-sided)	MSE
TKW (g)	42,636	0,9521	2,85%	2,3412	0,9938	0.0000	0,928	2,3558	1,48141
Grain Length (mm)	4,728	0,9179	1,47%	0,1335	0,0567	0.0000	0,876	2,3558	0,00481
Specific Weight (kg/hl)	84,033	0,9138	0,67%	0,9137	0,4617	0.0000	0,868	1,9790	0,31972
Grain Area (mm²)	8,424	0,9065	2,77%	0,4491	0,1906	0.0000	0,859	2,3558	0,05450
Days to heading	105,104	0,8977	2,48%	4,2102	2,1275	0.0000	0,846	1,9790	6,78935
Grain Width (mm)	2,314	0,8835	1,94%	0,0861	0,0366	0.0000	0,825	2,3558	0,00201
Plant Height (mm)	726,124	0,645	6,63%	77,7892	39,2786	0.0000	0,438	1,9804	2314,21559
Yield (kg/ha)	1,913	0,6568	25,16%	0,7777	0,393	0.0000	0,345	1,9790	0,23167
Spike Length (mm)	80,594	0,5815	9,36%	12,1988	6,1596	0.0000	0,313	1,9804	56,91177
Grain Number	48,263	0,5631	15,40%	12,0153	6,067	0.0001	0,286	1,9804	55,21222
Floret Fertility (%)	0,725	0,5363	7,63%	0,0895	0,0452	0.0002	0,284	1,9806	0,00307
Protein (dry weight basis) (%)	14,103	0,5235	6,72%	1,5320	0,7741	0.0002	0,270	1,9790	0,89896
Harvest Index (%)	0,485	0,5372	6,80%	0,0533	0,0269	0.0004	0,266	1,9806	0,00109
Spike Number	15,805	0,5926	7,47%	1,9095	0,9642	0.0004	0,262	1,9806	1,39445
Tiller Number	3,614	0,4845	19,44%	1,1362	0,5737	0.0201	0,158	1,9804	0,49375
Grain Weight (g)	6,981	0,5019	26,68%	3,0124	1,521	0.0743	0,110	1,9806	3,46996

4.3.2.   Phenotyping with the use of image-based analysis

4.3.2.1. Validation of SeedCounter application

The validation of the SeedCounter application imaging was executed by capturing 10 digital images of the same orientation of three different seed entries, in order to identify whether there was any significant difference between the measured results for each image in order to confirm repeatability of imaging between measurements.

The measurement generated from each image captured with the SeedCounter (Andriod) application for the grain length, width and area were analysed in Microsoft Excel (2016) by performing a single factor ANOVA for each trait, which resulted in p-values that were > 0.05 (Table 4.22) (Microsoft, 2016). Therefore, no significant difference existed between the results of the measurements of the different images captured with the application. Thus, SeedCounter was validated to be able to use as a phenotyping tool for measuring grain length, width and area, which is in agreement with the findings by Komyshev *et al.* (2017).

Table 4.22: The p-value results obtained from the ANOVA for validation of the SeedCounter application

Entry	Orientation of the Seeds								
	1			2			3		
	Length	Width	Area	Length	Width	Area	Length	Width	Area
1	1.0000	1.0000	0.9996	1.0000	1.0000	0.9999	0.9998	1.0000	0.9996
2	0.9996	1.0000	0.9980	1.0000	1.0000	0.9998	1.0000	1.0000	1.0000
3	0.9981	0.9989	0.9736	0.9586	0.7317	0.2950	1.0000	1.0000	1.0000

4.3.2.2. SeedCounter and 1KK application results

SeedCounter was used to measure the high-yielding genotypes in three replicates. The measurements that are analysed through the images captured are stored in a tab separated values (tsv) file. SeedCounter gives each trait measurement for each seed counted on the image. This information was then combined into a table containing the average of grain length, width, area and TKW for each replicate of each genotype.

The measurements for grain length, width, area and TKW were analysed using Agrobase Generation II version 34.4.38 (Agronomix Software, Winnipeg, Canada). The CV values for the traits were quite low, with values ranging between 1.47% and 2.85% (Table 4.21). The R^2

values and H^2 values were quite high, where the R^2 values ranged between 0.8835 to 0.9521 and the H^2 values ranged between 0.825 to 0.928. Therefore, the data obtained is quite reliable when considering the values for CV, R^2 and H^2. Thus, this data was used to correlate with the molecular markers associated with yield-determining traits that were previously identified.

The 1KK application is known to be able to analyse the average grain length, width, area and seed count of an image captured. The Elane USB scale can also be coupled with the application to weight the seeds in order to calculate the TKW. However, an error in the application was found during the testing of the application. The seed count was discovered to be inaccurate for the summarised and raw data. Although, the seed count on the actual image identifies the seed count accurately as can be seen on figure 4.12, but the summarised data calculates the seed count as 6-fold more than on the actual image (Table 4.23).

With this discovery; communication with Trevor Rife, project designer and programmer of 1KK, was exchanged through the 1KK GitHub profile and it was confirmed that the awareness of the inaccuracy of the seed count is known. However, as per communication, the development of an improved counting algorithm is currently in progress and the estimated time for the updated software is within the first few months of 2018, but an exact time could not be given. 1KK could not be validated to be used presently as a measurement for grain morphology and thus, the validation and use of 1KK as a phenotyping tool will have to be delayed until the new software update is made. This is however an important observation since it highlights the need for validation of new software.

Table 4.23: The summarised data set obtained from the 1KK application.

Sample ID	Seed count	Weight	Length avg	Length var	Length cv	Width avg	Width var	Width cv	Area avg	Area var	Area cv
1.1	615	4.0	6,014	0,196	0,074	3,481	0,116	0,098	15,550	4,186	0,132
1.2	619	4.0	5,883	0,191	0,074	3,321	0,159	0,120	14,715	6,001	0,166

Figure 4.12: An image that was generated with 1KK application for one sample that was tested.

4.3.3.   The relationship between the molecular markers and yield-determining traits

The data analysed by Agrobase for the yield-determining traits was first compared to the data of the molecular markers associated with specific yield-determining traits. No correlation or pattern could be identified and further analyses were then executed as some molecular markers were assumed to be correlated to more than one trait. Further measurements of grain length, width, area and TKW (once again) were measured with the SeedCounter application.

The data of the yield-determining traits and the molecular marker data were then evaluated in Agrobase to identify the correlations between trait and marker. The correlation matrix was analysed on Agrobase for the following traits: specific weight, days to heading, tiller number, grain length, grain width, grain area and TKW, as these traits are assumed to be associated to these molecular markers: *Ppd-D1, tin1, TaGS5, TaGS_D1, TaGW2* markers (Table 4.24). To validate the correlations that were analysed through Agrobase, the same dataset was analysed with Data Analysis on Microsoft Excel (data not shown) (Microsoft, 2016). The correlation values for both methods produced similar values.

The correlations of the traits to its respective molecular markers are shown in table 4.25. The days to heading trait displayed a non-significant negative correlation with the *Ppd_D1* marker. The tiller number trait displayed a significant negative correlation with the *tin1* marker. This

90

suggests that an increase in tiller number would result in a decrease in the allele frequency of the *tin1* gene, but this is the desired result as this marker is associated with lower tiller number (Kumar *et al*., 2015).

The TKW trait displayed non-significant negative correlations with the *TaGS5* and the *TaGW2* markers, but displayed a significant negative correlation with the *TaGS_D1* marker. This significant negative correlation suggests that an increase in TKW would result in a decrease in the allele frequency of the *TaGS-D1a* allele. This contradicts Zhang *et al*. (2014b) that reported, increased TKW is associated with the *TaGS-D1a* favourable allele. Another study displayed that the frequency for the favourable allele within Pakistan wheat landraces was quite low, but within the historical wheat cultivars higher frequencies were found for the favourable allele. The selection pressure for the favourable allele within the historical wheat cultivars during breeding resulted in increased TKW and successfully contributing to cultivar improvement (Rasheed *et al*., 2016). However, within this study, the allele displays a negative association which shows that the favourable allele identified by literature is not favourable within these genotypes investigated. The study investigated by Zhang *et al.* (2014b) used cultivars distributed across Chinese wheat regions to conclude the association between the TKW trait and *TaGS-D1a* allele. However, the global geographic distribution of the *TaGS-D1* gene was also investigated in various countries, but no cultivars were included from Southern Africa (Zhang *et al.,* 2014b). Therefore, once again, no data is available for this allele's prevalence within South African cultivars for comparison to this study.

Specific weight displayed non-significant positive correlations with *TaGS5* (0.0661) and *TaGS_D1* (0.0236) markers but displayed a non-significant negative correlation with the *TaGW2* (-0.0698) marker (Table 4.25). Therefore, no conclusion can be drawn from how these markers are associated with specific weight.

The grain length trait displayed a non-significant negative correlation to the *TaDS_D1* marker and the grain width trait displayed a non-significant positive correlation to the *TaGW2* marker. However, grain width displayed a significant negative correlation to the *TaGS5* marker, suggesting that an increase in grain width results in a decrease in the allele frequency of the *TaGS5-3A-T* favourable allele. The *TaGS5-3A* gene is an ortholog of the rice gene, *OsGS5*, which is associated with increased grain width and since it was identified that high homology exist between these two genes, similar functions were suggested (Ma *et al*., 2016). The study by Wang *et al.* (2015), identified that the *TaGS5-3A-T* allele (named *TaGS5-A1b* in that study)

91

is associated with wider grain width. However, within this study, the allele displays a negative association with grain width, which suggests that the favourable allele identified by literature is not favourable for this trait within these genotypes investigated.

