Mapping and survey sequencing of *Dn* resistance genes in *Triticum aestivum* L.

by

Anandi Bierman

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

Department of Genetics,
Stellenbosch University,
Private Bag X1, Matieland 7602, South Africa.

Supervisor: Prof. A-M. Oberholster

March 2015

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: .....

Abstract

Mapping and survey sequencing of *Dn* resistance genes in *Triticum aestivum* L.

A. Bierman

Department of Genetics,

Stellenbosch University,

Private Bag X1, Matieland 7602, South Africa.

Thesis: PhD (Gen)

December 2014

Diuraphis noxia Kurdjumov (Russian Wheat Aphid; RWA) is a pest of wheat and barley that has spread from its home range in the fertile crescent to most wheat producing countries except Australia. Since its first introduction to South Africa and the USA in the late 20th century, breeding programs for wheat phenotypes resistant to the aphid were put in place. Conventional breeding practices rely on phenotypic screening to verify traits carried by offspring and genetic tools such as marker assisted selection (MAS) have greatly aided this process in speed and accuracy. The size and complexity of the wheat genome, its allopolyploid nature and repetitive elements have, however, posed a challenge

to studies on the genetics of this cereal crop. Many studies have focused on chromosome 3B which is the largest of the wheat chromosomes and easily separated from the redundant genomic background by techniques such as flow cytometry. The similarity in size of the remaining chromosomes however, limits the application of flow cytometry to their isolation. Databases such as GrainGenes (http://wheat.pw.usda.gov/GG2/index.shtml) house marker data from various mapping studies for all wheat chromosomes and in 2014 the International Wheat Genome Sequencing Consortium (IWGSC) completed the draft genome sequence of wheat categorized by chromosome. Sources of resistance (*Dn* resistance genes) against RWA are located on chromosome 7D. but despite the marker and sequence data available currently, mapping studies specific for the *Dn* resistance genes are few. Additionally, sequence data available is derived from cultivars susceptible to RWA and is not comprehensively annotated and assembled in many cases. In this study, we demonstrate a novel, combined approach to isolate and characterize the *Dn* resistance genes through the use of a genetic map constructed from Amplified Fragment Length Polymorphism (AFLP), Expressed Sequence Tag (EST) and microsatellite markers and a physical map constructed from Next Generation Sequencing (NGS) data of ditelosomic chromosomes (7DS and 7DL) isolated by microdissection on the PALM microbeam system. A 122.8 cM genetic map was produced from 38 polymorphic AFLP markers and two ESTs with the microsatellite *Xgwm111* as anchor to related genetic maps. Through comparison to maps available on GrainGenes the location of the *Dn1* resistance gene was narrowed down to a deletion bin (7DS5-0.36-0.62) on the short arm of chromosome 7D with an AFLP marker (E-ACT/M-CTG_0270.84) mapping closely at 3.5 cM and two ESTs mapping at 15.3 cM and 15.9 cM from *Dn1*. Isolation of

individual chromosome arms 7DS and 7DL using the PALM Microbeam system allowed sequencing of the chromosome without the redundancy of the remainder of the hexaploid genome. Through isolating the chromosome arms in this way, a >80-fold reduction in genome size was achieved as well as a major reduction in repetitive elements. Analysis of the sequencing data confirmed that 7DL is the physically shorter arm of the chromosome though it contains the majority of protein coding sequences.

Uittreksel

Kartering en basispaarvolgordebepaling van *Dn* weerstandsgene in *Triticum aestivum* L.

("Mapping and survey sequencing of Dn resistance genes in Triticum aestivum L.")

A. Bierman

Departement Genetika,

Universiteit Stellenbosch,

Privaatsak X1, Matieland 7602, Suid Afrika.

Tesis: PhD (Gen)

Desember 2014

Diuraphis noxia Kurdjumov (Russiese koring-luis; RWA) is ń pes wat op koring en gars voorkom. Die pes het vanaf sy tuiste in die midde Ooste na meeste koring-produserende lande behalwe Australië versprei. Sedert die eerste bekendstelling van RWA in Suid Afrika en die VSA in die vroeë 20ste eeu is teelprogramme ten gunste van koring lyne met weerstand teen RWA begin. Tradisionele teelprogramme maak op fisieise observasie van die fenotipe staat om te verifieer of plante in die nageslag die gewenste eienskap dra. Genetiese metodes soos merkerondersteunde seleksie (MAS) versnel hierdie selekteringsproses grootliks. Die grootte en

kompleksiteit van die koring genoom asook die polyploïde en herhalende natuur daarvan is ń groot hindernis vir genetiese studies van hierdie graangewas. Baie studies het op chromosoom 3B gefokus wat die grootste van die koring chromosome is en dus maklik vanaf die res van die oorbodige genomiese agtergond deur tegnieke soos vloeisitometrie geskei word. Die ooreenkoms in grootte tussen die res van die chromosome bemoeilik die toepassing van vloeisitometrie om hulle te isoleer. Databasisse soos GrainGenes (http://wheat.pw.usda.gov/GG2/index.shtml) bevat merker data vanaf verskeie karterings-studies vir al die chromosome en in 2014 het die "International Wheat Genome Sequencing Consortium"(IWGSC) die voorlopige basispaarvolgorde van die koring genoom bekendgestel, gekategoriseer volgens chromosoom. Weerstandsbronne (*Dn* weerstandsgene) teen RWA kom meestal op chromosoom 7D voor. Ten spyte van merker en basispaarvolgorde data tans beskikbaar is karterings-studies spesifiek tot die *Dn* gene skaars en basispaarvolgorde data is vanaf kultivars afkomstig wat nie weerstandbiedend teen RWA is nie en waarvan die annotasie en samestelling baie keer nie goed is nie. In hierdie studie demonstreer ons ń nuwe, gekombineerde aanslag om die *Dn* weerstandsgene te isoleer en karakteriseer deur van ń genetiese kaart opgestel met "Amplified Fragment Length Polymorphism"(AFLP), "Expressed Sequence Tag"(EST) en mikrosatelliet merkers asook ń fisiese kaart saamgestel deur die volgende-generasie-basispaarvolgordebepaling van ditelosomiese chromosome (7DS en 7DL) geïsoleer deur mikrodisseksie met die "PALM Microbeam"sisteem gebruik te maak. ń Genetiese kaart van 122.8 cM was met 38 polimorfiese AFLP merkers en twee EST merkers geskep. Die mikrosatelliet, *Xgwm111*, is ook ingesluit en het as anker vir verwante genetiese-kaarte gedien. Deur vergelyking met genetiese-kaarte op GrainGenes is die posisie van die *Dn1* weerstandsgeen vernou na ń delesie bin

(7DS5-0.36-0.62) op die kort arm van chromosoom 7D met ń AFLP merker (E-ACT/M-CTG_0270.84) wat ongeveer 3.5 cM vanaf die geen karteer. Die twee EST merkers is 15.3 cM en 15.9 cM vanaf die geen gekarteer. Isolering van die individuele chromosoom arms, 7DS en 7DL, deur van die "PALM Microbeam"sisteem gebruik te maak het basispaarvolgordebepaling van die chromosoom toegelaat sonder die oortolligheid van die res van die hexaploïde genoom. Deur die chromosoom so te isoleer is ń >80-maal verkleining in genoom grootte bereik insluitend ń groot reduksie in herhalende elemente. Analise van die data vanaf basispaarvolgordebepaling het bevestig dat chromosoom 7D die fisiese kleiner chromosoom is maar dat dit die meerderheid van proteïn koderende basispaarvolgordes bevat.

Acknowledgements

I would like to express my sincere gratitude to the following people and organisations :

Prof. Anna-Maria Botha-Oberholster for giving me space in her lab, space to grow, make mistakes and get back up again. The list of things I need to thank you for is too long for this thesis!

The NRF and WCT for funding - without their financial support furthering my education would have been impossible.

Friends and colleagues in the Cereal Genomics Lab for keeping me sane, laughing and well hydrated through good times and bad.

The Aqua Lab, Department of Genetics, Stellenbosch University for the use of their computer and JoinMap software.

Dirk Swanevelder at ARC, Onderstepoort for the use of the PALM Microbeam system - thank you for making me a part of your lab and allowing me to play with your amazingly expensive toys!

Computations were performed using facilities provided by the University of Cape Town's ICTS High Performance Computing team: http://hpc.uct.ac.za and particular thanks should go to Andrew Lewis for millions of E-mails answered and guidance given to this first-timer bioinformaticist.

Contents

List of Figures

List of Tables

Nomenclature

List of abbreviations

AAMP – aphid associated molecular patterns

AFLP – amplified fragment length polymorphism

AM – association mapping

AP-PCR – arbitrarily primed polymerase chain reaction

Avr gene – avirulence gene

BAC – bacterial artificial chromosome

BGA – blue green aphid

bp – base pair

BWT – burrows wheeler transform

CAPS – cleaved amplified polymorphic sites

cDNA-AFLP – complementary DNA AFLP

cM – centi Morgan

DArT – diversity arrays technology

DH – double haploid

EST – expressed sequence tag

FISH – fluorescent *in situ* hybridization

Gbp – gigabase pair

GCE – genomics character estimator

gDNA – genomic deoxyribonucleic acid

GO – gene ontology

HR – hypersensitive response

ITMI – International Triticeae Mapping Initiative

IWGSC– International Wheat Genome Sequencing Consortium

JA – jasmonic acid

Kbp – kilobase pair

LCM – laser capture microdissection

LOD – logarithm of the odds

LPC – laser pressure catapult

LRR – leucine rich repeat

LTR – long terminal repeat

MAS – marker assisted selection

MBC – map based cloning

Mbp – megabase pair

MYA – million years ago

NBS–LRR – nucleotide binding site leucine rich repeat

NGS – next generation sequencing

NIL – near isogenic line

PA – pea aphid

PCG – protein coding genes

PCR – polymerase chain reaction

PE – paired–end

PR – pathogen resistance

QTL – quantitative trait locus

R gene – resistance gene

RAPD – random amplified polymorphic DNA

RFLP – restriction fragment length polymorphism

RIL – recombinant inbred lines

ROS – reactive oxygen species

RWA – Russian wheat aphid

SA – salicylic acid

SAR – systemic acquired resistance

SCAR – sequence characterized amplified regions

SDF – single dose fragments

SNP – single nucleotide polymorphism

SSR – simple sequence repeat

STS – sequence tagged site

TREP – Triticeae repeat sequence database

Chapter 1

Introduction

There are 1.2 billion people living in countries classified as wheat-dependent and 2.5 billion people living in countries classified as wheat-consuming (Rosegrant and Agcaoili, 2010). Wheat is a source of livelihood for 30 million wheat producers across the world (Rosegrant and Agcaoili, 2010). Despite the high yielding properties of wheat (South African Grain Laboratory; http://www.sagl.co.za/, 2011/2012) global yields are decreasing (Long and Ort, 2010) and food production needs to increase by 70 to 100% by the year 2050 in order to meet global demands (Godfray *et al.*, 2010). Decreasing losses caused by biotic and abiotic stresses are vital components in the endeavour towards more sustainable farming as land area used for agriculture is unlikely to increase (Von Braun, 2007; Godfray *et al.*, 2010).

One such biotic stress is *Diuraphis noxia* Kurdjumov, commonly known as the Russian wheat aphid (RWA). This phloem feeding insect is a pest of wheat and barley that originated in the fertile crescent and today this invasive pest species is present in all wheat producing countries except Australia (Shea *et al.*, 2000; Stary *et al.*, 2003). *Diuraphis noxia* was first reported in South Africa in 1978 (Walters *et al.*, 1980) and by 1981 had reached North America (Gilchrist *et al.*, 1984). Wheat

landraces from the middle east and eastern Europe with resistance against *D. noxia* were used in breeding programs in South Africa to produce resistant commercial cultivars (Du Toit, 1987; 1988; 1990; Du Toit *et al.*, 1995). The first resistance genes (*Dn1* and *Dn2*) were derived from Iranian and Bulgarian wheat lines and in 1992, the first commercially available RWA resistant wheat, Tugela-*Dn1* was released in South Africa. By 2003 nearly 25% of Colorado winter wheat planted in the USA consisted of wheat varieties containing a resistance gene to *D. noxia* called *Dn4* and three years later farmers had access to 27 RWA resistant wheat cultivars (Tolmay *et al.*, 2007). Currently, there are 14 *Dn* resistance genes against *D. noxia*. Many of these genes are clustered on chromosome 7D of wheat (*Dn1*, *Dn2*, *Dn5*, *Dn6*, *Dnx*, *Dn8*, *Dn626580* and *DnCI2401*). *Dn7* and *DnCI2414* are resistance genes introduced to wheat through a wheat/rye translocation on chromosome 1RS and 1BL of rye and wheat respectively. The arms race between plant and pest never ceases though. With the availability of new sources of resistance against *D. noxia*, novel aphid biotypes were soon observed. Biotypes are defined as aphid populations showing virulence to wheat cultivars containing *Dn* genes which used to provide resistance (Smith *et al.*, 2004).

Faster and more efficient wheat breeding approaches are needed to keep ahead of developing virulent aphid biotypes and to allow the pyramiding of different resistance genes into single cultivars. Marker assisted selection (MAS) is such an option as it negates the need for physical screening of plantlets and allows resistance to be detected using genetic tests in the offspring of breeding programs for new cultivars resistant to *D. noxia* (Liu *et al.*, 2002). Initiatives such as the International Wheat Genome Sequencing Consortium (IWGSC; http://www.wheatgenome.org/; Gill *et al.*, 2004) and The International Triticeae Mapping Initiative (ITMI;http://-

wheat.pw.usda.gov/ITMI/; Gupta *et al.*, 2008) contribute molecular markers, genetic maps and physical mapping data that are vital to gene identification and characterization.

Mapping studies, cloning and characterization of genes in wheat are challenging tasks due to the size (17 Giga base pair (Gbp)) and complexity (hexaploid with the majority consisting of repetitive and transposable elements (Gill *et al.*, 2004; Brenchley *et al.*, 2012)) of its genome. There is still uncertainty regarding the chromosomal location and relationship of three of the first *Dn* genes clustered on chromosome 7D, namely *Dn1*, *Dn2* and *Dn5*. Whether these genes are found on the long or short arm of chromosome 7D is in dispute (Marais and Du Toit, 1993; Liu *et al.*, 2001) as well as whether they are, in fact, three individual genes or whether *Dn1* and *Dn5* might be alleles of the same locus (Marais and Du Toit, 1993; Saidi and Quick, 1996; Liu *et al.*, 2001). Werner *et al.* (1992) showed that the physically or cytologically longer arm of 7D is actually the genetically shorter arm as it is homoeologous to 7AS and 7BS. In some instances, authors investigating ditelosomic lines, 7DL Dt, were actually working with 7DS Dt. This was established using chromosome banding (Werner *et al.*, 1992). *Dn1*, *Dn2* and *Dn5* could thus have been wrongly assigned to 7DL by several authors according to Liu *et al.* (2005).

With the discrepancies surrounding the location of the *Dn* resistance genes and markers closely linked to them as well as variability observed in the way marker data is interpreted within different genetic backgrounds, MAS would be more accurate if screens were to be done for the actual gene of interest, rather than closely linked markers. Sequencing the wheat genome however has proven a monumental task. The IWGSC completed the draft sequence of individual chromosomes in

2014 (Mayer *et al.*, 2014) through isolation via flow cytometry of aneuploid lines. Focusing on single chromosomes reduced the redundancy of such a large genome.

The aim of this study was to map the *Dn1* resistance gene to a specific location on chromosome 7D and to explore technologies that will enable for the reduction of genomic complexities.

To address the aims of the project two main objectives were set: to use a genetic mapping approach with established and novel markers in order to place *Dn1* on either the long or short arm of chromosome 7D and then to assess different technologies of chromosomal isolation that will enable for a reduction in redundant/repetitive genomic regions and enrichment of genic information. This will pave the way for cloning of the *Dn1* and possibly, *Dn5* genes.

The hypothesis of this study was that the *Dn1* resistance gene is located on chromosome 7D. Because of the size and complexity of the wheat genome, innovative approaches to circumvent its redundancy are required in order to provide a physically close location of the *Dn1* resistance gene.

The outline of this study consists of five chapters. The research chapters are comprised of chapters three and four which are divided into an abstract, introduction, materials and methods, results and discussion. The content of each chapter is as follows:

Chapter 2 is a survey of current and previous literature relevant to the study and focuses on wheat and the nature of its genome. The techniques used throughout this study are also reviewed. Genetic mapping and markers employed are given particular attention. The pest, *Diuraphis noxia* is discussed along with resistance genes associated with it and how it affects its host plant.

Chapter 3 gives a discription of the construction of a saturated genetic map in order to position the *Dn1* resistance gene. The mapping population and markers used to construct the map are described.

Chapter 4 explains the preparation of mitotic metaphase chromosome squashes as well as how the 7DS and 7DL ditelosomic chromosome arms were isolated via microdissection in order to sequence the individual chromosomes. This chapter proceeds to discuss data analysis of the sequencing results such as sequencing quality, *de novo* assembly and mapping to the IWGSC scaffolds and gene sets as well as analyzing the repetitiveness of the obtained data using the Triticeae repeat sequence (TREP) database.

Chapter 5 is a general conclusion to this thesis. The difficulties associated with mapping and gene characterization in wheat is reiterated and the aim of the study is stated again. The main findings from the two research chapters are given. Suggestions for future work toward characterization of the *Dn* resistance genes are also given.

Outputs associated with this project include the following paper and poster presentations:

- Bierman, Anandi, Swanevelder, Dirk Z H and Botha, Anna-Maria (2014), "Mapping and characterization of selected *Diuraphis noxia* resistance genes in *Triticum aestivum*", *21st Biennial International Plant Resistance to Insects Workshop*, Marrakech, Morocco. (PAPER)

- Bierman, Anandi and Botha, Anna-Maria (2014), "Mapping and survey se-

quencing of *Dn* resistance genes in *Triticum aestivum* L.", *Winter Cereal Trust Annual Meeting*, CSIR, Pretoria, South Africa. (PAPER)

- Bierman, Anandi, Swanevelder, Dirk Z H and Botha, Anna-Maria (2013), "Using laser capture microdissection to excise chromosome 7DS and 7DL from wheat (*Triticum aestivum* L.)", *12th International Wheat Genetics Symposium*, Yokohama, Japan. (POSTER)

- Bierman, Anandi and Botha, Anna-Maria (2013), "Mapping and characterization of selected *Diuraphis noxia* resistance genes in *Triticum aestivum*", *Winter Cereal Trust Annual Meeting*, CSIR, Pretoria, South Africa. (PAPER)

- Bierman, Anandi and Botha, Anna-Maria (2012), "Mapping and characterization of selected *Diuraphis noxia* resistance genes in *Triticum aestivum*", *20th Biennial International Plant Resistance to Insects Workshop*, Minneapolis, Minnesota. (PAPER)

- Bierman, Anandi and Botha, Anna-Maria (2012), "Mapping and characterization of selected *Diuraphis noxia* resistance genes in *Triticum aestivum*", *Winter Cereal Trust Annual Meeting*, CSIR, Pretoria, South Africa. (PAPER)

- Bierman, Anandi, Loos, Ben and Botha, Anna-Maria (2012), "Mapping and characterization of selected *Diuraphis noxia* resistance genes in *Triticum aestivum*", *Biennial South African Genetics Society Conference*, Stellenbosch, South Africa. (POSTER)

- Bierman, Anandi and Botha, Anna-Maria (2011), "Mapping and characterization of selected *Diuraphis noxia* resistance genes in *Triticum aestivum*", *Winter Cereal Trust Annual Meeting*, CSIR, Pretoria, South Africa. (PAPER)

1.1 List of References

Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G.L.A., Amore, R.D., Allen, A.M., Mckenzie, N., Kramer, M., Kerhornou, A., Bolser, D., Kay, S., Waite, D., Trick, M., Bancroft, I., Gu, Y., Huo, N., Luo, M., Sehgal, S., Gill, B., Kianian, S., Anderson, O., Kersey, P., Dvorak, J., Mccombie, W.R., Hall, A., Mayer, K.F.X., Edwards, K.J., Bevan, M.W. and Hall, N. (2012). Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature*, vol. 491, pp. 705–710.

Du Toit, F. (1987). Resistance in wheat (*Triticum aestivum*) to *Diuraphis noxia* (Homoptera: Aphididae). *Cereal Research Communications*, vol. 15, pp. 175–179.

Du Toit, F. (1988). Another source of Russian wheat aphid (*Diuraphis noxia*) resistance in *Triticum aestivum*. *Cereal Research Communications*, vol. 16, pp. 105–106.

Du Toit, F. (1990). Field resistance in three bread wheat lines to the Russian wheat aphid, *Diuraphis noxia* (Hemiptera: Aphididae). *Crop Protection*, vol. 9, pp. 255–258.

Du Toit, F., Wessels, W.G. and Marais, G.F. (1995). The chromosome arm location of the Russian wheat aphid resistance gene, *Dn5*. *Cereal Research Communications*, vol. 23, pp. 15–17.

Gilchrist, L.I., Rodriguez, R. and Burnett, P.A. (1984). The extent of freestate streak and *Diuraphis noxia* in Mexico. *Barley yellow dwarf, proceedings of workshop*. pp. 157–163. Mexico City: International Maize and Wheat Improvement Center (CIMMYT).

Gill, B.S., Appels, R., Botha-Oberholster, A.-M., Buell, C.R., Bennetzen, J.L., Chalhoub, B., Chumley, F., Dvorak, J., Iwanaga, M., Keller, B., Li, W., Mccombie, W.R., Ogihara, Y., Quetier, F. and Sasaki, T. (2004). A workshop report on wheat genome sequencing: international genome research on wheat consortium. *Genetics*, vol. 168, pp. 1087–1096.

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M. and Toulmin, C. (2010). Food security: the challenge of feeding 9 billion people. *Science*, vol. 327, pp. 812–818.

Gupta, P.K., Mir, R.R., Mohan, A. and Kumar, J. (2008). Wheat genomics: present status and future prospects. *International Journal of Plant Genomics*, vol. 2008, pp. 36.

Liu, M., Smith, M. and Gill, S. (2002). Identification of microsatellite markers

linked to Russian wheat aphid resistance genes *Dn4* and *Dn6*. *Theoretical and Applied Genetics*, vol. 104, pp. 1042–1048.

Liu, X. M., Smith, C.M., Gill, B.S. and Tolmay, V. (2001). Microsatellite markers linked to six Russian wheat aphid resistance genes in wheat. *Theoretical and Applied Genetics*, vol. 102, pp. 504–510.

Liu, X.M., Smith, C.M., Friebe, B.R. and Gill, B.S. (2005). Molecular mapping and allelic relationships of Russian wheat aphid-resistance genes. *Crop Science*, vol. 45, p. 2273.

Long, S.P. and Ort, D.R. (2010). More than taking the heat: crops and global change. *Current Opinions in Plant Biology*, vol. 13, pp. 1–8.

Marais, G.F. and Du Toit, F. (1993). A monosomic analysis of Russian wheat aphid resistance in the common wheat PI 294994. *Plant Breeding*, vol. 111, pp. 246–248.

Mayer, K.F.X., Rogers, J., Doležel, J., Pozniak, C., Eversole, K., Feuillet, C., Gill, B., Friebe, B., Lukaszewski, A. J., Sourdille, P., Endo, T.R., Kubaláková, M., Ihalikova, J., Dubska, Z., Vrana, J., Perkova, R., Imkova, H., Febrer, M., Clissold, L., McLay, K., Singh, K., Chhuneja, P., Singh, N.K., Khurana, J., Akhunov, E., Choulet, F., Alberti, A., Barbe, V., Wincker, P., Kanamori, H., Kobayashi, F., Itoh, T., Matsumoto, T., Sakai, H., Tanaka, T., Wu, J., Ogihara, Y., Handa, H., Maclachlan, P.R., Sharpe, A., Klassen, D., Edwards, D., Batley,

J., Olsen, O.-A., Sandve, S.R., Lien, S., Steuernagel, B., Wulff, B., Caccamo, M., Ayling, S., Ramirez-Gonzalez, R.H., Clavijo, B.J., Wright, J., Pfeifer, M., Spannagl, M., Martis, M.M., Mascher, M., Chapman, J., Poland, J.A., Scholz, U., Barry, K., Waugh, R., Rokhsar, D. S., Muehlbauer, G.J., Stein, N., Gundlach, H., Zytnicki, M., Jamilloux, V., Quesneville, H., Wicker, T., Faccioli, P., Colaiacovo, M., Stanca, A. M., Budak, H., Cattivelli, L., Glover, N., Pingault, L., Paux, E., Sharma, S., Appels, R., Bellgard, M., Chapman, B., Nussbaumer, T., Bader, K. C., Rimbert, H., Wang, S., Knox, R., Kilian, A., Alaux, M., Alfama, F., Couderc, L., Guilhot, N., Viseux, C., Loaec, M., Keller, B., Praud, S. and IWGSC (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science*, vol. 345, pp. 1251788–1251788.

Rosegrant, M.W. and Agcaoili, M. (2010). Global food demand, supply, and price prospects to 2010. Washington D C.

Saidi, A. and Quick, J.S. (1996). Inheritance and allelic relationships among Russian wheat aphid resistance genes in winter wheat. *Crop Science*, vol. 36, pp. 256–258.

Shea, G., Botha, J. and Hardie, D. (2000). Russian wheat aphid. *Agriculture Western Australia*.

Smith, C.M., Belay, T., Stauffer, C., Stary, P., Kubeckova, I. and Starkey, S. (2004). Identification of Russian wheat aphid (Homoptera: Aphididae) popu-

lations virulent to the *Dn4* resistance gene. *Journal of Economic Entomology*, vol. 97, pp. 1112–1117.

Stary, P., Basky, Z., Tanigoshi, L.K. and Tomanovicc, Z. (2003). Distribution and history of Russian wheat aphid, *Diuraphis noxia* (Kurdj.) in the Carpathian Basin (Hom., Aphididae). *Journal of Pest Science*, vol. 76, pp. 17–21.

Tolmay, V.L., Lindeque, R.C. and Prinsloo, G.J. (2007). Preliminary evidence of a resistance-breaking biotype of the Russian wheat aphid, *Diuraphis noxia* (Kurdjumov) (Homoptera: Aphididae), in South Africa. *African Entomology*, vol. 15, pp. 228–230.

Von Braun, J. (2007). The world food situation: new driving forces and required actions. Washington D C.

Walters, M.C., Penn, F., Du Toit, F., Botha, T.C., Aalbersberg, K., Hewitt, P.H. and Broodryk, S.W. (1980). The Russian wheat aphid. *Farming in South Africa, Leaflet series*, vol. wheat G3, pp. 1–6.

Werner, J.E., Endo, T.R. and Gill, B.S. (1992). Toward a cytogenetically based physical map of the wheat genome. *Proceedings of the National Academy of Sciences*, vol. 89, pp. 11307–11311.

Chapter 2

Literature Review

2.1 Mapping Overview

Genetic mapping makes use of the recombination frequency between genes and/or markers on a chromosome. The frequency of recombination events between two markers depends on their distance from one another on the chromosome (Morgan, 1916). The closer two genes are to one another, the lower the frequency of crossovers. This frequency is converted to a percentage value and measured in units of centi Morgan (cM). Anything that can alter the frequency of crossovers will also affect mapping data and needs to be taken into account (Kosambi, 1943). Underestimation of map distances is often the result of double crossovers in two-factor mapping. However, when markers are close together (less than 5 cM) the probability of double crossovers occurring is close to zero (Haldane, 1919). Another factor that can affect mapping distances is sampling error. The relationship between phenotypic recombination frequencies and crossover frequencies is curvilinear. Mapping functions such as Kosambi's or Haldane's mapping functions are used to correct phenotypic recombination frequencies to approximate crossover

frequencies. Haldane's mapping function, however, does not take interference into account (Haldane, 1919). Interference is an instance where a crossover event interferes with the initiation of another crossover in its vicinity. Kosambi's mapping function accounts for these events during recombination (Kosambi, 1943).

Recombinant chromosomes are chromosomes where crossing over occurs between two linked markers while non-recombinance implies that there is no crossover (Schwarzacher, 2003). Coupling conformation implies that two dominant alleles are on the same chromosome of a homologous pair while repulsion conformation is the opposite, wherein two dominant alleles are on opposite homologous chromosomes (Myburg *et al.*, 1998).

Bansal *et al.* (2003) describes linkage disequilibrium as a statistical measure of a lack of independence between alleles at two independent loci. It exists between linked loci which can be defined as loci that occur at the same haplotype more often than would be expected by chance. A marker in linkage disequilibrium with its causal variant (disease for example) provides a flag for its location (Bansal *et al.*, 2003). Linkage of two loci manifests when the association between two phenotypic traits or markers deviate from independent assortment. This is seen as a deviation from a phenotypic ratio of 1:2:1 for single, co-dominant genes (Ma *et al.*, 1998) or 9:3:3:1 for more than one locus. The Chi-square test can be used to determine whether this deviation is significant (Lancaster and Seneta, 1969). Mapping approaches make use of linkage disequilibrium by either establishing linkage through variances in the phenotype or through quantitative trait locus mapping (QTL mapping) where statistical methods are used to establish linkage between QTLs and marker loci.

As more loci are added to a mapping experiment, the number of possible genotypes doubles. In this instance, manual calculations and counting of recombinant offspring is no longer feasible and computers able to run Chi-square contingency table analysis perform the necessary calculations. The first step is to map genetic markers to linkage groups or chromosomal segments containing linked loci. The Chi-square statistical test determines two-point linkage between markers, which can then form a basis for constructing linkage groups. Unfortunately, as the number of markers begins to grow, this approach becomes increasingly unsuited for comparing possible orders and choosing the best order of markers. Mapping software such as Mapmaker (Lander *et al.*, 1987) is based on the concept of the LOD score, the log of the odds-ratio (Morton, 1955) which indicates the log (10) of the ratio between the odds of one hypothesis (for example, linkage between two loci) versus an alternative hypothesis (no linkage in this example) (Young, 2000).

Ultimately, mapping experiments allow us to construct genetic maps that show the relative locations of genes on a chromosome as determined by the recombination frequencies between linked genes. Genetic map distances do not, however, represent a physical map, i.e. the physical distances in base pairs (bp) on a chromosome (Sturtevant, 1913). With a saturated genetic map in place, a physical map can be constructed from distances between markers and genes in bp. Physical maps are often used as the first step toward isolating and characterizing genes that have been placed on genetic maps (Raats *et al.*, 2013).

A physical mapping approach using bacterial artificial chromosome (BAC) clones as probes in Fluorescent *in situ* Hybridization (FISH) was used by Lapitan *et al.* (1997) to saturate regions with markers and build contigs spanning those regions. Clones of known genes could be used to screen the BAC library which could

then be localized to a chromosomal location using FISH (Lapitan *et al.*, 1997). Comparative mapping aims to identify conserved regions or regions of synteny between organisms (Berkman *et al.*, 2011). Finding conserved regions between organisms with well characterized genomes and organisms without, aids in the transferability of marker data to an otherwise uncharacterized genome (Ishikawa *et al.*, 2009). Chromosome walking can serve as an alternative to comparative mapping. Chromosome walking is the reconstruction of a section of DNA from many shorter, cloned segments starting with a linked DNA marker and sequencing the DNA to approach the gene of interest. However, even this method is complicated significantly by the hexaploid and repetitive nature of wheat (Stein *et al.*, 2000).

Another mapping approach focusing on the correlation between genotype and phenotype on the basis of linkage disequilibrium is association mapping (AM) (Peng *et al.*, 2009). Being able to use unrelated individuals is an advantage though a saturated genetic map is still required. In addition, it should be noted that marker alleles identified through AM are only correlated with alleles and aren't entirely predictive of these alleles. For mapping studies in wheat, however, there are already many microsatellite markers available on all chromosomes (Röder *et al.*, 1998; Somers *et al.*, 2004; Peng *et al.*, 2009).

Association mapping (Bansal *et al.*, 2003) relies on the presence of differences in allele frequencies between test and control samples. Differences observed do not always directly imply causality as there are factors such as population history that may affect allele frequencies. Associations observed do, however, provide incentive for further study and can often be interpreted as being due to the marker being physically close to the gene of interest. Many variations on this basic mapping

approach are available and, especially, useful in plants. Deletion mapping and Radiation hybrid mapping (Gupta *et al.*, 2008) have proven useful in wheat which can tolerate the generation of deletion mutants lacking entire chromosome sets or the generation of hybrids with other species. Deletion mapping in wheat began with generating aneuploid stock (Sears, 1954) where each cell is missing at least one chromosome or has an added chromosome. This allowed for the mapping of genes to individual chromosomes. Deletion lines (Endo and Gill, 1996) allowed the mapping of genes to physical segments of chromosomes. Radiation hybrid mapping (Gupta *et al.*, 2008) uses the addition and substitution of individual chromosomes from a donor (progenitor or other species) in order to physically map traits to specific chromosomes (Cox *et al.*, 1990).

Other, more customized mapping approaches are also available. BAC based physical mapping (Gupta *et al.*, 2008) has been beneficial in wheat where the genome is large and complicated in its hexaploid form. BAC libraries of diploid progenitor species have been constructed, fingerprinted and assembled into contigs allowing the physical mapping of genes in both the progenitor and hexaploid wheat (http://wheat.pw.usda.gov/PhysicalMapping/
index.html). *In silico* mapping (Gupta *et al.*, 2008) is a novel mapping approach made possible by the abundance of mapping data accumulated to date and allows for markers with known sequences to be mapped to wheat chromosomes through similarity searches in Expressed Sequence Tag (EST) databases (http://wheat.pw.usda.gov/GG2/blast.shtml). The advantage of such a mapping strategy is that a mapping population or genotyping is not a requirement.

Even though there are many approaches to physical mapping, Map Based Cloning (MBC) is the most widely used option for physical mapping and gene

isolation, especially when studying large, complex genomes such as barley and wheat (Feuillet *et al.*, 2003). Gene isolation in large grass genomes could be done through cross genome MBC using rice as a model because of its smaller genome, but comparative genetics at the micro level shows rearrangements between the grasses that would complicate the method (Sorrels *et al.*, 2003; Gill *et al.*, 2004).

In choosing a mapping population, sufficient polymorphism between the parent lines is required, as without polymorphism, segregation analysis and ultimate linkage mapping is impossible. The simplest populations to use for mapping are F_2 derived from F_1 hybrids and backcross populations. F_1 plants in a backcross population will be classified as parental or recombinant (Heyns, 2005). Recombinants are needed as the frequency of recombination between the gene of interest and a linked marker is indicative of the distance between the two loci. The only major drawback to these populations is that they are not true breeding, so having enough sample for future work could become a problem if provisions are not made.

Sample size is the next critical factor as the ability to determine the order of markers and map resolution is dependant on sample size. Mapping populations of less than 50 individuals are insufficient and most often, especially in plants, populations range into the thousands. A strategy for targeting a specific region of the genome for mapping is to use Near Isogenic Lines (NILs). NILs consist of pairs of parents similar throughout most of their genomes except for the region surrounding a selected gene. Near Isogenic Lines make it easy to determine the location of a marker relative to the target gene. This is in contrast to genetic mapping in other populations where it would be necessary to test every clone against the entire mapping population to determine whether it mapped near the gene of interest (Young, 2000). NIL populations allow for a popular polyploid

mapping method namely single dose fragments (SDF) which relies on a marker, present in single copy, in one of the parents (Cervantes-Flores *et al.*, 2008). NILs, however, have a low localization resolution compared to other mapping populations and the mapping power of a NIL population lies in replication number, rather than population size (Sharma *et al.*, 2011).

In the same way that NIL lines are used because of limited and specific variation in the parental stock, other mapping populations also attempt to reduce unwanted variation. Recombinant Inbred Lines (RILs) (Bansal *et al.*, 2003) make use of a single seed descent inbreeding program of F_2 progeny in order to obtain progeny homozygous for a chosen allele. Song *et al.* (2005) used a RIL population for genetic linkage mapping of microsatellite markers in wheat. Double haploid (DH) populations also rely on generating homozygous lines in as short a time as possible (Amrani *et al.*, 1993). Haploid wheat plants can be generated through ovary or anther culture or through chromosome elimination in intergeneric crosses with for eg. maize (Kisana *et al.*, 1993). Double haploid plants can be created from haploid stock by chromosome doubling with colchicine (Heyns, 2005; Oleszczuk *et al.*, 2011). Double haploid populations are used extensively in genetic studies in wheat, ranging from QTL mapping (Zhang *et al.*, 2008) to the sequencing of the wheat genome (Mayer *et al.*, 2014).

2.2 Genetic Markers

A genetic marker can be defined as an amplified locus that is informative, because it shows polymorphism between individuals of a population and can be visualized by some method (Meudt *et al.*, 2007). Genetic markers are classified by type: genes with known functions being type I and anonymous or unidentified DNA

fragments being type II. Type II markers make up the majority of marker systems such as Amplified Fragment Length Polymorphism (AFLP), Random Amplified Polymorphic DNA (RAPD), Simple Sequence Repeats (SSRs), etc. (Emara and Kim, 2003).

Genetic markers fulfill a dual purpose: they are used to create genetic maps and discover the positions of genes and QTLs, and they are applied to incorporate these genes into commercial crops via MAS (Song *et al.*, 2005). Marker assisted selection is an indispensible tool to plant breeders as they require a reliable method of selecting plants with the desired trait. Marker assisted selection then not only allows for the selection of the desired gene in offspring plants but also allows pyramiding of multiple genes (Venter and Botha, 2000).

Phenotypic selection can be fairly straight forward but it is still faced with several limitations such as being time consuming and subject to environmental factors. With regards to the RWA, phenotypic screening is generally done during the cooler winter months as aphid mortality increases at temperatures above 20 °C (Michels and Behle, 1988). Environmental influences on symptom expression can also result in inaccurate scoring with typical error rates for greenhouse screening of up to 10%. Therefore, employing a screening technique based on genetic markers instead of the phenotype, is faster and more accurate (Miller *et al.*, 2001).

Botstein *et al.* (1980) state that for marker application during mapping, four parameters need to be established: i) determination of the least number of markers needed to construct a genetic map; ii) the polymorphism level of each marker; iii) the required number of families to establish linkage; and iv) the level of polymorphism within the sample population. These parameters all point to a single goal: finding an informative marker tightly linked to the gene of interest, which in turn

will allow the accurate prediction of the genotype. More markers mean greater coverage and high levels of polymorphism which are a prerequisite for markers associated with a gene to be informative. The level of polymorphism within the family or sample population is important as populations with low variability contain fewer informative markers.

2.2.1 Marker types

2.2.1.1 Restriction Fragment Length Polymorphism (RFLP)

Restriction Fragment Length Polymorphisms (RFLP) were developed in 1974 (Grodzicker *et al.*, 1974), even though the first human mapping study to use this marker was only published in 1980 (Botstein *et al.*, 1980). Restriction Fragment Length Polymorphisms are based on the digestion of genomic DNA (gDNA) by specific endonucleases, yielding fragments of differing lengths. Polymorphisms are observed as differences in electrophoretic mobility on a gel. In order to identify the specific DNA fragment underlying an RFLP fragment of interest, hybridization by Southern blotting has to be performed (Southern, 1975). Restriction Fragment Length Polymorphisms are co-dominant markers meaning that homozygotes and heterozygotes can be differentiated.