These molecular markers do not display the same function for TKW and other related traits as was identified in literature, therefore additional literature regarding TKW should be identified. A recent genome-wide association study (GWAS) was reported by Sukumaran *et al.* (2017) in which identification of positive association between TKW and grain yield was investigated as well as the identification of specific markers that can be utilised to maximise TKW and grain number (as these two traits are usually negatively correlated). A common locus found in chromosome 6A, was identified for grain yield and TKW as well as two common loci found in chromosome 5A and 2B, for TKW and grain number. It was also identified that a combination of markers was optimised that could explain 32% of grain yield as compared to the 27% of the marker for grain yield only. This combination of markers indicates that grain yield components such as TKW and grain number are important when selecting for grain yield within breeding programmes (Sukumaran *et al.*, 2017).

Table 4.24: Correlations between various yield-determining traits and high-yielding molecular markers generated by Agrobase

	AREA	DAYS TO HEAD	GRAIN LENGTH	PPD_D1	SPECIFIC WEIGHT	TAGS5	TAGS_D1	TAGW2	TKW	TIN1	TILLER NUMBER	GRAIN WIDTH
DAYS TO HEAD	-0.0299	1										
GRAIN LENGTH	-0.5541***	0.0786	1									
PPD-D1	-0.0340	-0.0725	-0.0195	1								
SPECIFIC WEIGHT	-0.0432	-0.1356	0.1338	0.0092	1							
TAGS5	0.2744***	-0.1039	-0.1070	-0.0685	0.0661	1						
TAGS_D1	0.0524	0.0804	-0.0110	0.1231	0.0236	0.0281	1					
TAGW2	0.0224	0.1387	-0.1745*	-0.1809*	-0.0698	0.1442*	-0.0881	1				
TKW	-0.4193***	-0.0938	0.6242***	0.1564*	0.1409	-0.1334	-0.1436*	-0.0026	1			
TIN1	0.0575	0.1692*	-0.2388***	0.0864	-0.0787	-0.1682*	-0.0881	0.2262**	0.0412	1		
TILLER NUMBER	-0.1075	0.1091	0.2658***	0.2692***	0.2278**	-0.0406	0.1661*	-0.1827*	0.3843***	-0.1921**	1	
GRAIN WIDTH	-0.3665***	-0.1098	0.3783***	0.2396***	0.0998	-0.1642*	-0.1232	0.0954	0.8289***	0.1699*	0.4126***	1

* Significant at <0.05 level, ** Significant at <0.01 level, ***Significant at <0.001 level

Table 4.25: Correlation of the specific high-yielding traits to its respective molecular marker

Trait	Markers				
	PPD-D1	TIN1	TAGS5	TAGS_D1	TAGW2
DAYS TO HEAD	-0,0725	0,1692*	-0,1039	0,0804	0,1387
TILLER NUMBER	0,2692***	-0,1921**	-0,0406	0,1661*	-0,1827*
TKW (g)	0,1564*	0,0412	-0,1334	-0,1436*	-0,0026
SPECIFIC WEIGHT (kg/hl)	0,0092	-0,0787	0,0661	0,0236	-0,0698
GRAIN LENGTH (mm)	-0,0195	-0,2388***	-0,1070	-0,0110	-0,1745*
GRAIN WIDTH (mm)	0,2396***	0,1699*	-0,1642*	-0,1232	0,0954

* Significant at <0.05 level, ** Significant at <0.01 level, ***Significant at <0.001 level

-Highlighted cells are the high-yielding traits related to its corresponding molecular marker.

Chapter 5:   Conclusion

The aim of this study was the assessment of high-yielding genotypes through validating yield-determining traits by utilising genotypic and phenotypic screening. These high-yielding genotypes were also selected as male crossing parents within the male-sterility marker-assisted mediated recurrent selection breeding (MS-MARS) scheme for the improvement of grain yield.

The identification of yield-determining traits as well as molecular markers associated with some of the yield-determining traits were identified through reviewing literature. The molecular markers were validated through genotypic screening and each yield-determining trait was phenotypically screened and statistically analysed for identification of best performing genotypes. The relationship between marker and trait was also analysed through data obtained from genotypic and phenotypic screening. Image-based technologies were also validated for future use as a phenotyping tool to measure grain morphology. The MS-MARS scheme used the high-yielding genotypes as male crossing parents for the transfer of high-yielding traits into the recurrent population to achieve increased grain yield.

The recurrent population of the MS-MARS cycle 1 and 2 were successfully characterised with the durable rust resistance molecular markers. The allele frequency increased slightly for the *Lr34* gene between successive cycles but decreased for the *Sr2* gene. This is most probably due to the male population used within these MS-MARS cycles as this population are dominant for high-yielding traits and not for rust resistance which is evident from the low frequencies observed within the male population for the rust resistance genes. The low frequency for the *Sr2* gene could be due to deselecting wheat lines that display the phenotypic characteristic of *PBC* which is associated with the *Sr2* gene. This phenotype is associated with a decrease in yield and are therefore deselected by breeders when selecting for yield, thus explaining the low presence of this durable rust resistance gene within the male population.

The male population was also molecularly characterised with selected molecular markers associated with yield-determining traits identified through literature. The molecular markers could successfully discriminate between the presence and absence of the desired gene that is associated with its' respective yield-determining trait. One molecular marker could not be validated for its' association with tiller number. However, when observing the statistical analysis of the phenotypic data, tiller number was identified to be strongly influenced by genotype x environment interaction and had low heritability values. Therefore, further

94

investigation on a reliable molecular marker for this trait would not be recommended. The validated molecular markers for the *TaGS5*, *TaGW2*, *TaGS-D1* and *Ppd-D1* genes can be used in MAS to identify the yield-determining traits; TKW, grain length, grain width and earlier flowering within any given population.

Selection of male sterile (female) plants was made from the 1:1 segregating recurrent population for cross-pollination with the high-yielding genotypes as the male population. The successful cross-pollination for the MS-MARS cycle 1 and 2 averaged to 23,88% and 23,95%, respectively. This low average could be due to various factors which include the selection of tillers at the wrong reproductive time and the selection of tillers with narrow glumes instead of wide glumes. The height at which the male tillers are placed could also influence the low cross-pollination and therefore, male tillers should not be placed too high above the female tillers. Although, the average for cross-pollination was low, a large number of F_1 seeds were produced for both cycles. This could be due to the large number of female tillers that were cut during both cycles, therefore even though little seeds were produced per female tiller, many seeds were produced among the larger set of female tillers per cutting sessions.

The 1:1 ratio for male sterile and male fertile plants deviated from the goodness fit to ratio for the overall benches of both cycles. This could be explained by an error that occurred when recording the ratios of male fertile to male sterile plants per pot. Plants could possibility have been falsely selected for fertile tillers as some sterile plants may have been fertilised by fertile male plants within the greenhouse. The possibility of partial sterile plants could also have occurred due to elevated temperatures which could also result in falsely recording sterile plants as fertile plants. The confirmation of the *Ms3* gene controlling male sterility could not be confirmed since the goodness fit to ratio displayed a p-value of less than 0.05. Nevertheless, the *Ms3* gene has been confirmed to be a single dominant gene that controls male sterility by previous studies. Therefore, when executing the MS-MARS scheme, proper recordings of the ratio for male sterile to male fertile plants should be taken.

The use of image-based technology was executed and validated for two applications, SeedCounter and 1KK. Both these applications can be used to measure grain length, width and area; but only SeedCounter was validated to be used as a phenotyping tool for grain length, width and area. The 1KK application, however, could not be validated for use as the seed count was discovered to be inaccurate. Therefore, the results obtained with the SeedCounter

application was used as the measurement for grain length, width and area of the high-yielding genotypes.

The phenotypic assessment of the high-yielding genotypes was statistically analysed and the yield results were used to select four high-yielding genotypes as male crossing parents for cycle 2 of the MS-MARS scheme. However, for the selection of the male crossing parents for the next cycle, the statistically analysed results were used to identify the best performing genotypes within this study and whether the previous four genotypes should remain as the male crossing parents. It was evident from the stable traits that were highly heritable and displayed little genotype x phenotype interactions that the common entries were 36, 37 and 64. These genotypes are selected as male crossing parents that possess desired high-yielding traits to be used in the next MS-MARS cycle to achieve increased grain yield.

The relationship between the molecular marker and its' associated yield-determining trait was statistically analysed and the correlation results showed that many of the molecular markers were significantly negatively correlated with its' associated yield-determining trait. These negative correlations could be due to many circumstances and since geographic distribution investigations on some of the high-yielding genes did not include any Southern Africa cultivars, comparisons with the conclusions obtained in this study could not be made. However, the best performing genotypes (entries 36, 37 and 64) showed to only carry only one or two of the genes (not the same ones) associated with TKW while these genotypes were the best performing genotypes for TKW. This suggests that these high-yielding genes could have a different function within this set of genotypes as compared to literature.

The aim and objectives of this study was achieved as the assessment of the high-yielding genotypes with genotypic and phenotypic screening was successfully executed, as well as the successful execution of the MS-MARS cycles.

The use of image-based technology is an effective, efficient and an economical tool that should be further exploited within the breeding industry and therefore future studies should include validations of more such tools that could largely contribute to breeding programmes. Thus, further studies should also include the validation of the 1KK application once again after the release of the updated application.

Future studies concerning yield should include investigations on the relationship between yield and the molecular markers reported in this study in order to identify which markers to select for and against when molecularly screening future genotypes as the favoured alleles are not

necessarily favoured within the South African cultivars. Once this is confirmed, these molecular markers can form part of the standard panel of markers deployed in the SU-PBL.

Future studies should also include QTL mapping and the identification of additional molecular markers associated with high yield that are prevalent within South African cultivars. Therefore, more studies such as the study by Sukumaran *et al.* (2017) should be implemented in order to contribute to the understanding of the underlying genetic control of the desired phenotypes that contribute to higher grain yield.

Chapter 6:   References

Acosta-Gamboa, L.M., Liu, S., Langley, E., Campbell, Z., Castro-Guerrero, N., Mendoza-Cozatl, D. and Lorence, A., 2017. Moderate to severe water limitation differentially affects the phenome and ionome of Arabidopsis. *Functional Plant Biology*, *44*(1), pp.94-106.

Afonnikov, D.A., Genaev, M.A., Doroshkov, A.V., Komyshev, E.G. and Pshenichnikova, T.A., 2016. Methods of high-throughput plant phenotyping for large-scale breeding and genetic experiments. *Russian journal of genetics*, *52*(7), pp.688-701.

Agarwal, P.K., Agarwal, P., Reddy, M.K. and Sopory, S.K., 2006. Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. *Plant cell reports*, *25*(12), pp.1263-1274.

Akhunov, E.D., Akhunova, A.R., Anderson, O.D., Anderson, J.A., Blake, N., Clegg, M.T., Coleman-Derr, D., Conley, E.J., Crossman, C.C., Deal, K.R. and Dubcovsky, J., 2010. Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes. *BMC genomics*, *11*(1), p.702.

Akinwale, M.G., Gregorio, G., Nwilene, F., Akinyele, B.O., Ogunbayo, S.A. and Odiyi, A.C., 2011. Heritability and correlation coefficient analysis for yield and its components in rice (Oryza sativa L.). *African Journal of plant science*, *5*(3), pp.207-212.

Allen, A.M., Winfield, M.O., Burridge, A.J., Downie, R.C., Benbow, H.R., Barker, G.L., Wilkinson, P.A., Coghill, J., Waterfall, C., Davassi, A. and Scopes, G., 2017. Characterization of a Wheat Breeders' Array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (Triticum aestivum). *Plant biotechnology journal*, *15*(3), pp.390-401.

Anderson, C., Khan, M.A., Catanzariti, A.M., Jack, C.A., Nemri, A., Lawrence, G.J., Upadhyaya, N.M., Hardham, A.R., Ellis, J.G., Dodds, P.N. and Jones, D.A., 2016. Genome analysis and avirulence gene cloning using a high-density RADseq linkage map of the flax rust fungus, Melampsora lini. *BMC genomics*, *17*(1), p.667.

Araus, J.L. and Cairns, J.E., 2014. Field high-throughput phenotyping: the new crop breeding frontier. *Trends in plant science*, *19*(1), pp.52-61.

Atkinson, N.J. and Urwin, P.E., 2012. The interaction of plant biotic and abiotic stresses: from genes to the field. *Journal of experimental botany*, *63*(10), pp.3523-3543.

Badea, A., Eudes, F., Graf, R.J., Laroche, A., Gaudet, D.A. and Sadasivaiah, R.S., 2008. Phenotypic and marker-assisted evaluation of spring and winter wheat germplasm for resistance to fusarium head blight. *Euphytica*, *164*(3), pp.803-819.

Barlow, K.M., Christy, B.P., O'leary, G.J., Riffkin, P.A. and Nuttall, J.G., 2015. Simulating the impact of extreme heat and frost events on wheat crop production: A review. *Field Crops Research*, *171*, pp.109-119.

Beales, J., Turner, A., Griffiths, S., Snape, J.W. and Laurie, D.A., 2007. A pseudo-response regulator is misexpressed in the photoperiod insensitive Ppd-D1a mutant of wheat (Triticum aestivum L.). *Theoretical and Applied Genetics*, *115*(5), pp.721-733.

Beckmann, J.S. and Soller, M., 1986. Restriction fragment length polymorphisms in plant genetic improvement. *Oxford surveys of plant molecular and cell biology*.

Bello, O.B., Ige, S.A., Azeez, M.A., Afolabi, M.S., Abdulmaliq, S.Y. and Mahamood, J., 2012. Heritability and genetic advance for grain yield and its component characters in maize (Zea mays L.). *International Journal of Plant Research*, *2*(5), pp.138-145.

Bennett, D., Reynolds, M., Mullan, D., Izanloo, A., Kuchel, H., Langridge, P. and Schnurbusch, T., 2012. Detection of two major grain yield QTL in bread wheat (Triticum aestivum L.) under heat, drought and high yield potential environments. *Theoretical and Applied Genetics*, *125*(7), pp.1473-1485.

Bierman, A. and Botha, A.M., 2017. A Review of Genome Sequencing in the Largest Cereal Genome, Triticum aestivum L. *Agricultural Sciences*, *8*(02), p.194.

Blum, A., 2005. Drought resistance, water-use efficiency, and yield potential—are they compatible, dissonant, or mutually exclusive?. *Australian Journal of Agricultural Research*, *56*(11), pp.1159-1168.

Boehm Jr, J.D., Ibba, M.I., Kiszonas, A.M., See, D.R., Skinner, D.Z. and Morris, C.F., 2018. Genetic analysis of kernel texture (grain hardness) in a hard red spring wheat (Triticum aestivum L.) bi-parental population. *Journal of Cereal Science*, *79*, pp.57-65.

Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G.L., D'Amore, R., Allen, A.M., McKenzie, N., Kramer, M., Kerhornou, A., Bolser, D. and Kay, S., 2012. Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature*, *491*(7426), pp.705-710.

Brozynska, M., Furtado, A. and Henry, R.J., 2016. Genomics of crop wild relatives: expanding the gene pool for crop improvement. *Plant biotechnology journal*, *14*(4), pp.1070-1085.

Campbell, B.T. and Bauer, P.J., 2007. Improving the precision of cotton performance trials conducted on highly variable soils of the southeastern USA coastal plain. *Plant breeding*, *126*(6), pp.622-627.

Cao, M., Chen, G., Wang, C., Zhen, S., Li, X., Zhang, W., Zeller, F.J., Hsam, S.L.K., Hu, Y. and Yan, Y., 2015. 1Sl (1B) Chromosome substitution in Chinese Spring wheat promotes starch granule development and starch biosynthesis. *Crop and Pasture Science*, *66*(9), pp.894-903.

Chao, S., Dubcovsky, J., Dvorak, J., Luo, M.C., Baenziger, S.P., Matnyazov, R., Clark, D.R., Talbert, L.E., Anderson, J.A., Dreisigacker, S. and Glover, K., 2010. Population-and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (Triticum aestivum L.). *BMC genomics*, *11*(1), p.727.

Cheeseman, J., 2016. Food Security in the Face of Salinity, Drought, Climate Change, and Population Growth. *Halophytes for Food Security in Dry Lands*, pp.111-123.

Chichti, E., George, M., Delenne, J.Y. and Lullien-Pellerin, V., 2015. Changes in the starch-protein interface depending on common wheat grain hardness revealed using atomic force microscopy. *Plant Science*, *239*, pp.1-8.

Collard, B.C. and Mackill, D.J., 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *363*(1491), pp.557-572.

Cossani, C.M. and Reynolds, M.P., 2012. Physiological traits for improving heat tolerance in wheat. *Plant physiology*, *160*(4), pp.1710-1718.

Cox, T.S., Sears, R.G. and Gill, B.S., 1991. Registration of KS87UP9, a winter wheat germplasm segregating for a dominant male-sterility gene. *Crop science*, *31*(1), pp.247-247.

Cox, T.S., Raupp, W.J. and Gill, B.S., 1994. Leaf rust-resistance genes Lr41, Lr42, and Lr43 transferred from Triticum tauschii to common wheat. *Crop science*, *34*(2), pp.339-343.

Cui, F., Zhao, C., Ding, A., Li, J., Wang, L., Li, X., Bao, Y., Li, J. and Wang, H., 2014. Construction of an integrative linkage map and QTL mapping of grain yield-determining traits using three related wheat RIL populations. *Theoretical and applied genetics*, *127*(3), pp.659-675.

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*.

Diaz-Lago, J.E., Stuthman, D.D. and Abadie, T.E., 2002. Recurrent selection for partial resistance to crown rust in oat. *Crop science*, *42*(5), p.1475.

Doyle, J.J. and Doyle, J.L., 1990. Isolation of plant DNA from fresh tissue. *Focus*, *12*, pp. 13-15.

Dreisigacker, S., 2012. Genetic marker systems in wheat breeding. *Physiological breeding I: Interdisciplinary approaches to improve crop adaptation*, pp.140-152.

DuPont, F.M. and Altenbach, S.B., 2003. Molecular and biochemical impacts of environmental factors on wheat grain development and protein synthesis. *Journal of cereal science*, *38*(2), pp.133-146.

Dvorak, J., Luo, M.C. and Akhunov, E.D., 2011. NI Vavilov's theory of centres of diversity in the light of current understanding of wheat diversity, domestication and evolution. *Czech Journal of Genetics and Plant Breeding*, *47*(Special Issue).

Dyck, P.L., Samborski, D.J. and Anderson, R.G., 1966. Inheritance of adult-plant leaf rust resistance derived from the common wheat varieties Exchange and Frontana. *Canadian Journal of Genetics and Cytology*, *8*(4), pp.665-671.

Ellis, M., Spielmeyer, W., Gale, K., Rebetzke, G. and Richards, R., 2002. " Perfect" markers for the Rht-B1b and Rht-D1b dwarfing genes in wheat. *TAG Theoretical and Applied Genetics*, *105*(6), pp.1038-1042.

Falconer, D. and Mackay, T., 1996. *Introduction to quantitative genetics*. Harlow: Pearson, Prentice Hall.

Falk, D.E., 2002. The theory, methods, and results of using recurrent selection in breeding barley. *Proceeding of the EUCARPIA, Cereal Section: From biodiversity to genomics: breeding strategies for small grain cereals in the third millennium., Salsomaggiore, Italy*, pp.21-25.