Restriction Fragment Length Polymorphism-derived marker loci are highly polymorphic and are well spaced across a genome (Botstein *et al.*, 1980). Restriction Fragment Length Polymorphisms is not a Polymerase Chain Reaction (PCR)-based method and detection of markers requires hybridization with radioactive probes, however this has been negated by the advent of fluorescent technology. Automating RFLPs is difficult and the amount of DNA needed is fairly large (1-10 μg) but Southern blots prepared from RFLP fragments can be re-probed many

times which makes the technique slightly more feasible (Rafalski and Tingey, 1993).

One of the obstacles in mapping wheat populations is the lack of polymorphism. Restriction Fragment Length Polymorphisms used in combination with deletion or aneuploid wheat populations provide a suitable alternative as mapping is done in hemizygous or homozygous form and any probe can be used without identifying polymorphism (Werner *et al.*, 1992). A setback for the use of RFLPs in MAS is that it is expensive, time consuming and to reduce costs these markers need to be converted to a PCR-based system (Ma *et al.*, 1998).

Through reverse genetics, RFLPs can be used to detect genes governing important phenotypic traits. Ma *et al.* (1998) used RFLPs in conjunction with aneuploid wheat stocks to map the *Dn2* and *Dn4* resistance genes in the donor parents PI262660 and PI372129 (Table 2.1). In 2004, the same authors used 212 RFLPs to map QTLs against net blotch in barley (Ma *et al.*, 2004). Additional RFLP markers have been identified linked to the RWA resistance gene, *Dn7* (Table 2.1).

Table 2.1: Restriction Fragment Length Polymorphisms linked to RWA resistance genes.

Resistance gene	**RFLP marker**	**Distance**	**Reference**
Dn2	*Ksua1*	9.8 cM	Ma *et al.* (1998)
Dn2	*Xksua1*	9.9 cM	Miller *et al.* (2001)
Dn4	*Xabc156*	11.6 cM	Ma *et al.* (1998)
Dn4	*Xksue18*	16 cM	Ma *et al.* (1998)
Dn4	*Xksud14*	32.5 cM	Ma *et al.* (1998)
Dn7	*Xmwg2062*	10.6 cM	Anderson *et al.* (2003)
Dn7	*Xwrga2*	5.3 cM	Anderson *et al.* (2003)
Dn7	*Xbcd1434*	1.4 cM	Anderson *et al.* (2003)
Dn7	*Xksud14*	7.4 cM	Anderson *et al.* (2003)
Dn7	*Xmwg36*	8.6 cM	Anderson *et al.* (2003)
Dn7	*XksyF43*	17.6 cM	Anderson *et al.* (2003)

2.2.1.2 Random Amplified Polymorphic DNA (RAPD)

One of the first papers to describe RAPD analysis was by Williams *et al.* (1990) who describe the technique as Arbitrarily Primed Polymerase Chain Reaction or AP-PCR. Devos and Gale (1992) stated that RAPDs will be of limited use in the linkage mapping of wheat but perhaps it will be useful for the characterization of introgressed chromosome segments. However, the technique did gain popularity because of its ease of use and simplicity (Rafalski and Tingey, 1993).

Random Amplified Polymorphic DNAs are PCR-based markers but are distinct in that they rely on amplification of gDNA with single primers of which the nucleotide sequence is arbitrary (Welsh and McClelland, 1990). Random Amplified Polymorphic DNA is a dominant marker system and cannot distinguish between homozygotes and heterozygotes. It relies on the detection of polymorphisms in the form of nucleotide mismatches (Myburg *et al.*, 1998). RAPD amplifications can be viewed on an agarose gel with simple stains such as ethidium bromide. No hybridization or radioactive labelling is required (Rafalski and Tingey, 1993). However, RAPD markers need to be converted into more stringent markers to increase specificity. Such converted markers are known as Sequence Characterized Amplified Regions (SCAR). The conversion to SCAR markers is often unsuccessful as a polymorphism based on short arbitrary primers may result in the loss of the initial polymorphism. Another option to SCAR markers however, is the generation of PCR-RFLP as it is cheaper and involves no sequencing of internal bases (Venter and Botha, 2000).

Interestingly, RAPD markers had been identified that segregate with a RWA resistance phenotype though the amplicon is absent in both parents. These are known as non parental fragments, co-segregating with the gene of interest (Myburg

et al., 1998). Another anomaly observed with RAPD markers is the occurence of repulsion-phase markers, which is the absence of an amplicon in the heterozygous resistant offspring. This suggests the inability of the RAPD primers to prime at their target loci in the presence of the respective resistance alleles due to possible template competition effects (Myburg *et al.*, 1998). This phenomenon has been observed for all repulsion-phase markers linked to *Dn* resistance genes against RWA (Myburg *et al.*, 1998).

Random Amplified Polymorphic DNA markers are sensitive enough to detect single base changes though RAPDs are poorly reproducible and difficult to transfer between laboratories and genetic backgrounds (Qi and Lindhout, 1997; Myburg *et al.*, 1998; Venter and Botha, 2000). This marker system requires only small amounts of DNA (15-25 ng), a non radioactive and simple setup and is a quick and efficient way to screen for sequence polymorphisms in large numbers of loci (Rafalski and Tingey, 1993). Unlike RFLPs, no prior sequence information is required for RAPD analysis and this marker system has the added advantage of employing a universal set of primers. Each RAPD is comparable to a Sequence Tagged Site (STS) and determining genotypes within a population can be automated, something to which the RFLP is less amenable (Myburg *et al.*, 1998; Venter and Botha, 2000).

Some of the most documented marker systems in wheat are RFLPs and RAPDs (Devos and Gale, 1992; Schachermayr *et al.*, 1994; Demeke *et al.*, 1996). To date, many RAPD markers have been linked to RWA resistance genes. Venter *et al.* (1998) identified a marker 43.7 cM from *Dn1* and another marker 4.4 cM from *Dn2*. Myburg *et al.* (1998) identified four RAPD markers from an initial set of 2 700, closely linked to *Dn2* (OPB10880c at 3.3 cM; OPM91600r at 3.3 cM; OPN1400r

at 3.3 cM and OPO11900c at 4.4 cM) though only two of these markers were successfully converted to SCAR markers. Random amplified polymorphic DNA markers have also been used to successfully tag several other resistance genes in wheat ranging from leaf rust (*Lr9* and *Lr24*) (Schachermayr *et al.*, 1994) to powdery mildew (*Pm21*) (Hartl *et al.*, 1993; Qi *et al.*, 1996), the *Bt-10* common bunt resistance gene (Demeke *et al.*, 1996) and a wheat streak mosaic virus resistance gene (Talbert *et al.*, 1996).

2.2.1.3 Amplified Fragment Length Polymorphism (AFLP)

Amplified Fragment Length Polymorphism (AFLP) was first described by Vos *et al.* (1995) as a new method of DNA fingerprinting. Vos *et al.* (1995) state that an ideal fingerprinting method should not require investment in sequence characterization or primer design and AFLP, like RAPD, conforms to this criterion.

The principle behind AFLP is based on the detection of gDNA restriction fragments of varying lengths by PCR amplification. Two restriction enzymes, a rare and frequent cutter, are used. The frequent cutter is expected to produce smaller fragments that amplify well and can be optimally separated. The rare cutter reduces the number of fragments produced. Amplified Fragment Length Polymorphisms produce a range of fragment lengths. Polymorphisms are observed as the absence of a fragment of a certain size in one sample, that is present in another sample. It is suitable for DNA of varying complexity and origin. What sets AFLP apart from other whole genome marker systems is the ability to filter the number of fragments detected by using specific selective primers (Vos *et al.*, 1995). Amplified Fragment Length Polymorphism is a dominant marker system.

Amplified Fragment Length Polymorphisms allow for the simultaneous identifi-

cation of a large number of amplification products (Qi and Lindhout, 1997) which gives it the possibility of producing an infinite number of markers. It is highly reproducible however it is sensitive to reaction conditions and the quality of DNA used. Amplified Fragment Length Polymorphisms are less suitable to single locus studies including MAS, MBC and allele frequency studies. In these cases, there is a need to convert AFLPs into single locus markers such as Cleaved Amplified Polymorphic Sites (CAPS) or SCAR, as is also the case with RAPDs. However, creating single locus markers from AFLP fragments is not simple as often, multiple fragments are hidden within a single AFLP band (Brugmans *et al.*, 2003).

Amplified Fragment Length Polymorphisms have been applied in wheat mapping projects though to date there are no recorded AFLP markers linked to any of the RWA resistance genes. Penner *et al.* (1998) created a molecular map based on 325 AFLP and microsatellite markers using a DH population, while Boyko *et al.* (1999) used AFLPs to construct a high density genetic map of *Aegilops tauschii*. Peng *et al.* (2000) made use of AFLPs to construct a molecular map of wild emmer wheat (*Triticum dicoccoides*).

In a study by Zaayman *et al.* (2009), complementary DNA-AFLP (cDNA-AFLP) was used to identify transcripts differentially expressed in resistant and susceptible wheat lines infested by different RWA biotypes. Though not a mapping study in itself, this paper shows the capability of the AFLP technique to identify candidate genes associated with specific phenotypic traits without any prior sequence knowledge.

Amplified Fragment Length Polymorphism as a marker system was selected for use in this study because of its capacity as a whole-genome marker system that does not require any sequence information in order to generate markers spanning

the entire genome. Unlike RFLPs and RAPDs, AFLPs are highly reproducible and require fairly little starting material of moderate quality. AFLPs are also PCR-based, a criterion that made RFLPs unsuitable for this study. In addition, the number of polymorphic loci detected by AFLP is highest among the techniques already discussed (Rafalski and Tingey, 1993).

2.2.1.4 Sequence Tagged Site (STS)

An STS is defined as a short stretch of DNA that is unique in that it is amplified from only one site in the genome and it is detected by PCR (Green and Green, 1991). By simply sequencing any mapping landmark and designing primers to amplify the fragment a STS marker can be generated from almost any DNA sample. The size of an operational STS in the human genome equates to 200-500 bp (Olson *et al.*, 1989). Sequence Tagged Site markers can be derived from already-informative DNA fragments such as BAC end sequences or EST libraries (Šimková *et al.*, 2011).

Sequence Tagged Site markers are easy to transfer between laboratories as only primer sequences, PCR setup and fragment sizes are required to amplify the marker from gDNA (Olson *et al.*, 1989). However, STS markers are not suitable for high-throughput screening (Green and Green, 1991).

An example of success with EST-STS markers comes from the mapping of a Greenbug resistance gene *Gb3* at 0.08 cM (Šimková *et al.*, 2011). Sequence Tagged Sites along with SSRs were the markers of choice to map one of the most recently described RWA resistance genes, *Dn2414*. Three markers *Xiag95*, *Xksu951* and *Xrems-cw* were found tightly linked to the gene (Peng *et al.*, 2007).

2.2.1.5 Expressed Sequence Tag (EST)

Expressed Sequence Tags (ESTs), like STSs are usually short fragments of 200-800 bp and are created from unedited, randomly selected single pass sequence reads from a cDNA library. These markers can be generated at high throughput fairly inexpensively (Nagaraj *et al.*, 2007). As a mapping tool in wheat, ESTs can be useful due to the clustered nature of the genic regions in the wheat genome. Expressed Sequence Tags have been used in comparative sequence analysis of both rice and wheat (La Rota and Sorrells, 2004). Qi *et al.* (2004) used ESTs to construct a chromosome bin map of 16 000 markers in order to distribute genes among the three genomes of bread wheat. Swanepoel *et al.* (2003) mapped two EST markers derived from diploid progenitors of wheat, to the *Dn1* resistance gene at 7.41 cM (*AMO00SSHL1*; GenBank Accession AF4446141.1) and 3.15 cM (*NBS-RGA2*; GenBank Accession AF326781).

2.2.1.6 Microsatellites

Microsatellites or SSRs are an example of hypervariable markers. These markers have variable lengths within a population and different alleles are discriminated based on differing sizes. This variability make microsatellites highly informative in linkage studies (Nakamura *et al.*, 1987). Microsatellites are made up of repeat sequence motifs ranging from fewer than ten to hundreds of bases in their total length. The size of the repeat motifs range from two to six bp. The repeating motifs can be categorized as simple ((CA)n); compound (two or more microsatellites found in close proximity) or complex (containing repeat units of several nucleotides), either of which may be interrupted or not (Kofler *et al.*, 2008). Microsatellites are co-dominant markers that can be easily visualized on gel based

systems (Song *et al.*, 2005). Microsatellites found in genes or ESTs are referred to as eSSRs. These eSSR markers are physically associated with coding regions and can enhance the role of markers during the evaluation of germplasm. Peng and Lapitan (2005) constructed a consensus chromosome map using eSSRs in wheat.

Microsatellites are easily automated on high throughput systems (Somers *et al.*, 2004) and are generally more specific compared to techniques such as RFLPs (Song *et al.*, 2005). Automating microsatellite analysis also alleviates any difficulties in genotyping these markers which often arises in dinucleotide repeats due to strand slippage (Song *et al.*, 2005). Microsatellites is a popular marker system and has been employed extensively in mapping projects as well as breeding programs in wheat, despite the fact that the large genome size, polyploidy and repetitive nature of the wheat genome makes microsatellite development time consuming and difficult (Song *et al.*, 2005).

Microsatellites have been used to anchor the physical map of the largest of the wheat chromosomes, 3B (Paux *et al.*, 2008) and between authors such as Song *et al.* (2005) and Röder *et al.* (1998), 534 microsatellites were developed in wheat with many more currently available (Graingenes database). Table 2.2 provides a list of the microsatellite markers that have been linked to RWA resistance genes.

Table 2.2: Microsatellite markers linked to RWA resistance genes.

Resistance gene	SSR marker	Distance	Reference
Dn7	*xbcd14341*	1.4 cM	Anderson *et al.* (2003)
Dn2	*Xgwm437*	2.8 cM	Miller *et al.* (2001)
Dn2	*xgwm44*	12.7 cM	Miller *et al.* (2001)
Dn2	*xpsp3123*	8.1 cM	Miller *et al.* (2001)
Dn2	*xpsp3113*	21.7 cM	Miller *et al.* (2001)
Dn2	*xgwm111*	3.2 cM	Miller *et al.* (2001); Liu *et al.* (2001)
Cl2401	*xgwm111*	3.2 cM	Valdez *et al.* (2012)
Cl2401	*xbarc214*	0.8 cM	Valdez *et al.* (2012)
Cl2401	*xgwm437*	1.2 cM	Valdez *et al.* (2012)
Dn626580	*xgwm214*	1.8 cM	Valdez *et al.* (2012)
Dn626580	*xgwm473*	5 cM	Valdez *et al.* (2012)
Dn626580	*xgwm437*	8.2 cM	Valdez *et al.* (2012)
Dn6	*xgwm111*	3.35 cM	Liu *et al.* (2002)
Dn6	*xgwm111*	2.82 cM	Liu *et al.* (2002)
Dn6	*xgwm44*	14.63 cM	Liu *et al.* (2002)
Dn1	*xgwm111*	3.82 cM	Liu *et al.* (2001)
Dn2	*xgwm111*	3.05 cM	Liu *et al.* (2001)
Dn5	*xgwm111*	3.2 cM	Liu *et al.* (2001)
Dnx	*xgwm111*	1.52 cM	Liu *et al.* (2001)
Dn8	*xgwm635*	3.2 cM	Liu *et al.* (2001)
Dn5	*xgwm437*	28.6 cM	Heyns (2005)
Dn5	*xgwm111*	25.4 cM	Heyns (2005)
Dn5	*xgwm44*	16.08 cM	Heyns (2005)
Dn5	*xgwm111*	26.5 cM	Heyns (2005)
Dn5	*barc26*	28 cM	Heyns (2005)
Dn5	*xgwm437*	29.03 cM	Heyns (2005)
Dn5	*barc172*	35.95 cM	Heyns (2005)
Dn5	*xwmc94*	38.03 cM	Heyns (2005)
Dn5	*xgdm46*	39.12 cM	Heyns (2005)
Dn5	*xgdm67*	47.97 cM	Heyns (2005)
Dn5	*xwmc157*	78.7 cM	Heyns (2005)
Dn5	*xgwm37*	107.43 cM	Heyns (2005)
Dn5	*barc76*	111.69 cM	Heyns (2005)
DnCl2401	*xcfd68*	2.7 cM	Šimková *et al.* (2011)
DnCl2401	*xbarc214*	2.7 cM	Šimková *et al.* (2011)
DnCl2401	*xgwm473*	2.7 cM	Šimková *et al.* (2011)
Dn1	*xgwm111*	3.15 cM	Swanepoel *et al.* (2003)
Dn1	*xgwm635*	7.41 cM	Swanepoel *et al.* (2003)

2.2.1.7 Single Nucleotide Polymorphism (SNP)

Single Nucleotide Polymorphisms (SNPs) are some of the most commonly occuring polymorphisms (Brookes, 1999; Deschamps and Campbell, 2010) and are defined by single base sequence differences. Single Nucleotide Polymorphisms are biallelic markers, which makes them less informative than hypervariable markers such as microsatellites when viewed individually but their abundance makes up for this shortcoming. Single Nucleotide Polymorphisms in large numbers allow for the construction of high density genetic maps (Brumfield *et al.*, 2003). Individual SNPs can still directly contribute toward phenotypic variation, especially if found in intragenic regions or promoter regions where they can be used as perfect markers for phenotypic traits (Beales *et al.*, 2005; Konishi *et al.*, 2006).

There are different options available for identifying SNPs such as resequencing of PCR amplicons; electronic SNP discovery in genomic libraries and eSNPs from EST libraries. In addition there are many different SNP assays available and the choice of assay depends on cost, throughput, equipment available, difficulty of assay development and multiplexing potential (Rafalski, 2002).

Some of the advantages of SNPs are that they are amenable to high throughput (Gut, 2001) and do not depend on sizing differences which negates the need for standardization amongst different labratories (Chao *et al.*, 2009). A drawback of SNPs specifically with regards to polyploid species such as wheat, is that most SNPs are sequence variants between homeologous gene sequences rather than being allelic variants. The presence of multi-copy sequences and paralogs adds to the difficulty in correctly scoring SNPs at any one locus between homeologous genomes (Akhunov *et al.*, 2009).

Single Nucleotide Polymorphism discovery in wheat had a slow start (Somers *et al.*, 2003) due to lack of sequence data and low polymorphism in the wheat genome as well as its polyploid and repetitive nature (Edwards and Batley, 2010). Single Nucleotide Polymorphism densities in plants are variable and tend to be low in self-pollinating species. A study comparing 21 gene sequences across 26 wheat germplasm accessions revealed that on average one SNP per 330 bp can be expected in genic regions (Ravel *et al.*, 2006), while other authors working with different germplasm samples (smaller sample set and less diverse) found one eSNP per 540 bp in wheat EST regions (Somers *et al.*, 2003).

Blake *et al.* (2004) used intronic SNP detection to identify SNPs associated with starch biosynthesis in wheat (http://wheat.pw.usda. gov/SNP) and Qi *et al.* (2004) generated a chromosome bin map of ESTs that serve as a valuable source for SNP analysis (eSNPs). With the sequencing of the wheat genome (Mayer *et al.*, 2014) 13.3 million SNPs were identified. To date, no SNP markers have been closely associated to RWA resistance in wheat.

2.2.1.8 Diversity Arrays Technology (DArT)

The proof of concept for DArT markers was reported by Jaccoud *et al.* (2001). Diversity Arrays Technology is a microarray-based, high throughput marker system often used in combination with other marker systems (Gupta *et al.*, 2008). This system reduces complexity (Wenzl *et al.*, 2004) and relies on hybridization as the basis for detecting polymorphisms on solid state platforms. Diversity Arrays Technology allows for high throughput that many other popular marker systems, such as microsatellites, lack. The DArT marker system is competitive in its costs and time (Kilian *et al.*, 2005) and can generate hundreds of biallelic, dominant

markers in a single experiment. DArT markers have been extensively used in wheat mapping projects. The physical map of chromosome 3B was constructed using, among others, DArT markers (Paux *et al.*, 2008) and there are dedicated genotyping platforms for bread wheat (Akbari *et al.*, 2006). Though no DArT markers have been shown with association to RWA resistance genes, a study by Crossa *et al.* (2007) used 242 DArT markers in an AM project for resistance genes against stem rust, leaf rust, yellow rust and powdery mildew as well as QTLs for grain yield (Gupta *et al.*, 2008).

2.3 Sequencing

2.3.1 Sequencing platforms

Sequencing based on chain-termination methods were first published by Sanger *et al.* (1977) and remains a commonly used sequencing technique to this day. Next Generation Sequencing (NGS) technologies were introduced in 2005 and have since revolutionized genomic research. NGS applications, however, extend beyond sequencing and re-sequencing of genomes to applications such as discovery of transcription factor binding sites, as well as coding and noncoding RNA expression profiling (Morozova and Marra, 2008). There are three sequencing platforms that are most frequently employed: the Genome Sequencer FLX from 454 Life Sciences/Roche, Illumina's Genome Analyzer, and Applied Biosystems' SOLiD system (Lister *et al.*, 2009). All three platforms are capable of yielding millions of reads per run in a time frame ranging from ten hours to a few days. Table 2.3 lists the major NGS platforms with their advantages and disadvantages.

Table 2.3: Advantages and disadvantages of various NGS platforms based on Metzker, 2010 and Glenn, 2011.

Sequencing platform	**Advantages**	**Disadvantages**
Roche 454	Longer 330-700 bp reads. Low cost /experiment	High cost per Mb. High reagent cost. High error rate in homopolymer repeats.
Illumina	Low cost instrument and runs. Fast run times. Widely used.	Short 100-150 bp reads.
SOLiD	Low cost per Gbp. High throughput. High accuracy. SOLiD 500 outputs in bases, not colour space.	Unusual informatics with colour space. Short 50-75 bp reads. More gaps in data and less even data distribution. Long run time.
Polonator	Least expensive. Open source to adapt alternative sequencing chemistries.	Users must maintain and quality control reagents. Shortest NGS read length (26 bp).
Helicos	Large number of reads directly from single molecule. Non biased representation of templates for genome and sequence based applications	Service platform available only. Questionable longevity of company. High error rates compared with other reversible terminator chemistries. Short 35 bp reads
PacBio	Single molecule, real-time sequencing. Longest available read length: 869-1100 bp	Error rates. Low total number of reads per run. High cost per Mb. High capital cost. Many methods still in development.
Ion Torrent	Low cost instrument upgraded through disposable chips. Very simple machine. Low cost per sample for smaller studies.	New platform with a variety of unknowns. Highest cost per Mbp of all platforms. Sample preparation takes a long time.

The Illumina sequencing platform is currently the most widely used. Berkman *et al.* (2011; 2012) sequenced the short arms of chromosomes 7D and 7B in wheat using the Illumina GAIIx. The authors obtained reasonably high sequencing coverage of 34 x and 30 x for the chromosome arms respectively which negates the short read length of 101 bp generated by the sequencing platform they employed. However, due to the repetitive nature of the wheat genome, underrepresentation of AT and GC rich regions are likely to occur
using the Illumina platforms (Dohm *et al.*, 2008; Harismendy, 2009) most likely as a result of amplification bias during the preparation of the template.

Platforms such as Illumina, known to provide shorter read lengths, may create difficulties during *de novo* assembly (Morozova and Marra, 2008).
Vitulo *et al.* (2011) used this platform to sequence chromosome 5A in wheat but stated that since they only obtained two times coverage of short reads they will not be able to *de novo* assemble the entire chromosome. In the present study, the Illumina® HiSeq^{TM} 2000 platform was applied for sequencing the Dt7DS and Dt7DL chromosomes.

.

2.3.2 Next Generation Sequencing (NGS) Analysis

Next Generation Sequencing offers the ability to produce more data in a shorter time span than has ever been possible (Metzker, 2010). Table 2.4 lists the most popularly applied software tools for analysis of sequence data.

Table 2.4: List of popular software packages and algorithms for NGS data analysis.

Objective	**Approach**	**Software**	**Reference**
de novo contig assembly	de Bruijn graphs	Velvet	Zerbino and Birney (2008)
		ALLPATHS	Butler *et al.* (2008)
		EULER-SR	Chaisson and Pevzner (2008)
		ABySS	Simpson *et al.* (2009)
		SOAPdenovo	Luo *et al.* (2012)
Similarity searches and asigning gene ontology	Local alignment	FASTA	Pearson and Lipman (1988)
	Gene ontology	BLAST	Altschul *et al.* (1997)
		MegaBLAST	Morgulis *et al.* (2008)
		SSAHA2	Ning *et al.* (2001)
		BLAT	Kent (2002)
		PatternHunter	Ma *et al.* (2002)
Alignment of larger scale sequence comparisons	Hashing of the reads and scanning through the reference sequence provided	RMAP	Smith *et al.* (2008)
		MAQ	Li *et al.* (2008*a*)
		ZOOM	Lin *et al.* (2008)
		SeqMap	Jiang and Wong (2008)
		SHRiMP	(http://compbio.cs.toronto.ed/shrimp)
	Algorithms that hash the reference genome	SOAPv1	Li *et al.* (2008*b*)
		NovoALign	Li and Durbin (2009)
		ReSEQ	(http://code.google.com/p/re-seq)
		Mosaik	(http://bioinformatics.bc.edu)
		BFAST	(http://genome.ucla.edu/bfast)
	Merge-sorting of the reference and read sequences		Malhis *et al.* (2008)
	Burrows Wheeler Transform (BWT)	SOAPv2	(http://soap.genomics.org.cn)
		Bowtie	Langmead *et al.* (2009)
		BWA	Li and Durbin (2009)

Short read assembly is based on de Bruijn graphs and was pioneered by Pevzner and Tang (2001) and Pevzner *et al.* (2001). Approaches other than the de Bruijn graph-based method include prefix tree-based approaches (Warren *et al.*, 2007) and overlap-layout-consensus methods (Hernandez *et al.*, 2008). The speed at which these programs function depend greatly on the size of the genome and the sequencing error rate (Li and Durbin, 2009). High levels of repetitive elements found in large eukaryotic genomes such as wheat, are stumbling blocks during sequence assembly using algorithms that make use of the de Bruijn graph principle, as repeats cause "tangles" in the graph that are not easy to undo. More unique sequences have clear paths on the graph and therefore assemble much better (Berkman *et al.*, 2011).

Data volume as produced by the larger eukaryotic genomes is another hampering factor for *de novo* assembly, and alignment-based methods were often implemented instead. For this reason authors such as Simpson *et al.* (2009) have developed assemblers that can run the assembly process in parallel, increasing the memory available for assembly and allowing handling of much larger data sets. Alignment-based methods and comparisons to known reference genomes are another approach for constructing useable sequence scaffolds from shorter reads. Burrows Wheeler Transform (BWT) algorithms for sequence alignments and comparisons have proven most efficient as exact repeats are collapsed together and reads are not aligned against simpler copies (Li and Durbin, 2009).

An important parameter in sequence data analysis is the k-mer. Marçais and Kingsford (2011) defines a k-mer as a substring of length k in which a string, S occurs. These length k-substrings are the k-mers and determining their occurence in numbers, is called k-mer counting. Counting the k-mers is important as many

genome assemblers use this parameter to find overlaps in sequence and k-mer count statistics can be used to estimate the genome size (Marçais and Kingsford, 2011). The frequencies of k-mers are used to assess the likelihood of sequencing errors (Kelley *et al.*, 2010) and in *de novo* repeat annotation k-mer frequencies are used to identify candidate regions (Campagna *et al.*, 2004; Healy *et al.*, 2003). Popular k-mer counting algorithms include Jellyfish (Marçais and Kingsford, 2011) and KmerGenie (Chikhi and Medvedev, 2013).

Estimation of the genes present in a sequencing data set is a valuable outcome and several strategies are available. Vossen *et al.* (2013) describe an experimental approach wherein degenerate primers are designed for specific motifs associated with, for example, resistance genes. This motif directed profiling allows gene candidates to be amplified and either analyzed on a gel, cut out and sequenced or PCR products are directly sequenced using NGS platforms. As an alternative, protein coding genes (PCG) can be searched using software packages such as FunCat (Functional Catalogue) for the systematic classification of proteins from whole genome data (Ruepp *et al.*, 2004) or Gene Ontology (GO; Ashburner *et al.*, 2000). Several sequencing studies also make use of manual data mining using BLAST algorithms to compare genome sequence data to known genes in databases such as Ensembl (http://ensemblgenomes.org/; Thole *et al.*, 2012), SwissProt (Vitulo *et al.*, 2011), harvEST (Mayer *et al.*, 2011; Brenchley *et al.*, 2012) and GenBank UniGene (Vitulo *et al.*, 2011).

2.4 Isolation of Wheat Chromosomes

The wheat genome is large with a complex organization and a high proportion of repetitive DNA (75%) (Vedel and Delseny, 1987; Wicker *et al.*, 2003). These repetitive regions are interspersed by clusters of unique, low-copy DNA along the length of the chromosome (Akhunov *et al.*, 2003). Additionally, the wheat genome contains 42 chromosomes, many of which are similar in size (Mayer *et al.*, 2014), making karyotyping, mapping and gene cloning challenging (Šafář *et al.*, 2004).

Flow cytomtery is a widely applied automated technique used during karyotyping and sorting of chromosomes. Cells used for flow cytometry can be isolated from different tissues though a high mitotic index is required in order to produce the best chromosome suspensions (Metezeau *et al.*, 1993). Chromosomes are stained just before they are analyzed and there are various fluorochromes to choose from based on specificity, quantum yield and matching of their optical properties to the instrument's wavelength. Ethidium bromide and propidium iodide label AT and GC bp indifferently and are excited at 488 nm. Dyes such as Chromomycin A3 is specific to GC bp and excites at 458 nm while Hoescht stain is specific to AT bp and excites at 351 to 364 nm (Metezeau *et al.*, 1993). The fluorophore selected for labelling chromosomes allow classification of the individual chromosomes according to DNA content. Flow cytometry has the ability to separate large numbers of chromosomes. Between 30 and 50 chromosomes can be separated per second (Vrána *et al.*, 2012).

Initially flow cytometry was used in aneuploid or ditelosomic wheat lines or wheat-rye addition lines (Pfosser *et al.*, 1995) . Flow cytometry observes genetic anomalies resulting from variations in DNA content (such as deletions, insertions and translocations (Metezeau *et al.*, 1993)) and by using aneuploids no chromoso-

mal separation is needed and flow cytometry can be sufficiently sensitive to determine overall chromosome-related differences within a genome eg. determination of the composition of wheat-rye addition lines (Bashir *et al.*, 1993) .

Kubaláková *et al.* (2002) demonstrated that wheat chromosomes could not only be flow sorted intact but also separated. However, due to the lack of size differences between chromosomes, only 3B could be isolated with high purity. Estimated at 995 Mega base pair (Mbp) (Gill *et al.*, 1991; Mayer *et al.*, 2014) chromosome 3B is twice as large as the entire rice genome (370 Mbp; Itoh *et al.*, 2007), making chromosome 3B an attractive candidate for initial attempts at physical mapping and sequencing (Vrána *et al.*, 2000). As flow cytometric techniques improved, smaller chromosomes could be separated using telosomic wheat lines and lines carrying isochromosomes (Kubaláková *et al.*, 2002)). Berkman *et al.* (2011; 2012) isolated and sequenced chromosomes 7DS and 7BS using this method, and the reported sizes of these chromosomes are estimated at 381 Mbp and 360 Mbp for 7DS and 7BS respectively (Mayer *et al.*, 2014).

The major drawback of flow cytometry, besides the limitation in the size separation, is the requirement of high concentrations of DNA which equates to millions or billions of isolated chromosomes (Vrána *et al.*, 2012). Though wheat lines containing chromosome aberrations technically allow for the separation of any of the wheat chromosomes, breeding such lines is time consuming and not always feasible.

Microdissection is a viable alternative for isolating individual chromosomes that circumvents the "similarity of size" problem faced by flow cytometry. Scalenghe *et al.* (1981) were the first to develop microdissection as a technique to isolate fragments of chromosomes for the purpose of isolating DNA for cloning. The first report on microdissection in plants came from the isolation of the B chromosome

from rye in 1991 (Sandery *et al.*, 1991). Though microdissection is often followed by cloning of some form in order to increase yield and allow for downstream work, direct amplification from microdissected product has also been demonstrated (Maças *et al.*, 1993).

Initially, microdissection involved dissecting the chromosome with glass microneedles (Scalenghe *et al.*, 1981) which is a technically challenging task. The development of laser microbeams in microdissection (Monajembashi *et al.*, 1986) eased the isolation process but still involved the use of the glass microneedles to collect the specimen isolated via the laser (Zhou and Hu, 2007). With the introduction of the PALM® Robot-Microbeam system, that uses a laser to cut tissue from microscope slides and subsequently uses that same laser to catapult fragments into a collection tube (Olofsson *et al.*, 2012), the glass needles were negated and accuracy improved.

Zhou and Hu (2007) stated that the use of microdissection can definitely benefit genomic research on organisms with large and complex genomes such as wheat. However, the authors stated that at that time microdissection could not meet the requirements for sequencing complete chromosomes because of low coverage and small size of the DNA fragments. Microdissection in combination with deletion lines was applied in the present study.

2.5 Wheat

2.5.1 The wheat genome

Wheat is an allopolyploid cereal (2n = 6x = 42) in that it has more than two genomes brought together into the same organism by hybridization events followed

by chromosome doubling (Feldman and Levy, 2005). The first paper published, covering the entire wheat genome sequence, was by Brenchley *et al.* (2012) who produced a draft genome sequence of wheat with 5 x coverage of the 17 Gbp genome. Subsequently, in 2014, a draft genome sequence of wheat derived from the sequencing of individual chromosomes was published and therein is reported the estimated sizes of every chromosome in wheat (Mayer *et al.*, 2014).

2.5.2 Evolution of wheat from diploid and tetraploid progenitors

Wheat and rice split from a common ancestor 55 to 75 million years ago (MYA) (Gill *et al.*, 2004). Hexaploid wheat evolved from the addition of the D genome from *Aegilops tauschii* to a tetraploid (AABB) ancestor, *Triticum turgidum*. The tetraploid ancestor evolved from amphiploidy between *Triticum urartu* (AA) and diploids similar to *Aegilops speltoides* (BB). Figure 2.1 depicts this graphically. Since there are some discrepancies between the genome of *A. speltoides*, the possible B genome donor, and the actual B genome of modern bread wheat, it is debated as to whether *A. speltoides* is in fact the donor of the B genome. However, the lack of similarity might also be attributed to rearrangements in the genome of the tetraploids. Albeit, it is also possible that the original donor plant of the B genome went extinct (Sarkar and Stebbins, 1956). The sub-genomes in modern hexaploid wheat differ significantly from one another and many genes are not present in triplicate, but are chromosome specific (Gill *et al.*, 2004). Genes are also not distributed randomly, but are clustered in gene-rich regions, particularly at the distal regions of chromosomes (Gupta *et al.*, 2008).

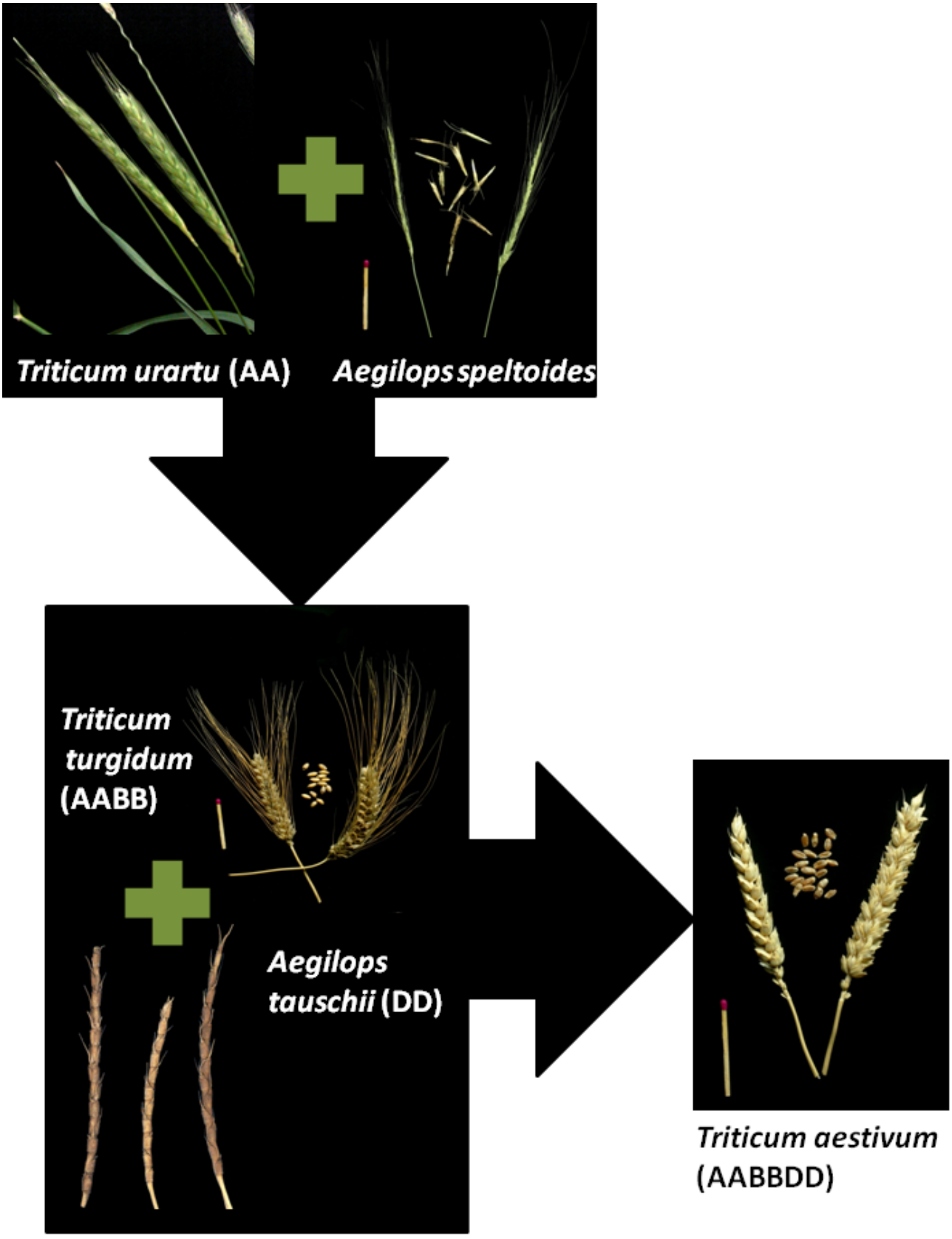


Figure 2.1: Diagram depicting the origins of hexaploid wheat from its diploid and tetraploid progenitors. Images adapted from: http:// commons. wikimedia. org, http:// www. sortengarten. ethz. ch and http:// www. k-state. edu.

2.5.3 Comparitive mapping

Paux *et al.* (2008) aligned the physical map of chromosome 3B from wheat to the rice genome and identified four inversions and non-colinear genes. Their data confirms rearrangements observed on the most conserved wheat and rice chromosomes (La Rota and Sorrells, 2004), and suggest that these local rearrangements have occurred in global colinear regions since rice and wheat diverged (Paux *et al.*, 2008).

Vitulo *et al.* (2011) also aligned their sequence data from chromosome 5A to related species, in this case *Brachypodium distachyon*, rice and sorghum. The majority of coding sequences had homologues in all three related species but several thousand reads from wheat showed unique homologues in only one related species - the most of these having a homologue only in *B. distachyon*. Syntenic regions with *B. distachyon* are distributed over several chromosomes, particularly for the long arm of chromosome 5A confirming a low conservation in gene order (Sorrells *et al.*, 2003). Despite this distribution, regions of synteny could be identified with chromosome 5AS related to *B. distachyon* chromosome four, rice chromosome 12 and sorghum chromosome eight. Chromosome 5AL from wheat showed relatedness to *B. distachyon* chromosomes four and one, rice chromosome nine and three and sorghum chromosome one and two.

Berkman *et al.* (2011) aligned their sequence data from chromosome 7DS to *B. distachyon* and rice and found regions aligning on chromosome one and three of *B. distachyon* and chromosome six and eight of rice. Further comparisons showed conservation within annotated genes between the three species but, except for conserved regions of repetitive elements, there was little conservation in the regions outside of the annotated genes. Berkman *et al.* (2012) reported on the

translocation event between chromosomes 7BS and 4AL in wheat and showed that 13% of genes from 7BS had been translocated to chromosome 4AL and 13 genes were observed on chromosome 7BS that seemed to have originated from chromosome 4AL.

Brenchley *et al.* (2012) showed a high degree of similarity between gene sets in *B. distachyon* and wheat and pointed out chromosomes one and four from *B. distachyon* as areas of lower conservation. The authors confirmed the insertions and translocations of gene blocks as observed in other studies (Hernandez *et al.*, 2012) and identified a previously undescribed syntenic alignment of wheat group five chromosomes and *B. distachyon* chromosome three genes as depicted in Figure 2.2.

Brenchley *et al.* (2012) also observed an overall reduction in gene family sizes in modern hexaploid wheat compared to its diploid progenitors. Despite this, expanded gene families were still present and were found to be common to wheat and *Ae. tauschii*. These gene families included ribosome proteins, photosystem II components, storage and transposon-related proteins, cytochrome P450s, NB-ARC domain proteins and F-box proteins. In addition, the authors identified 233 000 gene fragments that are potential pseudogenes. Brenchley *et al.* (2012) also reported an overall gene loss of between 10 000 to 16 000 genes in wheat when compared to the three diploid progenitor species.

2.5.4 Sequencing of the wheat genome and its progenitors

2.5.4.1 Sequencing of individual chromosomes

Sequencing of the wheat genome started with isolated, individual chromosomes. Vitulo *et al.* (2011) sequenced chromosome 5A with an estimated size of 857.8 Mbp

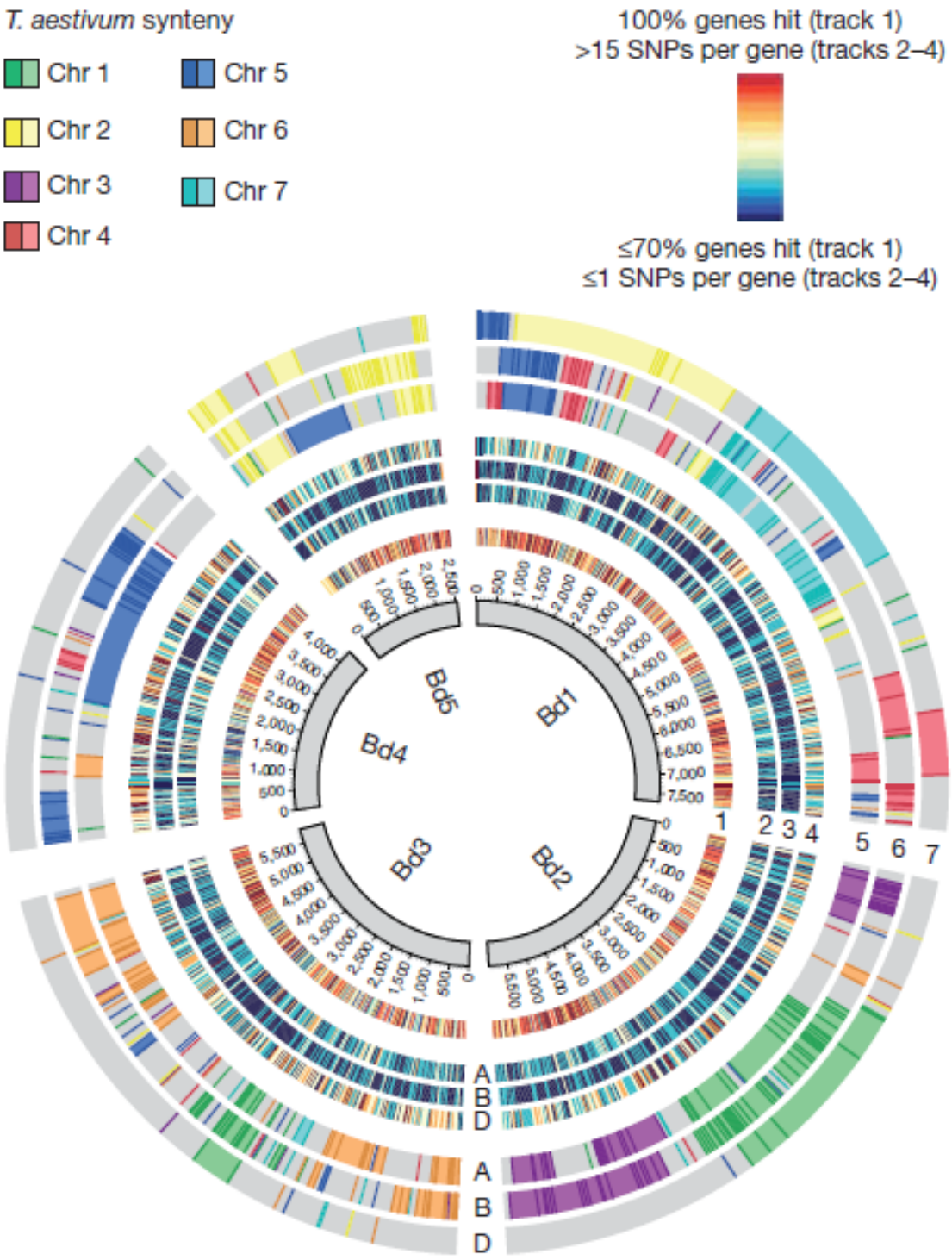


Figure 2.2: Alignment of wheat 454 sequencing reads, SNPs and genetic maps to the *B. distachyon* genome taken from Brenchley *et al.* (2012). The inner-most circle is representative of gene order on the five *B. distachyon* chromosomes. Track one shows 454 reads and *B. distachyon* gene conservation, as a window of genes present in wheat. Tracks two to four depict SNP density in the A (track two), B (track three) and D (track four) genomes of wheat. Tracks five to seven indicate wheat synteny with *B. distachyon* for the A (track five), B (track six) and D (track seven) genomes. Genetic markers (shown in darker colours) are colour-coded by wheat chromosome. Gaps between markers are filled in to show synteny (lighter colours).

on the Roche-454 sequencing platform. The coding fraction of each of the chromosome arms was found to be 1.08% and 1.3% for 5AS and 5AL respectively. When comparing their data to the Triticeae repeat sequence database (TREP) they observed that repetitive elements constituted 76.13% of the short arm and 82.23% of the long arm of the chromosome. In addition, 195 candidate miRNAs were identified.

Berkman *et al.* (2011; 2012) sequenced chromosomes 7DS and 7BS with estimated sizes of 381 Mbp and 360 Mbp respectively (Šafář *et al.*, 2010) using the Illumina® HiSeq^{TM} 2000. The authors enriched for low copy regions as repetitive elements are difficult to assemble and obtained chromosome coverage of 40% and 49% for 7DS and 7BS respectively. They observed a translocation between 4AL and 7BS which could be delimited based on sequence homology to 7DS. Chromosome 4A was sequenced by Hernandez *et al.* (2012) who identified 9 571 genes on the 856 Mbp chromosome.

2.5.4.2 Sequencing of the diploid D genome progenitor, *Aegilops tauschii*

Jia *et al.* (2013) published a draft genome sequence of the diploid D genome progenitor, *Ae. tauschii*, and obtained an assembly representative of 97% of the genome which is 4.36 Gbp in size, using the Roche-454 sequencing platform. The percentage of repetitive elements are slightly lower in *Ae. tauschii* than in wheat, with only 65.9% repetitive elements reported (Jia *et al.*, 2013). The estimated number of genes in the genome of *Ae. tauschii*, is 34 498 which does correspond, roughly, to a third of the total number of genes for the hexaploid bread wheat genome (Mayer *et al.*, 2014).

2.5.4.3 Whole genome sequence

Apart from the size and complexity of the wheat genome, the large fraction of repetitive elements has hindered sequencing attempts. Identifying and classifying these repeat elements are a part of every sequencing study in wheat and often, repetitive elements are masked or removed in order to perform further downstream analyses without their confounding effects. Retrotransposons alone constitute more than 67% (Paux *et al.*, 2006) of the genome and such tandemly repeated sequences can easily cause misassembly of BAC clones (Paux *et al.*, 2008) or sequencing reads (Berkman *et al.*, 2011).

In 2012 with the publication of the draft genome sequence for the entire wheat genome, Brenchley *et al.* (2012) stated that repeat elements (from comparisons to TREP) accounted for 79% of the genome and that most of these elements consisted of retrotransposons. This estimation was higher than that reported by Paux *et al.* (2006) but the class of repeat elements found was consistent with previous studies (Paux *et al.*, 2006; Choulet *et al.*, 2010).

The wheat genome sequence data as reported by the IWGSC (Mayer *et al.*, 2014) found 76.6% repetitive elements in their assembled sequences, a number close to that described by Brenchley *et al.* (2012) in the draft genome sequence. This number was, however, much higher in the raw sequence reads (81%). The IWGSC found retroelements to be most abundant in the A genome and least so in the D genome while DNA transposons were found to display the opposite, being most abundant in the D genome and least so in the A genome. Unclassifiable long terminal repeat (LTR) retrotransposons showed a gradient of abundance in the three sub-genomes, being most prevalent in the A genome and least prevalent in the B genome. Mayer *et al.* (2014) stated that these retrortansposons might

represent older, more deteriorated elements. This suggests that the B genome progenitor contained a lower number of these retroelements and that their activity after polyploidization introduced a high proportion of recent amplifications in the B sub-genome.

Brenchley *et al.* (2012) estimated the total number of genes in the wheat genome to be between 94,000 and 96,000. Two years later, with the publication of the IWGSC's draft genome sequence the estimated number was lower, estimated at around 59 307 (Mayer *et al.*, 2014).

The draft genome sequence as published by Brenchley *et al.* (2012) focused on sub-genome-specific characterization as the sequencing data was not separated into chromosomal origin. The draft genome sequence published by the IWGSC, however, could be delineated into chromosomal origins (Mayer *et al.*, 2014). Their chromosomal sequence assemblies contain a high number of the total gene content of wheat as their gene sets mapped to more than 96% of HarvEST (publicly available wheat ESTs) sequences. The B sub-genome was shown to contain the highest number of genes (35%) with the D sub-genome containing the least (32%). This distribution was not mirrored at the chromosomal level however homeologous chromosomal groups showed varying degrees of gene distribution. As an example, homeologous group three had the highest number of genes on 3A while homeologous group seven had the highest number of genes on 7D. The authors speculated that this is due to pre-existing variability in the sub-genomes or that drivers governing genome composition can be said to act regionally and not on the sub-genome level (Brenchley *et al.*, 2012; Mayer *et al.*, 2014).

Gene loss was reported to be mostly associated with expanded gene families (Mochida *et al.*, 2009; Brenchley *et al.*, 2012). The most recent addition to the

poliploid genome (sub-genome D) showed gene loss at lower levels than the older sub-genomes, A and B. Compared to the diploid and tetraploid progenitor species, hexaploid wheat contained a high number of substitutions that impacted protein functionality which suggests gene redundancy. This, in turn, allows for sequence evolution, potentially, toward proteins with novel functions (Brenchley *et al.*, 2012; Mayer *et al.*, 2014).

2.6 Russian Wheat Aphid

2.6.1 Background

Aphids employ complex feeding strategies and have adapted to a wide range of hosts and habitats. Along with the emergence of insecticide resistance, their adaptability make them successful pests (Walling, 2008). For RWA this success has resulted in estimated losses (direct and indirect) of more than $800 million in the western USA between 1987 and 1993 (Haley *et al.*, 2004) and South Africa has seen 21 to 92% losses on susceptible cultivars since the introduction of the RWA in 1978 (Walters *et al.*, 1980; Hewitt, 1988). Resistant wheat cultivars promise the most effective strategy against RWA as aphids hide in the rolled leaves rendering even broad spectrum insecticides ineffective (Basky, 2003).

Aphid food ingestion is passive and driven by the pressure in sieve tubes (Will *et al.*, 2008). Aphids use modified mouthparts called stylets to navigate the host plant's cuticle layer, epidermis and mesophyll in order to reach the phloem sieve elements from which it draws its food. The paths of the stylet through the host plant is multi-branched and aphids make use of multiple feeding sites. When a sieve element is pierced by the aphid stylet, it must be plugged to avoid losing phloem

sap. Plants achieve this by depositing callose and proteins, however, aphids have developed ways to inactivate this function and instead use saliva that cement their stylet sheath to the sieve element (Tjallingii and Hogen Esch, 1993; Will *et al.*, 2007).

Aphids secrete two types of saliva: sheath saliva and watery saliva. Sheath saliva is secreted as the aphid stylet pierces the epidermis and mesophyll of its host. It limits damage to epidermal cells and contact with extracellular plant defences and reduces the loss of phloem sap through the wound. Sheath saliva is rapidly gelling and forms a thickening at the leaf surface to avoid slippage of the stylet. Watery saliva is secreted when aphids start piercing the sieve elements and is secreted until successful penetration is achieved and feeding can commence. It is hypothesized that watery saliva also interferes with sieve plate-occlusion by binding calcium in the sieve elements and preventing calcium from reaching a threshold level. Watery saliva may be just as important in preventing coagulation in the food canal of the aphid (Will *et al.*, 2007).

Sheath saliva is composed of proteins, phospholipids and conjugated carbohydrates while watery digestive saliva is a complex mixture of enzymes and other components (Miles, 1999). Calcium ion (Ca^{2+}) binding proteins that antagonize protein deposits have been observed in other aphid species (Walling, 2008). It is known that compounds within aphid saliva are responsible for eliciting defence responses in plants.

Russian wheat aphid originated in Central Asia through the Middle East and is now present in all cereal producing countries except Australia (Shea *et al.*, 2000; Stary *et al.*, 2003). Though it occurs in Hungary, the Czech Republic and Germany it is not yet considered a pest of central Europe as damage caused by aphids differs

between regions (Basky, 2003). In South Africa, RWA was first recorded in 1978 (Walters *et al.*, 1980). The first report of RWA in North America (Mexico) was in 1981 (Gilchrist *et al.*, 1984). By 1986 RWA was reported in the USA (Stoetzel, 1987; Webster *et al.*, 1987).

2.6.2 Russian wheat aphid Biotypes

Plant resistance and RWA virulence undergo an adaptation and counter adaptation kind of evolution (Botha, 2013). Aphids adapt to an existing resistant host by forming new biotypes that are morphologically similar to the original but differ in their behavior, such as preference for certain host genotypes (Dreyer and Campbell, 1987). New aphid biotypes have little nuclear and mitochondrial sequence variation (Shufran *et al.*, 2007; De Jager *et al.*, 2014; Botha *et al.*, 2014) and are not distinguished based on morphology but by their ability to overcome host plant resistance, their fecundity and the damage they cause to different cultivars (Smith *et al.*, 1991). The designation of a biotype is based strictly on phenotypic response as a result of aphid feeding.

The first new biotype identified in the USA was described in Colorado in 2003 (Haley *et al.*, 2004). This biotype was particularly concerning as it was virulent to the resistance gene *Dn4* which was the major resistance gene used commercially at the time. In addition it was also resistant to eight of the other resistance genes available. Burd *et al.* (2006) identified three new biotypes in Texas and Wyoming based on chlorosis scores and leaf rolling. Russian wheat aphid infestation on commercially resistant cultivars of wheat in the eastern Free State in South Africa was reported during the 2005 growth season. This was one of the first incidences of the emergence of new biotypes in SA (Tolmay *et al.*, 2007). It is clear that

there must be genetic differences between biotypes to enable them to feed on the different cultivars and to induce the varied responses seen in the plants.

Reasons for differences observed between biotypes have included variation in aphid genetics and in the obligatory aphid symbiont, *Buchnera aphidicola* (Swanevelder *et al.*, 2010). This bacterial endosymbiont is found inside cells produced by the aphid, called mycetocytes or bacteriocytes, and is maintained transovarially and maternally between generations (Munson *et al.*, 1991; Baumann *et al.*, 1995; Dixon *et al.*, 1998; Douglas, 1998). The symbiotic relationship is obligatory to both aphid and bacterium as removal of the bacteria often result in sterile aphid offspring (Munson *et al.*, 1991; Douglas, 1998). Aphids are able to exploit nutritionally poor food sources such as phloem because of their symbionts. The gram negative bacterium produces and recycles specific amino acids that do not occur in sufficient quantities in the aphid's diet (Mittler, 1971; Douglas and Prosser, 1992; Douglas, 1998). *B. aphidicola* produces these amino acids in the quantities needed by duplicating genes or whole pathways to single or multi copy plasmids (Lai *et al.*, 1994; Van Ham *et al.*, 1997; Baumann *et al.*, 1999). These plasmids are known to vary in copy number between aphid biotypes (Swanevelder *et al.*, 2010). Another variation associated with the plasmids, a CCC insert in the leucine plasmid, was observed in four of ten biotypes. However, despite these differences in plasmid copy number and sequence, the genetic diversity of *B. aphidicola* alone does not explain the ecological diversity observed between aphid biotypes (Swanevelder *et al.*, 2010).

2.6.3 Symptoms of RWA feeding

Leaf rolling, chlorotic streaking and trapping of the grain heads in barley and wheat are symptoms caused by RWA infestation (Ma *et al.*, 1998). The physical damage caused by RWA feeding is minimal in comparison to chewing insects, however RWA feeding does up-regulate photosynthetic related genes (Botha *et al.*, 2006) and introduces effectors to its plant host that elevate pathogenesis response gene products (Lapitan *et al.*, 2007).

Brigham (1992) extracted soluble compounds from ground-up RWA and introduced these extracts to susceptible wheat and barley which subsequently exhibited leaf rolling symptoms associated with susceptibility. Lapitan *et al.* (2007) performed a similar experiment but divided the RWA extract up into whole extract, metobalites and proteins. It was found that metabolites only, had no effect on susceptible plants, however the protein extract mimicked symptoms of infestation. From this data they concluded that it is a protein elicitor in the aphid that is recognized by a plant receptor according to the gene-for-gene model proposed by Flor (1971).

2.7 Host plant resistance

2.7.1 Host defence responses

Plant resistance can involve compatible or incompatible interactions between host and pest. A virulent pest utilizing a susceptible host leads to a compatible interaction while incompatible interactions involve a resistance response from the host plant with the pest, therefore, rendering the pest avirulent (Kaloshian, 2004). One of the first resistance responses initiated is the Hypersensitive Response (HR).

The purpose of this is to limit spread of the pathogen by toughening the cell wall and forming tissue lesions and restrictions around the infection site (Levine *et al.*, 1996). The oxidative burst is part of this early stress response and is characterized by the accumulation of reactive oxygen species (ROS) which induce the HR and related cell death through an influx of calcium ions (Levine *et al.*, 1996).

Plant defence signaling pathways activated by aphid feeding are dependent on both salicylate and jasmonate signaling molecules (Botha *et al.*, 2006; 2010; 2014; Smith and Boyko, 2007). Salicylic acid (SA) and its methyl conjugate are utilized in SA-dependent cascades that trigger the expression of defence response genes in the host plant. Salicylic acid is vital to Systemic Acquired Resistance (SAR) which is a longer term defence response (Botha *et al.*, 2010). Jasmonic Acid (JA) is a plant defence response signaling molecule associated with wounding (Smith *et al.*, 2010). The exact mechanism of activation of the JA and SA pathways in response to feeding remains to be fully elucidated.

Localized and plant-wide increases of Pathogen Resistance (PR) mRNA or PR proteins like peroxidases, chitinases and glucanases are also elicited in response to aphid feeding. In wheat PR proteins are induced to higher levels in resistant than in susceptible plants (van der Westhuizen *et al.*, 1998a; 1998b).

2.7.2 Basal defence response

The basal defence strategy involves pre-formed defences that are always in place, such as waxy cuticular outer layers (Bahlmann *et al.*, 2003) and pre-formed anti-microbial compounds that provide protection in a passive manner (Botha, 2013). Wound responses are activated by chewing insects and physical damage to the plant. These responses trigger the production of protease inhibitors and alkaloids

that deter pathogens. Sap feeding insects, in turn, tend to inflict developmental responses in their hosts (Dangl and Jones, 2001). This process is similar to the basal stress response seen in association with phytopathogens. This basal response occurs in both susceptible and resistant plants while the gene-for-gene interaction only occurs in resistant plants.

2.7.3 Recognition of effectors

Plants recognise microbe- or pathogen-associated molecular patterns (also known by the terms damage-associated molecular patterns or herbivore-associated molecular patterns) and a basal defence response is elicited (Lotze *et al.*, 2007; Boller and Felix, 2009; Tor *et al.*, 2009). In the case of aphid-wheat interactions, the host plant will recognize aphid-associated molecular patterns which result in the activation of defence responses using structural and chemical components to deter the aphid (Botha *et al.*, 2014). In addition to these molecular patterns it has been shown that a protein elicitor from the aphid is detected by the host plant (Lapitan *et al.*, 2007). Botha *et al.* (2014) have shown that virulent aphid biotypes' salivary effectors are not recognized by the host plant, producing a susceptible response to aphid feeding, despite the plant being categorized as RWA resistant.

The gene-for-gene model describes the interaction and recognition between proteins from plant and pest. A pathogenic or insect elicitor interacts with a resistance gene product from the plant and elicits a specific resistance response (Flor, 1971). This model attempts to explain the interaction between plant and pathogen on a molecular level. In short, there is recognition and subsequent interaction between insect-derived elicitors and resistance gene (R gene) products from the host plant. The resulting compatible interaction will lead to resistance against the insect. In

the absence of a *R* gene in the host plant there will be no recognition and disease would occur. There is seldom a direct interaction between eliciting proteins from the pathogen, known as avirulence (*Avr*) gene products, and *R* gene products from the host plant. The guard hypothesis helps explain why some plant resistance proteins need intermediary proteins in order to activate defence (Jones and Takemoto, 2004). This hypothesis proposes that proteins (such as nucleotide-binding site-leucine rich repeat (NBS-LRRs)) guard targets in the host plant against effectors from the pathogen (Lacock *et al.*, 2003; Botha *et al.*, 2006). This guarding interaction is required in order to activate defence upon recognition of the pathogenic *Avr* protein (Dangl and Jones, 2001). Thus, whether protein recognition is non-specific via aphid-associated molecular patterns or specific via aphid effector proteins, a defence response will be elicited by the host plant (Agrawal and Fishbein, 2006; Botha *et al.*, 2014).

2.7.4 Resistance genes (*R* genes)

Elicited or induced defence responses are specific to the pest or pathogen and function according to the gene-for-gene model (Flor, 1971). Insect effectors are generally proteins found in the insect's saliva (Walling, 2008) while plant resistance genes can be cytoplasmic or transmembrane (Dangl and Jones, 2001) and are divided up into four categories: the serine-threonine kinases (Martin *et al.*, 1993; Ritter and Dangl, 1996), putative transmembrane receptors with extracellular leucine rich repeat (LRR) domains (Jones *et al.*, 1994; Dixon *et al.*, 1998), receptor-like kinases and NBS-LRR resistance genes (Dangl and Jones, 2001). By coding for multiple protein complexes, these resistance genes can recognize more than one *Avr* gene (Dangl and Jones, 2001).

Mi-1.2, a gene from wild tomato confers resistance to the potato aphid and three species of root knot nematode. It was the first insect resistance gene to be cloned (Rossi *et al.*, 1998). This gene is an NBS-LRR Class II disease and nematode resistance gene. None of the resistance genes associated with the wheat/RWA interaction have been cloned, though in wheat infested with RWA the up regulation of a *Pto*-like serine/threonine kinase as well as a *Pto*-interactor-like kinase gene have been reported (Boyko *et al.*, 2006) and leucine zipper NBS-LRRs have been identified in resistant wheat cultivars (Lacock *et al.*, 2003). Other close associations have been made such as the linkage between a LZ-NBS-LRR gene and a RWA resistance gene identified by Swanepoel *et al.* (2003). By far the most abundant class of resistance genes is the NBS-LRR (Ellis *et al.*, 2000; Sandhu and Gill, 2002) which has been shown to associate with the plasma membrane. The carboxy terminal (LRR) functions as a site of interaction and binding between proteins, peptides and ligands or proteins and carbohydrates. It also shows variability in number which contrasts with the nucleotide-binding site which is more conserved. Serine/threonine kinases are another major *R* gene group involved in phosphorylation cascades (Dangl and Jones, 2001).

Often the association of genes or pathways is not a simple matter of increased expression and correlation to resistance. In a microarray experiment with *Arabidopsis thaliana* and *Myzus persicae*, oxidative stress genes were found to be both positively and negatively affected by aphid infestation. Glutathione-S-transferase levels increased as did cytosolic superoxide dismutase. On the other hand, another form of superoxide dismutase (FeSOD) and a peroxidase gene decreased while other glutathione-S-transferases associated genes were not altered. This same dual induction and repression also occurred in a group of Ca^{2+}/calmodulin-

related signaling genes (Moran *et al.*, 2002). A microarray experiment on the interaction between RWA and wheat was conducted by Botha *et al.* (2010) who also found glutathione-S-transferase significantly up regulated along with esterase, actin, ATPase and a putative nucleotide binding protein.

2.7.5 Resistance responses to the aphid

In a study by Manickavelu *et al.* (2010) the authors screened EST libraries from resistant and susceptible wheat cultivars looking for wheat-pathogen interaction genes. The authors compiled a list of wheat-pathogen interaction genes and classified them as resistant, susceptible or common genes. The most abundant genes in the resistance group was proton ATPase and calcium binding proteins. In addition, proteinase inhibitors and sequences related to detoxification were also identified. Genes involved in the alteration of plant cell wall composition (cysteine proteinase, phenylalanine ammonia-lyase, plasma membrane ATPase and chalcone synthase) and signal transduction genes (CDP and MAP kinases) were also found in the resistance group. In the susceptible category the authors classified many genes with functions related to oxidative burst (glutathione-S-transferase, peroxidase and oxalate oxidase). Genes classified to both the susceptible and resistant groups included transcription factors such as WRKY and other pathogen response-related sequences such as receptor-like kinase, GTP-binding protein, RING finger protein, cytochrome P450, LRR, PR protein, G-box binding protein, STAR-related lipid transfer protein and starch synthase.

In addition to the genes that have been found to play a role in resistance against aphids, a further complication arises when one considers the genetic background that a resistance gene is found in. Gao *et al.* (2008) studied the effects of two closely

related aphid species on various cultivars of wheat: Blue green aphid (BGA) and Pea aphid (PA). Co-segregation of BGA and PA resistance as observed by the authors might imply that a single gene confers resistance to both aphid species. The resistance gene to BGA is a single dominant gene and maps to a cluster of *R*-genes (Klingler *et al.*, 1998) and PA resistance was found to segregate with this cluster. Unfortunately no distinction could be made as to whether it is a single gene that mediates resistance to both aphids or whether closely linked genes mediate aphid-specific resistance. What is interesting, however, is the effects of varying genetic backgrounds on aphid resistance. In a genetic background containing an *R* gene for BGA resistance, the resistance to PA differs and in the cultivar 'Jester' the downstream defence response to both PA and BGA differ. In this cultivar the resistance to PA is more moderate than BGA (Gao *et al.*, 2008). Similar trends were observed between different wheat cultivars after infestation with RWA (Van Der Westhuizen *et al.*, 1998a; 1998b).

There is a multitude of RWA-induced plant sequences identifiable in aphid-infested wheat (Botha *et al.*, 2006; Lacock *et al.*, 2003) and it is clear that there is more than one mechanism involved in the interactions between plants and RWA. This variety may lead to specific plant differences in early defence signaling and defence response pathways. Many defence responses that plants deploy result in direct damage to themselves and, in response, plants produce increased levels of housekeeping gene products for processes such as photosynthesis, photorespiration, protein synthesis, antioxidant production, detoxification and maintenance of cell homeostasis. Aphid infested celery, sorghum and wheat plants up-regulate sequences associated with strengthening the cell wall and cell membrane in addition to redox homeostasis and detoxification (Botha *et al.*, 2006; Smith and Boyko,

2007).

Numerous plant sequences and plant expression profiles (Botha *et al.*, 2006; Boyko *et al.*, 2006) associated with RWA infestation have been identified. These components include sequences involved in signaling, protein synthesis, modification and degradation, maintenance of cell structure and homeostasis, and secondary metabolism or sequences that encode proteins functioning in direct plant defence and signaling, oxidative burst, cell wall degradation, cell maintenance, photosynthesis, and energy production.

2.8 Russian wheat aphid resistance genes

Resistance genes against RWA are designated *Dn* genes. Table 2.5 lists the 14 *Dn* resistance genes, their accessions and countries of origin as well as their chromosomal locations.

Resistance responses to RWA such as antibiosis, antixenosis and tolerance or a combination thereof hint at the presence of varying resistance genes in different wheat cultivars (Painter, 1951). Antibiosis is quantified by a significant reduction in aphid fecundity or number of nymphs per aphid while antixenosis is known as non-preference and is used to describe instances where the plant is an unsuitable host for the insect due to morphological and/or chemical factors that influence the insect and cause it to select a different host. Tolerance is the ability of a plant to continue growth under aphid infestation (Painter, 1951). Ratios of resistant to susceptible plants can allude to the presence of a dominant or recessive resistance gene, or perhaps more than one resistance gene (Marais and Du Toit, 1993).

Table 2.5: The fourteen *Dn* resistance genes identified to date.

Resistance gene	Origin	Chromosome	Accession	Reference
Dn4	Russia	1D	PI372129	Nkongolo *et al.* (1991) Saidi and Quick (1996)
Dn1	Iran/Hungary	7D	PI137739	Du Toit (1987; 1988)
Dn2	Russia	7D	PI262660	Du Toit (1987; 1988)
Dn5	Bulgaria	7D	PI294994	Marais and Du Toit (1993)
Dn6	Iran	7D	PI243781/ CI6501	Saidi and Quick (1996)
Dnx	Afghanistan	7D	PI220127	Harvey and Martin (1990)
Dny	USA cultivar 'Stanton'		PI220350	Smith *et al.* (2004)
Dn7	*Secale cereale*	1RS.1BL translocation	94M370	Marais *et al.* (1994)
dn3	*Ae. tauschii*		SQ24	Nkongolo *et al.* (1991)
Dn8, Dn9		7D, 1D	PI294994 derivatives	Liu *et al.* (2001)
Dn2414	USDA-ARS breeding line	1RS.1BL translocation	ST-ARS 02RWA2414-11	Peng *et al.* (2007)
Dn626580	Iran	7D	PI626580	Valdez *et al.* (2012)
DnCI2401	Tajikistan	7D	CI2401	Fazel-Najafabadi *et al.* (2014) (unpublished)

In South Africa breeding wheat cultivars resistant to RWA was initiated by Du Toit in the late 1980's after the discovery of natural resistance to the aphid in host plants (Du Toit, 1987; 1988; 1990). The first resistant genes were derived from wheat lines SA 1684 (PI 1377739) and SA 2199 (PI 262660) which were designated as carrying the *Dn1* and *Dn2* resistance genes respectively (Du Toit, 1987). In 1992 Tugela-*Dn1* was released (Van Niekerk, 2001) and by 2006 there were 27 RWA resistant cultivars available to South African farmers (Tolmay *et al.*, 2007).

The RWA resistance gene *Dn5* was the third resistance gene identified, world-wide, in 1987 from SA 463 (PI294994), a Bulgarian wheat accession (Du Toit, 1988). Du Toit (1990) stated that the segregation ratios of *Dn5* was unclear and speculated that the resistant phenotype might be under the control of more than one gene.

Dn genes originate from species other than wheat as well: a recessive gene *dn3*, present in the *Ae. tauschii* line SQ24 (Nkongolo *et al.*, 1991) and *Dn7*, a gene derived from a rye accession transferred to the short arm of the 1RS 1BL translocation in the wheat cultivar Gamtoos (Marais *et al.*, 1994; 1998) are two examples.

Dn4 is the resistance gene present in all but one resistant cultivar grown in Colorado (Haley *et al.*, 2004). Resistant wheat cultivars were widely adopted by Colorado winter wheat producers in the USA and made up roughly 25% of the winter wheat acreage in the 2003-2004 growing season.

The resistance genes, *Dn8* and *Dn9*, were identified in near-isogenic wheat lines derived from the *Dn5* progenitor (Liu *et al.*, 2001). Two other known dominant RWA resistance genes, *Dn4* and *Dn6*, originated from Russian and Iranian bread wheat accessions respectively (Nkongolo *et al.*, 1991; Saidi and Quick, 1996). Two

of the latest and least studied of the *Dn* genes originated from Afghanistan (*Dnx*, Harvey and Martin, 1990) and a USA cultivar called Stanton (*Dny*; Smith *et al.*, 2004). Resistance conferred by the cereal introduction CI2401 consists of two resistance genes, one allelic to *Dn4* and the other a novel resistance gene (Dong and Quick, 1995; Collins *et al.*, 2005; Voothuluru *et al.*, 2006). A study by Valdez *et al.* (2012) described another novel resistance gene discovered in an Iranian landrace and located on chromosome 7D designated *Dn626580*, while the latest resistance gene, *Dn2414* was described by Fazel-Najafabadi 2014 (unpublished).

2.8.1 Molecular markers against *Dn* genes

There have been many attempts at finding markers closely linked to the *Dn* resistance genes. Some of the most successful markers used to date include *Xgwm111*, *Xgwm44*, *Xgwm437* and numerous RAPD markers (Myburg *et al.*, 1998; Venter *et al.*, 1998; Venter and Botha, 2000; Liu *et al.*, 2001; 2002; 2005; Miller *et al.*, 2001; Heyns, 2005). *Xgwm111* has been linked to *Dn1, Dn2, Dn5, Dnx* and *Dn6* at distances between 1.52 and 3.0 cM (Liu *et al.*, 2001; 2002) while other authors have linked *Dn5* and *Xgwm111* at a much greater distance of 25.4 and 26.5 cM (Heyns, 2005). *Xgwm44* and *Xgwm437* have both been linked to *Dn2* and *Dn5* at distances ranging from 2.8 to 29.03 cM (Miller *et al.*, 2001; Heyns, 2005) and *Xgwm44* has been linked to *Dn6* at 14.63 cM (Liu *et al.*, 2002). *Dn8* is linked to *Xgwm635* on the distal portion near the telomere of wheat chromosome 7DS and *Dn9* is tightly linked to *Xgwm642* on chromosome 1DL (Liu *et al.*, 2001).

Other, non-microsatellite markers have also been closely associated with *Dn* genes. *Dn4* is loosely linked (11.6 cM) to the RFLP marker ABC156 on wheat chromosome 1DS (Ma *et al.*, 1998). Random Amplified Polymorphic DNA markers

have shown linkage to *Dn5* at 2.2 cM (Venter and Botha, 2000), *Dn2* at 3.3 and 4.4 cM (Myburg *et al.*, 1998) and to *Dn1* at 43.7 cM (Venter *et al.*, 1998). The markers that have been reported linked to the *Dn* genes are depicted in Table 2.6.

2.8.2 *Dn1*, *Dn2* and *Dn5* resistance gene cluster on chromosome 7D

Mapping studies rely on phenotyping and segregation ratios to establish the inheritance pattern of resistance genes. To date, several authors reported segregation ratios for the relationship between resistance genes *Dn1, Dn2* and *Dn5*. There are three schools of thought regarding the relationship between these three resistance genes describing them as either non-allelic, allelic or linked. Non-allelic, independent inheritance has been shown by Du Toit (1990) who found that two of the wheat genome progenitors (PI137739 and PI262660 from Iran and Bulgaria respectively) showed single, dominant gene inheritance. These progenitors were subsequently identified as the sources of the *Dn1* and *Dn2* resistance genes (Table 2.5).

Dn5 has proven particularly controversial amongst the *Dn* resistance genes clustered on chromosome 7D (Table 2.7). There are many opinions as to the exact mode of inheritance and as many different segregation patterns have been observed. In the *Dn5* progenitor (PI294994 from Bulgaria), Marais and Du Toit (1993) observed single dominant inheritance and stated that it was not allelic to *Dn1* or *Dn2* as *Dn1* and *Dn5* segregated independently from *Dn2*. Instead, *Dn1* and *Dn5* are probably linked but separate genes. However, others reported that *Dn1*, *Dn2* and *Dn5* are allelic (Saidi and Quick, 1996) or tightly linked or both (Liu *et al.*, 2001).

Table 2.6: Prominently linked markers thought to be clustered near the centromere on chromosome 7D. Numbers depict the number of publications listing the markers as occurring on either the short or long arm of chromosome 7D.

Closely linked markers	**7DS**	**7DL**	**Linked resistance genes**	**Reference**
Xgwm111	5	2	*CI2401, Dn1, Dn2, Dn5, Dn6, Dnx*	Adhikari *et al.* (2004); Liu *et al.* (2002); Liu *et al.* (2005); Liu *et al.* (2001); Groenewald (2001); Heyns (2005); Röder *et al.* (1998)
Xgwm437	1	4	*CI2401, Dn2, Dn5, Dn626580*	Valdez *et al.* (2012); Röder *et al.* (1998); Somers *et al.* (2004); Groenewald (2001); Heyns (2005)
Xgwm44	4	1	*Dn2, Dn5, Dn6,*	Röder *et al.* (1998); Liu *et al.* (2002); Liu *et al.* (2001); Somers *et al.* (2004); Heyns (2005)
Xbarc214	1	1	*CI2401, Dn626580*	Röder *et al.* (1998); Valdez *et al.* (2012); Somers *et al.* (2004); Song *et al.* (2005)
Xgwm635	1		*Dn1, Dn8*	Swanepoel *et al.* (2003); Röder *et al.* (1998); Liu *et al.* (2001)
xksua1	1	1	*Dn2*	Ma *et al.* (1998), Liu *et al.* (2002); Liu *et al.* (2001); Miller *et al.* (2001); Boyko *et al.* (1999)
Xgwm37	0	1	*Dn5*	Heyns (2005)
Xgwm473	1	0	*Dn626580*	Valdez *et al.* (2012); Somers *et al.* (2004)

Table 2.7: Three *Dn* genes and the number of publications listing the genes as occurring on either the short or long arm of chromosome 7D.