FAO, Food and Agriculture Organization of the United Nations., 2017. *FAO Cereal Supply and Demand Brief.* [Online] Available at: http://www.fao.org/worldfoodsituation/csdb/en/ [Accessed 7 December 2017].

Farooq, M., Bramley, H., Palta, J.A. and Siddique, K.H., 2011. Heat stress in wheat during reproductive and grain-filling phases. *Critical Reviews in Plant Sciences*, *30*(6), pp.491-507.

Fasoula, V.A. and Tokatlidis, I.S., 2012. Development of crop cultivars by honeycomb breeding. *Agronomy for sustainable development*, *32*(1), pp.161-180.

Feldman, M. and Sears, E.R., 1981. The wild gene resources of wheat. *Scientific American*, *244*(1), pp.102-113.

Figlan, S., Baloyi, T.A., Hlongoane, T., Terefe, T.G., Shimelis, H. and Tsilo, T.J., 2017. Adult plant resistance of selected Kenyan wheat cultivars to leaf rust and stem rust diseases. *Cereal Research Communications*, *45*(1), pp.68-82.

Figlan, S., Le Roux, C., Terefe, T., Botes, W., Visser, B., Shimelis, H. and Tsilo, T., 2014. Wheat stem rust in South Africa: Current status and future research directions. *African Journal of Biotechnology*, *13*(44), pp.4188-4199.

Figueroa, M., Upadhyaya, N.M., Sperschneider, J., Park, R.F., Szabo, L.J., Steffenson, B., Ellis, J.G. and Dodds, P.N., 2016. Changing the game: using integrative genomics to probe virulence mechanisms of the stem rust pathogen Puccinia graminis f. sp. tritici. *Frontiers in plant science*, *7*.

Fischer, R.A., 2011. Wheat physiology: a review of recent developments. *Crop and Pasture Science, 62*(2), pp.95-114.

Foulkes, M.J., Reynolds, M.P. and Sylvester-Bradley, R., 2009. Genetic improvement of grain crops: yield potential. *Crop physiology: applications for genetic improvement and agronomy*, pp.355-385.

Fu, D., Uauy, C., Distelfeld, A., Blechl, A., Epstein, L., Chen, X., Sela, H., Fahima, T. and Dubcovsky, J., 2009. A kinase-START gene confers temperature-dependent resistance to wheat stripe rust. *science*, *323*(5919), pp.1357-1360.

Gao, F., Wen, W., Liu, J., Rasheed, A., Yin, G., Xia, X., Wu, X. and He, Z., 2015. Genome-wide linkage mapping of QTL for yield components, plant height and yield-determining physiological traits in the Chinese wheat cross Zhou 8425B/Chinese Spring. *Frontiers in plant science*, *6*.

Garcia, G.A., Hasan, A.K., Puhl, L.E., Reynolds, M.P., Calderini, D.F. and Miralles, D.J., 2013. Grain yield potential strategies in an elite wheat double-haploid population grown in contrasting environments. *Crop Science*, *53*(6), pp.2577-2587.

Garcia, G.A., Serrago, R.A., González, F.G., Slafer, G.A., Reynolds, M.P. and Miralles, D.J., 2014. Wheat grain number: identification of favourable physiological traits in an elite doubled-haploid population. *Field Crops Research*, *168*, pp.126-134.

Gegas, V.C., Nazari, A., Griffiths, S., Simmonds, J., Fish, L., Orford, S., Sayers, L., Doonan, J.H. and Snape, J.W., 2010. A genetic framework for grain size and shape variation in wheat. *The Plant Cell*, *22*(4), pp.1046-1056.

Girma, K., Martin, K.L., Anderson, R.H., Arnall, D.B., Brixey, K.D., Casillas, M.A., Chung, B., Dobey, B.C., Kamenidou, S.K., Kariuki, S.K. and Katsalirou, E.E., 2006. Mid-season prediction of wheat-grain yield potential using plant, soil, and sensor measurements. *Journal of Plant Nutrition*, *29*(5), pp.873-897.

Golbach, F., Kootstra, G., Damjanovic, S., Otten, G. and van de Zedde, R., 2016. Validation of plant part measurements using a 3D reconstruction method suitable for high-throughput seedling phenotyping. *Machine Vision and Applications*, *27*(5), pp.663-680.

Goutam, U., Kukreja, S., Yadav, R., Salaria, N., Thakur, K. and Goyal, A.K., 2015. Recent trends and perspectives of molecular markers against fungal diseases in wheat. *Frontiers in microbiology*, *6*.

GraphPad software., 2018. *QuickCalcs* [Online] Available at: https://www.graphpad.com/quickcalcs/contingency2/ [Accessed 3 January 2018].

Guo, Z. and Schnurbusch, T., 2015. Variation of floret fertility in hexaploid wheat revealed by tiller removal. *Journal of experimental botany*, *66*(19), pp.5945-5958.

Guo, Z., Chen, D., Alqudah, A.M., Röder, M.S., Ganal, M.W. and Schnurbusch, T., 2017. Genome-wide association analyses of 54 traits identified multiple loci for the determination of floret fertility in wheat. *New Phytologist*, *214*(1), pp.257-270.

Gupta, P.K., Varshney, R.K., Sharma, P.C. and Ramesh, B., 1999. Molecular markers and their applications in wheat breeding. *Plant breeding*, *118*(5), pp.369-390.

Gupta, S.K., Charpe, A., Koul, S., Prabhu, K.V. and Haq, Q.M.R., 2005. Development and validation of molecular markers linked to an Aegilops umbellulata–derived leaf-rust-resistance gene, Lr9, for marker-assisted selection in bread wheat. *Genome*, *48*(5), pp.823-830.

Guzman, C., Peña, R.J., Singh, R., Autrique, E., Dreisigacker, S., Crossa, J., Rutkoski, J., Poland, J. and Battenfield, S., 2016. Wheat quality improvement at CIMMYT and the use of genomic selection on it. *Applied & translational genomics*, *11*, pp.3-8.

Hardy, O.J., 2003. Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. *Molecular Ecology*, *12*(6), pp.1577-1588.

Hartl, D.L. and Clark, A.G., 1998. Principles of population genetics.

Haghighattalab, A., Pérez, L.G., Mondal, S., Singh, D., Schinstock, D., Rutkoski, J., Ortiz-Monasterio, I., Singh, R.P., Goodin, D. and Poland, J., 2016. Application of unmanned aerial systems for high throughput phenotyping of large wheat breeding nurseries. *Plant methods*, *12*(1), p.35.

Hallauer A.R., 1981. Selection and breeding methods, in: Frey K.J. ed., *Plant Breeding II*. Iowa State University Press, Ames, Iowa 50010*,* USA, pp 3–55.

Hawkesford, M.J., Araus, J.L., Park, R., Calderini, D., Miralles, D., Shen, T., Zhang, J. and Parry, M.A., 2013. Prospects of doubling global wheat yields. *Food and Energy Security*, *2*(1), pp.34-48.

Hedden, P., 2003. The genes of the Green Revolution. *TRENDS in Genetics*, 19(1), pp.5-9.

Hou, G.G. ed., 2011. *Asian noodles: Science, technology, and processing*. John Wiley & Sons.

Hou, J., Jiang, Q., Hao, C., Wang, Y., Zhang, H. and Zhang, X., 2014. Global selection on sucrose synthase haplotypes during a century of wheat breeding. *Plant Physiology*, *164*(4), pp.1918-1929.

Hu, M.J., Zhang, H.P., Liu, K., Cao, J.J., Wang, S.X., Jiang, H., Wu, Z.Y., Lu, J., Zhu, X.F., Xia, X.C. and Sun, G.L., 2016. Cloning and characterization of TaTKW-7A gene associated with grain weight in wheat via SLAF-seq-BSA. *Frontiers in plant science*, *7*.

Huang, X.Q., Cloutier, S., Lycar, L., Radovanovic, N., Humphreys, D.G., Noll, J.S., Somers, D.J. and Brown, P.D., 2006. Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (Triticum aestivum L.). *Theoretical and Applied Genetics*, *113*(4), pp.753-766.

Huang, Y.Y. and Deng, J.Y., 1988. Preliminary analyses of the effectiveness of utilization of Taigu genetic male sterile wheat in recurrent selection and complex crossing.

Igrejas, G., Faucher, B., Bertrand, D., Guibert, D., Leroy, P. and Branlard, G., 2002. Genetic analysis of the size of endosperm starch granules in a mapped segregating wheat population. *Journal of Cereal Science*, *35*(1), pp.103-107.

Inostroza-Blancheteau, C., Soto, B., Ibáñez, C., Ulloa, P., Aquea, F., Arce-Johnson, P. and Reyes-Díaz, M., 2010. Mapping aluminum tolerance loci in cereals: A tool available for crop breeding. *Electronic Journal of Biotechnology*, *13*(4), pp.11-12.

Iqbal, M., Navabi, A., Yang, R.C., Salmon, D.F. and Spaner, D., 2007. Molecular characterization of vernalization response genes in Canadian spring wheat. *Genome*, *50*(5), pp.511-516.

Iqbal, M., Shahzad, A. and Ahmed, I., 2011. Allelic variation at the Vrn-A1, Vrn-B1, Vrn-D1, Vrn-B3 and Ppd-D1a loci of Pakistani spring wheat cultivars. *Electronic Journal of Biotechnology*, *14*(1), pp.1-2.

Iqbal, A.M., Lone, A.A., Wani, S.A., Wani, S.H. and Nehvi, F.A., 2013. Pre-breeding and Population Improvement. *LS: International Journal of Life Sciences*, *2*(3), pp.188-197.