Resistance gene	**7DS**	**7DL**	**Reference**
Dn1	2	1	Liu *et al.* (2001; 2002); Marais and Du Toit (1993)
Dn2	2	1	Liu *et al.* (2001; 2002); Ma *et al.* (1998)
Dn5	2	6	Du Toit (1988; 1990); Liu *et al.* (2001; 2002); Marais and Du Toit (1993); Heyns *et al.* (2006); Zhang *et al.* (1998); Marais *et al.* (1998)

Ma *et al.* (1998) observed segregation of *Dn5* as a single gene and Marais and Du Toit (1993) agreed with this observation by stating that there was a dominant resistance gene on chromosome 7D not allelic to *Dn1* or *Dn2* but possibly linked to *Dn1* (Marais and Du Toit, 1993; Ma *et al.*, 1998). Many other authors described a different scenario, where two or three separate genes were observed in the *Dn5* progenitor. Zhang *et al.* (1998) observed three classes of genotypes for resistance in the *Dn5* progenitor. This was observed by Saidi and Quick (1996) as well, who stated that these three genes are made up of two dominant genes and one allelic to *Dn1* and *Dn2*. Other authors, such as Elsidaig and Zwer (1993) described two genes in the progenitor, one dominant and one recessive. This study was supported by Dong and Quick (1995). A study by Zhang *et al.* (1998) gave chi-square data that showed ratios which could correlate with three models of either two dominant genes, one dominant gene or one dominant and one recessive gene. A reason for the discrepancy surrounding the *Dn5* resistance gene might be that there is heterogeneity present in the original progenitor line used (Liu *et al.*, 2005). In addition, inconsistencies in expression are observed when crosses are made between *Dn5* parental lines and various other genetic backgrounds (Marais and Du Toit, 1993). These show the discrepancies on the chromosomal location of the *Dn1*, *Dn2* and *Dn5* genes as reported in literature.

It is clear that there is a fair amount of uncertainty regarding the location of *Dn1, Dn2* and *Dn5*. Data on the microsatellite marker, *Xgwm111*, which has been closely linked to these genes is also uncertain. Using chromosome banding, Werner *et al.* (1992) showed that the physically or cytologically longer arm of 7D is actually the genetically shorter arm as it is homoeologous to 7AS and 7BS suggesting that reports on 7DS may refer to 7DL and *vice versa*. In some instances,

authors analyzing ditelosomic lines erronously claimed to use Dt 7DL lines, when instead they were using Dt 7DS. This was established using chromosome banding (Liu *et al.*, 2005; Werner *et al.*, 1992).

List of References

Adhikari, T.B., Cavaletto, J.R., Dubcovsky, J., Gieco, J.O., Schlatter, A.R. and Goodwin, S.B. (2004). Molecular mapping of the *Stb4* gene for resistance to *Septoria tritici* blotch in wheat. *Phytopathology*, vol. 94, pp. 1198–1206.

Agrawal, A.A. and Fishbein, M. (2006). Plant defence syndromes. *Ecology (Suppl)*, vol. 87, pp. 132–149.

Akbari, M., Wenzl, P., Caig, V., Carling, J., Xia, L., Yang, S., Uszynski, G., Mohler, V., Lehmensiek, A., Kuchel, H., Hayden, M.J., Howes, N., Sharp, P., Vaughan, P., Rathmell, B., Huttner, E. and Kilian, A. (2006). Diversity arrays technology (DArT) for high- throughput profiling of the hexaploid wheat genome. *Theoretical and Applied Genetics*, vol. 113, pp. 1409–1420.

Akhunov, E.D., Akhunova, A.R., Linkiewicz, A.M., Dubcovsky, J., Hummel, D., Lazo, G., Chao, S., Anderson, O.D., David, J., Qi, L., Echalier, B., Gill, B.S., Miftahudin, Gustafson, J.P., La Rota, M., Sorrells, M.E., Zhang, D., Nguyen, H.T., Kalavacharla, V., Hossain, K., Kianian, S., Peng, J., Lapitan, N.L.V., Wennerlind, E.J., Nduati, V., Anderson, J.A., Sidhu, D., Gill, K.S., McGuire, P.E., Qualset, C.O. and Dvorak, J. (2003). Synteny perturbations between wheat homoeologous chromosomes caused by locus duplications and deletions correlate

with recombination rates. *Proceedings of the National Academy of Sciences*, vol. 100, pp. 543–549.

Akhunov, E.D., Nicolet, C. and Dvorak, J. (2009). Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina Golden Gate assay. *Theoretical and Applied Genetics*, vol. 119, pp. 507–517.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, vol. 25, pp. 3389–3402.

Amrani, N., Sarrafi, A. and Alibert, G. (1993). Genetic variability for haploid production in crosses between tetraploid and hexaploid wheats with maize. *Plant Breeding*, vol. 110, pp. 123–128.

Anderson, G.R., Papa, D., Peng, J., Tahir, M. and Lapitan, N.L.V. (2003). Genetic mapping of *Dn7* , a rye gene conferring resistance to the Russian wheat aphid in wheat. *Theoretical and Applied Genetics*, vol. 107, pp. 1297–1303.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S. and Eppig, J.T. (2000). Gene ontology: tool for the unification of biology. *Nature Genetics*, vol. 25, pp. 25–29.

Bahlmann, L., Govender, P. and Botha, A.-M. (2003). The leaf epicuticular wax ultrastructure and trichome presence on Russian wheat aphid (*Diuraphis noxia*) resistant and susceptible leaves. *African Entomology*, vol. 11, pp. 59–64.

Bansal, A., Boyd, P.R. and Mcginnis, R. (2003). Tools for statistical analysis of genetic data. In: Barnes, M.R. and Gray, I.C. (eds.), *Bioinformatics for Geneticists*, chap. 11, pp. 218–245. John Wiley & Sons Ltd., Essex.

Bashir, A., Auger, J.A. and Rayburn, A.L. (1993). Flow cytometric DNA analysis of wheat-rye addition lines. *Cytometry*, vol. 14, pp. 843–847.

Basky, Z. (2003). Biotypic and pest status differences between Hungarian and South African populations of Russian wheat aphid, *Diuraphis noxia* (Kurdjumov) (Homoptera: Aphididae). *Pest Management Science*, vol. 59, pp. 1152–1158.

Baumann, L., Baumann, P., Moran, N.A., Sandstrom, J. and Thao, M.L. (1999). Genetic characterization of plasmids containing genes encoding enzymes of leucine biosynthesis in endosymbionts *(Buchnera)* of aphids. *Journal of Molecular Evolution*, vol. 48, pp. 77–85.

Baumann, P., Lai, C.-Y., Baumann, L., Rouhbakhsh, D. and Clark, A.M. (1995). Genetics, physiology and evolutionary relationships of the genus *Buchnera*: intracellular symbionts of aphids. *Annual Reviews in Microbiology*, vol. 94, pp. 55–94.

Beales, J., Laurie, D.A. and Devos, K.M. (2005). Allelic variation at the linked AP1 and PhyC loci in hexaploid wheat is associated but not perfectly correlated with vernalization response. *Theoretical and Applied Genetics*, vol. 110, pp. 1099–1107.

Berkman, P.J., Skarshewski, A., Lorenc, M.T., Lai, K., Duran, C., Ling, E.Y.S., Stiller, J., Smits, L., Imelfort, M., Manoli, S., McKenzie, M., Kubaláková, M.,

Šimková, H., Batley, J., Fleury, D., Doležel, J. and Edwards, D. (2011). Sequencing and assembly of low copy and genic regions of isolated *Triticum aestivum* chromosome arm 7DS. *Plant Biotechnology Journal*, vol. 9, pp. 768–75.

Berkman, P.J., Skarshewski, A., Manoli, S., Lorenc, M.T., Stiller, J., Smits, L., Lai, K., Campbell, E., Kubaláková, M., Šimková, H., Batley, J., Doležel, J., Hernandez, P. and Edwards, D. (2012). Sequencing wheat chromosome arm 7BS delimits the 7BS/4AL translocation and reveals homoeologous gene conservation. *Theoretical and Applied Genetics*, vol. 124, pp. 423–32.

Blake, N.K., Sherman, J.D., Dvorak, J. and Talbert, L.E. (2004). Genome-specific primer sets for starch biosynthesis genes in wheat. *Theoretical and Applied Genetics*, vol. 109, pp. 1295–1302.

Boller, T. and Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual Review in Plant Biology*, vol. 60, pp. 379–406.

Botha, A.-M. (2013). A coevolutionary conundrum: the arms race between *Diuraphis noxia* (Kurdjumov) a specialist pest and its host *Triticum aestivum* (L.). *Arthropod Plant Interactions*, vol. 7, pp. 359–372.

Botha, A.-M., Burger, N.F.V. and Van Eck, L. (2014). Hypervirulent *Diuraphis noxia* (Hemiptera: Aphididae) biotype SAM avoids triggering defenses in its host (*Triticum aestivum*) (Poales: Poaceae) during feeding. *Environmental Entomology*, vol. 43, pp. 672–681.

Botha, A.-M., Lacock, L., Van Niekerk, C., Matsioloko, M.T., Du Preez, F.B., Lootsa, S., Venter, E., Kunert, K.J. and Cullis, C.A. (2006). Is photosynthetic

transcriptional regulation in *Triticum aestivum* L. cv. TugelaDN a contributing factor for tolerance to *Diuraphis noxia* (Homoptera: Aphididae)? *Plant Cell Reports*, vol. 25, pp. 41–54.

Botha, A.-M., Swanevelder, Z.H. and Lapitan, N.L.V. (2010). Transcript profiling of wheat genes expressed during feeding by two different biotypes of *Diuraphis noxia*. *Environmental Entomology*, vol. 39, pp. 1206–1231.

Botstein, D., White, R.L., Skolnick, M. and Davis, R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, vol. 32, pp. 314–331.

Boyko, E.V., Gill, K.S., Mickelson-Young, L., Nasuda, S., Raupp, W.J., Ziegle, J.N., Singh, S., Hassawi, D.S., Fritz, A.K., Namuth, D., Lapitan, N.L.V. and Gill, B.S. (1999). A high-density genetic linkage map of *Aegilops tauschii*, the D-genome progenitor of bread wheat. *Theoretical and Applied Genetics*, vol. 99, pp. 16–26.

Boyko, E.V., Smith, C.M., Thara, V.K., Bruno, J.O.N.M., Deng, Y., Starkey, S.R. and Klaahsen, D.L. (2006). Molecular basis of plant gene expression during aphid invasion: wheat *Pto*- and *Pti*-like sequences are involved in interactions between wheat and Russian wheat aphid (Homoptera: Aphididae). *Journal of Economic Entomology*, vol. 99, pp. 1430–1445.

Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G.L.A., Amore, R.D., Allen, A.M., Mckenzie, N., Kramer, M., Kerhornou, A., Bolser, D., Kay, S., Waite, D., Trick, M., Bancroft, I., Gu, Y., Huo, N., Luo, M.-C., Sehgal, S., Gill, B., Kianian, S., Anderson, O., Kersey, P., Dvorak, J., Mccombie, W.R., Hall, A., Mayer, K.F.X., Edwards, K.J., Bevan, M.W. and Hall, N. (2012). Analysis of

the bread wheat genome using whole-genome shotgun sequencing. *Nature*, vol. 491, pp. 705–710.

Brigham, D.L. (1992). *Chemical ecology of the Russian wheat aphid: host selection and phytotoxic effects.* Master's thesis, Colorado State University.

Brookes, A.J. (1999). The essence of SNPs. *Gene*, vol. 234, pp. 177–186.

Brugmans, B., Van Der Hulst, R.G.M., Visser, R.G.F., Lindhout, P. and Van Eck, H.J. (2003). A new and versatile method for the successful conversion of AFLP markers into simple single locus markers. *Nucleic Acids Research*, vol. 31, p. e55.

Brumfield, R.T., Beerli, P., Nickerson, D.A. and Edwards, S.V. (2003). The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecological Evolution*, vol. 18, pp. 249–256.

Burd, J.D., Porter, D.R., Puterka, G.J., Haley, S.D. and Peairs, B. (2006). Biotypic variation among North American Russian wheat aphid (Homoptera: Aphididae) populations. *Journal of Economic Entomology*, vol. 99, pp. 1862–1866.

Butler, J., MacCallum, I., Kleber, M., Shlyakhter, I.A., Belmonte, M.K., Lander, E.S., Nusbaum, C. and Jaffe, D.B. (2008). ALLPATHS: *De novo* assembly of whole-genome shotgun microreads. *Genome Research*, vol. 18, pp. 810–820.

Campagna, D., Romualdi, C., Vitulo, N., Del Favero, M., Lexa, M., Cannata, N. and Valle, G. (2004). RAP: a new computer program for *de novo* identification of repeated sequences in whole genomes. *Bioinformatics*, vol. 21, pp. 582–588.

Cervantes-Flores, J.C., Yencho, G.C., Kriegner, A., Pecota, K.V., Faulk, M.A., Mwanga, R.O.M. and Sosinski, B.R. (2008). Development of a genetic link-

age map and identification of homologous linkage groups in sweetpotato using multiple-dose AFLP markers. *Molecular Breeding*, vol. 21, pp. 511–532.

Chaisson, M.J. and Pevzner, P.A. (2008). Short read fragment assembly of bacterial genomes. *Genome Research*, vol. 18, pp. 324–330.

Chao, S., Zhang, W., Akhunov, E., Sherman, J., Ma, Y., Ming-Cheng, L. and Dubcovsky, J. (2009). Analysis of gene-derived SNP marker polymorphism in US wheat (*Triticum aestivum* L .) cultivars. *Molecular Breeding*, vol. 23, pp. 23–33.

Chikhi, R. and Medvedev, P. (2013). Informed and automated k-mer size selection for genome assembly. *Bioinformatics*, p. btt310.

Choulet, F., Wicker, T., Rustenholz, C., Paux, E., Salse, J., Leroy, P., Schlub, S., Le Paslier, M.-C., Magdelenat, G., Gonthier, C., Couloux, A., Budak, H., Breen, J., Pumphrey, M., Liu, S., Kong, X., Jia, J., Gut, M., Brunel, D., Anderson, J.A., Gill, B.S., Appels, R., Keller, B. and Feuillet, C. (2010). Megabase level sequencing reveals contrasted organization and evolution patterns of the wheat gene and transposable element spaces. *The Plant Cell*, vol. 22, pp. 1686–1701.

Collins, M.B., Haley, S.D., Randolph, T.L., Peairs, F.B. and Rudolph, J.B. (2005). Comparison of *Dn4*- and *Dn7*-carrying spring wheat genotypes artificially infested with Russian wheat aphid (Homoptera: Aphididae) biotype 1. *Journal of Economic Entomology*, vol. 98, pp. 1698–1703.

Cox, D.R., Burmeister, M., Price, E.R., Kim, S. and Myers, R.M. (1990). Radiation hybrid mapping: a somatic cell genetic method for constructing high-resolution maps of mammalian chromosomes. *Science*, vol. 250, pp. 245–250.

Crossa, J., Burgueno, J. and Dreisigacker, S. (2007). Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. *Genetics*, vol. 177, pp. 1889–1913.

Dangl, J.L. and Jones, J.D.G. (2001). Plant pathogens and integrated defence responses to infection. *Nature*, vol. 411, pp. 826–833.

De Jager, L., Burger, N. and Botha, A.-M. (2014). Complete mitochondrial genome of *Diuraphis noxia* (hemiptera: Aphididae) from nine populations, snp variation between populations, and comparison with other aphididae species. *African Entomology*, vol. 22, pp. 847–862.

Demeke, T., Laroche, A. and Gaudet, D.A. (1996). A DNA marker for the *Bt-10* common bunt resistance gene in wheat. *Genome*, vol. 39, pp. 51–55.

Deschamps, S. and Campbell, M. (2010). Utilization of next-generation sequencing platforms in plant genomics and genetic variant discovery. *Molecular Breeding*, vol. 25, pp. 553–570.

Devos, K.M. and Gale, M.D. (1992). The use of random amplified polymorphic DNA markers in wheat. *Theoretical and Applied Genetics*, vol. 84, pp. 567–572.

Dixon, M.S., Hatzixanthis, K., Jones, D.A., Harrison, K. and Jones, J.D.G. (1998). The tomato *Cf-5* disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. *Plant Cell*, vol. 10, pp. 1915–1925.

Dohm, J.C., Lottaz, C., Borodina, T. and Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Research*, vol. 36, p. e105.

Dong, H. and Quick, J.S. (1995). Inheritance and allelism of resistances to the Russian wheat aphid in seven wheat lines. *Euphytica*, vol. 81, pp. 229–303.

Douglas, A.E. (1998). Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria *Buchnera*. *Annual Review in Entomology*, vol. 43, pp. 17–37.

Douglas, A.E. and Prosser, W.A. (1992). Synthesis of the essential amino acid tryptophan in the pea aphid (*Acyrthosiphon pisum*) symbiosis. *Journal of Insect Physiology*, vol. 38, pp. 565–568.

Dreyer, D.L. and Campbell, B.C. (1987). Chemical basis of host-plant resistance to aphids. *Plant Cell Environment*, vol. 10, pp. 353–361.

Du Toit, F. (1987). Resistance in wheat (*Triticum aestivum*) to *Diuraphis noxia* (Homoptera: Aphididae). *Cereal Research Communications*, vol. 15, pp. 175–179.

Du Toit, F. (1988). Another source of Russian wheat aphid (*Diuraphis noxia*) resistance in *Triticum aestivum*. *Cereal Research Communications*, vol. 16, pp. 105–106.

Du Toit, F. (1990). Field resistance in three bread wheat lines to the Russian wheat aphid, *Diuraphis noxia* (Hemiptera: Aphididae). *Crop Protection*, vol. 9, pp. 255–258.

Edwards, D. and Batley, J. (2010). Plant genome sequencing: applications for crop improvement. *Plant Biotechnology Journal*, vol. 8, pp. 2–9.

Ellis, J., Dodds, P. and Pryor, T. (2000). Structure, function and evolution of plant disease resistance genes. *Current Opinion in Plant Biology*, vol. 3, pp. 278–284.

Elsidaig, A.A. and Zwer, P.K. (1993). Genes for resistance to RWA in PI294994 wheat. *Crop Science*, vol. 33, pp. 998–1001.

Emara, M.G. and Kim, H. (2003). Genetic markers and their application in poulrty breeding. *Poultry Science*, vol. 82, pp. 952–957.

Endo, T.R. and Gill, B.S. (1996). The deletion stocks of common wheat. *Journal of Heredity*, vol. 87, pp. 295–307.

Feldman, M. and Levy, A.A. (2005). Allopolyploidy: a shaping force in the evolution of wheat genomes. *Cytogenetic and Genome Research*, vol. 109, pp. 250–258.

Feuillet, C., Travella, S., Stein, N., Albar, L., Nublat, A. and Keller, B. (2003). Map-based isolation of the leaf rust disease resistance gene *Lr10* from the hexaploid wheat (*Triticum aestivum* L.) genome. *Proceedings of the National Academy of Sciences*, vol. 100, pp. 15253–15258.

Flor, H. (1971). Current status of the gene-for-gene concept. *Annual Review of Phytopathology*, vol. 9, pp. 275–296.

Gao, L.-L., Klingler, J.P., Anderson, J.P., Edwards, O.R. and Singh, K.B. (2008). Characterization of pea aphid resistance in *Medicago trunculata*. *Plant Physiology*, vol. 146, pp. 996–1009.

Gilchrist, L.I., Rodriguez, R. and Burnett, P.A. (1984). The extent of freestate streak and *Diuraphis noxia* in Mexico. In: *Barley yellow dwarf, proceedings*

of workshop, pp. 157–163. International Maize and Wheat Improvement Center (CIMMYT), Mexico City.

Gill, B.S., Appels, R., Botha-Oberholster, A.-M., Buell, C.R., Bennetzen, J.L., Chalhoub, B., Chumley, F., Dvorak, J., Iwanaga, M., Keller, B., Li, W., Mccombie, W.R., Ogihara, Y., Quetier, F. and Sasaki, T. (2004). A workshop report on wheat genome sequencing: international genome research on wheat consortium. *Genetics*, vol. 168, pp. 1087–1096.

Gill, B.S., Friebe, B. and Endo, T.R. (1991). Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). *Genome*, vol. 34, pp. 830–839.

Glenn, T.C. (2011). Field guide to next-generation DNA sequencers. *Molecular Ecology Resources*, vol. 11, pp. 759–769.

Green, E.D. and Green, P. (1991). Sequence-tagged site (STS) content mapping of human chromosomes: theoretical considerations and early experiences. *Genome Research*, vol. 1, pp. 77–90.

Grodzicker, T., Williams, J., Sharp, P. and Sambrook, J. (1974). Physical mapping of temperature–sensitive mutations of adenoviruses. *Quantitative Biology*, vol. 39, pp. 439–446.

Groenewald, I. (2001). *Tagging and mapping of prominent structural genes on chromosome arm 7DL of common wheat.* Ph.D. thesis, University of Stellenbosch.

Gupta, P.K., Mir, R.R., Mohan, A. and Kumar, J. (2008). Wheat genomics: present status and future prospects. *International Journal of Plant Genomics*, vol. 2008, pp. 1–36.

Gut, I.G. (2001). Automation in genotyping of single nucleotide polymorphisms. *Human Mutation*, vol. 17, pp. 475–492.

Haldane, J.B.S. (1919). The combination of linkage values and the calculation of distances between the loci of linked factors. *Journal of Genetics*, vol. 8, pp. 299–309.

Haley, S.D., Peairs, F.B., Walker, C.B., Rudolph, J.B. and Randolph, T.L. (2004). Occurrence of a new Russian wheat aphid biotype in Colorado. *Crop Science*, vol. 44, pp. 1589–1592.

Harismendy, O. (2009). Evaluation of next generation sequencing platforms for population targeted sequencing studies. *Genome Biology*, vol. 10, p. R32.

Hartl, L., Weiss, H., Zeller, F.J. and Jahoor, A. (1993). Use of RFLP markers for the identification of alleles of the *Pm3* locus conferring powdery mildew resistance in wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, vol. 86, pp. 959–963.

Harvey, T.L. and Martin, T.J. (1990). Resistance to Russian wheat aphid, *Diuraphis noxia*, in wheat (*Triticum aestivum*). *Cereal Research Communications*, vol. 18, pp. 127–129.

Healy, J., Thomas, E.E., Schwartz, J.T. and Wigler, M. (2003). Annotating large genomes with exact word matches. *Genome Research*, vol. 13, pp. 2306–2315.

Hernandez, D., Francois, P., Farinelli, L., Osteras, M. and Schrenzel, J. (2008). *De novo* bacterial genome sequencing: millions of very short reads assembled on a desktop computer. *Genome Research*, vol. 18, pp. 802–809.

Hernandez, P., Martis, M., Dorado, G., Pfeifer, M., Gálvez, S., Schaaf, S., Jouve, N., Šimková, H., Valárik, M., Doležel, J. and Mayer, K.F.X. (2012). Next-generation sequencing and syntenic integration of flow-sorted arms of wheat chromosome 4A exposes the chromosome structure and gene content. *The Plant Journal*, vol. 69, pp. 377–386.

Hewitt, P.H. (1988). The South African experience with the Russian wheat aphid. In: Peairs, F.B. and Pilcher, S.D. (eds.), *Proceedings of the second Russian wheat aphid workshop*, pp. 1–3. Denver, Colorado.

Heyns, I., Groenewald, E., Marais, F., Du Toit, F. and Tolmay, V. (2006). Chromosomal location of the Russian wheat aphid resistance gene, *Dn5*. *Crop Science*, vol. 46, pp. 630–636.

Heyns, I.C. (2005). *Mapping of chromosome arm 7DL of Triticum aestivum L.* Master's thesis, University of Stellenbosch; MSc.

Ishikawa, G., Nakamura, T., Ashida, T., Saito, M., Nasuda, S., Endo, T.R., Wu, J. and Matsumoto, T. (2009). Localization of anchor loci representing five hundred annotated rice genes to wheat chromosomes using PLUG markers. *Theoretical and Applied Genetics*, vol. 118, pp. 499–514.

Itoh, T., Tanaka, T., Barrero, R.A., Yamasaki, C., Fujii, Y., Hilton, P.B., Antonio, B.A., Aono, H., Apweiler, R., Bruskiewich, R., Bureau, T., Burr, F., Costa de Oliveira, A., Fuks, G., Habara, T., Haberer, G., Han, B., Harada, E., Hiraki,

A.T., Hirochika, H., Hoen, D., Hokari, H., Hosokawa, S., Hsing, Y.-I., Ikawa, H., Ikeo, K., Imanishi, T., Ito, Y., Jaiswal, P., Kanno, M., Kawahara, Y., Kawamura, T., Kawashima, H., Khurana, J.P., Kikuchi, S., Komatsu, S., Koyanagi, K.O., Kubooka, H., Lieberherr, D., Lin, Y.-C., Lonsdale, D., Matsumoto, T., Matsuya, A., McCombie, W.R., Messing, J., Miyao, A., Mulder, N., Nagamura, Y., Nam, J., Namiki, N., Numa, H., Nurimoto, S., O'Donovan, C., Ohyanagi, H., Okido, T., Oota, S., Osato, N., Palmer, L.E., Quetier, F., Raghuvanshi, S., Saichi, N., Sakai, H., Sakai, Y., Sakata, K., Sakurai, T., Sato, F., Sato, Y., Schoof, H., Seki, M., Shibata, M., Shimizu, Y., Shinozaki, K., Shinso, Y., Singh, N.K., Smith-White, B., Takeda, J.-I., Tanino, M., Tatusova, T., Thongjuea, S., Todokoro, F., Tsugane, M., Tyagi, A.K., Vanavichit, A., Wang, A., Wing, R.A., Yamaguchi, K., Yamamoto, M., Yamamoto, N., Yu, Y., Zhang, H., Zhao, Q., Higo, K., Burr, B., Gojobori, T. and Sasaki, T. (2007). Curated genome annotation of *Oryza sativa* ssp. *japonica* and comparative genome analysis with *Arabidopsis thaliana*. *Genome Research*, vol. 17, pp. 175–183.

Jaccoud, D., Peng, K., Feinstein, D. and Kilian, A. (2001). Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Research*, vol. 29, p. e25.

Jia, J., Zhao, S., Kong, X., Li, Y., Zhao, G., He, W., Appels, R., Pfeifer, M., Tao, Y., Zhang, X., Jing, R., Zhang, C., Ma, Y., Gao, L., Gao, C., Spannagl, M., Mayer, K.F.X., Li, D., Pan, S., Zheng, F., Hu, Q., Xia, X., Li, J., Liang, Q., Chen, J., Wicker, T., Gou, C., Kuang, H., He, G., Luo, Y., Keller, B., Xia, Q., Lu, P., Wang, J., Zou, H., Zhang, R., Xu, J., Gao, J., Middleton, C., Quan, Z., Liu, G., Wang, J., Yang, H., Liu, X., He, Z., Mao, L. and Wang, J.

(2013). *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature*, vol. 496, pp. 91–95.

Jiang, H. and Wong, W.H. (2008). SeqMap: mapping massive amount of oligonucleotides to the genome. *Bioinformatics*, vol. 24, pp. 2395–2396.

Jones, D.A. and Takemoto, D. (2004). Plant innate immunity - direct and indirect recognition of general and specific pathogen-associated molecules. *Current Opinion in Immunology*, vol. 16, pp. 48–62.

Jones, D.A., Thomas, C.M., Hammons-Kosack, K.E., Balint-Kurto, P.J. and Jones, J.D.G. (1994). Isolation of the tomato *Cf-9* gene for resistance to *Cladosporum fulvum* by transposon tagging. *Science*, vol. 266, pp. 789–793.

Kaloshian, I. (2004). Gene-for gene disease resistance: bridging insect pest and pathogen defence. *Journal of Chemical Ecology*, vol. 30, pp. 2419–2438.

Kelley, D.R., Schatz, M.C. and Salzberg, S.L. (2010). Quake: quality-aware detection and correction of sequencing errors. *Genome Biology*, vol. 11, p. R116.

Kent, W.J. (2002). BLAT: the BLAST-like alignment tool. *Genome Research*, vol. 12, pp. 656–664.

Kilian, A., Huttner, E., Wenzl, P., Jaccoud, D., Carling, J. and Caig, V. (2005). The fast and the cheap: SNP and DArT-based whole genome profiling for crop improvement. In: Phillips, R.L. and Gale, M. (eds.), *Proceedings of the international congress in the wake of the double helix: from the green revolution to the gene revolution*, pp. 443–461. Tuberosa, Italy.

Kisana, N.S., Nkongolo, K.K., Quick, J.S. and Johnson, D.L. (1993). Production of doubled haploids by anther culture and wheat x maize method in a wheat breeding program. *Plant Breeding*, vol. 110, pp. 96–102.

Klingler, J., Powell, G., Thomson, G.A. and Isaacs, R. (1998). Phloem specific aphid resistance in *Cucumis melo* line AR 5: effects on feeding behaviour and performance of *Aphis gossypii*. *Entomologia Experimentalis et Applicata*, vol. 86, pp. 79–88.

Kofler, R., Schlotterer, C., Luschutzky, E. and Lelley, T. (2008). Survey of microsatellite clustering in eight fully sequenced species sheds light on the origin of compound microsatellites. *BMC Genomics*, vol. 9, pp. 1–14.

Konishi, S., Izawa, T., Lin, S.Y., Ebana, K., Fukuta, Y., Sasaki, T. and Yano, M. (2006). A SNP caused loss of seed shattering during rice domestication. *Science*, vol. 312, pp. 1392–1396.

Kosambi, D.D. (1943). The estimation of map distances from recombination values. *Annals of Eugenics*, vol. 12, pp. 172–175.

Kubaláková, M., Vrána, J., Číhalíková, J., Šimková, H. and Doležel, J. (2002). Flow karyotyping and chromosome sorting in bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, vol. 104, pp. 1362–1372.

La Rota, M. and Sorrells, M.E. (2004). Comparative DNA sequence analysis of mapped wheat ESTs reveals the complexity of genome relationships between rice and wheat. *Functional Integrated Genomics*, vol. 4, pp. 34–46.

Lacock, L., Van Niekerk, C., Loots, S., Du Preez, F. and Botha, A.-M. (2003). Functional and comparative analysis of expressed sequences from *Diuraphis*

noxia infested wheat obtained utilizing the conserved Nucleotide Binding Site. *African Journal of Biotechnology*, vol. 2, pp. 75–81.

Lai, C.-Y., Baumann, L. and Baumann, P. (1994). Amplification of trpEG: adaptation of *Buchnera aphidicola* to an endosymbiotic association with aphids. *Proceedings of the National Academy of Sciences*, vol. 91, pp. 3819–3823.

Lancaster, H.O. and Seneta, E. (1969). *Chi-square distribution*. Wiley Online Library.

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

Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, vol. 10, p. R25.

Lapitan, N.L.V., Brown, S.E., Kennard, W., Stephens, J.L. and Knudson, D.L. (1997). FISH physical mapping with barley BAC clones. *The Plant Journal*, vol. 11, pp. 149–156.

Lapitan, N.L.V., Peng, J. and Sharma, V. (2007). A high-density map and PCR markers for Russian wheat aphid resistance gene on chromosome 1RS/1BL. *Crop Science*, vol. 47, p. 811.

Levine, A., Pennell, R.I., Alvarez, M.E., Palmer, R. and Lamb, C. (1996). Calcium-mediated apoptosis in a plant hypersensitive disease resistance response. *Current Biology*, vol. 6, pp. 427–437.

Li, H. and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, vol. 25, pp. 1754–1760.

Li, H., Ruan, J. and Durbin, R. (2008*a*). Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Research*, vol. 18, pp. 1851–1858.

Li, R., Li, Y., Kristiansen, K. and Wang, J. (2008*b*). SOAP: short oligonucleotide alignment program. *Bioinformatics*, vol. 24, pp. 713–714.

Lin, H., Zhang, Z., Zhang, M.Q., Ma, B. and Li, M. (2008). ZOOM! Zillions of oligos mapped. *Bioinformatics*, vol. 24, pp. 2431–2437.

Lister, R., Gregory, B.D. and Ecker, J.R. (2009). Next is now: new technologies for sequencing of genomes, transcriptomes, and beyond. *Current Opinion in Plant Biology*, vol. 12, pp. 107–118.

Liu, M., Smith, M. and Gill, S. (2002). Identification of microsatellite markers linked to Russian wheat aphid resistance genes *Dn4* and *Dn6*. *Theoretical and Applied Genetics*, vol. 104, pp. 1042–1048.

Liu, X.M., Smith, C.M., Friebe, B.R. and Gill, B.S. (2005). Molecular mapping and allelic relationships of Russian wheat aphid-resistance genes. *Crop Science*, vol. 45, pp. 2273–2280.

Liu, X.M., Smith, C.M., Gill, B.S. and Tolmay, V. (2001). Microsatellite markers linked to six Russian wheat aphid resistance genes in wheat. *Theoretical and Applied Genetics*, vol. 102, pp. 504–510.

Lotze, M.T., Zeh, H.J., Rubartelli, A., Sparvero, L.J., Amoscato, A.A., Washburn, N.R., Devera, M.E., Liang, X., Tor, M. and Billiar, T. (2007). The grateful

dead: damage- associated molecular pattern molecules and reduction/oxidation regulate immunity. *Immunological Review*, vol. 220, pp. 60–81.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Liu, Y., Yu, C., Wang, B., Lu, Y., Han, C., Cheung, D.W., Yiu, S.-M., Peng, S., Xiaoqian, Z., Liu, G., Liao, X., Li, Y., Yang, H., Wang, J., Lam, T.-W. and Wang, J. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *GigaScience*, vol. 1, p. 18.

Ma, B., Tromp, J. and Li, M. (2002). PatternHunter: faster and more sensitive homology search. *Bioinformatics*, vol. 18, pp. 440–445.

Ma, Z., Lapitan, N.L. and Steffenson, B. (2004). QTL mapping of net blotch resistance genes in a doubled-haploid population of six-rowed barley. *Euphytica*, vol. 137, pp. 291–296.

Ma, Z., Saidi, A., Quick, J.S. and Lapitan, N.L.V. (1998). Genetic mapping of Russian wheat aphid resistance genes *Dn2* and *Dn4* in wheat. *Genome*, vol. 41, pp. 303–306.

Maças, J., Weschke, W., Baumlein, H., Pich, U., Houben, A., Wobus, U. and Schubert, I. (1993). Localization of vicilin genes via polymerase chain reaction of microisolated field bean chromosomes. *Plant Journal*, vol. 3, pp. 883–886.

Malhis, N., Butterfield, Y.S.N., Ester, M. and Jones, S.J.M. (2008). Slider: maximum use of probability information for alignment of short sequence reads and SNP detection. *Bioinformatics*, vol. 25, pp. 6–13.

Manickavelu, A., Kawaura, K., Oishi, K., Shin-I, T., Kohara, Y., Yahiaoui, N., Keller, B., Suzuki, A., Yano, K. and Ogihara, Y. (2010). Comparative gene expression analysis of susceptible and resistant near-isogenic lines in common wheat infected by *Puccinia triticina*. *DNA Research*, vol. 17, pp. 211–222.

Marais, G.F. and Du Toit, F. (1993). A monosomic analysis of Russian wheat aphid resistance in the common wheat PI 294994. *Plant Breeding*, vol. 111, pp. 246–248.

Marais, G.F., Horn, M. and Du Toit, F. (1994). Intergeneric transfer (rye to wheat) of a gene(s) for Russian wheat aphid resistance. *Plant Breeding*, vol. 113, pp. 265–271.

Marais, G.F., Wessels, W.G., Horn, M. and Du Toit, F. (1998). Association of a stem rust resistance gene (*Sr45*) and two Russian wheat aphid resistance genes (*Dn5* and *Dn7*) with mapped structural loci in common wheat. *South African Journal of Plant and Soil*, vol. 15, pp. 67–71.

Marçais, G. and Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*, vol. 27, pp. 764–770.

Martin, G.B., Brommonschenkel, S.H., Chunwongse, J., Frary, A., Ganal, M.W., Spivey, R., Wu, T., Earle, E.D. and Tanksley, S.D. (1993). Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science*, vol. 262, pp. 1432–1436.

Mayer, K.F.X., Martis, M., Hedley, P.E., Šimková, H., Liu, H., Morris, J.A., Steuernagel, B., Taudien, S., Roessner, S., Gundlach, H., Kubala, M., Sucha, P., Murat, F., Felder, M., Nussbaumer, T., Graner, A. and Salse, J. (2011).

Unlocking the barley genome by chromosomal and comparative genomics. *The Plant Cell*, vol. 23, pp. 1249–1263.

Mayer, K.F.X., Rogers, J., Doležel, J., Pozniak, C., Eversole, K., Feuillet, C., Gill, B., Friebe, B., Lukaszewski, A.J., Sourdille, P., Endo, T.R., Kubaláková, M., Ihalikova, J., Dubska, Z., Vrána, J., Perkova, R., Imkova, H., Febrer, M., Clissold, L., McLay, K., Singh, K., Chhuneja, P., Singh, N.K., Khurana, J., Akhunov, E., Choulet, F., Alberti, A., Barbe, V., Wincker, P., Kanamori, H., Kobayashi, F., Itoh, T., Matsumoto, T., Sakai, H., Tanaka, T., Wu, J., Ogihara, Y., Handa, H., Maclachlan, P.R., Sharpe, A., Klassen, D., Edwards, D., Batley, J., Olsen, O.-A., Sandve, S.R., Lien, S., Steuernagel, B., Wulff, B., Caccamo, M., Ayling, S., Ramirez-Gonzalez, R.H., Clavijo, B.J., Wright, J., Pfeifer, M., Spannagl, M., Martis, M.M., Mascher, M., Chapman, J., Poland, J.A., Scholz, U., Barry, K., Waugh, R., Rokhsar, D.S., Muehlbauer, G.J., Stein, N., Gundlach, H., Zytnicki, M., Jamilloux, V., Quesneville, H., Wicker, T., Faccioli, P., Colaiacovo, M., Stanca, A.M., Budak, H., Cattivelli, L., Glover, N., Pingault, L., Paux, E., Sharma, S., Appels, R., Bellgard, M., Chapman, B., Nussbaumer, T., Bader, K.C., Rimbert, H., Wang, S., Knox, R., Kilian, A., Alaux, M., Alfama, F., Couderc, L., Guilhot, N., Viseux, C., Loaec, M., Keller, B. and Praud, S. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science*, vol. 345, p. 1251788.

Metezeau, P., Schmitz, A. and Frelat, G. (1993). Analysis and sorting of chromosomes by flow cytometry: new trends. *Biological Cell*, vol. 78, pp. 31–39.

Metzker, M.L. (2010). Sequencing technologies - the next generation. *Nature Reviews: Genetics*, vol. 11, pp. 31–46.

Meudt, H.M. and Clarke, A.C. (2007). Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends in Plant Science*, vol. 12, pp. 106–117.

Michels, G.J.J. and Behle, R.J. (1988). Reproduction and development of *Diuraphis noxia* (Homoptera: Aphididae) at constant temperatures. *Journal of Economic Entomology*, vol. 81, pp. 1097–1101.

Miles, P.W. (1999). Aphid saliva. *Biological Review*, vol. 74, pp. 41–85.

Miller, C.A., Altinkut, A. and Lapitan, N.L.V. (2001). A microsatellite marker for tagging *Dn2*, a wheat gene conferring resistance to the Russian wheat aphid. *Crop Science*, vol. 41, pp. 1584–1589.

Mittler, T.E. (1971). Dietary amino acid requirements of the aphid *Myzus persicae* affected by antibiotic uptake. *The Journal of Nutrition*, vol. 101, pp. 1023–1028.

Mochida, K., Yoshida, T., Sakurai, T., Ogihara, Y. and Shinozaki, K. (2009). TriFLDB: A database of clustered full-length coding sequences from Triticeae with applications to comparative grass genomics. *Plant Physiology*, vol. 150, pp. 1135–1146.

Monajembashi, S., Cremer, C., Cremer, T., Wolfrum, J. and Greulich, K.O. (1986). Microdissection of human chromosomes by a laser microbeam. *Expressive Cell Research*, vol. 167, pp. 262–265.

Moran, P.J., Cheng, Y., Cassell, J.L. and Thompson, G.A. (2002). Gene expression profiling of *Arabidopsis thaliana* in compatible plant-aphid interactions. *Archives of Insect Biochemistry and Physiology*, vol. 51, pp. 182–203.

Morgan, T.H. (1916). *Sex-linked inheritance in Drosophila*, vol. 237. Carnegie Institution of Washington.

Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T.L., Agarwala, R. and Schaffer, A.A. (2008). Database indexing for production megablast searches. *Bioinformatics*, vol. 24, pp. 1757–1764.

Morozova, O. and Marra, M.A. (2008). Genomics applications of next-generation sequencing technologies in functional genomics. *Genomics*, vol. 92, pp. 255–264.

Morton, N.E. (1955). Sequential tests for the detection of linkage. *American Journal of Human Genetics*, vol. 7, pp. 277–318.

Munson, M.A., Baumann, P. and Kinsey, M.G. (1991). *Buchnera* gen. nov. and *Buchnera aphidicola* sp. nov., a taxon consisting of the mycetocyte-associated, primary endosymbionts of aphids. *International Journal of Systems Bacteriology*, vol. 41, pp. 566–568.

Myburg, A.A., Cawood, M., Wingfield, B.D. and Botha, A.-M. (1998). Development of RAPD and SCAR markers linked to the Russian wheat aphid resistance gene *Dn2* in wheat. *Theoretical and Applied Genetics*, vol. 96, pp. 1162–1169.

Nagaraj, S.H., Gasser, R.B. and Ranganathan, S. (2007). A hitchhiker's guide to expressed sequence tag (EST) analysis. *Briefings In Bioinformatics*, vol. 8, pp. 6–21.

Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M. and Kumlin, E. (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science*, vol. 235, pp. 1616–1622.

Ning, Z., Cox, A.J. and Mullikin, C. (2001). SSAHA: a fast search method for large DNA databases. *Genome Research*, vol. 11, pp. 1725–1729.

Nkongolo, K.K., Quick, J.S., Peairis, F.B. and Meyer, W.L. (1991). Inheritance of resistance of PI 372129 wheat to the Russian wheat aphid. *Crop Science*, vol. 31, pp. 905–907.

Oleszczuk, S., Rabiza-Swider, J., Zimny, J. and Lukaszewski, A.J. (2011). Aneuploidy among androgenic progeny of hexaploid triticale (X *Triticosecale* Wittmack). *Plant Cell Reports*, vol. 30, pp. 575–586.

Olofsson, L., Lundgren, A. and Brodelius, P.E. (2012). Trichome isolation with and without fixation using laser microdissection and pressure catapulting followed by RNA amplification: expression of genes of terpene metabolism in apical and sub-apical trichome cells of *Artemisia annua* L. *Plant Science*, vol. 183, pp. 9–13.

Olson, M., Hood, L., Cantor, C. and Botstein, D. (1989). A common language for physical mapping of the human genome. *Science*, vol. 245, pp. 1434–1435.

Painter, R.H. (1951). Insect resistance in crop plants. vol. 72, p. 481. LWW.

Paux, E., Roger, D., Badaeva, E., Gay, G., Bernard, M., Sourdille, P. and Feuillet, C. (2006). Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. *The Plant Journal: for cell and molecular biology*, vol. 48, pp. 463–474.

Paux, E., Sourdille, P., Salse, J., Saintenac, C., Choulet, F., Leroy, P., Korol, A., Michalak, M., Kianian, S., Spielmeyer, W., Lagudah, E., Somers, D., Kilian, A., Alaux, M., Vautrin, S., Bergès, H., Eversole, K., Appels, R., Šafář, J., Šimková,

H., Doležel, J., Bernard, M. and Feuillet, C. (2008). A physical map of the 1-gigabase bread wheat chromosome 3B. *Science*, vol. 322, pp. 101–104.

Pearson, W.R. and Lipman, D.J. (1988). Improved tools for biological sequence comparison. *Proceedings of the National Academy of Sciences*, vol. 85, pp. 2444–2448.

Peng, J., Korol, B., Fahima, T., Röder, M.S., Ronin, Y.I., Li, Y.C. and Nevo, E. (2000). Molecular genetic maps in wild emmer wheat, *Triticum dicoccoides*: genome-wide coverage, massive negative interference, and putative quasi-linkage. *Genome Research*, vol. 10, pp. 1509–1531.

Peng, J., Wang, H., Haley, S.D., Peairs, F.B. and Lapitan, N.L. (2007). Molecular mapping of the Russian wheat aphid resistance gene *Dn2414* in wheat. *Crop Science*, vol. 47, pp. 2418–2429.

Peng, J.H., Bai, Y., Haley, S.D. and Lapitan, N.L.V. (2009). Microsatellite-based molecular diversity of bread wheat germplasm and association mapping of wheat resistance to the Russian wheat aphid. *Genetica*, vol. 135, pp. 95–122.

Peng, J.H. and Lapitan, N.L.V. (2005). Characterization of EST-derived microsatellites in the wheat genome and development of eSSR markers. *Functional and Integrative Genomics*, vol. 5, pp. 80–96.

Penner, G.A., Zirino, M., Kruger, S. and Townley-Smith, F. (1998). Accelerated recurrent parent selection in wheat with microsatellite markers. In: *Proceedings of the 9th International Wheat Genetics Symposium*, pp. 131–134. University of Saskatchewan, Saskatoon, Canada.

Pevzner, P.A. and Tang, H. (2001). Fragment assembly with double-barreled data. *Bioinformatics (Suppl. 1)*, vol. 17, pp. S225–S233.

Pevzner, P.A., Tang, H. and Waterman, M.S. (2001). An Eulerian path approach to DNA fragment assembly. *Proceedings of the National Academy of Sciences*, vol. 98, pp. 9748–9753.

Pfosser, M., Amon, A., Lelley, T. and Heberle-Bors, E. (1995). Evaluation of sensitivity of flow cytometry in detecting aneuploidy in wheat using disomic and ditelosomic wheat-rye addition lines. *Cytometry*, vol. 21, pp. 387–393.

Qi, L., Cao, M., Chen, P., Li, W. and Liu, D. (1996). Identification, mapping and application of polymorphic DNA associated with resistance gene *Pm21* of wheat. *Genome*, vol. 39, pp. 191–197.

Qi, L.L., Echalier, B., Chao, S., Lazo, G.R., Butler, G.E., Anderson, O.D., Akhunov, E.D., Dvorák, J., Linkiewicz, A.M., Ratnasiri, A., Dubcovsky, J., Bermudez-Kandianis, C.E., Greene, R.A., Kantety, R., La Rota, C.M., Munkvold, J.D., Sorrells, S.F., Sorrells, M.E., Dilbirligi, M., Sidhu, D., Erayman, M., Randhawa, H.S., Sandhu, D., Bondareva, S.N., Gill, K.S., Mahmoud, A.A., Ma, X.-F., Miftahudin, Gustafson, J.P., Conley, E.J., Nduati, V., Gonzalez-Hernandez, J.L., Anderson, J.A., Peng, J.H., Lapitan, N.L.V., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Pathan, M.S., Zhang, D.S., Nguyen, H.T., Choi, D.-W., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O. and Gill, B.S. (2004). A chromosome bin map of 16 000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. *Genetics*, vol. 168, pp. 701–712.

Qi, X. and Lindhout, P. (1997). Development of AFLP markers in barley. *Molecular General Genetics*, vol. 254, pp. 330–336.

Raats, D., Frenkel, Z., Krugman, T., Dodek, I., Sela, H., Simková, H., Magni, F., Cattonaro, F., Vautrin, S., Bergès, H., Wicker, T., Keller, B., Leroy, P., Philippe, R., Paux, E., Doležel, J., Feuillet, C., Korol, A. and Fahima, T. (2013). The physical map of wheat chromosome 1BS provides insights into its gene space organization and evolution. *Genome Biology*, vol. 14, p. R138.

Rafalski, A. (2002). Applications of single nucleotide polymorphisms in crop genetics. *Current Opinion in Plant Biology*, vol. 5, pp. 94–100.

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

Ravel, C., Praud, S., Murigneux, A., Canaguier, A., Sapet, F., Samson, D., Balfourier, F., Dufour, P., Chalhoub, B., Brunel, D., Beckert, M. and Charmet, G. (2006). Single-nucleotide polymorphisms (SNPs) frequency in a set of selected lines of bread wheat (*Triticum aestivum* L.). *Genome*, vol. 49, pp. 1131–1139.

Ritter, C. and Dangl, J.L. (1996). Interference between two specific pathogen recognition events mediated by distinct plant disease resistance genes. *Plant Cell*, vol. 8, pp. 251–257.

Röder, M.S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M.-H., Leroy, P. and Ganal, M.W. (1998). A microsatellite map of wheat. *Genetics*, vol. 149, pp. 2007–2023.

Rossi, M., Giggin, F.L., Milligan, S.B., Kaloshian, I., Ullman, D.E. and Williamson, V.M. (1998). The nematode resistance gene *Mi* of tomato con-

fers resistance against potato aphid. *Proceedings of the National Academy of Sciences*, vol. 95, pp. 9750–9754.

Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mannhaupt, G., Mokrejs, M., Tetko, I., Ulrich, G., Martin, M. and Mewes, H.W. (2004). The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Research*, vol. 32, pp. 5539–5545.

Šafář, J., Bartoš, J., Janda, J., Bellec, A., Kubaláková, M., Valarik, M., Pateyron, S., Weiserova, J., Tuskova, R., Číhalíková, J., Vrána, J., Šimková, H., Faivre-Rampant, P., Sourdille, P., Caboche, M., Bernard, M., Doležel, J. and Chalhoub, B. (2004). Dissecting large and complex genomes: flow sorting and BAC cloning of individual chromosomes from bread wheat. *The Plant Journal*, vol. 39, pp. 960–968.

Šafář, J., Šimková, H., Kubaláková, M., Hihalikova, J., Suchánková, P., Bartoš, J. and Doležel, J. (2010). Development of chromosome-specific BAC resources for genomics of bread wheat. *Cytogenetic Genome Research*, vol. 129, pp. 211–223.

Saidi, A. and Quick, J.S. (1996). Inheritance and allelic relationships among Russian wheat aphid resistance genes in winter wheat. *Crop Science*, vol. 36, pp. 256–258.

Sandery, M.J., Forster, J.W., Macadam, S.R., Blunden, R., Jones, R.N. and Brown, S.D.M. (1991). Isolation of a sequence common to A- and B-chromosomes of rye (*Secale cereale*) by microcloning. *Plant Molecular Biology Reporter*, vol. 9, pp. 21–30.

Sandhu, D. and Gill, K.S. (2002). Gene-containing regions of wheat and the other grass genomes. *Plant Physiology*, vol. 128, pp. 803–811.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, vol. 74, pp. 5463–5467.

Sarkar, P. and Stebbins, G.L. (1956). Morphological evidence concerning the origin of the B genome in wheat. *American Journal of Botany*, vol. 43, pp. 297–304.

Scalenghe, F., Turco, E., Edstrom, J.E., Pirrotta, V. and Melli, M. (1981). Microdissection and cloning of DNA from a specific region of *Drosophila melanogaster* polytene chromosomes. *Chromosoma*, vol. 82, pp. 205–216.

Schachermayr, G., Siedler, H., Gale, M.D., Winzeler, H., Winzeler, M. and Keller, B. (1994). Identification and localization of molecular markers linked to the *Lr9* leaf rust resistance gene of wheat. *Theoretical and Applied Genetics*, vol. 88, pp. 110–115.

Schwarzacher, T. (2003). Meiosis, recombination and chromosomes: a review of gene isolation and fluorescent in situ hybridization data in plants. *Journal of Experimental Botany*, vol. 54, pp. 11–23.

Sears, E.R. (1954). The aneuploids of common wheat. Bulletin 572, University of Missouri Agriculture Experiment Station.

Sharma, S., Xu, S., Ehdaie, B., Hoops, A., Close, T.J., Lukaszewski, A.J. and Waines, J.G. (2011). Dissection of QTL effects for root traits using a chromosome arm-specific mapping population in bread wheat. *Theoretical and Applied Genetics*, vol. 122, pp. 759–769.

Shea, G., Botha, J. and Hardie, D. (2000). Russian wheat aphid. Newsletter, Agriculture Western Australia.

Shufran, K.A., Kirkman, L.R. and Puterka, G.J. (2007). Absence of mitochondrial DNA sequence variation in Russian wheat aphid (Hemiptera: Aphididae) populations consistent with a single introduction into United States. *Journal of the Kansas Entomological Society*, vol. 80, pp. 319–326.

Šimková, H., Šafář, J., Kubaláková, M., Suchánková, P., Číhalíková, J., Robert-Quatre, H., Azhaguvel, P., Weng, Y., Peng, J., Lapitan, N.L.V., Ma, Y., You, F.M., Luo, M.-C., Bartovs, J. and Doležel, J. (2011). BAC libraries from wheat chromosome 7D: efficient tool for positional cloning of aphid resistance genes. *Journal of Biomedicine & Biotechnology*, vol. 2011, pp. 1–11.

Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J.M. and Birol, I. (2009). ABySS: a parallel assembler for short read sequence data. *Genome Research*, vol. 19, pp. 1117–1123.

Smith, A.D., Xuan, Z. and Zhang, M.Q. (2008). Using quality scores and longer reads improves accuracy of Solexa read mapping. *BMC Bioinformatics*, vol. 9, p. 128.

Smith, C.M., Belay, T., Stauffer, C., Stary, P., Kubeckova, I. and Starkey, S. (2004). Identification of Russian wheat aphid (Homoptera: Aphididae) populations virulent to the *Dn4* resistance gene. *Journal of Economic Entomology*, vol. 97, pp. 1112–1117.

Smith, C.M. and Boyko, E.V. (2007). The molecular bases of plant resistance and

defense responses to aphid feeding: current status. *Entomologia Experimentalis et Applicata*, vol. 122, pp. 1–16.

Smith, C.M., Liu, X.M., Wang, L.J., Liu, X., Chen, M.S., Starkey, S. and Bai, J. (2010). Aphid feeding activates expression of a transcriptome of oxylipin-based defense signals in wheat involved in resistance to herbivory. *Journal of Chemical Ecology*, vol. 36, pp. 260–276.

Smith, C.M., Schotzko, D., Zemetra, R.S., Souza, E.J. and Schroeder-Teeter, S. (1991). Identification of Russian wheat aphid (Homoptera: Aphididae) resistance in wheat. *Journal of Economic Entomology*, vol. 84, pp. 328–332.

Somers, D.J., Isaac, P. and Edwards, K. (2004). A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, vol. 109, pp. 1105–1114.

Somers, D.J., Kirkpatrick, R., Moniwa, M. and Walsh, A. (2003). Mining single-nucleotide polymorphisms from hexaploid wheat ESTs. *Genome*, vol. 437, pp. 431–437.

Song, Q.J., Shi, J.R., Singh, S., Fickus, E.W., Costa, J.M., Lewis, J., Gill, B.S., Ward, R. and Cregan, P.B. (2005). Development and mapping of microsatellite (SSR) markers in wheat. *Theoretical and Applied Genetics*, vol. 110, pp. 550–560.

Sorrells, M.E., La Rota, M., Bermudez-Kandianis, C.E., Greene, R.A., Kantety, R., Munkvold, J.D., Mahmoud, A., Ma, X., Gustafson, P.J., Qi, L.L., Echalier, B., Gill, B.S., Matthews, D.E., Lazo, G.R., Chao, S., Anderson, O.D., Edwards, H., Linkiewicz, A.M., Dubcovsky, J., Akhunov, E.D., Dvorak, J., Zhang, D.,

Nguyen, H.T., Peng, J., Lapitan, N.L.V., Gonzalez-Hernandez, J.L., Anderson, J.A., Hossain, K., Kalavacharla, V., Kianian, S.F., Choi, D.-W., Close, T.J., Dilbirligi, M., Gill, K.S., Steber, C., Walker-Simmons, M.K., McGuire, P.E. and Qualset, C.O. (2003). Comparative DNA sequence analysis of wheat and rice genomes. *Genome Research*, vol. 13, pp. 1818–1827.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*, vol. 98, pp. 503–517.

Stary, P., Basky, Z., Tanigoshi, L.K. and Tomanovicc, Z. (2003). Distribution and history of Russian wheat aphid, *Diuraphis noxia* (Kurdj.) in the Carpathian Basin (Hom., Aphididae). *Journal of Pest Science*, vol. 76, pp. 17–21.

Stein, N., Feuillet, C., Wicker, T., Schlagenhauf, E. and Keller, B. (2000). Subgenome chromosome walking in wheat: a 450-kb physical contig in *Triticum monococcum* L . spans the *Lr10* resistance locus in hexaploid wheat (*Triticum aestivum* L.). *Proceedings of the National Academy of Sciences*, vol. 97, pp. 13436–13441.

Stoetzel, M.B. (1987). Information on and identification of *Diuraphis noxia* (Homoptera: Aphididae) and other aphid species colonizing leaves of wheat and barley in the United States. *Journal of Economic Entomology*, vol. 80, pp. 696–704.

Sturtevant, A.H. (1913). The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *Journal of Experimental Zoology*, vol. 14, pp. 43–59.

Swanepoel, E., Lacock, L., Myburg, A. and Botha, A. (2003). A leucine rich homolog to *Aegilops tauschii* from bread wheat line PI137739 obtained by suppression subtractive hybridization shows linkage to russian wheat aphid resistance gene *Dn1*. In: *Proceedings of the 10th International Wheat Genetics Symposium*, vol. 3, pp. 1263–1265. Paestum, Italy.

Swanevelder, Z.H., Surridge, A.K.J., Venter, E. and Botha, A.-M. (2010). Limited endosymbiont variation in *Diuraphis noxia* (Hemiptera: Aphididae) biotypes from the United States and South Africa. *Journal of Economic Entomology*, vol. 103, pp. 887–897.

Talbert, L.E., Bruckner, P.L., Smith, L.Y., Sears, R. and Martin, T.J. (1996). Development of PCR markers linked to resistance to wheat streak mosaic virus in wheat. *Theoretical and Applied Genetics*, vol. 93, pp. 463–467.

Thole, V., Peraldi, A., Worland, B., Nicholson, P., Doonan, J.H. and Vain, P. (2012). T-DNA mutagenesis in *Brachypodium distachyon*. *Journal of Experimental Botany*, vol. 63, pp. 567–576.

Tjallingii, W.F. and Hogen Esch, T.H. (1993). Fine structure of aphid stylet routes in plant tissues in correlation with EPG signals. *Physiological Entomology*, vol. 18, pp. 317–328.

Tolmay, V.L., Lindeque, R.C. and Prinsloo, G.J. (2007). Preliminary evidence of a resistance-breaking biotype of the Russian wheat aphid, *Diuraphis noxia* (Kurdjumov) (Homoptera: Aphididae), in South Africa. *African Entomology*, vol. 15, pp. 228–230.

Tor, M., Lotze, M.T. and Holton, N. (2009). Receptor-mediated signalling in plants: molecular patterns and programmes. *Journal of Experimental Biology*, vol. 60, pp. 3645–3654.

Valdez, V.A., Byrne, P.F., Lapitan, N.L.V., Peairs, F.B., Bernardo, A., Bai, G. and Haley, S.D. (2012). Inheritance and genetic mapping of Russian wheat aphid resistance in Iranian wheat landrace accession PI 626580. *Crop Science*, vol. 52, pp. 676–682.

Van der Westhuizen, A.J., Qian, X.-M. and Botha, A.-M. (1998*a*). B-1,3-Glucanases in wheat and resistance to the Russian wheat aphid. *Physiologia Plantarum*, vol. 103, pp. 125–131.

Van der Westhuizen, A.J., Qian, X.-M. and Botha, A.-M. (1998*b*). Differential induction of apoplastic peroxidase and chitinase activities in susceptible and resistant wheat cultivars by Russian wheat aphid infestation. *Plant Cell Reports*, vol. 18, pp. 132–137.

Van Ham, R.C.H.J., Moya, A. and Latorre, A. (1997). Putative evolutionary origin of plasmids carrying the genes involved in leucine biosynthesis in *Buchnera aphidicola* (endosymbiont of aphids). *Journal of Bacteriology*, vol. 179, pp. 4768–4777.

Van Niekerk, H.A. (2001). Southern Africa wheat pool. In: Bonjean, A.P. and Angus, W.J. (eds.), *The world wheat book: the history of wheat breeding*, pp. 923–936. Lavoisier Publishing, Paris.

Vedel, F. and Delseny, M. (1987). Repetitivity and variability of higher-plant genomes. *Plant Physiology and Biochemistry*, vol. 25, pp. 191–210.

Venter, E. and Botha, A.-M. (2000). Development of markers linked to *Diuraphis noxia* resistance in wheat using a novel PCR-RFLP approach. *Theoretical and Applied Genetics*, vol. 100, pp. 965–970.

Venter, E., Myburg, A.A. and Botha, A.-M. (1998). Characterisation of sequences linked to Russian wheat aphid resistance genes *Dn1*, *Dn2* and *Dn5*. In: *Proceedings of the 9th International Wheat Genetics Symposium*, vol. 3, pp. 332–335. Saskatoon Saskatchewan, Canada.

Vitulo, N., Albiero, A., Forcato, C., Campagna, D., Dal Pero, F., Bagnaresi, P., Colaiacovo, M., Faccioli, P., Lamontanara, A., Šimková, H., Kubaláková, M., Perrotta, G., Facella, P., Lopez, L., Pietrella, M., Gianese, G., Doležel, J., Giuliano, G., Cattivelli, L., Valle, G. and Stanca, A.M. (2011). First survey of the wheat chromosome 5A composition through a next generation sequencing approach. *PloS ONE*, vol. 6, p. e26421.

Voothuluru, P., Meng, J.Y., Khajuria, C., Louis, J., Zhu, L.C., Starkey, S., Wilde, G.E., Baker, C.A. and Smith, C.M. (2006). Categories and inheritance of resistance to Russian wheat aphid (Homoptera: Aphididae) biotype 2 in a selection from wheat cereal introduction 2401. *Journal of Economic Entomology*, vol. 99, pp. 1854–1861.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van De Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, vol. 23, pp. 4407–4414.

Vossen, J.H., Dezhsetan, S., Esselink, D., Arens, M., Sanz, M.J., Verweij, W., Verzaux, E. and Van Der Linden, G.C. (2013). Novel applications of motif-

directed profiling to identify disease resistance genes in plants. *Plant Methods*, vol. 9, p. 37.

Vrána, J., Kubaláková, M., Šimková, H., Číhalíková, J., Lysák, M.A. and Doležel, J. (2000). Flow sorting of mitotic chromosomes in common wheat (*Triticum aestivum* L.). *Genetics*, vol. 156, pp. 2033–2041.

Vrána, J., Šimková, H., Kubaláková, M., Čihalíková, J. and Doležel, J. (2012). Flow cytometric chromosome sorting in plants: The next generation. *Methods*, vol. 57, pp. 331–337.

Walling, L.L. (2008). Avoiding effective defenses: strategies employed by phloem-feeding insects. *Plant Physiology*, vol. 146, pp. 859–866.

Walters, M.C., Penn, F., Du Toit, F., Botha, T.C., Aalbersberg, K., Hewitt, P.H. and Broodryk, S.W. (1980). The Russian wheat aphid. Leaflet series wheat G3, Farming in South Africa.

Warren, R.L., Sutton, G.G., Jones, S.J. and Holt, R.A. (2007). Assembling millions of short DNA sequences using SSAKE. *Bioinformatics*, vol. 23, pp. 500–501.

Webster, J.A., Starks, K.J. and Burton, R.L. (1987). Plant resistance studies with *Diuraphis noxia* (Homoptera: Aphididae), a new United States wheat pest. *Journal of Economical Entomology*, vol. 80, pp. 944–949.

Weiland, A.A., Peairs, F.B., Randolph, T.L., Rudolph, J.B., Haley, S.D. and Puterka, G.J. (2008). Biotypic diversity in Colorado Russian wheat aphid (Hemiptera: Aphididae) populations. *Journal of Economic Entomology*, vol. 102, pp. 569–574.

Welsh, J. and McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research*, vol. 18, pp. 7213–7218.

Wenzl, P., Carling, J., Kudrna, D., Jaccoud, D., Huttner, E., Kleinhofs, A. and Kilian, A. (2004). Diversity arrays technology (DArT) for whole-genome profiling of barley. *Proceedings of the National Academy of Sciences*, vol. 101, pp. 9915–9920.

Werner, J.E., Endo, T.R. and Gill, B.S. (1992). Toward a cytogenetically based physical map of the wheat genome. *Proceedings of the National Academy of Sciences*, vol. 89, pp. 11307–11311.

Wicker, T., Yahyaoui, N., Gyuot, R., Schlagenhauf, E., Liu, Z.-D., Dubkovsky, J. and Keller, B. (2003). Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and Am genomes of wheat. *Plant Cell*, vol. 15, pp. 1186–1197.

Will, T., Hewer, A. and Van Bel, A.J.E. (2008). A novel perfusion system shows that aphid feeding behaviour is altered by decrease of sieve-tube pressure. *Entemologia Experimentalis et Applicata*, vol. 127, pp. 237–245.

Will, T., Tjallingii, W.F., Thonnessen, A. and Van Bel, A.J.E. (2007). Molecular sabotage of plant defense by aphid saliva. *Proceedings of the National Academy of Sciences*, vol. 104, pp. 10–15.

Williams, J.G., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, vol. 18, pp. 6531–6535.

Young, N.D. (2000). Constructing a plant genetic linkage map with DNA markers. In: *Constructing a plant genetic linkage map with DNA markers*, chap. 3, pp. 31–47. Minnesota.

Zaayman, D., Lapitan, N.L.V. and Botha, A.-M. (2009). Dissimilar molecular defense responses are elicited in *Triticum aestivum* after infestation by different *Diuraphis noxia* biotypes. *Physiologia Plantarum*, vol. 136, pp. 209–222.

Zerbino, D.R. and Birney, E. (2008). Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Research*, vol. 18, pp. 821–829.

Zhang, K., Tian, J., Zhao, L. and Wang, S. (2008). Mapping QTLs with epistatic effects and QTL environment interactions for plant height using a doubled haploid population in cultivated wheat. *Journal of Genetics and Genomics*, vol. 35, pp. 119–127.

Zhang, Y., Quick, J.S. and Liu, S. (1998). Genetic variation in PI 294994 wheat for resistance to Russian wheat aphid. *Crop Science*, vol. 38, pp. 527–530.

Zhou, R.-N. and Hu, Z.-M. (2007). The development of chromosome microdissection and microcloning technique and its applications in genomic research. *Current Genomics*, vol. 8, pp. 67–72.

Chapter 3

Genetic mapping

A high-density *Triticum aestivum* L. chromosome 7DS map in the region of the *Dn1* resistance gene

Submitted for review to Theoretical and Applied Genetics

3.1 Abstract

Diuraphis noxia (Kurdjumov, Hemiptera: Aphididae), commonly known as the RWA, is a devastating pest of wheat and barley. Even though fourteen sources of *Dn* resistance have been identified to date, none have been cloned. In the present study, we constructed a saturated genetic map in the region of the *Dn1* gene and confirmed the position of the gene on chromosome 7D of wheat using a $F_{3/4}$ Tugela x Tugela-*Dn1* segregating population consisting of 581 individuals, and Chinese Spring deletion lines. The 122.8 cM genetic map constructed of the *Dn1* resistance gene placed the gene on chromosome 7DS near the centromere, in bin 7DS5-0.36-0.62 proportionally to the physical Chinese Spring deletion 7D map. The map contains 38 new AFLP markers, one microsatellite (*Xgwm111*) and two EST markers. Of all markers, AFLP E-ACT/M-CTG_0270.84 showed the closest linkage to *Dn1* and mapped at 3.5 cM, while EST markers RGA2-29_30 and SSH-RGA2 mapped respectively at 15.3 cM and 15.9 cM, from *Dn1*.

3.2 Introduction

Diuraphis noxia (Kurd.) commonly known as the Russian wheat aphid is a devastating pest of *Triticum aestivum* L. that occurs in most wheat producing countries. However, its economic impact on wheat was reported mainly in South Africa and the USA, with reported yield losses of up to 80% in susceptible cultivars (Marasas *et al.*, 1997), and an estimated economic burden of $893 million in the USA from 1987 to 1993 (Morrison and Peairs, 1998). Chlorotic streaking, leaf rolling and head trapping are symptomatic of *D. noxia* feeding in susceptible varieties with consequences including loss in photosynthesis, loss of turgor and death (Fouché *et al.*, 1984; Heng-Moss *et al.*, 2003; Botha, 2013). The modes offered by sources of resistance against *D. noxia* have been defined as antibiosis, antixenosis and tolerance or a combination thereof that hint at the presence of varying resistance genes in different cultivars (Painter, 1951, 1958; Botha *et al.*, 2014). Tolerance is seen as a lack of plant height reduction despite feeding; antixenosis is the non-preference of a cultivar as host; while antibiosis is observed when the plant reduces the reproductive fitness of aphids feeding on it (Painter, 1951, 1958). Several studies indicate that *Dn1* affords antibiosis (Du Toit, 1990; Smith *et al.*, 1992; Unger and Quisenberry, 1997; Budak *et al.*, 1999; Wang *et al.*, 2004).

Since the first reports of this invasive pest, efforts to control the insect relied mainly on chemical spraying, which proved a costly, environmentally unfriendly and ineffective approach, due to the aphids' feeding preference – i.e. feeds on the new, uncurled inner parts of the leaves thus providing it with shelter. The use of parasitic wasps as biological control measure has also been rendered ineffective through the natural protection of the aphid by the uncurled leaves (Prinsloo *et al.*, 2002). Until recently, *D. noxia* infestation was managed by planting resistant

cultivars, modified cultural practices and intermittent insecticide dosage (Smith and Clement, 2012), but this integrated control management strategy was rendered ineffective with the development of new *D. noxia* biotypes in the USA in 2003 (Haley *et al.*, 2004) and South Africa in 2006 (Tolmay *et al.*, 2006; Jankielsohn, 2011).

To date, 14 sources of resistance against *D. noxia* have been identified, with *Dn1* and *Dn2* being the first sources identified and originating from Iran (PI137739) and Bulgaria (PI262660) respectively (Du Toit, 1987, 1988, 1990). These resistance sources showed single, dominant gene inheritance (Du Toit, 1987, 1990; Marais and Du Toit 1993). Since then, another twelve sources of resistance against *D. noxia* have been identified based on observed resistance responses of seedlings in the greenhouse or adult plants in field trials, including *dn3* and *Dn4* (Nkongolo *et al.*, 1991); *Dn5* (Marais and Du Toit, 1993); *Dn6* (Saidi and Quick, 1996); *Dn7* (Marais *et al.*, 1998; Anderson *et al.*, 2003); *Dn8*, *Dn9* and *Dnx* (Harvey and Martin, 1990; Liu *et al.*, 2001); *Dny* (Smith *et al.*, 2004); *Dn2414* (Peng *et al.*, 2007); *Dn626580* (Valdez *et al.*, 2012) and *DnCI2401* (Fazel-Najafabadi *et al.*, 2014 unpublished results). Even though the aphid associated molecular patterns (AAMPs) associated with *D. noxia* feeding have been elucidated (Botha *et al.*, 2014), none of the genes conferring resistance to the aphid has been cloned and sequence characterized. Although genes that confer resistance to *D. noxia* has been placed on chromosome 1B for *Dn7* (Anderson *et al.*, 2003) and most of the other genes on chromosome 7D (Liu *et al.*, 2002), the exact chromosomal location (either 7DS or 7DL) of the *Dn1*, *Dn2* and *Dn5* genes remains a contentious issue. Reports in literature place these genes on both arms of chromosome 7D, with two reports of *Dn1* occurring on 7DS (Liu *et al.*, 2001, 2002) and one that places the

gene on 7DL (Marais and Du Toit, 1993).

Although *Dn1* is not widely used in commercial wheat varieties as the only source of resistance to *D. noxia*, it is still present in many of the South African wheat cultivars, where the gene in combination with other *Dn* resistance sources (i.e., *Dn2* and *Dn5*), provides a reasonable level of resistance in planted commercial varieties against South African *D. noxia* biotypes. However, *Dn1* was applied in the present study as it serves as a useful model to study as it is mapped on chromosome 7D, clustered with most of the other *Dn* genes and could therefore serve as a starting point for their characterization. In addition, the molecular expression patterns of the *Dn1* resistance gene are defined by its hypersensitive response (Botha *et al.*, 2014), which is clearly visible in the resistance phenotype through the expression of necrotic lesions after aphid infestation, unlike that of *Dn2* (i.e., tolerance) and *Dn5* (antixenosis) (Botha *et al.*, 2014).

There have been many attempts at finding markers closely linked to the *Dn* resistance genes for their use as molecular breeding tools and for cloning. Some of the best-described and more useful markers include microsatellites and RAPD markers (Myburg *et al.*, 1998; Venter *et al.*, 1998; Venter and Botha, 2000; Liu *et al.*, 2001, 2005; Miller *et al.*, 2001). Microsatellite marker *Xgwm111* has been linked to *Dn1* at distances of 3.15 cM and 3.82 cM (Liu *et al.*, 2001; Swanepoel *et al.*, 2003), while RAPD markers have been linked to *Dn1* at 43.7 cM (Venter *et al.*, 1998). An as yet unexplored marker system in genetic mapping studies of the *D. noxia* resistance genes in wheat is AFLPs. Although AFLPs are dominant markers (Vos *et al.*, 1995), the fact that they offer genome wide coverage with high throughput, made them attractive for use in the present study. Thus, the objective of this study was to construct a saturated genetic map in the region of

Dn1 in order to firstly, identify markers that can be used for the characterization of the *Dn1* gene, and secondly, resolve the location of this gene (i.e., either on 7DS or 7DL) on chromosome 7D of *Triticum aestivum* L. For this purpose, a mapping population of 581 $F_{3/4}$ individuals segregating for *Dn1* was produced, and analyzed with different markers to assess linkage.

3.3 Materials and Methods

3.3.1 Plant material and segregation analysis

The mapping population was developed from single seed decent after crossing the NILs, Tugela and Tugela-*Dn1* (Tugela/*5 SA1684). Tugela-*Dn1* is derived from a cross between Tugela (susceptible) and SA1684 (PI 137739, 'Gandum I Fasai', Du Toit, 1987; Table 3.1). The resulting F_1 progeny was self-pollinated, and segregation analysis was conducted on the F_2 progeny. The obtained F_2 phenotypic segregation ratio of 3:1 (data not shown) confirmed previous reports that the *Dn1* gene is a single, dominant gene (Du Toit, 1987, 1989). The resistant offspring was then self polinated producing a progeny of 581 individuals. Seeds were sown into five pots for each cultivar and thinned to three seedlings per pot after five days. Plants were grown for 14 days (2-3 leaf stage) under greenhouse conditions in a 1:2:2:1 mixture of perlite (Chemserve, Olifantsfontein, South Africa), sifted bark compost, loam and sand at 25 °C ± 2 °C, before infestation and phenotypic screening.

Table 3.1: Genotype and plant symptoms after infestation with *D. noxia* biotype SA1 based on chlorosis scores, streaking, leaf rolling and virulence score. Virulence scores and final virulence profiles calculated according to Weiland *et al.* (2008), where resistant (R; RR, homozygote resistant) = 1-3; intermediate (R; Rr, heterozygote resistant) = 4-6; and susceptible (S; rr, homozygote susceptible) = 7-10. For leaf streaking and rolling: 'Y' = visible, 'N' = none visible. For chlorosis: 1 = 20% or less chlorosis per leaf and 4 = 100% chlorosis (dead leaf/tissue).

Plant genotype	**Genotype and/or *Dn* resistance source**	**Streaking**	**Leaf rolling**	**Chlorosis**	**Virulence score**	**Virulence profile**
Tugela	*Dn0* (PI 634771)	Y	Y	3	9/10	S (rr)
Tugela–*Dn1*[1]	Tugela/*5 SA1684	N	N	1	1/2	R (RR)
Tugela–*Dn2*[2]	Tugela/*5 SA2199	N	N	1	1	R (RR)
Tugela–*Dn5*[3]	Tugela/*5 SA463	N	N	1	1	R (RR)
Palmiet–*Dn5*[4]	Palmiet/*5 SA463	N	N	1	1	R (RR)
SA1684	*Dn1* (PI 137739, 'Gandum I Fasai')	N	N	1	1/2	R (RR)
SA2199	*Dn2* (PI 262660, 'Turtsikum')	N	N	1	1	R (RR)
SA463	*Dn5* (PI 294994, 'Strelinskaya Mestnaya')	N	N	1	1	R (RR)
*W*237[5]	Chinese Spring 7DS dt (40 + $2t^{7DS}$)	N	N	1	1	R (RR)
*W*1318/1376[5]	Chinese Spring 7DL dt	N	N	1	1	R (RR)

[1] Du Toit (1987); [2] Du Toit (1988); [3] Marais and Du Toit (1993); [4] Courtesy of Dr. H. Van Niekerk, Small Grains Institute, Bethlehem, South Africa (unpublished results); [5] Courtesy of Dr. G.F. Marais, Genetics Department, Stellenbosch University (unpublished results).

Aphids used for inducing plant responses were adult, apterous *D. noxia* of the South African biotype SA1, obtained from a colony established from field-collected parthenogenetic females at the Agricultural Research Council-Small Grains Institute and maintained on the susceptible cv. Tugela.

Each individual plant within the segregating population was phenotypically evaluated using a scoring system as described by Weiland *et al.* (2008) after infestation with *D. noxia* biotype SA1 (Table A.1). This phenotypic scoring scale is based on rating the extent of leaf chlorosis and leaf rolling. Chlorosis is measured using a modified scale from one to ten (Webster *et al.*, 1987) with one being indicative of a healthy plant and ten indicating a severely streaked or dead plant. Leaf rolling is measured on a one to three scale with one being a flat leaf and three being tightly rolled. The combination of these two scores is indicative of the plant's resistance. Resistant plants have leaf chlorosis scores of one to three and leaf rolling scores of one while susceptible plants have leaf chlorosis scores ranging from seven to ten and a leaf rolling score of three. In this study, only resistant samples scoring one for chlorosis and leaf rolling, and susceptible samples scoring ten and three respectively, were used for the construction of the genetic map. All other samples, i.e. leaf chlorosis scores of two to nine and leaf rolling scores of two, were omitted during the genotypic analyses resulting in a mapping population of 343.

To determine the location of *Dn1*, marker *Xgwm111* (Table A.2) and Chinese Spring ditelosomic lines (W237, Chinese Spring 7DS dt (40 + 2t^{7DS}) and W1318/W1376, Chinese Spring 7DL dt) were included in the study (Table 3.1). Other plant material also used in the analysis included *D. noxia* susceptible and resistant hexaploid wheat (*Triticum aestivum* L.) NILs, progenitors and donors

of *Dn1*, *Dn2* and *Dn5* resistance (Table 3.1). Progenitors included SA1684 (PI 137739, 'Gandum I Fasai' – donor of *Dn1*, Du Toit, 1987), SA2199 (PI 262660, 'Turtsikum' – donor of *Dn2*, Du Toit, 1988) and SA463 (PI 294994, 'Strelinskaya Mestnaya' – donor of *Dn5*, Marais and Du Toit, 1993; Table 3.1). Wheat lines containing *Dn2* or *Dn5* (i.e., Tugela-*Dn2*, Tugela-*Dn5* and Palmiet-*Dn5*) were included in the study, as these genes have previously been found to be allelic to *Dn1* and located within the same gene cluster on chromosome 7D (Liu *et al.*, 2002).