IWGSC, International Wheat Genome Sequencing Consortium., 2017. *IWGSC Reference Sequence v1.0 assembly now available at URGI*. [Online] Available at: http://www.wheatgenome.org/News/Latest-news/RefSeq-v1.0-URGI. [Accessed 5 November 2017].

Jensen, N.F., 1970. A diallel selective mating system for cereal breeding. *Crop Science*, *10*(6), pp.629-635.

Jia, J., Zhao, S., Kong, X., Li, Y., Zhao, G., He, W., Appels, R., Pfeifer, M., Tao, Y., Zhang, X. and Jing, R., 2013. Aegilops tauschii draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature*, *496*(7443), pp.91-95.

Jiang, Y., Jiang, Q., Hao, C., Hou, J., Wang, L., Zhang, H., Zhang, S., Chen, X. and Zhang, X., 2014. A yield-associated gene TaCWI, in wheat: its function, selection and evolution in global breeding revealed by haplotype analysis. *Theoretical and applied genetics*, *128*(1), pp.131-143.

Jiang, Y., van Groenigen, K.J., Huang, S., Hungate, B.A., van Kessel, C., Hu, S., Zhang, J., Wu, L., Yan, X., Wang, L. and Chen, J., 2017. Higher yields and lower methane emissions with new rice cultivars. *Global Change Biology*.

Jones, H., Gosman, N., Horsnell, R., Rose, G.A., Everest, L.A., Bentley, A.R., Tha, S., Uauy, C., Kowalski, A., Novoselovic, D. and Simek, R., 2013. Strategy for exploiting exotic germplasm using genetic, morphological, and environmental diversity: the Aegilops tauschii Coss. example. *Theoretical and applied genetics*, *126*(7), pp.1793-1808.

Juliana, P., Rutkoski, J.E., Poland, J.A., Singh, R.P., Murugasamy, S., Natesan, S., Barbier, H. and Sorrells, M.E., 2015. Genome-wide association mapping for leaf tip necrosis and pseudo-black chaff in relation to durable rust resistance in wheat. *The Plant Genome*, *8*(2).

Kamran, A., Iqbal, M. and Spaner, D., 2014. Flowering time in wheat (Triticum aestivum L.): a key factor for global adaptability. *Euphytica*, *197*(1), pp.1-26.

Kang, H.Y., Lin, L.J., Song, Z.J., Yuan, J.Y., Zhong, M.Y., Zhang, H.Q., Fan, X., Sha, L.N., Wang, Y., Xu, L.L. and Zeng, J., 2012. Identification, fine mapping and characterization of Rht-dp, a recessive wheat dwarfing (reduced height) gene derived from Triticum polonicum. *Genes & Genomics*, pp.1-8.

Kaur, J., Fellers, J., Adholeya, A., Velivelli, S.L., El-Mounadi, K., Nersesian, N., Clemente, T. and Shah, D., 2017. Expression of apoplast-targeted plant defensin MtDef4. 2 confers resistance to leaf rust pathogen Puccinia triticina but does not affect mycorrhizal symbiosis in transgenic wheat. *Transgenic research*, *26*(1), pp.37-49.

Kempe, K., Rubtsova, M. and Gils, M., 2014. Split-gene system for hybrid wheat seed production. *Proceedings of the National Academy of Sciences*, *111*(25), pp.9097-9102.

Kirigwi, F.M., Van Ginkel, M., Brown-Guedira, G., Gill, B.S., Paulsen, G.M. and Fritz, A.K., 2007. Markers associated with a QTL for grain yield in wheat under drought. *Molecular Breeding*, *20*(4), pp.401-413.

Kiseleva, A.A., Shcherban, A.B., Leonova, I.N., Frenkel, Z. and Salina, E.A., 2016. Identification of new heading date determinants in wheat 5B chromosome. *BMC plant biology*, *16*(Suppl 1).

Komyshev, E., Genaev, M. and Afonnikov, D., 2017. Evaluation of the SeedCounter, A Mobile Application for Grain Phenotyping. *Frontiers in plant science*, *7*, p.1990.

Krasensky, J. and Jonak, C., 2012. Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks. *Journal of experimental botany*, *63*(4), pp.1593-1608.

Krattinger, S.G., Lagudah, E.S., Spielmeyer, W., Singh, R.P., Huerta-Espino, J., McFadden, H., Bossolini, E., Selter, L.L. and Keller, B., 2009. A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science*, *323*(5919), pp.1360-1363.

Kuchel, H., Fox, R., Reinheimer, J., Mosionek, L., Willey, N., Bariana, H. and Jefferies, S., 2007. The successful application of a marker-assisted wheat breeding strategy. *Molecular Breeding*, *20*(4), pp.295-308.

Kumar, P., Yadava, R.K., Gollen, B., Kumar, S., Verma, R.K. and Yadav, S., 2011. Nutritional contents and medicinal properties of wheat: a review. *Life Sci Med Res*, *22*, pp.1-10.

Kumar, N., Belhumeur, P.N., Biswas, A., Jacobs, D.W., Kress, W.J., Lopez, I.C. and Soares, J.V., 2012. Leafsnap: A computer vision system for automatic plant species identification. In *Computer Vision–ECCV 2012* (pp. 502-516). Springer, Berlin, Heidelberg.

Kumar, S., Singh, S.S., Mishra, C.N., Saroha, M., Gupta, V., Sharma, P., Tiwari, V. and Sharma, I., 2015. Assessment of Tiller Inhibition (tin) Gene Molecular Marker for its Application in Marker-Assisted Breeding in Wheat. *National Academy Science Letters, 38*(6), pp.457-460.

Kumar, A., Mantovani, E.E., Seetan, R., Soltani, A., Echeverry-Solarte, M., Jain, S., Simsek, S., Doehlert, D., Alamri, M.S., Elias, E.M. and Kianian, S.F., 2016. Dissection of genetic factors underlying wheat kernel shape and size in an elite× nonadapted cross using a high density SNP linkage map. *The plant genome*, *9*(1).

Kunert, A., Naz, A.A., Dedeck, O., Pillen, K. and Léon, J., 2007. AB-QTL analysis in winter wheat: I. Synthetic hexaploid wheat (T. turgidum ssp. dicoccoides× T. tauschii) as a source of favourable alleles for milling and baking quality traits. *Theoretical and Applied Genetics*, *115*(5), pp.683-695.

Lado, B., Battenfield, S., Guzmán, C., Quincke, M., Singh, R.P., Dreisigacker, S., Peña, R.J., Fritz, A., Silva, P., Poland, J. and Gutiérrez, L., 2017. Strategies for Selecting Crosses Using Genomic Prediction in Two Wheat Breeding Programs. *The Plant Genome, 10*(2).

Laghari, K.A., Sial, M.A., Arain, M.A., Mirbahar, A.A., Pirzada, A.J., Dahot, M.U. and Mangrio, S.M., 2010. Heritability studies of yield and yield associated traits in bread wheat. *Pak. J. Bot, 42*(1), pp.111-115.

Lagudah, E.S., Krattinger, S.G., Herrera-Foessel, S., Singh, R.P., Huerta-Espino, J., Spielmeyer, W., Brown-Guedira, G., Selter, L.L. and Keller, B., 2009. Gene-specific markers for the wheat gene Lr34/Yr18/Pm38 which confers resistance to multiple fungal pathogens. *Theoretical and Applied Genetics*, *119*(5), pp.889-898.

Lan, C., Zhang, Y., Herrera-Foessel, S.A., Basnet, B.R., Huerta-Espino, J., Lagudah, E.S. and Singh, R.P., 2015. Identification and characterization of pleiotropic and co-located resistance loci to leaf rust and stripe rust in bread wheat cultivar Sujata. *Theoretical and applied genetics*, *128*(3), pp.549-561.

Lan, C., Hale, I.L., Herrera-Foessel, S.A., Basnet, B.R., Randhawa, M.S., Huerta-Espino, J., Dubcovsky, J. and Singh, R.P., 2017. Characterization and Mapping of Leaf Rust and Stripe Rust Resistance Loci in Hexaploid Wheat Lines UC1110 and PI610750 under Mexican Environments. *Frontiers in plant science*, *8*, p.1450.

Lande, R. and Thompson, R., 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*, *124*(3), pp.743-756.

Langridge, P., 2014. Reinventing the green revolution by harnessing crop mutant resources. *Plant physiology*, *166*(4), pp.1682-1683.

Legates, D.R. and McCabe, G.J., 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water resources research*, *35*(1), pp.233-241.

Lewis, S., Faricelli, M.E., Appendino, M.L., Valárik, M. and Dubcovsky, J., 2008. The chromosome region including the earliness per se locus Eps-Am1 affects the duration of early developmental phases and spikelet number in diploid wheat. *Journal of Experimental Botany*, *59*(13), pp.3595-3607.

Li, C., Wang, Z., Li, C., Bowden, R., Bai, G., Li, C., Li, C., Su, Z. and Carver, B.F., 2017. Mapping of Quantitative Trait Loci for Leaf Rust Resistance in the Wheat Population Ning7840× Clark. *Plant Disease*, *101*(12), pp.1974-1979.

Lillemo, M., Asalf, B., Singh, R.P., Huerta-Espino, J., Chen, X.M., He, Z.H. and Bjørnstad, Å., 2008. The adult plant rust resistance loci Lr34/Yr18 and Lr46/Yr29 are important determinants of partial resistance to powdery mildew in bread wheat line Saar. *Theoretical and Applied Genetics*, *116*(8), pp.1155-1166.

Liu, J., Liu, D., Tao, W., Li, W., Wang, S., Chen, P., Cheng, S. and Gao, D., 2000. Molecular marker-facilitated pyramiding of different genes for powdery mildew resistance in wheat. *Plant Breeding*, *119*(1), pp.21-24.

Liu, J., Liu, L., Hou, N., Zhang, A. and Liu, C., 2007. Genetic diversity of wheat gene pool of recurrent selection assessed by microsatellite markers and morphological traits. *Euphytica*, *155*(1-2), pp.249-258.

Liu, Y., He, Z., Appels, R. and Xia, X., 2012. Functional markers in wheat: current status and future prospects. *Theoretical and Applied Genetics*, *125*(1), pp.1-10.

Lopes, M.S., Reynolds, M.P., Jalal-Kamali, M.R., Moussa, M., Feltaous, Y., Tahir, I.S.A., Barma, N., Vargas, M., Mannes, Y. and Baum, M., 2012. The yield correlations of selectable physiological traits in a population of advanced spring wheat lines grown in warm and drought environments. *Field Crops Research*, *128*, pp.129-136.

Lopes, M.S., Rebetzke, G.J. and Reynolds, M., 2014. Integration of phenotyping and genetic platforms for a better understanding of wheat performance under drought. *Journal of experimental botany*, *65*(21), pp.6167-6177.

Ma, L., Li, T., Hao, C., Wang, Y., Chen, X. and Zhang, X., 2016. TaGS5-3A, a grain size gene selected during wheat improvement for larger kernel and yield. *Plant biotechnology journal*, *14*(5), pp.1269-1280.

Maan, S.S., Carlson, K.M., Williams, N.D. and Yang, T., 1987. Chromosomal arm location and gene-centromere distance of a dominant gene for male sterility in wheat. *Crop science*, *27*(3), pp.494-500.

Mago, R., Brown-Guedira, G., Dreisigacker, S., Breen, J., Jin, Y., Singh, R., Appels, R., Lagudah, E.S., Ellis, J. and Spielmeyer, W., 2011. An accurate DNA marker assay for stem rust resistance gene Sr2 in wheat. *Theoretical and applied genetics*, *122*(4), pp.735-744.

Mallick, N., Sharma, J.B., Tomar, R.S., Sivasamy, M. and Prabhu, K.V., 2015. Marker-assisted backcross breeding to combine multiple rust resistance in wheat. *Plant Breeding*, *134*(2), pp.172-177.

Manley, M., Engelbrecht, M.L., Williams, P.C. and Kidd, M., 2009. Assessment of variance in the measurement of hectolitre mass of wheat, using equipment from different grain producing and exporting countries. *Biosystems engineering*, *103*(2), pp.176-186.

Maphosa, L., Langridge, P., Taylor, H., Parent, B., Emebiri, L.C., Kuchel, H., Reynolds, M.P., Chalmers, K.J., Okada, A., Edwards, J. and Mather, D.E., 2014. Genetic control of grain yield and grain physical characteristics in a bread wheat population grown under a range of environmental conditions. *Theoretical and applied genetics*, *127*(7), pp.1607-1624.

Marais, G.F. and Botes, W.C., 2009. Recurrent mass selection for routine improvement of common wheat: A review. In *Organic Farming, Pest Control and Remediation of Soil Pollutants* (pp. 85-105). Springer Netherlands.

Marais, G.F., Botes, W.C. and Louw, J.H., 2000. Recurrent selection using male sterility and hydroponic tiller culture in pedigree breeding of wheat. *Plant breeding*, *119*(5), pp.440-442.

Marsalis, M.A., and Goldberg, N.P., 2016. Leaf, Stem, and Stripe Rust Diseases of Wheat. *Guide, A*.

Marshall, D.R., Mares, D.J., Moss, H.J. and Ellison, F.W., 1986. Effects of grain shape and size on milling yields in wheat. II. Experimental studies. *Australian Journal of Agricultural Research*, *37*(4), pp.331-342.

Marza, F., Bai, G.H., Carver, B.F. and Zhou, W.C., 2006. Quantitative trait loci for yield and related traits in the wheat population Ning7840× Clark. *Theoretical and Applied Genetics*, *112*(4), pp.688-698.

Mason, R.E., Hays, D.B., Mondal, S., Ibrahim, A.M. and Basnet, B.R., 2013. QTL for yield, yield components and canopy temperature depression in wheat under late sown field conditions. *Euphytica*, *194*(2), pp.243-259.

Matsuoka, Y., 2011. Evolution of polyploid Triticum wheats under cultivation: the role of domestication, natural hybridization and allopolyploid speciation in their diversification. *Plant and cell physiology*, *52*(5), pp.750-764.

Mayer, K.F.X., Rogers, J., Pozniak, C., Eversole, K., Feuillet, C., Gill, B., Friebe, B., Lukaszewski, A.J., Sourdille, P., Endo, T.R. and Kubalakova, M., 2014. A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science*, *345*(6194), pp.1251788-1251788.

McIntyre, C.L., Mathews, K.L., Rattey, A., Chapman, S.C., Drenth, J., Ghaderi, M., Reynolds, M. and Shorter, R., 2010. Molecular detection of genomic regions associated with grain yield and yield-determining components in an elite bread wheat cross evaluated under irrigated and rainfed conditions. *Theoretical and Applied Genetics*, *120*(3), pp.527-541.

Meintjes, S.W., 2017. *Improvement of rust resistance and FHB resistance QTL in wheat through the application of an integrated biotechnology approach* (Masters dissertation, Stellenbosch: Stellenbosch University).

Meng, L., Liu, Z., Zhang, L., Hu, G. and Song, X., 2016. Cytological characterization of a thermo-sensitive cytoplasmic male sterile wheat line having K-type cytoplasm of Aegilops kotschyi. *Breeding science*, *66*(5), pp.752-761.

Microsoft., 2016. Microsoft Excel [computer software]. Redmond, Washington: Microsoft.

Miedaner, T. and Korzun, V., 2012. Marker-assisted selection for disease resistance in wheat and barley breeding. *Phytopathology*, *102*(6), pp.560-566.

Mikulikova, D., 2007. The effect of friabilin on wheat grain hardness. *Czech Journal of Genetics and Plant Breeding*, *43*(2), p.35.

Miller, T.D., 1999. *Growth stages of wheat: identification and understanding improve crop management. Texas Agricultural Extension Service, the Texas A&M University system*. SCS-1999-16.

Monneveux, P., Jing, R. and Misra, S.C., 2012. Phenotyping for drought adaptation in wheat using physiological traits. *Frontiers in physiology*, *3*.

Moore, J.W., Herrera-Foessel, S., Lan, C., Schnippenkoetter, W., Ayliffe, M., Huerta-Espino, J., Lillemo, M., Viccars, L., Milne, R., Periyannan, S. and Kong, X., 2015. A recently evolved hexose transporter variant confers resistance to multiple pathogens in wheat. *Nature genetics*, *47*(12), pp.1494-1498.

Muterko, A., Balashova, I., Cockram, J., Kalendar, R. and Sivolap, Y., 2015. The new wheat vernalization response allele Vrn-D1s is caused by DNA transposon insertion in the first intron. *Plant Molecular Biology Reporter*, *33*(2), pp.294-303.

Muthe, S.T., Kulwal, P.L., Gadekar, D.A. and Jadhav, A.S., 2016. Molecular marker based detection of leaf rust resistance gene Lr34 in Indian bread wheat (Triticum aestivum L.). *Australasian Plant Pathology*, *45*(4), pp.369-376.

Neilsen, M.L., Gangadhara, S.D. and Rife, T., 2016, September. Extending Watershed Segmentation Algorithms for High-Throughput Phenotyping. In *Proceedings of the 29th International Conference on Computer Applications in Industry and Engineering, Denver, CO*.

Nirmal, R.C., Furtado, A., Wrigley, C. and Henry, R.J., 2016. Influence of Gene Expression on Hardness in Wheat. *PloS one*, *11*(10), p.e0164746.

Nocente, F., Gazza, L. and Pasquini, M., 2007. Evaluation of leaf rust resistance genes Lr1, Lr9, Lr24, Lr47 and their introgression into common wheat cultivars by marker-assisted selection. *Euphytica*, *155*(3), pp.329-336.

Nyiraneza, J., Cambouris, A.N., Ziadi, N., Tremblay, N. and Nolin, M.C., 2012. Spring wheat yield and quality related to soil texture and nitrogen fertilization. *Agronomy journal*, *104*(3), pp.589-599.

NZFMA, New Zealand Flour Millers Association., 2017. *Wheat - The Structure & Physical Characteristics* [Online] Available at: http://www.flourinfo.co.nz/index.php/nzfma-knowledgebase/4-wheat/8-wheat-the-structure-a-physical-characteristics [Accessed 24 December 2017].

Ortiz, R., Sayre, K.D., Govaerts, B., Gupta, R., Subbarao, G.V., Ban, T., Hodson, D., Dixon, J.M., Ortiz-Monasterio, J.I. and Reynolds, M., 2008. Climate change: can wheat beat the heat?. *Agriculture, Ecosystems & Environment*, *126*(1), pp.46-58.

Pasha, I., Anjum, F.M. and Morris, C.F., 2010. Grain hardness: a major determinant of wheat quality. *Revista de Agaroquimica y Tecnologia de Alimentos*, *16*(6), pp.511-522.

Patel, R., 2013. The long green revolution. *The Journal of Peasant Studies*, *40*(1), pp.1-63.

Pedro, A., Savin, R. and Slafer, G.A., 2012. Crop productivity as related to single-plant traits at key phenological stages in durum wheat. *Field Crops Research*, *138*, pp.42-51.

Peng, J.H., Sun, D. and Nevo, E., 2011. Domestication evolution, genetics and genomics in wheat. *Molecular Breeding*, *28*(3), p.281.