3.3.2 DNA extraction

Genomic DNA (gDNA) was extracted from all resistant and susceptible segregating offspring as well as parental and deletion lines using DNAzol® (Molecular Research Centre Inc. USA) following the manufacturers' protocol. DNA sample quality and quantity to be used for mapping was assessed spectrophotometricly on the NanoDrop v. 1000 (ThermoScientific), as well as visually after electrophoretic separation on 1.2% (w/v) agarose gels.

3.3.3 Amplified fragment length polymorphism (AFLP) analysis

Amplified fragment length polymorphism template preparation was performed using the AFLP template preparation kit from LI-COR Biosciences (LI-COR, Lincoln, NE, USA) according to the manufacturers' instructions. Polymerase chain reactions (PCRs) were performed using a BIO-RAD T100 Thermal Cycler (BIO-RAD Laboratories, inc). The preselective amplification cycle profile was as follows: incubation for 10 s at 72 °C, followed by 30 cycles of denaturation for 10 s at 94 °C, annealing for 30 s at 56 °C, and extension for one minute at 72 °C with a 1 s per

cycle increasing extension time. Selective amplification was performed on 1:20 diluted (in SABAX water) preselective amplification products with the following cycling profile: 13 cycles of 2 min at 94 °C, 30 s at 65 °C (reduced by 0.7 °C per cycle), and 1 min at 72 °C; followed by 20 cycles of 10 s at 94 °C, 30 s at 56 °C, and 1 min (extended one second per cycle) at 72 °C. Initially, a subset of two resistant and two susceptible samples from the Tugela x Tugela-*Dn1* $F_{3/4}$ population, as well as the resistant Tugela-*Dn1* and susceptible Tugela parents were chosen to test the level of polymorphism and ambiguity of 26 AFLP primer combinations (i.e., E-ACC/M-CAA; E-ACG/M-CAA; E-ACT/M-CAC; E-ACA/M-CAG; E-ACC/M-CAG; E-AGC/M-CAT; E-AGG/M-CAT; E-AGC/M-CTA (Figure A1); E-AGG/M-CTA; E-AAC/M-CTC; E-AAG/M-CTC; E-ACA/M-CTC; E-AAC/M-CTG; E-AAG/M-CTG; E-ACA/M-CTG; E-AAC/M-CTT; E-AAG/M-CTT; E-ACA/M-CTT; E-AAC/M-CAA; E-ACA/M-CAA; E-AAC/M-CAC; E-ACA/M-CAC; E-ACC/M-CTT; E-ACG/M-CTT; E-ACT/M-CTG; E-AGC/M-CTG). The 26 AFLP primer sets used to screen for polymorphisms between resistant and susceptible plants yielded 10 600 loci in total. The selected three most informative primer sets yielded 38 polymorphic loci between resistant and susceptible plants across the total mapping population. This constitutes only 0.35% of the total number of loci observed and is indicative of the low diversity present in the mapping population (data not shown). Marker sets with the highest number of polymorphisms (i.e., E-ACC/M-CTT; E-ACT/M-CTG; E-ACA/M-CAG) were selected to screen the entire mapping population.

An equal volume of loading solution (LI-COR) was added to each selective amplification reaction. Samples were denatured at 95 °C for 3 min and placed on ice for 10 min before loading. A volume of 0.8 μl was loaded with an eight-channel

syringe (Hamilton, Reno, Nevada) onto 25 cm (0.25 mm thick) 8% (v/v) Long Ranger gels (BMA, Rockland, ME, USA) (Myburg *et al.*, 2001). Electrophoresis and detection of AFLP fragments were performed on LI-COR IR (model 4200S) automated DNA analyzers. The electrophoresis parameters were set to 1500 V, 40 mA, 40 W, 50 °C, and a scan speed of three. The run-time was set to four hours and gel images were saved as TIF files for further analysis. The gel images were scored using a binary scoring system that recorded the presence and absence of bands as 1 and 0, respectively. SAGA^{TM} GT/MX Automated AFLP analysis software (LI-COR BioSciences, Lincoln, Nebraska) or GelQuest Digital DNA Processing Software v 3.1.1 (2010 SequentiX, Germany) were used to size fragments and place them into marker bins. A locus was scored as polymorphic when the frequency of the most common allele (band present or absent) was less than 0.97 (absent or present in at least two individuals). Bands with the same mobility were considered as identical products (Waugh *et al.*, 1997), receiving equal values regardless of their fluorescence intensity. In addition to scoring the gels using computer software, all gels were manually scored as well for verification. A total of 38 polymorphic AFLP loci were identified for use in linkage mapping.

3.3.4 Microsatellite and EST analysis

In addition to the AFLP markers, 129 microsatellite markers and 23 EST markers were screened for polymorphism and possible linkage to *Dn1* in the Tugela x Tugela-*Dn1* $F_{3/4}$ population. Ten of the EST markers were derived from Affymetrix data (Botha *et al.*, 2014) with the remainder of the ESTs being derived from GrainGenes Physical EST map (Table A.2). Primers were also used for amplification of DNA templates from the parental lines (resistant Tugela-*Dn1* and susceptible

Tugela), progenitors of *Dn1*, *Dn2* and *Dn5*, Chinese Spring deletion lines and other resistant lines (Table 3.1).

PCR reactions were set up from the following working concentrations: 0.2 mM dNTP, 2.5 mM $MgCl_2$, 1x GoTaq Flexi Buffer, 1.25 U/μL GoTaq Flexi Taq (Promega), 1 μM forward primer and 1 μM reverse primer up to 50 μL with distilled water. DNA template concentrations ranged from 100 to 500 ng. The reaction was denatured for 5 min at 95 °C followed by 35 cycles of 95 °C for 1 min; primer Tm °C for 45 s and 72 °C for 1 min with a final extension of 10 min at 72 °C on the Applied Biosystems GeneAmp® PCR System 2700. PCR products were visualized on 8% (v/v) Polyacrylamide gels [TBE buffer (Trizma base, Boric Acid, EDTA, pH 8.0); 30% (m/v) acrylamide/bisacrylamide (Acrylamide: N, N'-methylene-bis-acrylamide, 29:1); N,N,N',N'-Tetramethylethylenediamine, and 10% (v/v) ammonium peroxodisulfate] run at 90V for 75 min and stained with 0.025% ethidium bromide in 1 x TBE buffer (pH 8.0) for 30 min.

3.3.5 Mapping parameters and software

All markers and the *Dn1* gene were subjected to Chi-square analysis to test for the segregation pattern, 1.667:1. For Chi-square analysis a theoretical value of 3.84 with one degree of freedom was accepted as the cutoff point for significance. JoinMap software v 4.1 (Kyazma BV, Wageningen, Netherlands) (Van Ooijen, 2011) was used to construct the linkage map. Linkage groups were constructed using Linkage LOD with Maximum Likelihood parameters. Linkage groups were considered significant only if linkages had a LOD greater than 3. For Maximum likelihood a maximum chain length of 2000 was used with a stopping criterion of 2000 and a cooling criterion of 0.0007. The Kosambi mapping function was

applied to calculate CentiMorgan units (Kosambi, 1943). Markers were discarded if Chi-square analysis showed them not to be significant. Maps were also drawn up iteratively: omitting markers and replacing them in order to confirm their positions.

To enable for the positioning of *Dn1* relative to the available Chinese Spring deletion 7D genetic and physical maps, the linkage of *Dn1* relative to marker *Xgwm111* was conducted. Specifically maps constructed by Sourdille *et al.* (2004) and an unpublished Chinese Spring deletion 7D genetic and physical map that is available on GrainGenes (http://wheat.pw.usda.gov/cgibin/cmap/viewer?mapMenu=1&featureMenu=1&corrMenu=1&displayMenu=1&advancedMenu=1&ref_map_accs=Chinese_Spring_Deletion_SSR_7D&sub=Draw+Selected+Maps&ref_map_set_acc=Chinese_Spring_Deletion_SSR_7D&data_source=GrainGenes) were applied for this purpose.

3.4 Results

3.4.1 Segregation analysis of the *Dn1* phenotype

The mapping population was derived after crossing Tugela (susceptible) and its NIL Tugela-*Dn1*. After phenotyping the 581 individuals of the Tugela x Tugela-*Dn1* $F_{3/4}$ mapping population, the segregation analysis revealed a segregation ratio of 1.667:1 confirming previous reports that the *Dn1* gene is a single, dominant gene (data not shown). We, however, observed a segregation distortion between homozygotic resistant (RR) versus heterozygotic resistant (Rr) plants, with the latter occurring at lower than expected numbers. This observation may be due to incorrect scoring of the intermediates, or to the fact that the parental lines used

in the development of the $F_{3/4}$ mapping population were NILs. Phenotypic analyses of the parents, progenitors, selected NILs, and the deletion lines containing *Dn5*, confirmed that all plants, with the exception of Tugela, expressed a resistant phenotype against feeding by *D. noxia* biotype SA1 (Table 3.1).

3.4.2 Genetic map surrounding the *Dn1* resistance gene

The genetic map constructed surrounding *Dn1* located on chromosome 7D covers a distance of 122.8 cM (Figure 3.1). The map contains 38 AFLP loci, one microsatellite (*Xgwm111*) and two EST markers (RGA2-29_30 and SSH-RGA2) (Table 3.2). None of the other SSR or EST markers screened showed any polymorphism and therefore, were not included in the construction of the genetic map. The closest marker, AFLP marker E-ACT/M-CTG_0270.84, mapped 3.5 cM from *Dn1*, followed by EST markers RGA2-29_30 and SSH-RGA2 that mapped 15.3 cM and 15.9 cM, respectively from *Dn1*. Microsatellite marker *Xgwm111*, that features on the CS deletion physical map, mapped 82 cM from *Dn1*.

After analysis of the resistant and susceptible samples from the Tugela x Tugela-*Dn1* $F_{3/4}$ mapping population with EST marker RGA2-29_30, all resistant samples shared the same 200 bp and 230 bp PCR amplicons that is absent in the susceptible samples (Figure 3.2). While analysis with *Xgwm111* revealed shared loci (i.e., 210 bp, 240 bp and 250 bp) between the resistant samples from the Tugela x Tugela-*Dn1* $F_{3/4}$ mapping population, Tugela-*Dn1*, SA1684 (*Dn1* progenitor), and deletion line W237, Chinese Spring 7DS dt (40 + $2t^{7DS}$). The shared loci of 210 bp, 240 bp and 250 bp were also present in Tugela-*Dn5*, Palmiet-*Dn5* and SA463 (*Dn5* progenitor). These loci were absent in the susceptible Tugela, the susceptible samples from the Tugela x Tugela-*Dn1* $F_{3/4}$ mapping population and

samples of the deletion line W1318/W1376, Chinese Spring 7DL dt (Figure 3.3).

3.4.3 Positioning *Dn1* on chromosome 7D using deletion lines

Since *Dn1*, *Dn2* and *Dn5* were previously reported to be allelic and within the same gene cluster (Liu *et al.*, 2005), deletion lines, parental lines, progenitors and NILs containing these genes were analyzed for co-segregation with the *Dn1* phenotype and individuals of the Tugela x Tugela-*Dn1* $F_{3/4}$ mapping population. For this purpose samples derived from the Tugela x Tugela-*Dn1* progeny were analyzed with the microsatellite marker *Xgwm111* on the constructed map (Figure 3.1 and Figure 3.3). Resistant samples from the Tugela x Tugela-*Dn1* $F_{3/4}$ mapping population shared loci 210 bp, 240 bp, and 250 bp with Tugela-*Dn1*, SA1684, SA463, Tugela-*Dn5*, Palmiet-*Dn5*, deletion line W237, Chinese Spring 7DS dt ($40 + 2t^{7DS}$), but not with the susceptible Tugela and susceptible samples from the Tugela x Tugela-*Dn1* $F_{3/4}$ mapping population and SA2199 (*Dn2* progenitor), after analyses with marker *Xgwm111* (Figure 3.3). None of these loci were present in samples of the deletion line W1318/W1376, Chinese Spring 7DL dt.

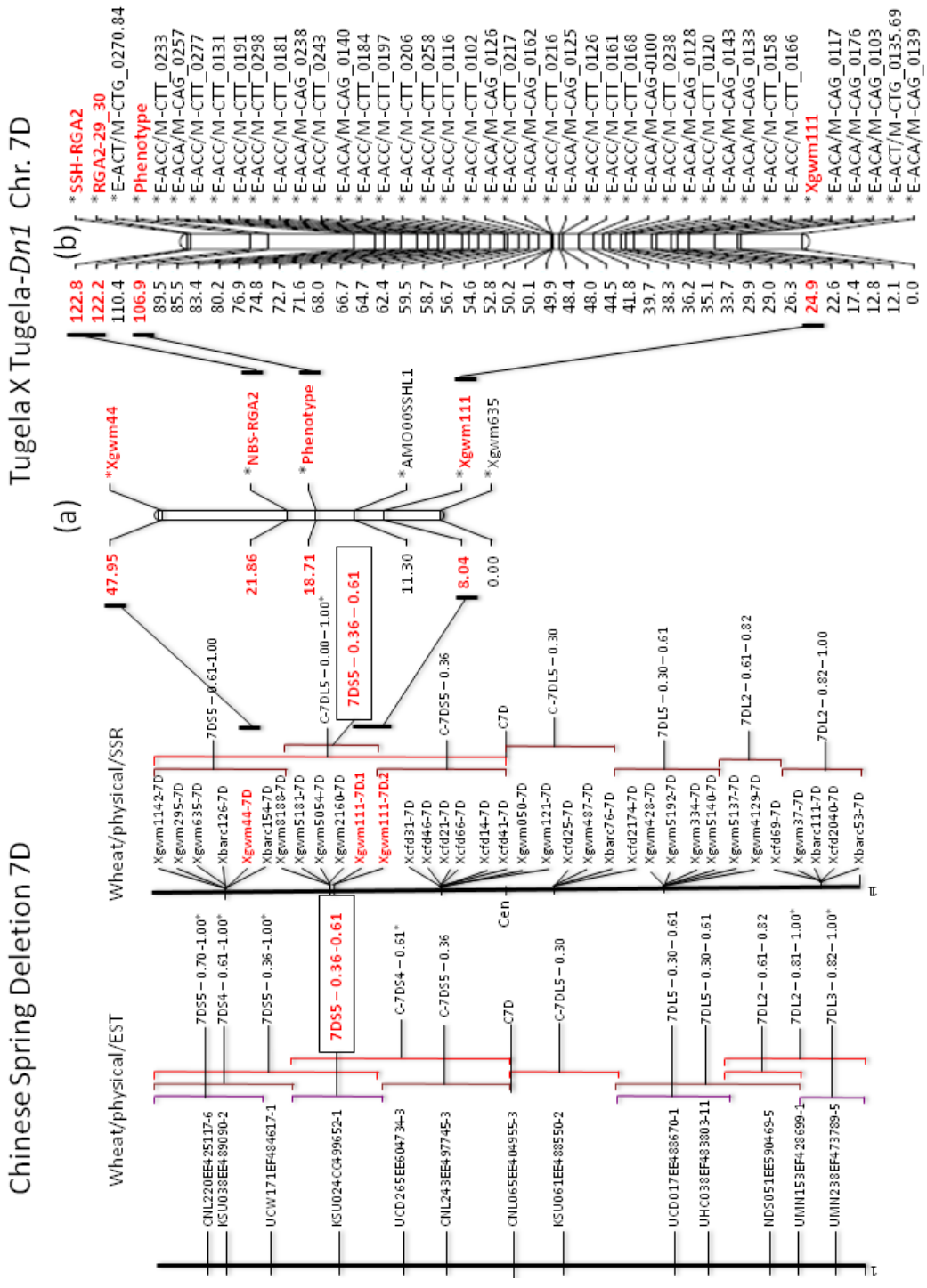


Figure 3.1: Genetic map of *Dn1* on chromosome 7D of *Triticum aestivum* (b) presented with reference to the maps of Sourdille *et al.* (2004) (Wheat/physical/SSR), Graingenes Wheat/physical/EST and Swanepoel *et al.* (2003) (a). Also indicated is the physical map location of marker *Xgwm111* used as anchor for positioning of the *Dn1* gene (http://-wheat.pw.usda.gov/cgibin/cmap/viewer?mapMenu=1&featureMenu=1 &corrMenu=1&displayMenu=1&advancedMenu=1-&ref_map_accs=Chinese_Spring_Deletion_SSR_7D&sub=Draw +Selected+Maps&ref_map_set_acc=Chinese_Spring-_Deletion_SSR_7D&data_source=GrainGenes).

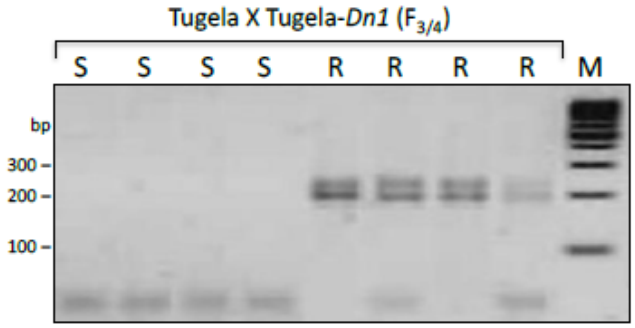


Figure 3.2: Comparison of DNA fragments amplified from DNA of susceptible (S, rr) and resistant (R, RR) samples of the Tugela x Tugela-*Dn1* $F_{3/4}$ population, using primer pair RGA2-29_30. M = 100 bp DNA ladder.

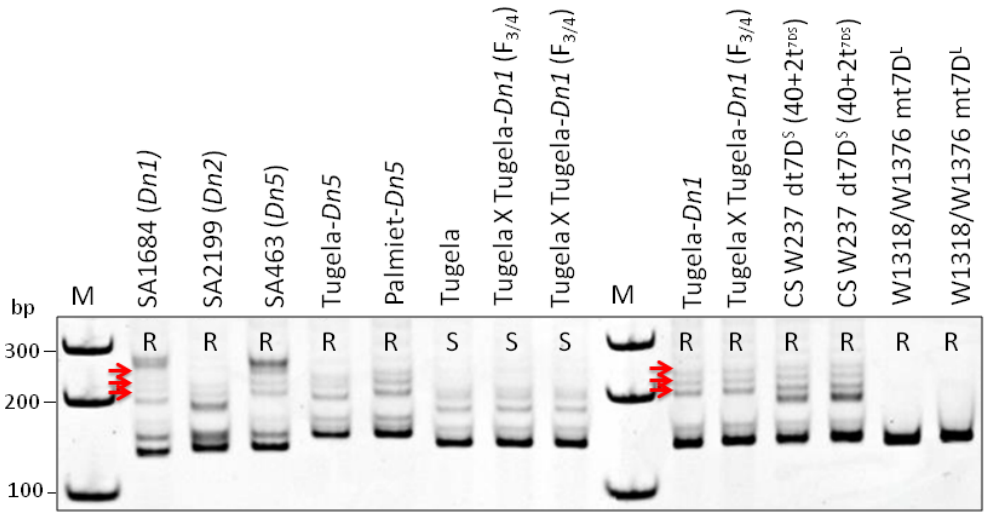


Figure 3.3: Comparison of DNA fragments amplified from DNA of the Tugela x Tugela-*Dn1* $F_{3/4}$ population, *Dn1*, *Dn2* and *Dn5* parental lines, the different *Dn* progenitors and ditelosomic deletion Chinese Spring lines using primer pair *Xgwm111*. M = 100 bp DNA ladder. Red arrows indicate 210 bp, 240 bp and 250 bp shared loci.

Table 3.2: Markers linked to *Dn1* and genetic distances from the gene.

Marker	**Type**	**LOD**	**Distance from *Dn1* (cM)**
RGA2-29_30	EST	5	15.3
SSH-RGA2	EST	5	15.9
Xgwm111	Microsatellite	5	82.0
E-ACA/M-CAG_100	AFLP	5	67.2
E-ACA/M-CAG_103	AFLP	5	94.1
E-ACA/M-CAG_117	AFLP	5	84.3
E-ACA/M-CAG_125	AFLP	5	58.5
E-ACA/M-CAG_126	AFLP	5	54.1
E-ACA/M-CAG_128	AFLP	5	70.7
E-ACA/M-CAG_133	AFLP	5	77.0
E-ACA/M-CAG_139	AFLP	5	106.9
E-ACA/M-CAG_140	AFLP	5	40.2
E-ACA/M-CAG_143	AFLP	5	73.2
E-ACA/M-CAG_162	AFLP	5	56.8
E-ACA/M-CAG_176	AFLP	5	89.5
E-ACA/M-CAG_238	AFLP	5	35.3
E-ACA/M-CAG_257	AFLP	5	21.4
E-ACC/M-CTT_102	AFLP	5	52.3
E-ACC/M-CTT_116	AFLP	5	50.2
E-ACC/M-CTT_120	AFLP	5	71.8
E-ACC/M-CTT_126	AFLP	5	58.9
E-ACC/M-CTT_131	AFLP	5	26.7
E-ACC/M-CTT_158	AFLP	5	77.9
E-ACC/M-CTT_161	AFLP	5	62.4
E-ACC/M-CTT_166	AFLP	5	80.6
E-ACC/M-CTT_168	AFLP	5	65.1
E-ACC/M-CTT_181	AFLP	5	34.2
E-ACC/M-CTT_184	AFLP	5	42.2
E-ACC/M-CTT_191	AFLP	5	30.0
E-ACC/M-CTT_197	AFLP	5	44.5
E-ACC/M-CTT_206	AFLP	5	47.4
E-ACC/M-CTT_216	AFLP	5	57.0
E-ACC/M-CTT_217	AFLP	5	56.7
E-ACC/M-CTT_233	AFLP	5	17.4
E-ACC/M-CTT_238	AFLP	5	68.6
E-ACC/M-CTT_243	AFLP	5	38.9
E-ACC/M-CTT_258	AFLP	5	48.2
E-ACC/M-CTT_277	AFLP	5	23.5
E-ACC/M-CTT_298	AFLP	5	32.1
E-ACT/M-CTG_135.69	AFLP	5	94.8
E-ACT/M-CTG_270.84	AFLP	5	3.5

3.5 Discussion

Wheat has a sizeable, 17 Gbp hexaploid genome, containing between 94,000 and 96,000 genes (Brenchley *et al.*, 2012) and is constituted of the genomes of three diploid progenitor species: *Aegilops tauschii* (D), *Triticum monococcum* (A) and possibly, *A. speltoides* (B) (Gill *et al.*, 2004; Marcussen *et al.*, 2014; Mayer *et al.*, 2014). The wheat genome harbors extensive repetitive elements and genes tend to occur in clusters. One such cluster near the centromere of chromosome 7D contains resistance genes against *D. noxia* and stem rust (Adhikari *et al.*, 2004; Liu *et al.*, 2005).

The genetic map constructed surrounding *Dn1* located on chromosome 7DS covers a distance of 122.8 cM (Figure 3.1). The map contains 38 new AFLP markers, one microsatellite *Xgwm111* and two EST markers (Figure 3.1; Table 3.2). Of all markers, AFLP marker E-ACT/M-CTG_0270.84 showed the closest linkage to *Dn1* and mapped 3.5 cM from the gene. EST markers RGA2-29_30 and SSH-RGA2 mapped 15.3 cm and 15.9 cM, respectively from *Dn1*. EST marker RGA2-29_30 is derived from an Affymetrix probe set, TaAffx.104444.1, that was previously shown to be significantly up-regulated in resistant Tugela NILs after *D. noxia* infestation, but not in the susceptible Tugela NILs (Table A.3; Botha *et al.*, 2014). SSH-RGA2 previously mapped 3.15 cM from *Dn1* (Swanepoel *et al.*, 2003; NBS-RGA2) in a subset of the mapping population that was used in this study. The discrepancy is likely the result of the smaller mapping population compared to the size of the mapping population used in the present study.

In this study, *Xgwm111* mapped 82 cM from *Dn1*, even though this marker previously mapped 3.82 cM (Liu *et al.*, 2001) and 10.67 cM (Swanepoel *et al.*, 2003) from *Dn1*. Microsatellite marker *Xgwm111* was used in this study, since

it provided a means to position *Dn1* on the Chinese Spring deletion 7D physical map. *Xgwm111* was physically mapped to chromosome 7D (Sourdille *et al.*, 2004; Song *et al.*, 2005). Sourdille *et al.* (2004)'s map which is demarcated by deletion bins and breakpoints, placed *Xgwm111* on the short arm of chromosome 7D very near the centromere. This is also in accordance with maps produced in the 90's by Cadalen *et al.* (1997) and an ITMI map by Van Deynze *et al.* (1995) who also placed the microsatellite marker on the short arm very near the centromere. Collectively, the data suggests that the *Dn1* resistance gene is located on chromosome 7DS near the centromere. Comparison to wheat physical maps constructed from EST and SSR markers respectively (wheat/physical/EST and wheat/physical/SSR from GrainGenes; Figure 3.1) suggests that *Dn1* is located on chromosome 7DS near the centromere. Results obtained after amplification of DNA obtained from the Chinese Spring deletion lines further provided support for this observation. The resistant samples from the Tugela x Tugela-*Dn1* $F_{3/4}$ mapping population shared loci 210 bp, 240 bp, and 250 bp with Tugela-*Dn1*, SA1684 (*Dn1* progenitor), and deletion line W237, Chinese Spring 7DS dt (40 + $2t^{7DS}$), but not with the susceptible Tugela, susceptible samples from the Tugela x Tugela-*Dn1* $F_{3/4}$ mapping population and samples of the deletion line W1318/W1376, Chinese Spring 7DL dt.

Marker *Xgwm44* lies in bin 7DS5-0.36-0.62 while *Xgwm111* is placed in between bins 7DS5-0.36 and 7DS5-0.36-0.62 on the physical Chinese Spring deletion 7D map. This marker also features on the Swanepoel *et al.* (2003) map and mapped 29.24 cM from *Dn1*. However, this marker mapped at a distance of 146.2 cM from *Dn1* in this study, and thus was not placed on the map.

Obtaining markers closely linked to the *Dn* resistance gene cluster is compli-

cated by the fact that many of the *Dn* genes are located close to the centromere. This is problematic when mapping by recombination, as there is strong suppression of recombination in the centromeric regions (Gupta *et al.*, 2008). Sequencing data from the IWGSC (Mayer *et al.*, 2014) have shown that these centromeric regions can be as large as 122 Mbp and completely deprived of recombination. In the present study, 38 AFLP markers were positioned on the genetic map, with the closest marker at 3.5 cM from *Dn1*, and a new EST marker for use in MAS i.e., RGA2-29_30 was identified. Genetic mapping in wheat has many tools at its disposal that facilitate the mapping process despite wheat's hexaploid nature, and with the aid of *Xgwm111* and *Xgwm44* we were also able to place the *Dn1* gene near the centromere within bin 7DS5-0.36-0.62 on the Chinese Spring deletion SSR physical map. SNPs were not used as a tool in the study as SNP discovery in wheat used to be slow (Somers *et al.*, 2003). Even though several studies were launched to identify SNPs in wheat associated with particular traits (Blake *et al.*, 2004; Qi *et al.*, 2004), none of these were shown to associate with *D. noxia* resistance. With the publication of the wheat draft genome sequence, 13.3 million SNPs were used to build an ultra dense genetic map (Mayer *et al.*, 2014). With such a large database at hand, the use of SNPs for RWA resistance studies might be more feasible.

3.6 List of References

Adhikari, T.B., Cavaletto, J.R., Dubcovsky, J., Gieco, J.O., Schlatter, A.R. and Goodwin, S.B. (2004). Molecular mapping of the *Stb4* gene for resistance to *Septoria tritici* blotch in wheat. *Phytopathology*, vol. 94, pp. 1198–1206.

Anderson, G.R., Papa, D., Peng, J., Tahir, M. and Lapitan, N.L.V. (2003). Genetic mapping of *Dn7*, a rye gene conferring resistance to the Russian wheat aphid in wheat. *Theoretical and Applied Genetics*, vol. 107, pp. 1297–1303.

Benjamini, Y., Hochberg, Y. and Kling, Y. (1994). False discovery rate controlling procedures for pairwise comparisons. *Unpublished manuscript.*

Blake, N.K., Sherman, J.D., Dvorak, J. and Talbert, L.E. (2004). Genome- specific primer sets for starch biosynthesis genes in wheat. *Theoretical and Applied Genetics*, vol. 109, pp. 1295–1302.

Botha, A.-M. (2013). A coevolutionary conundrum: the arms race between *Diuraphis noxia* (Kurdjumov) a specialist pest and its host *Triticum aestivum* (L.). *Arthropod Plant Interactions*, vol. 7, pp. 359–372.

Botha, A.-M., Van Eck, L., Burger, N.F.V. and Swanevelder, Z.H. (2014). Near-isogenic lines of *Triticum aestivum* with distinct modes of resistance exhibit dissimilar transcriptional regulation during *Diuraphis noxia* feeding. *Biology Open*, vol. 3, pp. 1116–1126.

Brenchley, R., Spanning, M., Pheifer, M., Barker, G.L.A., Amore, R.D., Allen, A.M., Mckenzie, N., Kramer, M., Kerhornou, A., Bolsen, D., Kay, S., Waite, D., Trick, M., Bancroft, I., Gu, Y., Huo, M-C., Sehgal, S., Gill, B., Kainian, S., Anderson, O., Kersey, P., Dvorak, J., Mccombie, W.R., Hall, A., Mayer, K.F.X., Edwards, K.J., Bevan, M.W. and Hall, N. (2012). Analysis of the bread

wheat genome using whole-genome shotgun sequencing. *Nature*, vol. 491, pp. 705–710.

Budak, S., Quisenberry, S.S. and Ni, X. (1999). Comparison of *Diuraphis noxia* resistance in wheat isolines and plant introduction lines. *Entomologia Experimentalis et Applicata*, vol. 92, pp. 157–164.

Cadalen, T., Boeuf, C., Bernard, S. and Bernard, M. (1997). An intervarietal molecular marker map in *Triticum aestivum* L. em. Thell. and comparison with a map from a wide cross. *Theoretical and Applied Genetics*, vol. 94, pp. 367–377.

Du Toit, F. (1987). Resistance in wheat (*Triticum aestivum*) to *Diuraphis noxia* (Homoptera: Aphididae). *Cereal Research Communication*, vol. 15, pp. 175–179.

Du Toit, F. (1988). Another source of Russian wheat aphid (*Diuraphis noxia*) resistance in *Triticum aestivum*. *Cereal Research Communication*, vol. 16, pp. 105–106.

Du Toit, F. (1989). Components of resistance in three bread wheat lines to *Diuraphis noxia* (Homoptera: Aphididae) in South Africa. *Journal of Economic Entomology*, vol. 82, pp. 1779–1781.

Du Toit, F. (1990). Field resistance in three bread wheat lines to the Russian

wheat aphid, *Diuraphis noxia* (Hemiptera: Aphididae). *Crop Protection*, vol. 9, pp. 255–258.

Fouché, A., Verhoeven, R.L., Hewitt, P.H., Walters, M.C., Kriel, C.F. and De Jager, J. (1984). Russian wheat aphid (*Diuraphis noxia*) feeding damage on wheat, related cereals and a Bromus grass species. Technical communication, Department of Agriculture, Johannesburg, Republic of South Africa. In: Walters, M.C. (ed.), *Progress in Russian wheat aphid (Diuraphis noxia Mordw.) research in the Republic of South Africa*, vol. 191, pp. 22–33.

Gill, B.S., Appels, R., Botha-Oberholster, A.-M., Buell, C.R., Bennetzen, J.L., Chalhoub, B., Chumley, F., Dvorak, J., Iwanaga, M., Keller, B., Li, W., Mccombie, W.R., Ogihara, Y., Quetier, F. and Sasaki, T. (2004). A workshop report on wheat genome sequencing: international genome research on wheat consortium. *Genetics*, vol. 168, pp. 1087–1096.

Gupta, P.K., Mir, R.R., Mohan, A. and Kumar, J. (2008). Wheat genomics: present status and future prospects. *International Journal of Plant Genomics*, vol. 2008, pp. 1–36.

Haley, S.D., Peairs, F.B., Walker, C.B., Rudolph, J.B. and Randolph, T.L. (2004). Occurrence of a new Russian wheat aphid biotype in Colorado. *Crop Science*, vol. 44, pp. 1589–1592.

Harvey, T.L. and Martin, T.J. (1990). Resistance to Russian wheat aphid, *Diu-*

raphis noxia, in wheat (*Triticum aestivum*). *Cereal Research Communication*, vol. 18, pp. 127–129.

Heng-Moss, T.M., Ni, X., Macedo, T., Markwell, J.P., Baxendale, F.P., Quisenberry, S.S. and Tolmay, V. (2003). Comparison of chlorophyll and carotenoid concentrations among Russian wheat aphid (Homoptera: Aphididae)-infested wheat isolines. *Journal of Economic Entomology*, vol. 96, pp. 475–481.

Heyns, I., Groenewald, E., Marais, F., Du Toit, F. and Tolmay, V. (2006). Chromosomal location of the Russian wheat aphid resistance gene, *Dn5*. *Crop Science*, vol. 46, pp. 630–636.

Jankielsohn, A. (2011). Distribution and diversity of Russian wheat aphid (Hemiptera: Aphididae) biotypes in South Africa and Lesotho. *Journal of Economic Entomology*, vol. 104, pp. 1736–1741.

Kosambi, D.D. (1943). The estimation of map distances from recombination values. *Annals of Eugenics*, vol. 12, pp. 172–175.

Liu, M., Smith, M. and Gill, S. (2002). Identification of microsatellite markers linked to Russian wheat aphid resistance genes *Dn4* and *Dn6*. *Theoretical and Applied Genetics*, vol. 104, pp. 1042–1048.

Liu, X.M., Smith, C.M., Friebe, B.R. and Gill, B.S. (2005). Molecular mapping and allelic relationships of Russian wheat aphid–resistance genes. *Crop Science*, vol. 45, p. 2273–2280.

Liu, X.M., Smith, C.M., Gill, B.S. and Tolmay, V. (2001). Microsatellite markers linked to six Russian wheat aphid resistance genes in wheat. *Theoretical and Applied Genetics*, vol. 102, pp. 504–510.

Marais, G.F. and Du Toit, F. (1993). A monosomic analysis of Russian wheat aphid resistance in the common wheat PI 294994. *Plant Breeding*, vol. 111, pp. 246–248.

Marais, G.F., Wessels, W.G., Horn, M. and Du Toit, F. (1998). Association of a stem rust resistance gene (*Sr45*) and two Russian wheat aphid resistance genes (*Dn5* and *Dn7*) with mapped structural loci in common wheat. *South African Journal of Plant and Soil*, vol. 15, pp. 67–71.

Marasas, C., Anandajayasekeram, P., Tolmay, V., Martella, D., Purchase, J. and Prinsloo, G. (1997). Socio-economic impact of the Russian wheat aphid control research program. *Southern African Centre for Cooperation in Agricultural and Natural Resources Research and Training*, pp. 1–103. Gaborone, Botswana.

Marcussen, T., Sandve, S.R., Heier, L., Spannagl, M., Pfeifer, M., Jakobsen, K.S., Wulff, B.B.H., Steuernagel, B., Mayer, K.F.X., Olsen, O-A., Rogers, J., Doležel, J., Pozniak, C., Eversole, K., Feuillet, C., Gill, B., Friebe, B.,

Lukaszewski, A.J., Sourdille, P., Endo, T.R., Kubaláková, M., Ihalikova, J., Dubska, Z., Vrána, J., Perkova, R., Imkova, H., Febrer, M., Clissold, L., McLay, K., Singh, K., Chhuneja, P., Singh, N.K., Khurana, J., Akhunov, E., Choulet, F., Alberti, A., Barbe, V., Wincker, P., Kanamori, H., Kobayashi, F., Itoh, T., Matsumoto, T., Sakai, H., Tanaka, T., Wu, J., Ogihara, Y., Handa, H., Maclachlan, P.R., Sharpe, A., Klassen, D., Edwards, D., Batley, J., Lien, S., Caccamo, M., Ayling, S., Ramirez-Gonzalez, R.H., Clavijo, B.J., Wright, J., Martis, M.M., Mascher, M., Chapman, J., Poland, J.A., Scholz, U., Barry, K., Waugh, R., Rokhsar, D.S., Muehlbauer, G.J., Stein, N., Gundlach, H., Zytnicki, M., Jamilloux, V., Quesneville, H., Wicker, T., Faccioli, P., Colaiacovo, M., Stanca, A.M., Budak, H., Cattivelli, L., Glover, N., Pingault, L., Paux, E., Sharma, S., Appels, R., Bellgard, M., Chapman, B., Nussbaumer, T., Bader, K.C., Rimbert, H., Wang, S., Knox, R., Kilian, A., Alaux, M., Alfama, F., Couderc, L., Guilhot, N., Viseux, C., Loaec, M., Keller, B. and Praud, S. (2014). Ancient hybridizations among the ancestral genomes of bread wheat. *Science*, vol. 80, p. 1250092.

Mayer, K.F.X., Rogers, J., Doležel, J., Pozniak, C., Eversole, K., Feuillet, C., Gill, B., Friebe, B., Lukaszewski, A.J., Sourdille, P., Endo, T.R., Kubaláková, M., Ihalikova, J., Dubska, Z., Vrána, J., Perkova, R., Imkova, H., Febrer, M., Clissold, L., McLay, K., Singh, K., Chhuneja, P., Singh, N.K., Khurana, J., Akhunov, E., Choulet, F., Alberti, A., Barbe, V., Wincker, P., Kanamori, H., Kobayashi, F., Itoh, T., Matsumoto, T., Sakai, H., Tanaka, T., Wu, J., Ogihara, Y., Handa, H., Maclachlan, P.R., Sharpe, A., Klassen, D., Edwards, D., Batley, J., Olsen, O.A., Sandve, S.R., Lien, S., Steuernagel, B., Wulff, B.,

Caccamo, M., Ayling, S., Ramirez-Gonzalez, R.H., Clavijo, B.J., Wright, J., Pfeifer, M., Spannagl, M., Martis, M.M., Mascher, M., Chapman, J., Poland, J.A., Scholz, U., Barry, K., Waugh, R., Rokhsar, D.S., Muehlbauer, G.J., Stein, N., Gundlach, H., Zytnicki, M., Jamilloux, V., Quesneville, H., Wicker, T., Faccioli, P., Colaiacovo, M., Stanca, A.M., Budak, H., Cattivelli, L., Glover, N., Pingault, L., Paux, E., Sharma, S., Appels, R., Bellgard, M., Chapman, B., Nussbaumer, T., Bader, K.C., Rimbert, H., Wang, S., Knox, R., Kilian, A., Alaux, M., Alfama, F., Couderc, L., Guilhot, N., Viseux, C., Loaec, M., Keller, B. and Praud, S. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science*, vol. 80, p. 1251788.

Miller, C.A., Altinkut, A. and Lapitan, N.L.V. (2001). A microsatellite marker for tagging *Dn2*, a wheat gene conferring resistance to the Russian wheat aphid. *Crop Science*, vol. 41, pp. 1584–1589.

Morrison, W.P. and Peairs, F.B. (1998). Response model concept and economic impact. In: Quisenberry, S.S. and Peairs, F.B. In: *A response model for an introduced pest – the Russian wheat aphid*, pp. 1–11. Thomas Say Publications in Entomology. Entomological Society of America, Lanham, MD.

Myburg, A.A., Cawood, M., Wingfield, B.D. and Botha, A.-M. (1998). Development of RAPD and SCAR markers linked to the Russian wheat aphid resistance gene *Dn2* in wheat. *Theoretical and Applied Genetics*, vol. 96, pp. 1162–1169.

Myburg, A.A., Remington, D.L., O'Malley, D.M., Sederoff, R.R. and Whetten,

R.W. (2001). High-throughput AFLP analysis using infrared dye-labeled primers and an automated DNA sequencer. *BioTechniques*, vol. 30, pp. 348–357.

Nkongolo, K.K., Quick, J.S., Peairis, F.B. and Meyer, W.L. (1991). Inheritance of resistance of PI 372129 wheat to the Russian wheat aphid. *Crop Science*, vol. 31, pp. 905–907.

Painter, R.H. (1951). Insect resistance in crop plants. vol. 72, p. 481. LWW.

Painter, R.H. (1958). Resistance of plants to insects. *Annual Review in Entomology*, vol. 3, pp. 267–290.

Peng, J., Wang, H., Haley, S.D., Peairs, F.B. and Lapitan, N.L.V. (2007). Molecular mapping of the Russian wheat aphid resistance gene *Dn2414* in wheat. *Crop Science*, vol. 47, p. 2418.

Pestova, E., Ganal, M., Röder, M. (2000). Isolation and mapping of microsatellite markers specific for the D genome of bread wheat. *Genome*, vol. 43, pp. 689–697.

Prinsloo, G., Chen, Y., Giles, K.L. and Greenstone, M.H. (2002). Release and recovery in South Africa of the exotic aphid parasitoid *Aphelinus hordei* Kurdjumov (Hymenptera: Aphelinidae) verified by the polymerase chain reaction. *BioControl*, vol. 47, pp. 127–136.

Qi, L.L., Echalier, B., Chao, S., Lazo, G.R., Butler, G.E., Anderson, O.D., Akhunov, E.D., Dvorák, J., Linkiewicz, A.M., Ratnasiri, A., Dubcovsky, J., Bermudez-Kandianis, C.E., Greene, R.A., Kantety, R., La Rota, C.M., Munkvold, J.D., Sorrells, S.F., Sorrells, M.E., Dilbirligi, M., Sidhu, D., Erayman, M., Randhawa, H.S., Sandhu, D., Bondareva, S.N., Gill, K.S., Mahmoud, A.A., Ma, X-F., Miftahudin, Gustafson, J.P., Conley, E.J., Nduati, V., Gonzalez-Hernandez, J.L., Anderson, J.A., Peng, J.H., Lapitan, N.L.V., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Pathan, M.S., Zhang, D.S., Nguyen, H.T., Choi, D-W., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O. and Gill, B.S. (2004). A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. *Genetics*, vol. 168, pp. 701–712.

Röder, M.S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M-H., Leroy, P. and Ganal, M.W. (1998). A microsatellite map of wheat. *Genetics*, vol. 149, pp. 2007–2023.

Saidi, A. and Quick, J.S. (1996). Inheritance and allelic relationships among Russian wheat aphid resistance genes in winter wheat. *Crop Science*, vol. 36, pp. 256–258.

Šimková, H., Šafář, J., Kubaláková, M., Suchánková, P., Číhalíková, J., Robert-Quatre, H., Azhaguvel, P., Weng, Y., Peng, J., Lapitan, N.L.V., Ma, Y., You, F.M., Luo, M.-C., Bartovs, J. and Doležel, J. (2011). BAC libraries from wheat chromosome 7D: efficient tool for positional cloning of aphid resistance genes. *Journal of Biomedicine & Biotechnology*, vol. 2011, pp. 1–11.

Smith, C.M., Schotzko, D.J., Zemetra, R.S. and Souza, E.J. (1992). Categories of resistance in plant introductions of wheat resistant to the Russian wheat aphid (Homoptera: Aphididae). *Journal of Economic Entomology*, vol. 85, pp. 1480–1484.

Smith, C.M., Belay, T., Stauffer, C., Stary, P., Kubeckova, I. and Starkey, S. (2004). Identification of Russian wheat aphid (Homoptera: Aphididae) populations virulent to the *Dn4* resistance gene. *Journal of Economic Entomology*, vol. 97, pp. 1112–1117.

Smith, C.M. and Clement, S.L. (2012). Molecular bases of plant resistance to arthropods. *Annual Review in Entomology*, vol. 57, pp. 309–328.

Somers, D.J., Kirkpatrick, R., Moniwa, M. and Walsh, A. (2003). Mining single-nucleotide polymorphisms from hexaploid wheat ESTs. *Genome*, vol. 437, pp. 431–437.

Somers, D.J., Isaac, P. and Edwards, K. (2004). A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, vol. 109, pp. 1105–1114.

Song, Q.J., Shi, J.R., Singh, S., Fickus, E.W., Costa, J.M., Lewis, J., Gill, B.S., Ward, R. and Cregan, P.B. (2005). Development and mapping of microsatellite (SSR) markers in wheat. *Theoretical and Applied Genetics*, vol. 110, pp. 550–560.

Sourdille, P., Singh, S., Cadalen, T., Brown-Guedira, G.L., Gay, G., Qi, L., Gill, B.S., Dufour, P., Murigneux, A. and Bernard, M. (2004). Microsatellite-based deletion bin system for the establishment of genetic-physical map relationships in wheat (*Triticum aestivum* L.). *Functional and Integrative Genomics*, vol. 4, pp. 12–25.

Swanepoel, E., Lacock, L., Myburg, A.A. and Botha, A-M. (2003). A leucine rich homolog to *Aegilops tauschii* from breadwheat line PI137739 obtained by suppression subtractive hybridization show linkage to Russian wheat aphid resistance gene *Dn1*, In: *Proceedings of the Tenth International Wheat Genetics Symposium*, vol. 3, pp. 1263–1265. Paestum, Italy.

Tolmay, V.L., Du Toit, F. and Smith, M.C. (2006). Registration of seven Russian wheat aphid resistant near isogenic lines developed in South Africa. *Crop Science*, vol. 46, p. 478.

Unger, L. and Quisenberry, S.S. (1997). Categorization of six wheat plant introduction lines for resistance to the Russian wheat aphid (Homoptera: Aphididae). *Journal of Economic Entomology*, vol. 90, pp. 1408–1413.

Valdez, V.A., Byrne, P.F., Lapitan, N.L.V., Peairs, F.B., Bernardo, A., Bai, G. and Haley, S.D. (2012). Inheritance and genetic mapping of Russian wheat aphid resistance in Iranian wheat landrace accession PI 626580. *Crop Science*, vol. 52, pp. 676–682.

Van Deynze, A.E., Dubcovsky, J., Gill, K.S., Nelson, J.C., Sorrels, M.E., Dvorak, J., Gill, B.S., Lagudah, E.S., McCouch, S.R. and Appels, R. (1995). Molecular-genetic maps for group 1 chromosomes of Triticeae species and their relation to chromosomes in rice and oat. *Genome*, vol. 38, pp. 45–59.

Van Ooijen, J.W. (2011). Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. *Genetics Research*, vol. 93, pp. 343–349.

Venter, E. and Botha, A.-M. (2000). Development of markers linked to *Diuraphis noxia* resistance in wheat using a novel PCR-RFLP approach. *Theoretical and Applied Genetics*, vol. 100, pp. 965–970.

Venter, E., Myburg, A.A. and Botha, A-M. (1998). Characterisation of sequences linked to Russian wheat aphid resistance genes *Dn1*, *Dn2* and *Dn5*. In: *Proceedings of the 9th International Wheat Genetics Symposium*, vol. 3, pp. 332–335. Saskatoon, Saskatchewan, Canada.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van De Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, vol. 23, pp. 4407–4414.

Wang, T., Quisenberry, S.S., Ni, X. and Tolmay, V. (2004). Aphid (Hemiptera: Aphididae) resistance in wheat near-isogenic lines. *Journal of Economic Entomology*, vol. 97, pp. 646–653.

Ward, R., Cregan, P., Song, Q.J., Shi, J.R., Gill, B. and Singh, S. (2003). 544 BARC SSRs from Rick Ward [Online]/ A. Available at http://wheat.pw.usda.gov/ggpages/whatsnew/2003.shtml

Waugh, R., Bonar, N., Baird, E., Thomas, B., Graner, A. and Hayes, P. (1997). Homology of AFLP products in three mapping populations of barley. *Molecular and General Genetics*, vol. 255, pp. 311–321.

Webster, J.A., Starks, K.J. and Burton, R.L. (1987). Plant resistance studies with *Diuraphis noxia* (Homoptera: Aphididae), a new United States wheat pest. *Journal of Economic Entomology*, vol. 80, pp. 944–949.

Weiland, A.A., Peairs, F.B., Randolph, T.L., Rudolph, J.B., Haley, S.D. and Puterka, G.J. (2008). Biotypic diversity in Colorado Russian wheat aphid (Hemiptera: Aphididae) populations. *Journal of Economic Entomology*, vol. 102, pp. 569–574.

Chapter 4

PALM Microbeam and NGS

Next generation sequencing of ditelosomic wheat (*Triticum aestivum* L.) chromosomes, 7DS and 7DL, after isolation using non-contact laser microdissection

Submitted for review to The Plant Genome

4.1 Abstract

The complex hexaploid genome of *Triticum aestivum* L. poses many challenges to NGS and bioinformatic data analyses of its genic content. In the present study, non-contact laser capture microdissection was applied for the isolation of ditelosomic chromosomes 7DS and 7DL from *T. aestivum* cv. Chinese Spring to reduce the redundancy in genic content of allohexaploid wheat for downstream NGS analysis. We found supporting evidence that 7DL is shorter in length than 7DS, but contains more protein coding sequences. Obtained sequences from the ditelosomic chromosomes matched roughly 10% repetitive sequences, and between 309-314 and 390-423 PCG mapped to 7DS dt and 7DL dt, respectively. Through using non-contact laser capture microdissection, a significant reduction ($>$ 80-fold) in genome size was achieved which also yielded a significantly lower proportion of repetitive elements, making in depth genomic studies of wheat more attractive.

4.2 Introduction

Common bread or hexaploid wheat (*Triticum aestivum* L., 2n = 6x = 42, AABBDD) has a large genome at 17 Gbp (Brenchley *et al.*, 2012), compared to the genome of the model monocotyledonous species *Oryza sativa* Japonica which currently stands at 500 Mbp (Itoh *et al.*, 2007). It is generally accepted that modern bread wheat is the result of a polyploidization event of several progenitor species. The first hybridization event occurred between *T. urartu* (AA, 2x) and *Aegilops speltoides* (BB, 2x) and resulted in the formation of *T. turgidum* (AABB, 4x). Thereafter, possibly assisted by human intervention, a hybridization event occurred between *T. turgidum* (AABB, 4x) and *A. tauschii* (DD, 2x) forming modern hexaploid wheat (*T. aestivum*, AABBDD, 6x) (Gill *et al.*, 2004). The fact that the wheat genome contains 42 chromosomes, of which many are similar in size (Mayer *et al.*, 2014), poses many challenges for karyotyping, mapping and gene cloning (Šafář *et al.*, 2004).

Early efforts at sequencing the wheat genome focussed primarily on PCGs which only accounts for approximately 2% of the genome. This resulted in millions of ESTs and full length cDNA sequences (Mochida *et al.*, 2009) being generated, along with many PCGs derived through comparative analysis between assembled wheat genome sequences and that of the draft genome sequences of related species (such as *Brachypodium distachyon* (Vogel *et al.*, 2010), *O. sativa* (Burr, 2005), *Sorghum bicolor* (Paterson *et al.*, 2009), and *Hordeum vulgare* (Mayer *et al.*, 2011)) and assembled RNA-sequensing data (Mayer *et al.*, 2014).

Brenchley *et al.* (2012) identified between 94,000 and 96,000 PCGs in their draft wheat genome which culminated in roughly 60,000 genes to be partitioned between the different sub-genomes. Collectively, these studies revealed that the

A and B sub-genomes are more similar to the D sub-genome, both in sequence identity and gene content, than they are to each other (Mayer *et al.*, 2014). To explain this, Marcussen *et al.* (2014) proposed a homoploid hybrid (hybridization without a change in chromosome number) origin for the D sub-genome lineage.

The wheat genome is known to contain a relatively high (75%) amount of repetitive sequences (Vedel and Delseny, 1987; Wicker *et al.*, 2003) that are interspersed by clustered genic regions and low-copy DNA sequences (Akhunov *et al.*, 2003; Choulet *et al.*, 2010), and it was suggested that these regions are under selective pressure to resist transposon insertion (Akhunov *et al.*, 2003). In a pilot sequencing study of 13 BAC clones of chromosome 3B, Choulet *et al.* (2014) further showed that gene density doubled moving from the centromere to the telomere. As an example, chromosome 3B contains 171 putative resistance genes of which 79% are located in the distal regions of the chromosome that only contains 33% of the entire gene set.

During a shotgun sequencing analysis of the bread wheat genome Brenchley *et al.* (2012) found that 79% of their sequences had matches to the TREP database which mainly consisted of retrotransposons and DNA transposons. In line with this, the IWGSC reported that 81% of the sequence reads that they produced (accounting for 76.6% of all assembled sequences) contained repetitive elements and that there was a bias in the amount of classes of repetitive elements present between the different sub-genomes.

Retroelements (class I) were found to be most abundant in the A genome while DNA transposons (class II) were most prevalent in the D genome (Mayer *et al.*, 2014). Choulet *et al.* (2014) stated that the B genome was shaped by transposable elements which were present before polyploidization of the genome which have,

since then, become less active. They further established that wheat intergenic size ranges are highly variable and although only 29% of intergenic regions are larger than 104 Kilo base pair (Kbp), they account for the majority of the genome, at 81%.

In 2014, a complete wheat genome sequence was released by the IWGSC. The authors used an aneuploid bread wheat line from a ditelosomic Chinese Spring cultivar and separated individual chromosomes through flow cytometry for sequencing. The study revealed the sizes and genic content of all chromosomes and showed that chromosome 7D is approximately 728 Mbp in size (Mayer *et al.*, 2014). Despite the myriad of sequencing data already available on the coding regions of the wheat genome, much remains unclear with regard to gene positions and distribution within the individual chromosomes, as well as the mechanism of evolution during polyploidization of the genome (Mayer *et al.*, 2014).

Therefore the objective in this study was firstly, to test the feasibility of isolating single *T. aestivum* L. cv Chinese Spring ditelosomic chromosomes 7DS and 7DL using non-contact Laser Capture Microdissection (LCM) as a means to reduce the redundancy in the sequence complement of allohexaploid wheat for downstream sequencing; and secondly to assess the quality of the obtained sequence data through bioinformatic analyses.

4.3 Materials and Methods

4.3.1 Plant material

Hexaploid wheat cv. Chinese Spring germplasm of ditelosomic lines, W237, 7DS dt (40 + 2t^{7DS}) and W1318/W1376, 7DL dt was obtained from the Stellenbosch

University germplasm collection. Seeds were germinated in petri dishes on filter paper without vernilization and root tips harvested after three days.

4.3.2 Mitotic metaphase preparations

Metaphase preparations for root squashes were conducted as previously described (Ostergren and Heneen, 1962). Root tips were cut 1 cm above the meristem and placed in distilled water at 4 °C for 29 hours. Water was replaced with a fixative (methanol:propionic acid, 3:1) for five days. After fixation, root tips were rinsed in distilled water for 30 min followed by 1 N HCl hydrolysis at 60 °C for 7 min. Roots were rinsed with distilled water again for 2 min after which they were stained with Feulgen (Lhotka and Davenport, 1949; Ostergren and Heneen, 1962) over night at 4 °C. The Feulgen was replaced with distilled water and roots were rinsed twice for 2 min before being placed in 7.5 mM NaOAc (pH 4.5) for 5 min. The roots were then placed in Pecticlear solution (2.5% Pecticlear prepared fresh in 7.5 mM NaOAc, pH 4.5) for 30 min at 37 °C after which roots were again placed in distilled water.

Root tip meristems were then ready to be cut from the root and placed on microscope slides in a drop of Rosner 1% (w/v) Aceto-carmine (Heyns, 2005). Squashes were made following the technique of Mirzaghaderi (2010). Slides were visualized on an Axio Scope.A1 Light Microscope (Zeiss) under an oil-immersion at 100x magnification to evaluate quality of the metaphase preparations.

4.3.3 Microdissection of 7DL dt and 7DS dt chromosome arms

Cover slips were removed from slides of mitotic metaphase preparations (Scalenghe *et al.*, 1981; Mirzaghaderi, 2010) before excision of ditelosomic chromosomes by submerging them twice in liquid nitrogen. Slides were then placed in absolute EtOH at room temperature with slight agitation until the coverslips floated off. The PALM® Robot-Microbeam system (P.A.L.M. Microlaser Technologies AG, Bernried, Germany) was used to remove the ditelosomic chromosomes from slides via LCM into specialized adhesive caps (AdhesiveCap opaque, Zeiss). Cut energy range was set between 25 and 35 with Laser Pressure Catapult (LPC) at 45. The dot function was used to catapult chromosome fragments into the adhesive cap above the slide. Presence of fragments inside the adhesive cap were directly visualized on the PALM® Robot-Microbeam system.

4.3.4 Next Generation Sequencing

Ditelosomic chromosomes isolated via LCM were amplified using the GenomePlex® Single Cell Whole Genome Amplification Kit (Sigma-Aldrich) according to the manufacturer's instructions. Human gDNA controls included in the kit were diluted to three different concentrations of 5 ng/μl, 0.5 ng/μl and 0.05 ng/μl and amplified along with the ditelosomic wheat chromosomes. Amplified samples were analyzed spectrophotometrically on the NanoDrop 1000 (Thermo Fisher Scientific Inc.) and underwent gel electrophoresis on a 2% (w/v) agarose gel in order to verify the efficacy of the amplification. A further amplification step followed using the TruSeq Nano DNA Sample Prep Kit (Illumina) before sequencing using the Illumina® HiSeq^{TM} 2000 system (Macrogen).

4.3.5 Analysis of sequencing results

A global overview of the in-house pipeline used for analyses of the NGS sequenced data sets are given in Figure 4.1. After trimming the reads, two different mapping strategies were followed. Firstly, the data was *de novo* assembled, whereafter *ab initio* gene calling was conducted. Genes were mapped against the full complement of the IWGSC scaffolds to provide gene identity and to determine the chromosomal location of genes in the wheat genome. The GO of the matched genes was then obtained. Reads were also mapped against the non-redundant TREP database and the trimmed reads were collapsed to assess the level of repetativeness. Secondly, reference mapping against the full complement of the IWGSC features were conducted, providing gene identity of matched genes, the complement of repeats, and RNAs in the data set. The resulting PCG sets were then compared. Reads with no matches to the TREP database and therefore not previously reported as repeats, were obtained.

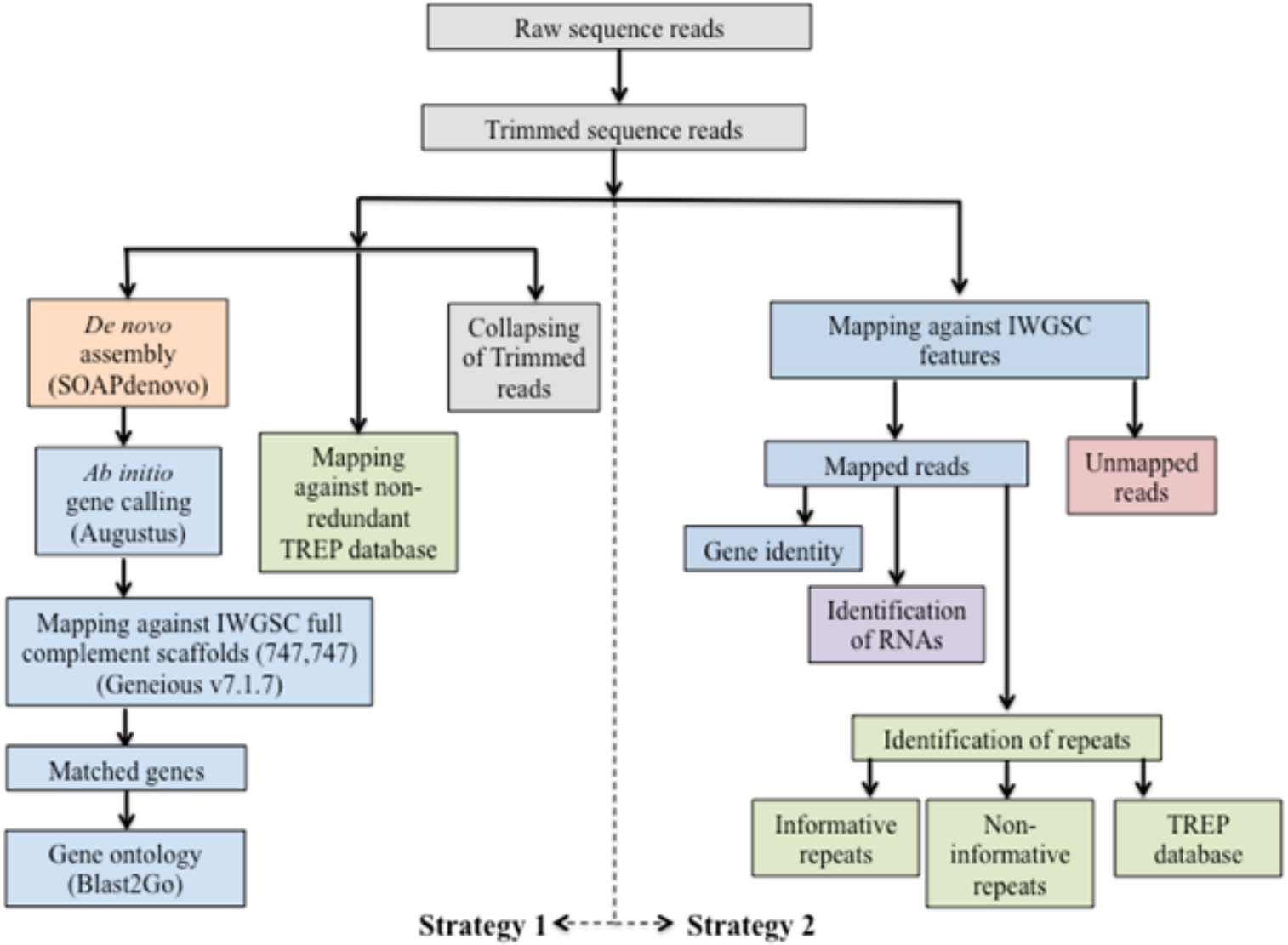


Figure 4.1: Global overview of the pipeline used for analysis of NGS data sets from ditelosomic chromosomes 7DS and 7DL.

Reads obtained from the Illumina® HiSeq^{TM} 2000 platform underwent scoring using the program FASTQC (http://www.bioinformatics.babraham.ac.uk) after which ends of poor quality reads were trimmed and filtered, using Geneious v7.1.7 (Kearse *et al.*, 2012), to obtain a data set where at least 98% of all bases sequenced had a phred score of at least Q20. Sequencing reads from the two data sets were collapsed using the fastx-collapser program included in the FASTX-toolkit suite (http://www.bioinformatics.babraham.ac.uk).

K-mer analysis was performed using Kmergenie (Chikhi and Medvedev, 2013) whilst k-mer counting was performed using DSK software (Rizk *et al.*, 2013). The predicted optimal k-mer value (obtained from Kmergenie) and the abundance of this k-mer size (obtained from DSK) was used in genomics character estimator (GCE, Liu *et al.*, 2013) to predict the genome size of the assembled data set.

Sequences obtained from the ditelosomic 7DS and 7DL chromosomes underwent *de novo* sequence assembly through use of the SOAPdenovo software package (http://soap.genomics.org.cn; Li and Durbin, 2010), which employs a de Bruijn graph algorithm in order to simplify the assembly and computational complexity. After contig assembly, realignment of quality trimmed reads onto contig sequences with paired-end (PE) information was used to link contigs and thus construct scaffolds.

Geneious v7.1.7 (Kearse *et al.*, 2012) was used to perform mapping of generated reads against that of the IWGSC scaffolds and gene set (http://www.ensembl.org, accessed on November 21, 2014). The ITMI TREP database (Wicker *et al.*, 2002; http://wheat.pw.usda.gov/ggpages/Repeats, accessed on November 7, 2014) was used in conjunction with all repeats identified within the IWGSC assembly to map repetitive elements.

4.4 Results

4.4.1 Microdissection of 7DL dt and 7DS dt chromosome arms using LCM-PALM

In order to isolate the Chinese Spring deletion chromosome 7DS and 7DL arms for sequence analysis, flow sorting via flow cytometry was initially attempted, however this proved an arduous task due to the limited size differences of the chromosomes in hexaploid wheat (data not shown). As an alternative, isolation of chromosome arms 7DL dt and 7DS dt using non-contact LCM was attempted. This resulted in the isolation of 25 chromosome arms for 7DS dt and seven chromosome arms for 7DL dt within one week (Figure 4.2A to C). The sizes of the ditelosomic chromosomes were also measured and 7DL dt was found to be shorter than 7DS dt, with the former being approximately 3.0±0.5 μm, and 7DS dt being approximately 4.6±0.5 μm in length (Figure 4.2).

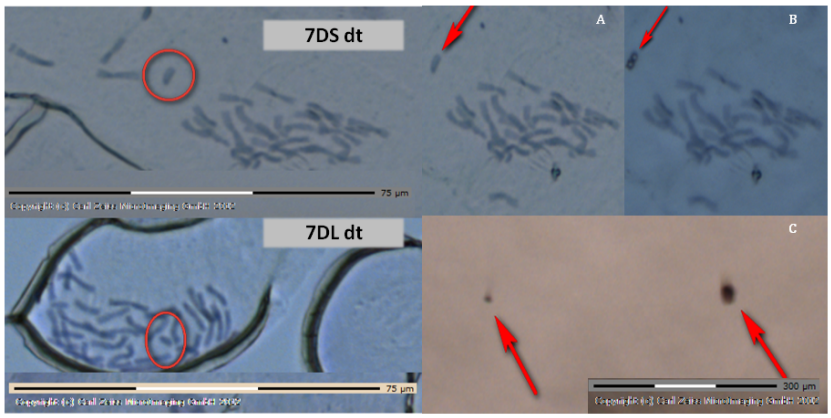


Figure 4.2: Metaphase preparation pre-microdissection showing the size of chromosome 7DS dt and the smaller size of chromosome 7DL dt. Metaphase preparation pre-microdissection (A) and post-microdissection (B) of the 7DS dt chromosome. (C) Cap-check view after microdissection to verify chromosomal fragments caught in the adhesive cap. Red arrows and circles indicate ditelosomic chromosomes.

4.4.2 Next Generation Sequencing

After whole genome amplification of the excised chromosomes, concentrations of both 7DS dt and 7DL dt were above 1,000 ng/μl, however samples contained impurities (i.e., ratios of 260/280 and 260/230 were well below 2, data not shown). This was also the case for the human DNA controls which all yielded similarly high concentrations, despite differences in starting concentration. The amplified samples, however yielded a smeared pattern when separated by agarose gel electrophoresis, with the highest yield around 400 bp, indicating successful amplification of the isolated chromosomes (Figure B.1). Sample concentrations measured with the Perkin Elmer Victor 3 model plate reader (PerkinElmer Inc.) using Quant-iT^{TM} PicoGreen® dsDNA Assay Kit (Life Technologies) yielded concentrations of 3.5 ng/μl and 3.1 ng/μl for 7DS dt and 7DL dt respectively which were insufficient for library construction. Further amplification using the TruSeq Nano DNA Sample Prep Kit yielded DNA concentrations of 89 ng/μl and 87 ng/μl for 7DS dt and 7DL dt respectively, and the library fragment sizes of 470 bp could subsequently be generated for use as template for sequencing using the Illumina® HiSeq^{TM} 2000 platform.

4.4.3 Analysis of sequencing results

Next Generation Sequencing using the Illumina® HiSeq^{TM} 2000 platform was conducted by Macrogen from a PE library with read lengths of 101bp. Sequencing produced 4,690,345,868 bases (46,439,068 reads) for the 7DS dt arm and 5,968,747,106 bases (59,096,506 reads) for the 7DL dt arm (Table 4.1, B.1). After base trimming, a total of 3,583,965,406 bases (35,484,806 reads) were obtained from 7DS dt and 4,515,561,126 bases (44,708,526 reads) from 7DL dt with a Q20 of 98.94% and

98.96% respectively, indicating that good quality sequence data was obtained (Table 4.1, Figure B.2). For both 7DS dt and 7DL dt reads, the percentage of bases called as N were fairly low, ranging from 0.0026 to 0.1704 (Table B.1). The GC content of the sequence data from 7DS dt and 7DL dt were 42.60% and 42.13%, respectively (Table 4.1).

The reads subjected to *de novo* sequence assembly using SOAPdenovo (http://soap.genomics.org.cn/soapdenovo.html) produced 38,409 and 78,861 contigs for 7DS dt and 7DL dt respectively, with the longest contigs being 2,755 (N50 = 215) and 3,755 (N50 = 223), respectively (Table 4.2, B.2, B.3, B.4 and Figure B.3, B.4 and B.5). The GC content of the built contigs from 7DS dt and 7DL dt were 44.85% and 42.70%, respectively (Table B.3; Figure B.6). A total of 37,093 (N50 = 219) and 77,068 (N50 = 227) scaffolds were obtained for the 7DS dt and 7DL dt, with the longest being 6,700 bp (Table B.2, B.3, B.4 and Figure B.7). The GC content of the built scaffolds from 7DS dt and 7DL dt were 44.23% and 42.39%, respectively (Table B.3; Figure B.8). Gap size distribution, an assembly statistic, depicting the gaps in PE reads not covered by sequence data was determined. The gap size distribution differed between scaffolds obtained from 7DS dt and 7DL dt with the latter giving a wider gap size distribution (Figure B.9), while the contig order in built scaffolds is presented in Figure B.10.

Table 4.1: Next generation sequencing platforms applied and number of obtained sequences.

Sequencing Platform	**Chromosome**	**Number of bases (untrimmed)**	**Number of bases (trimmed)**	**Q20 percentage (%)**	**GC content (%)**	**Genome coverage (Mbp)**[1]
Illumina® HiSeq^{TM} 2000	7DS	4,690,345,868	3,583,965,406	98.94	42.60	4.3x (5.5x)[2]
Illumina® HiSeq^{TM} 2000	7DL	5,968,747,106	4,515,561,126	98.96	42.13	5.8x (7.6x)[2]
TOTAL:	(7DS and 7DL)	10,659,092,974	8,099,526,532	98.95	42.37	–

[1] calculated using the predicted genome size (7DS = 381 Mbp; 7DL = 347 Mbp; Table 4.2, Mayer *et al.*, 2014)
[2] raw reads

Table 4.2: Available diploid and hexaploid wheat genome sequence data.

Species	Sequencing platform	Sequence coverage	Total genome size	Protein coding genes (or full length cDNAs)	Full genes (or anchored loci)	Contig number	Contig N50	Scaffold number	Scaffold N50	GC content (%)	Reference
Aegilops tauschii (DD)	Illumina® GAII and HiSeq™ 2000 and Roche/454	90x	4.36 Gbp	43,150	30,697	516,179 (>1 Kbp)	4,512	111,337 (>1 Kbp)	57.6	46% genic; 53% exon and 40% intron	Jia *et al.* (2013)
Triticum urartu (AA)	Illumina® HiSeq™ 2000	–	4.94Gbp	34,879	>18,000	385,430 (>2 Kbp)	3.42 Kbp	–	63.69 Kbp	40%	Ling *et al.* (2013)
Triticum aestivum (AABBDD)											
Chr 7DS dt	Illumina Hiseq2000	4.3 x[1] (5.5 x)[2]	–	309[3]–314[4]	–	38,409	215	37,093	218	42.6%	This study
Chr 7DL dt	Illumina Hiseq2000	5.8 x[1] (7.6 x)[2]	–	390[3]–423[4]	–	78,861	223	77,068	227	42.13%	This study
Chr 3B	Roche/454 paired-end sequencing	5x	886 Mbp	7,264	5,326	546,922 (293,806)	12 Kbp	2,808	892	46.16%	Choulet *et al.* (2014)
Chr 4AS	454 shotgun sequencing	2x	317 Mbp[5]	–	4,383	–	–	–	–	46.4%	Hernandez *et al.* (2012)
Chr 4AL	454 shotgun sequencing	1.7x	–	–	5,188	–	–	–	–	41%	Hernandez *et al.* (2012)
Chr 7DS	Whole genome sequencing after flow sorting	–	381 Mbp	(200)	(1,592)	36,701	5,031	–	–	44.2%	Mayer *et al.* (2014)
Chr 7DL	Whole genome sequencing	–	347 Mbp	(212)	(1,423)	26,737	7,399	–	–	45.4%	Mayer *et al.* (2014)
Chr 7DS	Illumina GAIIx	34x	381 Mbp[5]	–	–	571,038	1,159	–	–	–	Berkman *et al.* (2011)
Chr 7BS	Illumina GAIIx	30x	360 Mbp[5]	–	–	1,038,681	472	–	–	–	Berkman *et al.* (2012)
Whole genome	454 GS FLX Titanium and 454 GS FLX+ (Genomic fragment)	5x	17 Gbp	97,481	94,000 – 96, 000	532, 1847	884	–	–	48.25%	Brenchley *et al.* (2012)
Whole genome	Illumina® HiSeq™ 2000	30–241x	17 Gbp	976, 962	59, 307	1, 813, 412	–	–	–	–	Mayer *et al.* (2014)

– not determined/unknown; [1]calculated using trimmed sequence data and using the predicted genome size (7DS = 381 Mbp; 7DS = 347 Mbp; Mayer *et al.*, 2014); [2]calculated using raw sequence data and using the predicted genome size (7DS = 381 Mbp; 7DS = 347 Mbp; Mayer *et al.*, 2014); [3]calculated based on proteins of >33 amino acids or >100 bp equivalents; [4]calculated based on all possible proteins irrespective of amino acids length; [5]Genome sizes Šafář *et al.* (2010).

In order to estimate the sizes of the ditelosomic chromosomes, the NGS data set was used to construct an abundance histogram and calculate an optimal k-mer value of 17 bp and 19 bp for 7DS dt and 7DL dt respectively (Figure B.11). The optimal k-mer values were then used in GCE (Liu *et al.*, 2013) to calculate the estimated genomic sizes of the ditelosomic chromosomes. GCE makes use of Poisson occupancy models (Lander and Waterman, 1988; Arratia *et al.*, 1996; Wendl, 2006; Hooper *et al.*, 2010) and corrects for highly repetitive genomes like wheat (Vedel and Delseny, 1987; Akhunov *et al.*, 2003; Wicker *et al.*, 2003; Trebbi *et al.*, 2011; Mayer *et al.*, 2014). However, GCE was unable to provide reliable estimations for the genome sizes for 7DS dt and that of 7DL dt. Using the estimated genome sizes obtained from Mayer *et al.* (2014) for 7DS (381 Mbp) and 7DL (347 Mbp) we calculated that there was raw sequence coverage of 5.5 x and 7.6 x using raw sequence data, or 4.3 x and 5.8 x using trimmed sequence data (Table 4.1; Table B.2).

Through the use of BLASTn (Altschul *et al.*, 1990) comparisons made between the 7DS dt and 7DL dt scaffolds and the IWGSC assembly gene set, 309 PCGs for 7DS dt and 423 PCGs for 7DL dt were obtained (Table 4.2), making 7DL dt the more genic rich chromosome. To assign identity to the scaffolds matching PCGs, BLASTx (Altschul *et al.*, 1990) searches were conducted using the Blast2Go (Conesa and Götz, 2008) application with all available protein databases. After identity was assigned through BLASTx, GO mapping commenced to determine the molecular and biological functions of these PCGs along with annotations from the InterproScan (http://www.ebi.ac.uk/Tools/pfa/iprscan) databases (Figure 4.3). The PCG complement was divided into broad categories with the largest belonging to the uncharacterised protein category. The rest of the obtained PCGs

belonged to the cell cycle regulation, protein synthesis, energy metabolism and photosynthesis categories (Figure 4.3). Mapping against the IWGSC scaffolds also allowed for assigning chromosomal locations to mapped PCGs. Interestingly, the PCGs obtained from 7DS dt and 7DL dt mapped to PCGs that occurred on all chromosomes, and not only to those predicted to occur on 7DS and 7DL of wheat (Figure B.12).

To verify the number of PCGs obtained after *de novo* assembly and *ab initio* gene calling, reference feature mapping against the full complement of the IWGSC scaffolds were conducted and 314 PCGs for 7DS dt and 390 PCGs for 7DL dt were obtained (Table 4.2). A total of 7,088,109 reads (715,899,009 bp) from 7DS dt and 9,680,740 reads (977,754,740 bp) from 7DL dt, mapped to the IWGSC scaffolds, while 35,027,786 reads (3,537,806,386 base pairs) from 7DS dt and 28,396,697 reads (2,868,066,397 bp) from 7DL dt did not map to the IWGSC scaffolds. Based on the estimated genome size for 7D (728 Mbp, Mayer *et al.*, 2014), this suggested that we were only able to map approximately 27% of the IWGSC whole genome scaffolds. Since only a fraction of the sequence reads mapped to the IWGSC scaffolds, the repetitive nature of the 7DS dt and 7DL dt data sets were assessed. The reads were also mapped against the non-redundant TREP database and the trimmed reads were collapsed to assess the level of sequence redundancy and a limited level of redundancy was found (Figure B.13).

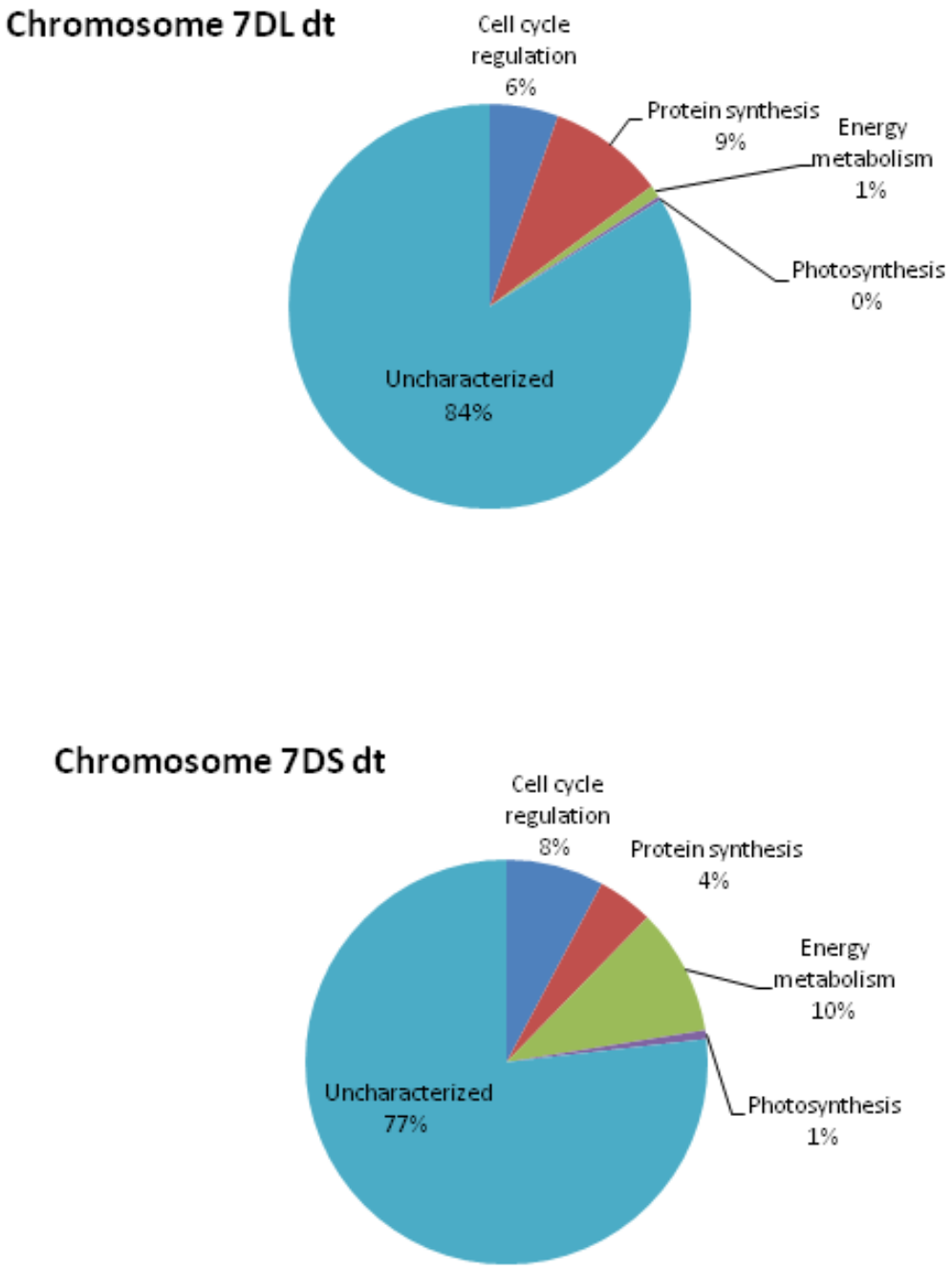


Figure 4.3: Predicted function of the mapped PCGs on ditelosomic chromosomes 7DS and 7DL after Blast2Go analysis.