Persoons, A., Hayden, K.J., Fabre, B., Frey, P., De Mita, S., Tellier, A. and Halkett, F., 2017. The escalatory Red Queen: Population extinction and replacement following arms race dynamics in poplar rust. *Molecular ecology*, *26*(7), pp.1902-1918.

Pethybridge, S.J. and Nelson, S.C., 2015. Leaf doctor: a new portable application for quantifying plant disease severity. *Plant Disease*, *99*(10), pp.1310-1316.

Pietragalla, J. and Pask, A., 2012. Grain yield and yield components. *Physiological breeding II: A Field Guide to Wheat Phenotyping. CIMMYT, Mexico DF*, pp.95-103.

Prabhu, K.V., Singh, A.K., Basavaraj, S.H., Cherukuri, D.P., Charpe, A., Gopala Krishnan, S., Gupta, S.K., Joseph, M., Koul, S., Mohapatra, T. and Pallavi, J.K., 2009. Marker assisted selection for biotic stress resistance in wheat and rice. *Indian J. Genet*, *69*(4), pp.305-314.

Qin, L., Hao, C., Hou, J., Wang, Y., Li, T., Wang, L., Ma, Z. and Zhang, X., 2014. Homologous haplotypes, expression, genetic effects and geographic distribution of the wheat yield gene TaGW2. *BMC plant biology*, *14*(1), p.107.

Racz, I., Kadar, R., Moldovan, V. and Has, I., 2015. Performance and stability of grain yield and yield components in some winter wheat varieties. *Romanian Agricuture Research*, *32*, pp.11-18.

Rafalski, J.A. and Tingey, S.V., 1993. Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. *Trends in Genetics*, *9*(8), pp.275-280.

Ramya, P., Chaubal, A., Kulkarni, K., Gupta, L., Kadoo, N., Dhaliwal, H.S., Chhuneja, P., Lagu, M. and Gupt, V., 2010. QTL mapping of 1000-kernel weight, kernel length, and kernel width in bread wheat (Triticum aestivum L.). *Journal of Applied Genetics*, *51*(4), pp.421-429.

Randhawa, H.S., Asif, M., Pozniak, C., Clarke, J.M., Graf, R.J., Fox, S.L., Humphreys, D.G., Knox, R.E., DePauw, R.M., Singh, A.K. and Cuthbert, R.D., 2013. Application of molecular markers to wheat breeding in Canada. *Plant Breeding*, *132*(5), pp.458-471.

Rasheed, A., Xia, X., Mahmood, T., Quraishi, U.M., Aziz, A., Bux, H., Mahmood, Z., Mirza, J.I., Mujeeb-Kazi, A. and He, Z., 2016. Comparison of economically important loci in landraces and improved wheat cultivars from Pakistan. *Crop Science*, *56*(1), pp.287-301.

Rebetzke, G.J., van Herwaarden, A.F., Jenkins, C., Weiss, M., Lewis, D., Ruuska, S., Tabe, L., Fettel, N.A. and Richards, R.A., 2008. Quantitative trait loci for soluble stem carbohydrate production in wheat. *Australian Journal of Agricultural Research, 59*, pp.891-905.

Reynolds, M., Manes, Y., Izanloo, A. and Langridge, P., 2009a. Phenotyping approaches for physiological breeding and gene discovery in wheat. *Annals of Applied Biology*, *155*(3), pp.309-320.

Reynolds, M., Foulkes, M.J., Slafer, G.A., Berry, P., Parry, M.A., Snape, J.W. and Angus, W.J., 2009b. Raising yield potential in wheat. *Journal of Experimental Botany*, *60*(7), pp.1899-1918.

Reynolds, M., Manes, Y. and Rebetzke, G., 2012. Application of physiology in breeding for heat and drought. *Physiological Breeding I: Interdisciplinary Approaches to Improve Crop Adaptation. CIMMYT, Mexico DF*, pp.18-32.

Reynolds, M. and Langridge, P., 2016. Physiological breeding. *Current opinion in plant biology*, *31*, pp.162-171.

Richards, R.A., Rebetzke, G.J., Watt, M., Condon, A.T., Spielmeyer, W. and Dolferus, R., 2010. Breeding for improved water productivity in temperate cereals: phenotyping, quantitative trait loci, markers and the selection environment. *Functional Plant Biology*, *37*(2), pp.85-97.

Rife, T.W., 2016. *Utilizing a historical wheat collection to develop new tools for modern plant breeding* (Doctoral dissertation, Kansas State University).

Rose, T., Nagler, S. and Kage, H., 2017. Yield formation of Central-European winter wheat cultivars on a large scale perspective. *European Journal of Agronomy*, *86*, pp.93-102.

Sacks, W.J., Deryng, D., Foley, J.A. and Ramankutty, N., 2010. Crop planting dates: an analysis of global patterns. *Global Ecology and Biogeography*, *19*(5), pp.607-620.

Sadras, V.O. and Lawson, C., 2011. Genetic gain in yield and associated changes in phenotype, trait plasticity and competitive ability of South Australian wheat varieties released between 1958 and 2007. *Crop and Pasture Science*, *62*(7), pp.533-549.

SAGIS, South African grain information service., 2017. [Online] Available at: http://www.sagis.org.za/cec_reports_2017.html [Accessed 24 December 2017].

Sakamoto, T. and Matsuoka, M., 2004. Generating high-yielding varieties by genetic manipulation of plant architecture. *Current Opinion in Biotechnology*, *15*(2), pp.144-147.

Seki, M., Chono, M., Nishimura, T., Sato, M., Yoshimura, Y., Matsunaka, H., Fujita, M., Oda, S., Kubo, K., Kiribuchi-Otobe, C. and Kojima, H., 2013. Distribution of photoperiod-insensitive allele Ppd-A1a and its effect on heading time in Japanese wheat cultivars. *Breeding science*, *63*(3), pp.309-316.

Sharma, S.N., Sain, R.S. and Sharma, R.K., 2003. Genetics of spike length in durum wheat. *Euphytica*, *130*(2), pp.155-161.

Sharma, S., Upadhyaya, H.D., Varshney, R.K. and Gowda, C.L.L., 2013. Pre-breeding for diversification of primary gene pool and genetic enhancement of grain legumes. *Frontiers in plant science*, *4*.

Shewry, P.R., 2009. Wheat. *Journal of experimental botany*, *60*(6), pp.1537-1553.

Shi, J.R., Xu, D.H., Yang, H.Y., Lu, Q.X. and Ban, T., 2008. DNA marker analysis for pyramided of Fusarium head blight (FHB) resistance QTLs from different germplasm. *Genetica*, *133*(1), pp.77-84.

Shiferaw, B., Smale, M., Braun, H.J., Duveiller, E., Reynolds, M. and Muricho, G., 2013. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Security*, *5*(3), pp.291-317.

Simmonds, J., Scott, P., Leverington-Waite, M., Turner, A.S., Brinton, J., Korzun, V., Snape, J. and Uauy, C., 2014. Identification and independent validation of a stable yield and thousand grain weight QTL on chromosome 6A of hexaploid wheat (Triticum aestivum L.). *BMC plant biology*, *14*(1), p.191.

Simmonds, J., Scott, P., Brinton, J., Mestre, T.C., Bush, M., Del Blanco, A., Dubcovsky, J. and Uauy, C., 2016. A splice acceptor site mutation in TaGW2-A1 increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. *Theoretical and Applied Genetics*, *129*(6), pp.1099-1112.

Singh, R., Datta, D., Singh, S. and Tiwari, R., 2004. Marker-assisted selection for leaf rust resistance genes Lr19 and Lr24 in wheat (Triticum aestivum L.). *Journal of applied genetics*, *45*(4), pp.399-404.

Singh, R.P., Hodson, D.P., Huerta-Espino, J., Jin, Y., Bhavani, S., Njau, P., Herrera-Foessel, S., Singh, P.K., Singh, S. and Govindan, V., 2011. The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. *Annual review of phytopathology*, *49*, pp.465-481.

Singh, A., Ganapathysubramanian, B., Singh, A.K. and Sarkar, S., 2016. Machine learning for high-throughput stress phenotyping in plants. *Trends in plant science*, *21*(2), pp.110-124.

Smit, H.A., Tolmay, V.L., Barnard, A., Jordaan, J.P., Koekemoer, F.P., Otto, W.M., Pretorius, Z.A., Purchase, J.L. and Tolmay, J.P.C., 2010. An overview of the context and scope of wheat (Triticum aestivum) research in South Africa from 1983 to 2008. *South African Journal of Plant and Soil*, *27*(1), pp.81-96.

Snape, J.W., Foulkes, M.J., Simmonds, J., Leverington, M., Fish, L.J., Wang, Y. and Ciavarrella, M., 2007. Dissecting gene× environmental effects on wheat yields via QTL and physiological analysis. *Euphytica*, *154*(3), pp.401-408.

Solis-Moya, E., Camacho-Casas, M.A., Alvarado-Padilla, J.I., Moreno, J.I., Borbón-Gracia, A., Molero, G. and Reynolds, M., 2017. Evaluation of the 3rd Wheat Yield Consortium Yield trial under irrigated growing conditions in Mexico during the 2015-2016 cycle. *TRIGO (Wheat) Yield Potential*, p.55.

Spielmeyer, W. and Richards, R.A., 2004. Comparative mapping of wheat chromosome 1AS which contains the tiller inhibition gene (tin) with rice chromosome 5S. *Theoretical and Applied Genetics*, *109*(6), pp.1303-1310.

Springfield, L.S., 2014. *Pyramiding of rust resistance genes in wheat utilizing male sterility mediated marker-assisted recurrent selection* (Masters dissertation, Stellenbosch: Stellenbosch University).

Sramkova, Z., Gregová, E. and Šturdík, E., 2009. Chemical composition and nutritional quality of wheat grain. *Acta Chimica Slovaca*, *2*(1), pp.115-138.

Stein, N., 2017. *The IWGSC Reference Genome Sequence*, paper presented to the International Plant and Animal Genome XXV Conference, San Diego, 14-18 January. [Online] Available at: https://www.wheatgenome.org/content/download/21978/374444/file/Stein_Wheat+sequencing_PAG2017.pdf [Accessed 5 November 2017].

Stratonovitch, P. and Semenov, M.A., 2015. Heat tolerance around flowering in wheat identified as a key trait for increased yield potential in Europe under climate change. *Journal of experimental botany*, *66*(12), pp.3599-3609.

Stroup, W.W. and Mulitze, D.K., 1991. Nearest neighbor adjusted best linear unbiased prediction. *The American Statistician*, *45*(3), pp.194-200.

Sukumaran, S., Lopes, M., Dreisigacker, S. and Reynolds, M., 2017. Genetic analysis of multi-environmental spring wheat trials identifies genomic regions for locus-specific trade-offs for grain weight and grain number. *Theoretical and Applied Genetics*, pp.1-14.

Sun, X., Marza, F., Ma, H., Carver, B.F. and Bai, G., 2010. Mapping quantitative trait loci for quality factors in an inter-class cross of US and Chinese wheat. *Theoretical and applied genetics*, *120*(5), pp.1041-1051.

Tamburic-Ilincic, L., Falk, D. and Schaafsma, A., 2011. Fusarium ratings in Ontario Winter Wheat Performance Trial (OWWPT) using an index that combines Fusarium head blight symptoms and deoxynivalenol levels. *Czech J. Genet. Plant Breed*, *47*, pp.S115-S122.

Tanabata, T., Shibaya, T., Hori, K., Ebana, K. and Yano, M., 2012. SmartGrain: high-throughput phenotyping software for measuring seed shape through image analysis. *Plant physiology*, *160*(4), pp.1871-1880.

Tanger, P., Klassen, S., Mojica, J.P., Lovell, J.T., Moyers, B.T., Baraoidan, M., Naredo, M.E.B., McNally, K.L., Poland, J., Bush, D.R. and Leung, H., 2017. Field-based high throughput phenotyping rapidly identifies genomic regions controlling yield components in rice. *Scientific Reports*, *7*, p.42839.

Tester, M. and Langridge, P., 2010. Breeding technologies to increase crop production in a changing world. *Science*, *327*(5967), pp.818-822.

Tyagi, S., Mir, R.R., Kaur, H., Chhuneja, P., Ramesh, B., Balyan, H.S. and Gupta, P.K., 2014. Marker-assisted pyramiding of eight QTLs/genes for seven different traits in common wheat (Triticum aestivum L.). *Molecular breeding*, *34*(1), pp.167-175.

Valkoun, J.J., 2001. Wheat pre-breeding using wild progenitors. *Euphytica*, *119*(1-2), pp.17-23.

Valluru, R., Reynolds, M.P. and Lafarge, T., 2015. Food security through translational biology between wheat and rice. *Food and Energy Security*, *4*(3), pp.203-218.

van Bueren, E.L., Backes, G., De Vriend, H. and Østergård, H., 2010. The role of molecular markers and marker assisted selection in breeding for organic agriculture. *Euphytica*, *175*(1), pp.51-64.

Wallace, D.H. and Yan, W., 1998. *Plant breeding and whole-system crop physiology: improving crop maturity, adaptation and yield*. CAB international.

Wang, S., Zhang, X., Chen, F. and Cui, D., 2015. A single-nucleotide polymorphism of TaGS5 gene revealed its association with kernel weight in Chinese bread wheat. *Frontiers in plant science*, *6*.

Wang, X.Y., Chen, P.D. and Zhang, S.Z., 2001. Pyramiding and marker-assisted selection for powdery mildew resistance genes in common wheat. *Yi chuan xue bao= Acta genetica Sinica*, *28*(7), pp.640-646.

Wanyera, R., Kinyua, M.G., Jin, Y. and Singh, R.P., 2006. The spread of stem rust caused by Puccinia graminis f. sp. tritici, with virulence on Sr31 in wheat in Eastern Africa. *Plant Disease*, *90*(1), pp.113-113.

Weather Underground., 2018. *Historical Weather* [Online] Available at: https://www.wunderground.com/history/ [Accessed 15 February 2018].

Wessels, E. and Botes, W.C., 2014. Accelerating resistance breeding in wheat by integrating marker-assisted selection and doubled haploid technology. *South African Journal of Plant and Soil*, *31*(1), pp.35-43.

Whitford, R., Fleury, D., Reif, J.C., Garcia, M., Okada, T., Korzun, V. and Langridge, P., 2013. Hybrid breeding in wheat: technologies to improve hybrid wheat seed production. *Journal of experimental botany*, *64*(18), pp.5411-5428.

Wiersma, J.J., Busch, R.H., Fulcher, G.G. and Hareland, G.A., 2001. Recurrent selection for kernel weight in spring wheat. *Crop Science*, *41*(4), pp.999-1005.

Wilhelm, E.P., Boulton, M.I., Al-Kaff, N., Balfourier, F., Bordes, J., Greenland, A.J., Powell, W. and Mackay, I.J., 2013. Rht-1 and Ppd-D1 associations with height, GA sensitivity, and days to heading in a worldwide bread wheat collection. *Theoretical and applied genetics*, *126*(9), pp.2233-2243.

Williams, K., Munkvold, J. and Sorrells, M., 2013. Comparison of digital image analysis using elliptic Fourier descriptors and major dimensions to phenotype seed shape in hexaploid wheat (Triticum aestivum L.). *Euphytica*, *190*(1), pp.99-116.

Wu, J., Yang, X., Wang, H., Li, H., Li, L., Li, X. and Liu, W., 2006. The introgression of chromosome 6P specifying for increased numbers of florets and kernels from Agropyron cristatum into wheat. *Theoretical and Applied Genetics*, *114*(1), pp.13-20.

Wu, X., Chang, X. and Jing, R., 2012. Genetic insight into yield-associated traits of wheat grown in multiple rain-fed environments. *PloS one*, *7*(2), p.e31249.

Wu, J.Q., Sakthikumar, S., Dong, C., Zhang, P., Cuomo, C.A. and Park, R.F., 2017. Comparative genomics integrated with association analysis identifies candidate effector genes corresponding to Lr20 in phenotype-paired Puccinia triticina isolates from Australia. *Frontiers in plant science*, *8*.

Xu, Y. and Crouch, J.H., 2008. Marker-assisted selection in plant breeding: from publications to practice. *Crop science*, *48*(2), pp.391-407.

Yan, L., Loukoianov, A., Blechl, A., Tranquilli, G., Ramakrishna, W., SanMiguel, P., Bennetzen, J.L., Echenique, V. and Dubcovsky, J., 2004. The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. *Science*, *303*(5664), pp.1640-1644.

Yang, Z., Bai, Z., Li, X., Wang, P., Wu, Q., Yang, L., Li, L. and Li, X., 2012. SNP identification and allelic-specific PCR markers development for TaGW2, a gene linked to wheat kernel weight. *Theoretical and Applied Genetics*, *125*(5), pp.1057-1068.

Yeo, F.K.S., Bouchon, R., Kuijken, R., Loriaux, A., Boyd, C., Niks, R.E. and Marcel, T.C., 2017. High-resolution mapping of genes involved in plant stage-specific partial resistance of barley to leaf rust. *Molecular Breeding*, *37*(4), p.45.

Younis, A., Siddique, M.I., Kim, C.K. and Lim, K.B., 2014. RNA interference (RNAi) induced gene silencing: a promising approach of hi-tech plant breeding. *International journal of biological sciences*, *10*(10), p.1150.

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

Yu, L.X., Lorenz, A., Rutkoski, J., Singh, R.P., Bhavani, S., Huerta-Espino, J. and Sorrells, M.E., 2011. Association mapping and gene–gene interaction for stem rust resistance in CIMMYT spring wheat germplasm. *Theoretical and Applied Genetics*, *123*(8), pp.1257-1268.

Yue, A., Li, A., Mao, X., Chang, X., Li, R. and Jing, R., 2015. Identification and development of a functional marker from 6-SFT-A2 associated with grain weight in wheat. *Molecular Breeding*, *35*(2), p.63.

Zhang, C., Gao, L., Sun, J., Jia, J. and Ren, Z., 2014a. Haplotype variation of Green Revolution gene Rht-D1 during wheat domestication and improvement. *Journal of integrative plant biology*, *56*(8), pp.774-780.

Zhang, Y., Liu, J., Xia, X. and He, Z., 2014b. TaGS-D1, an ortholog of rice OsGS3, is associated with grain weight and grain length in common wheat. *Molecular breeding*, *34*(3), pp.1097-1107.

Zhao, Y., Mette, M.F., Gowda, M., Longin, C.F.H. and Reif, J.C., 2014. Bridging the gap between marker-assisted and genomic selection of heading time and plant height in hybrid wheat. *Heredity*, *112*(6), pp.638-645.

Zikhali, M. and Griffiths, S., 2015. The effect of Earliness per se (Eps) genes on flowering time in bread wheat. In *Advances in Wheat Genetics: From Genome to Field* (pp. 339-345). Springer Japan.

Zou, J., Semagn, K., Chen, H., Iqbal, M., Asif, M., N'Diaye, A., Navabi, A., Perez-Lara, E., Pozniak, C., Yang, R.C. and Graf, R.J., 2017. Mapping of QTLs associated with resistance to common bunt, tan spot, leaf rust, and stripe rust in a spring wheat population. *Molecular Breeding*, 37(12), p.144.