4.4.4 Repetitive elements

To assess the amount and component of repetitive regions in the 7DS dt and 7DL dt sequence data set, the sequence set was mapped to the TREP database (Table 4.3, B.5) and the IWGSC assembly (Table 4.3, B.6). Mapping against the non-redundant TREP data base produced a total of 2,805 repetitive elements for chromosome 7DS dt and 3, 858 for 7DL dt, while 6,515 repetitive elements were obtained for 7DS dt and 16,078 for 7DL dt after mapping against the IWGSC TREP data set, suggesting more repeats on the smaller ditelosomic chromosome. These elements included LTR groups gypsy and copia; the interspersed repeat subfamily, DNA/CMC-EnSpm; rRNA; the DNA/PIF-Harbinger DNA transposon superfamily, Long Interspersed Nuclear Element/L1; the DNA /TcMar-Stowaway transposon family, Satellites, Transposon free regions, SSRs and DNA transposons (Table B.5 and B.6).

Table 4.3: Number of repetitive elements obtained from the sequence data set after mapping.

Chromosome	Number of repetitive elements[1]					
	Non-redundant TREP data base[1]			IWGSC TREP data set		
	Total	100% match[1]	>80% match[1]	Total	100% match[1]	>80% match[1]
7DS dt	2,805	31	2,785	6,515	139	6,408
7DL dt	3,858	68	3,842	16,078	301	15,838

[1] % of pairwise identity to the reference sequence.

4.4.5 RNA elements

To quantify the complement of RNA-like elements the sequence set was mapped to the IWGSC features (Table 4.4, B.5) and a full complement of RNA elements were obtained. These include several copies of 28S rRNA, 5.8 rRNAs, SSU rRNA eukarya, PK-G12rRNA, and other RNA-like elements. U2 and P27 were found in the data set of chromosome 7DS dt but not in the data set of 7DL dt, while five miRNAs and a snoR71 were found in the data set of 7DL dt but not in 7DS dt.

Chromosome 7DS dt contained 27 tRNA copies (tRNA-Ala (2), tRNA-Arg, tRNA-Asn, tRNA-Asp, tRNA-Cys, tRNA-Glu, tRNA-Gly, tRNA-His, tRNA-Ile (2), tRNA-Leu (2), tRNA-Met (4), tRNA-Phe, tRNA-Ser (5), tRNA-Trp, tRNA-Val (2)), while 7DL dt had only 26 tRNA copies (tRNA-Ala (2), tRNA-Arg (2), tRNA-Asn, tRNA-Glu, tRNA-His, tRNA-Ile (2), tRNA-Leu (3), tRNA-Met (2), tRNA-Phe (2), tRNA-Pro (2), tRNA-Ser (2), tRNA-Thr (3), tRNA-Tyr).

Table 4.4: Number of RNAs, miRNAs and other RNA-like elements obtained after mapping against the IWGSC scaffold data set.

Type of RNA element	**Chromosome**	
	7DS dt	**7DL dt**
28S rRNA	15	26
5.8S rRNA	5	7
miRNA	0	5
tRNAs	27	26
PK-G12rRNA	4	16
snoR71	0	1
SSU rRNA eukarya	29	31
U2	1	0
P27	1	0
Other RNA-like elements	39	29

4.5 Discussion

Kubaláková *et al.* (2002) demonstrated that wheat chromosomes could be flow sorted intact and separated. However, due to the lack of size differences between the different wheat chromosomes, in the past only chromosome 3B could be isolated with ease at high purity. The reason for this being that chromosome 3B is the largest wheat chromosome, estimated at 995 Mbp (Gill *et al.*, 1991; Mayer *et al.*, 2014), and twice the size of the entire rice genome (500 Mbp; Itoh *et al.*, 2007). Besides the limitation in size separation, to obtain high enough concentrations of DNA (i.e. 20-50 ng) from flow sorting for NGS, an estimated 2-11x10^4 isolated chromosomes must be flow sorted (Vrána *et al.*, 2012), making this process extremely laborious and not suitable for high throughput analysis.

Since we were unable to separate the *T. aestivum* L. chromosomes 7DS dt and 7DL dt using flow sorting to high purity, we had to seek an alternative method and applied the non-contact LCM for this purpose. The latter technology circumvents the "similarity of size" problem faced by flow cytometry. Even though microdissection historically involved the technically difficult procedure of dissecting the chromosome with glass microneedles (Scalenghe *et al.*, 1981), this isolation process has been eased with the use of lasers during microdissection using non-contact LCM technology (Olofsson *et al.*, 2012).

Werner *et al.* (1992) stated that chromosome 7DL is the morphologically shorter arm based on homology to 7BS and 7AS (Werner *et al.*, 1992; Liu *et al.*, 2005). This statement was confirmed after comparing the sizes of the 7DS dt and 7DL dt chromosomes on slides of the metaphase preparations. Due to the small size of chromosome 7DL dt non-contact LCM proved difficult, and therefore only a small number of ditelosomic chromosomes could be dissected for 7DL. For

NGS using the Illumina® HiSeq^{TM} 2000 platform as is the case in this study, a minimum of 50 ng starting material is required. Despite the limited number of ditelosomic chromosomes dissected, sufficient DNA template was obtained after additional whole genome amplification for further analysis.

Alhough some NGS data was lost in the filtering of reads, the quality scores of filtered reads were satisfactory, in that close to 99% of all sequences used had a miniPhred value of Q20 (Table B.1). Based on this, it was assumed that the NGS data was of good quality and suitable for further analysis. The obtained contig lengths were, however shorter than expected, although within the read length of the Illumina® HiSeq^{TM} 2000 platform (2x 101 bp). Despite less sample being isolated for 7DL dt, more NGS data was generated for this chromosome arm than for 7DS dt (Table B.1). The reason for the large difference in contig number and bases between chromosomes 7DS dt and 7DL dt is not clear. Due to a small insert size for PE reads, scaffolding was not as successful as with sequencing studies that makes use of multiple insert sizes for PE reads. However, the highly repetitive nature of the wheat genome has been reported as a major stumbling block during contig assembly and scaffolding of the genome (Brenchley *et al.*, 2012; Mayer *et al.*, 2014). Collectively, all these contibuted to the fact that GCE gave no genome size and it is assumed to be due to the low sequencing depth obtained from sequencing such limited material, and not due to the high level of repeats since limited sequence redundancy was obtained after collapsing the PE reads (Figure B.13). Berkman *et al.* (2011), using the Illumina® HiSeq^{TM} 2000, reported an estimated genome size of 381 Mbp for chromosome 7DS. The authors enriched for low copy regions as repetitive elements are difficult to assemble and obtained chromosome coverage of 34 x for 7DS, but despite this, they were only able to obtain 381 PCGs. Mayer *et*

al. (2014), after performing whole genome sequencing reported genome sizes of 381 Mbp and 347 Mbp for 7DS and 7DL respectively, 200 and 212 PCGs respectively, and GC contents of 44.2% and 45.4% respectively for these two chromosomes.

Despite these limitations, and the fact that GCE could not reliably infer genome sizes, we were able to place between 309 to 314 and 390 to 423 PCGs respectively and determine the GC content for 7DS dt (42.80%) and 7DL dt (42.13%) (Table 4.2) which is comparable to that reported by Berkman *et al.* (2011) and Mayer *et al.* (2014) for wheat chromosomes 7DS and 7DL. In 2012, chromosome 4A was sequenced by Hernandez *et al.* after flow sorting and a total of 9,571 PCGs reported for the 856 Mbp chromosome. In contrast, Vitulo *et al.* (2011) sequenced chromosome 5A with an estimated size of 857.8 Mb on the Roche-454 sequencing platform after flow sorting. They found that respectively, only 1.08% and 1.3% of 5AS and 5AL reads represented the coding fraction, with 76.13% of the short arm and 82.23% of the long arm of the chromosome constituting repetitive elements. Mayer *et al.* (2014) also reported a low genic content (i.e., 200 PCGs for chromosome 7DS and 212 PCGs for 7DL) following the same procedure (Table 4.2). Our NGS data provided further support for such low genic coverage (i.e., between 309-314 PCGs for 7DS dt and 390-423 PCGs for 7DL dt), yet we did not nearly match such a high amount of repetitive elements. In our data set, only 6,513 and 17,540 repetitive elements were obtained after mapping against the TREP database, and 6,515 and 16,078 after mapping against the IWGSC TREP features (Table 4.3), representing roughly 10% of our complete data set.

The observed locations of the obtained PCGs, however were an interesting finding, considering that these PCGs were obtained from isolated 7DS dt and 7DL dt chromosomes, and they showed a high degree of similarity to PCGs that are spread

throughout the hexaploid genome and not only to 7DS or 7DL. There are three plausible explanations for this observation. Firstly that the assigned position of these PCGs are incorrect in their placement on the IWGSC scaffolds, and/or that they represent examples of genomic rearrangements (i.e., insertions and translocations of gene blocks), followed by fast replacement rates with repetitive sequences to buffer these changes, to enable for adaptability to diverse environments in a relatively short evolutionary time (Dubcovsky and Dvorak, 2007; Brenchley *et al.*, 2012; Hernandez *et al.*, 2012). Secondly, it could be supportive of the homoploid hybridization origin theory (Hegde and Waines, 2004; Gross and Rieseberg, 2005) of the D genome and thus clarifying the seemingly random mapping. We also found highly repetitive sequences in our NGS data set (Table 4.3, B.6), albeit not nearly as many as previously reported. Lastly, it could show that the reiterative mapping of NGS reads and scaffolds against the IWGSC's scaffolds was incomplete and that these PCGs are the result of an introduced bias due to sample amplification. The latter seemed unlikely since the similarity to the IWGSC gene sequences was highly significant and that the amount of collapsed reads was minimal (Figure B.13).

In conclusion, non-contact LCM was demonstrated to be a feasible alternative to flow sorting during isolation of single *T. aestivum* L. chromosomes 7DS dt and 7DL dt from ditelosomic lines, when we were able to isolate single ditelosomic chromosomes using this method and conduct NGS. Non-contact LCM was shown as a means to reduce the redundancy in repetitive content of allohexaploid wheat for downstream NGS by isolating single ditelosomic *T. aestivum* L. chromosomes 7DS dt and 7DL dt. A significant reduction ($>$ 80-fold) in genome size was achieved making studies in wheat more feasible. As demonstrated, separating chromosome 7DS dt and 7DL dt from the genomic background significantly reduces size and

confounding homologous sequences and enabled NGS and analyses of the single ditelosomic chromosomes.

4.6 List of References

Akhunov, E.D., Akhunova, A.R., Linkiewicz, A.M., Dubcovsky, J., Hummel, D., Lazo, G., Chao, S., Anderson, O.D., David, J., Qi, L., Echalier, B., Gill, B.S., Miftahudin, Gustafson, J.P., La Rota, M., Sorrells, M.E., Zhang, D., Nguyen, H.T., Kalavacharla, V., Hossain, K., Kianian, S., Peng, J., Lapitan, N.L.V., Wennerlind, E.J., Nduati, V., Anderson, J.A., Sidhu, D., Gill, K.S., McGuire, P.E., Qualset, C.O. and Dvorak,J. (2003). Synteny perturbations between wheat homoeologous chromosomes caused by locus duplications and deletions correlate with recombination rates. *Proceedings of the National Academy of Sciences*, vol. 100, pp. 543–549.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, vol. 215, pp. 103–410.

Arratia, R., Martin, D., Reinert, G. and Waterman, M.S. (1996). Poisson process approximation for sequence repeats, and sequencing by hybridization. *Journal of Computational Biology*, vol. 3, pp. 425–463.

Berkman, P.J., Skarshewski, A., Lorenc, M.T., Lai, K., Duran, C., Ling, E.Y.S., Stiller, J., Smits, L., Imelfort, M., Manoli, S., McKenzie, M., Kubaláková, M., Šimková, H., Batley, J., Fleury, D., Doležel, J. and Edwards, D. (2011).

Sequencing and assembly of low copy and genic regions of isolated *Triticum aestivum* chromosome arm 7DS. *Plant Biotechnology Journal*, vol. 9, pp. 768–775.

Berkman, P.J., Skarshewski, A., Manoli, S., Lorenc, M.T., Stiller, J., Smits, L., Lai, K., Campbell, E., Kubaláková, M., Šimková, H., Batley, J., Doležel, J., Hernandez, P. and Edwards, D. (2012). Sequencing wheat chromosome arm 7BS delimits the 7BS/4AL translocation and reveals homoeologous gene conservation. *Theoretical and Applied Genetics*, vol. 124, pp. 423–432.

Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G.L.A., Amore, R.D., Allen, A.M., Mckenzie, N., Kramer, M., Kerhornou, A., Bolser, D., Kay, S., Waite, D., Trick, M., Bancroft, I., Gu, Y., Huo, N., Luo, M., Sehgal, S., Gill, B., Kianian, S., Anderson, O., Kersey, P., Dvorak, J., Mccombie, W.R., Hall, A., Mayer, K.F.X., Edwards, K.J., Bevan, M.W. and Hall, N. (2012). Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature*, vol. 491, pp. 705–710.

Burr, B. (2005). The map-based sequence of the rice genome. *Nature*, vol. 436, pp. 793–800.

Chikhi, R. and Medvedev, P. (2013). Informed and automated k-mer size selection for genome assembly. *Bioinformatics*, p. btt310.

Choulet, F., Wicker, T., Rustenholz, C., Paux, E., Salse, J., Leroy, P., Schlub, S., Le Paslier, M.-C., Magdelenat, G., Gonthier, C., Couloux, A., Budak, H.,

Breen, J., Pumphrey, M., Liu, S., Kong, X., Jia, J., Gut, M., Brunel, D., Anderson, J.A, Gill, B.S., Appels, R., Keller, B. and Feuillet, C. (2010). Megabase level sequencing reveals contrasted organization and evolution patterns of the wheat gene and transposable element spaces. *The Plant Cell*, vol. 22, pp. 1686–1701.

Choulet, F., Alberti, a., Theil, S., Glover, N., Barbe, V., Daron, J., Pingault, L., Sourdille, P., Couloux, A., Paux, E., Leroy, P., Mangenot, S., Guilhot, N., Le Gouis, J., Balfourier, F., Alaux, M., Jamilloux, V., Poulain, J., Durand, C., Bellec, A., Gaspin, C., Safar, J., Doležel, J., Rogers, J., Vandepoele, K., Aury, J.-M., Mayer, K., Berges, H., Quesneville, H., Wincker, P. and Feuillet, C. (2014). Structural and functional partitioning of bread wheat chromosome 3B. *Science*, vol. 345, p. 1249721.

Conesa, A. and Götz, S. (2008). Blast2GO: A comprehensive suite for functional analysis in plant genomics. *International Journal of Plant Genomics*, vol. 2008, pp. 1–12

Dubcovsky, J. and Dvorak, J. (2007). Genome plasticity a key factor of polyploid wheat under domestication. *Science*, vol. 316, pp. 1862–1866.

Gill, B.S., Friebe, B. and Endo, T.R. (1991). Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). *Genome*, vol. 34, pp. 830–839.

Gill, B.S., Appels, R., Botha-Oberholster, A.-M., Buell, C.R., Bennetzen, J.L., Chalhoub, B., Chumley, F., Dvorak, J., Iwanaga, M., Keller, B., Li, W., Mccombie, W.R., Ogihara, Y., Quetier, F. and Sasaki, T. (2004). A workshop report on wheat genome sequencing: international genome research on wheat consortium. *Genetics*, vol. 168, pp. 1087–1096.

Gross, B.L. and Rieseberg, L.H. (2005). The ecological genetics of homoploid hybrid speciation. *Journal of Heredity*, vol. 96, pp. 241–252.

Hegde, S.G. and Waines, J.G. (2004). Hybridization and introgression between bread wheat and wild and weedy relatives in North America. *Crop Science*, vol. 44, pp. 1145–1155.

Heyns, I.C. (2005). *Mapping of chromosome arm 7D of Triticum aestivum L.* Master's thesis, University of Stellenbosch; MSc.

Hernandez, P., Martis, M., Dorado, G., Pfeifer, M., Gálvez, S., Schaaf, S., Jouve, N., Šimková, H., ValÃąrik, M., Doležel, J. and Mayer, K.F.X. (2012). Next-generation sequencing and syntenic integration of flow-sorted arms of wheat chromosome 4A exposes the chromosome structure and gene content. *The Plant Journal*, vol. 69, pp. 377–386.

Hooper, S.D., Dalevi, D., Pati, A., Mavromatis, K., Ivanova, N.N. and Kyrpides, N.C. (2010). Estimating DNA coverage and abundance in metagenomes using gamma approximation. *Bioinformatics*, vol. 26, pp. 295–301.

Itoh, T., Tanaka, T., Barrero, R. A, Yamasaki, C., Fujii, Y., Hilton, P.B., Antonio, B.A, Aono, H., Apweiler, R., Bruskiewich, R., Bureau, T., Burr, F., Costa de Oliveira, A., Fuks, G., Habara, T., Haberer, G., Han, B., Harada, E., Hiraki, A.T., Hirochika, H., Hoen, D., Hokari, H., Hosokawa, S., Hsing, Y., Ikawa, H., Ikeo, K., Imanishi, T., Ito, Y., Jaiswal, P., Kanno, M., Kawahara, Y., Kawamura, T., Kawashima, H., Khurana, J.P., Kikuchi, S., Komatsu, S., Koyanagi, K.O., Kubooka, H., Lieberherr, D., Lin, Y.-C., Lonsdale, D., Matsumoto, T., Matsuya, A., McCombie, W.R., Messing, J., Miyao, A., Mulder, N., Nagamura, Y., Nam, J., Namiki, N., Numa, H., Nurimoto, S., O'Donovan, C., Ohyanagi, H., Okido, T., Oota, S., Osato, N., Palmer, L.E., Quetier, F., Raghuvanshi, S., Saichi, N., Sakai, H., Sakai, Y., Sakata, K., Sakurai, T., Sato, F., Sato, Y., Schoof, H., Seki, M., Shibata, M., Shimizu, Y., Shinozaki, K., Shinso, Y., Singh, N.K., Smith-White, B., Takeda, J., Tanino, M., Tatusova, T., Thongjuea, S., Todokoro, F., Tsugane, M., Tyagi, A.K., Vanavichit, A., Wang, A., Wing, R.A, Yamaguchi, K., Yamamoto, M., Yamamoto, N., Yu, Y., Zhang, H., Zhao, Q., Higo, K., Burr, B., Gojobori, T. and Sasaki, T. (2007). Curated genome annotation of *Oryza sativa* ssp. *japonica* and comparative genome analysis with *Arabidopsis thaliana*. *Genome Research*, vol. 17, pp. 175–183.

Jia, J., Zhao, S., Kong, X., Li, Y., Zhao, G., He, W., Appels, R., Pfeifer, M., Tao, Y., Zhang, X., Jing, R., Zhang, C., Ma, Y., Gao, L., Gao, C., Spannagl, M., Mayer, K.F.X., Li, D., Pan, S., Zheng, F., Hu, Q., Xia, X., Li, J., Liang, Q., Chen, J., Wicker, T., Gou, C., Kuang, H., He, G., Luo, Y., Keller, B., Xia, Q.,

Lu, P., Wang, J., Zou, H., Zhang, R., Xu, J., Gao, J., Middleton, C., Quan, Z., Liu, G., Wang, J., Yang, H., Liu, X., He, Z., Mao, L. and Wang, J. (2013). *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature*, vol. 496, pp. 91–95.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. and Drummond, A. (2012). Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, vol. 28, pp. 1647–1649.

Kubaláková, M., Vrána, J., Cíhalíková, J., Šimková, H. and Doležel, J. (2002). Flow karyotyping and chromosome sorting in bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, vol. 104, pp. 1362–1372.

Lander, E.S. and Waterman, M.S. (1988). Genomic mapping by fingerprinting random clones: a mathematical analysis. *Genomics*, vol. 2, pp. 231–239.

Lhotka, J.F. and Davenport, H.A. (1949). Deterioration of Schiff's reagent. Stain Technologies, vol. 24, pp. 237–239.

Li, H. and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, vol. 26, pp. 589–595.

Ling, H.-Q., Zhao, S., Liu, D., Wang, J., Sun, H., Zhang, C., Fan, H., Li, D.,

Dong, L., Tao, Y., Gao, C., Wu, H., Li, Y., Cui, Y., Guo, X., Zheng, S., Wang, B., Yu, K., Liang, Q., Yang, W., Lou, X., Chen, J., Feng, M., Jian, J., Zhang, X., Luo, G., Jiang, Y., Liu, J., Wang, Z., Sha, Y., Zhang, B., Wu, H., Tang, D., Shen, Q., Xue, P., Zou, S., Wang, X., Liu, X., Wang, F., Yang, Y., An, X., Dong, Z., Zhang, K., Zhang, X., Luo, M.-C., Dvorak, J., Tong, Y., Wang, J., Yang, H., Li, Z., Wang, D., Zhang, A. and Wang, J. (2013). Draft genome of the wheat A-genome progenitor *Triticum urartu*. *Nature*, vol. 496, pp. 87–90.

Liu, B., Shi, Y., Yuan, J., Hu, X., Zhang, H., Li, N., Li, Z., Chen, Y., Mu, D. and Fan, W. (2013). Estimation of genomic characteristics by analyzing k- mer frequency in *de novo* genome projects. *arXiv*, vol. 1308.2012, pp. 1–47.

Liu, X.M., Smith, C.M., Friebe, B.R. and Gill, B.S. (2005). Molecular mapping and allelic relationships of Russian wheat aphid-resistance genes. *Crop Science*, vol. 45, p. 2273.

Marcussen, T., Sandve, S.R., Heier, L., Spannagl, M., Pfeifer, M., Jakobsen, K.S., Wulff, B.B.H., Steuernagel, B., Mayer, K.F.X., Olsen, O.-A., Rogers, J., Doležel, J., Pozniak, C., Eversole, K., Feuillet, C., Gill, B., Friebe, B., Lukaszewski, A.J., Sourdille, P., Endo, T.R., Kubaláková, M., Ihalikova, J., Dubska, Z., Vrána, J., Perkova, R., Imkova, H., Febrer, M., Clissold, L., McLay, K., Singh, K., Chhuneja, P., Singh, N. K., Khurana, J., Akhunov, E., Choulet, F., Alberti, A., Barbe, V., Wincker, P., Kanamori, H., Kobayashi, F., Itoh, T., Matsumoto, T., Sakai, H., Tanaka, T., Wu, J., Ogihara, Y., Handa, H., Maclachlan, P.R., Sharpe, A., Klassen, D., Edwards, D., Batley, J., Lien, S.,

Caccamo, M., Ayling, S., Ramirez-Gonzalez, R.H., Clavijo, B.J., Wright, J., Martis, M.M., Mascher, M., Chapman, J., Poland, J.A., Scholz, U., Barry, K., Waugh, R., Rokhsar, D.S., Muehlbauer, G.J., Stein, N., Gundlach, H., Zytnicki, M., Jamilloux, V., Quesneville, H., Wicker, T., Faccioli, P., Colaiacovo, M., Stanca, A.M., Budak, H., Cattivelli, L., Glover, N., Pingault, L., Paux, E., Sharma, S., Appels, R., Bellgard, M., Chapman, B., Nussbaumer, T., Bader, K.C., Rimbert, H., Wang, S., Knox, R., Kilian, A., Alaux, M., Alfama, F., Couderc, L., Guilhot, N., Viseux, C., Loaec, M., Keller, B. and Praud, S. (2014). Ancient hybridizations among the ancestral genomes of bread wheat. *Science*, vol. 345, pp. 1250092.

Mayer, K.F.X., Martis, M., Hedley, P.E., Šimková, H., Liu, H., Morris, J.A., Steuernagel, B., Taudien, S., Roessner, S., Gundlach, H., Kubala, M., Sucha, P., Murat, F., Felder, M., Nussbaumer, T., Graner, A. and Salse, J. (2011). Unlocking the barley genome by chromosomal and comparative genomics. *The Plant Cell*, vol. 23, pp. 1249–1263.

Mayer, K.F.X., Rogers, J., Doležel, J., Pozniak, C., Eversole, K., Feuillet, C., Gill, B., Friebe, B., Lukaszewski, A.J., Sourdille, P., Endo, T.R., Kubaláková, M., Ihalikova, J., Dubska, Z., Vrána, J., Perkova, R., Imkova, H., Febrer, M., Clissold, L., McLay, K., Singh, K., Chhuneja, P., Singh, N.K., Khurana, J., Akhunov, E., Choulet, F., Alberti, A., Barbe, V., Wincker, P., Kanamori, H., Kobayashi, F., Itoh, T., Matsumoto, T., Sakai, H., Tanaka, T., Wu, J., Ogihara, Y., Handa, H., Maclachlan, P.R., Sharpe, A., Klassen, D., Edwards, D., Batley, J., Olsen, O.-A., Sandve, S.R., Lien, S., Steuernagel, B., Wulff, B.,

Caccamo, M., Ayling, S., Ramirez-Gonzalez, R.H., Clavijo, B.J., Wright, J., Pfeifer, M., Spannagl, M., Martis, M. M., Mascher, M., Chapman, J., Poland, J.A., Scholz, U., Barry, K., Waugh, R., Rokhsar, D.S., Muehlbauer, G.J., Stein, N., Gundlach, H., Zytnicki, M., Jamilloux, V., Quesneville, H., Wicker, T., Faccioli, P., Colaiacovo, M., Stanca, A.M., Budak, H., Cattivelli, L., Glover, N., Pingault, L., Paux, E., Sharma, S., Appels, R., Bellgard, M., Chapman, B., Nussbaumer, T., Bader, K. C., Rimbert, H., Wang, S., Knox, R., Kilian, A., Alaux, M., Alfama, F., Couderc, L., Guilhot, N., Viseux, C., Loaec, M., Keller, B., Praud, S. and IWGSC (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science*, vol. 345, pp. 1251788.

Mirzaghaderi, G. (2010). A simple metaphase chromosome preparation from meristematic root tip cells of wheat for karyotyping or in situ hybridization. *African Journal of Biotechnology*, vol. 9, pp. 314–318.

Mochida, K., Yoshida, T., Sakurai, T., Ogihara, Y. and Shinozaki, K. (2009). TriFLDB: A database of clustered full-length coding sequences from Triticeae with applications to comparative grass genomics. *Plant Physiology*, vol. 150, pp. 1135–1146.

Olofsson, L., Lundgren, A. and Brodelius, P.E. (2012). Trichome isolation with and without fixation using laser microdissection and pressure catapulting followed by RNA amplification: expression of genes of terpene metabolism in apical and sub-apical trichome cells of *Artemisia annua* L. *Plant Science*, vol. 183, pp. 9–13.

Ostergren, G. and Heneen, W.K. (1962). A squash technique for chromosome morphological studies. *Hereditas*, vol. 48, pp. 1–12.

Paterson, A.H., Bowers, J.E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellsten, U., Mitros, T., Poliakov, A., Schmutz, J., Spannagl, M., Tang, H., Wang, X., Wicker, T., Bharti, A. K., Chapman, J., Feltus, F.A., Gowik, U., Grigoriev, I.V., Lyons, E., Maher, C. A., Martis, M., Narechania, A., Otillar, R.P., Penning, B.W., Salamov, A.A., Wang, Y., Zhang, L., Carpita, N.C., Freeling, M., Gingle, A.R., Hash, C.T., Keller, B., Klein, P., Kresovich, S., McCann, M.C., Ming, R., Peterson, D.G., Ur-Rahman, M., Ware, D., Westhoff, P., Mayer, K.F.X., Messing, J. and Rokhsar, D.S. (2009). The *Sorghum bicolor* genome and the diversification of grasses. *Nature*, vol. 457, pp. 551–556.

Rizk, G., Lavenier, D. and Chikhi, R. (2013). DSK: k-mer counting with very low memory usage. *Bioinformatics*, vol. 29, pp. 652–653.

Šafář, J., Bartos, J., Janda, J., Bellec, A., Kubaláková, M., Valarik, M., Pateyron, S., Weiserova, J., Tuskova, R., Číhalíkova, J., Vrána, J., Šimková, H.,

Faivre-Rampant, P., Sourdille, P., Caboche, M., Bernard, M., Doležel, J. and Chalhoub, B. (2004). Dissecting large and complex genomes: flow sorting and BAC cloning of individual chromosomes from bread wheat. *The Plant Journal*, vol. 39, pp. 960–968.

Šafář, J., Šimková, H., Kubaláková, M., Hihalikova, J., Suchánková, P., Bartoš, J. and Doležel, J. (2010). Development of chromosome-specific BAC resources for genomics of bread wheat. *Cytogenetic Genome Research*, vol. 129, pp. 211–223.

Scalenghe, F., Turco, E., Edstrom, J.E., Pirrotta, V. and Melli, M. (1981). Microdissection and cloning of DNA from a specific region of *Drosophila melanogaster* polytene chromosomes. *Chromosoma*, vol. 82, pp. 205–216.

Trebbi, D., Maccaferri, M., De Heer, P., Sørensen, A., Giuliani, S., Salvi, S., Sanguineti, M.C., Massi, A., Van Der Vossen, E.A.G. and Tuberosa, R. (2011). High-throughput SNP discovery and genotyping in durum wheat (*Triticum durum* Desf.). *Theoretical and Applied Genetics*, vol. 123, pp. 555–569.

Vedel, F. and Delseny, M. (1987). Repetitivity and variability of higher-plant genomes. *Plant Physiology and Biochemistry*, vol. 25, pp. 191–210.

Vitulo, N., Albiero, A., Forcato, C., Campagna, D., Dal Pero, F., Bagnaresi, P., Colaiacovo, M., Faccioli, P., Lamontanara, A., Šimková, H., Kubaláková, M., Perrotta, G., Facella, P., Lopez, L., Pietrella, M., Gianese, G., Doležel, J.,

Giuliano, G., Cattivelli, L., Valle, G. and Stanca, A.M. (2011). First survey of the wheat chromosome 5A composition through a next generation sequencing approach. *PloSone*, vol. 6, pp. e26421.

Vogel, J.P., Garvin, D.F., Mockler, T.C., Schmutz, J., Rokshar, D. and Bevan, M.W. (2010). Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature*, vol. 463, pp. 763–768.

Vrána, J., Šimková, H., Kubaláková, M., Číhalíková, J. and Doležel, J. (2012). Flow cytometric chromosome sorting in plants: the next generation. *Methods*, vol. 57, pp. 331–337.

Wendl, M. (2006). Occupancy modeling of coverage distribution in the whole genome shotgun DNA sequencing. *Bulletin of Mathematical Biology*, vol. 68, pp. 179–196.

Werner, J.E., Endo, T.R. and Gill, B.S. (1992). Toward a cytogenetically based physical map of the wheat genome. *Proceedings of the National Academy of Sciences*, vol. 89, pp. 11307–11311.

Wicker, T., Matthews, D.E. and Keller, B. (2002). TREP: a database for Triticeae repetitive elements. *Proceedings of the National Academy of Sciences*, vol. 89, pp. 11307–11311.

Wicker, T., Yahyaoui, N., Gyuot, R., Schlagenhauf, E., Liu, Z.-D., Dubkovsky, J.

and Keller, B. (2003). Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and Am genomes of wheat. *Plant Cell*, vol. 15, pp. 1186–1197.

Chapter 5

Conclusions

The arms race between plant and pest never ceases and in the interaction between *Diuraphis noxia* and wheat, new biotypes are identified periodically which are virulent to many of the 14 *Dn* resistance genes identified to date. Faster and more efficient wheat breeding approaches are needed to keep ahead of developing virulence and to allow the pyramiding of different resistance genes into single cultivars. Conventionally, physical traits such as resistance need to be identified in plantlets through phenotypic screening for the trait - a process that is tedious and time consuming. Using genetic markers (MAS) linked to these physical traits negates the need for physical screening and speeds up the process of identifying suitable parent plants in breeding programs. However, screening for the actual gene of interest would be even faster and more accurate if MAS can be conducted using the gene of interest. This will negate the many discrepancies surrounding marker data in wheat and the location of the *Dn* resistance genes.

The aim of this study was to map the *Dn1* resistance gene from hexaploid wheat thereby confirming its exact location in the wheat genome. For the first objective a saturated genetic map in the region of *Dn1* was constructed to, firstly,

confirm its chromosomal location (7DS or 7DL) and secondly, to confirm the close linkage of reported markers to *Dn1*. Using a $F_{3/4}$ Tugela x Tugela-*Dn1* segregating population consisting of 581 individuals, and Chinese Spring deletion lines a 122.8 cM genetic map was constructed of the *Dn1* resistance gene which was placed on chromosome 7DS near the centromere, in bin 7DS5-0.36-0.62 proportional to the physical Chinese Spring deletion 7D map. The map contained 38 new AFLP markers, one microsatellite *Xgwm111* and two EST markers. Of all the markers, AFLP E-ACT/M-CTG_0270.84 showed the closest linkage to *Dn1* and mapped 3.5 cM from the gene, while EST markers RGA2-29_30 and SSH-RGA2 mapped respectively at 15.3 cM and 15.9 cM, from *Dn1*.

Cloning resistance genes and sequence characterization in wheat is challenging. Due to the large size and ploidy of the genome novel approaches for removing some of the redundancy, prior to sequencing need to be consideded. Strategies such as sequencing from diploid progenitor species and flow cytometry have been employed in many studies.

In order to accomplish the second objective of sequencing chromosome 7D, a novel approach of non-contact laser capture microdissection was applied for the isolation of ditelosomic chromosomes 7DS and 7DL from *Triticum aestivum* cv Chinese Spring. This served to reduce the redundancy of allohexaploid wheat for downstream NGS analysis. Supporting evidence was found confirming that 7DL is shorter in length than 7DS, but contains more protein coding sequences. Sequences obtained from sequencing the chromosome arms matched approximately 10% repetitive sequences and resulted in 309 and 423 PCG obtained from 7DS dt and 7DL dt, respectively. Through using non-contact laser capture microdissection, a significant reduction ($>$ 80-fold) in genome size was achieved which also

yielded a significantly lower proportion of repetitive elements making in-depth genomic studies of wheat more attractive.

One limitation brought about by Microdissection might be the small amount of starting material. Techniques such as flow cytometry typically yield chromosome numbers of 10^4 which is not a feasible number for isolation by microdissection. However, this study shows that informative sequencing data can be obtained from as little as 7 chromosomes. This study, therefore, showed the model of microdissection before sequencing to hold many benefits. Future work will build on this strategy by increasing the number of chromosomes used and sequencing target-specific wheat lines which will pave the way for cloning of resistance genes by comparaitive analysis with sequencing data from this project. It is anticipated that a substractive approach will yield sequences unique to RWA resistant wheat lines and in so doing, further narrow the search for these resistance genes.

In summary, benefits from this study included a saturated genetic map of the short arm of chromosome 7D, a possible marker closely linked to the *Dn* resistance gene for use in MAS and the development and application of a novel technique aimed at reducing the redundancy in the wheat genome for more targeted sequencing.

Appendices

Appendix A

Genetic Mapping

Table A.1: Plant symptoms after infestation with *D. noxia* biotype SA1. Virulence scores and final virulence profiles of a segregating population calculated according to Weiland *et al.* (2008), where resistant (RR, homozygote resistant) = 1-3; intermediate (Rr, heterozygote resistant) = 4-6; and susceptible (rr, homozygote susceptible) = 7-10.

Genotype	**Sample number**	**Virulence score**	**Virulence profile**
Tugela X Tugela-Dn1 ($F_{3/4}$)	1	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	2	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	3	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	4	8	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	5	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	6	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	7	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	8	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	9	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	10	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	11	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	12	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	13	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	14	9	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	15	9	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	16	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	17	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	18	10	rr

Tugela X Tugela-Dn1 ($F_{3/4}$)	19	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	20	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	21	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	22	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	23	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	24	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	25	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	26	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	27	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	28	8	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	29	9	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	30	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	31	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	32	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	33	3	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	34	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	35	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	36	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	37	4	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	38	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	39	9	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	40	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	41	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	42	1	RR

Tugela X Tugela-Dn1 ($F_{3/4}$)	43	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	44	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	45	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	46	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	47	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	48	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	49	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	50	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	51	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	52	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	53	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	54	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	55	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	56	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	57	8	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	58	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	59	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	60	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	61	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	62	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	63	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	64	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	65	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	66	10	rr

Tugela X Tugela-Dn1 ($F_{3/4}$)	67	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	68	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	69	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	70	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	71	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	72	8	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	73	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	74	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	75	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	76	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	77	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	78	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	79	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	80	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	81	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	82	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	83	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	84	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	85	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	86	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	87	4	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	88	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	89	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	90	1	RR

Tugela X Tugela-Dn1 ($F_{3/4}$)	91	3	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	92	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	93	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	94	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	95	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	96	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	97	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	98	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	99	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	100	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	101	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	102	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	103	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	104	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	105	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	106	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	107	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	108	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	109	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	110	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	111	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	112	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	113	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	114	1	RR

Tugela X Tugela-Dn1 ($F_{3/4}$)	115	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	116	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	117	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	118	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	119	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	120	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	121	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	122	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	123	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	124	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	125	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	126	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	127	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	128	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	129	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	130	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	131	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	132	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	133	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	134	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	135	6	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	136	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	137	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	138	10	rr

Tugela X Tugela-Dn1 ($F_{3/4}$)	139	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	140	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	141	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	142	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	143	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	144	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	145	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	146	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	147	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	148	9	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	149	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	150	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	151	8	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	152	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	153	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	154	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	155	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	156	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	157	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	158	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	159	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	160	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	161	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	162	10	rr

Tugela X Tugela-Dn1 ($F_{3/4}$)	163	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	164	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	165	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	166	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	167	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	168	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	169	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	170	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	171	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	172	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	173	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	174	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	175	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	176	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	177	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	178	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	179	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	180	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	181	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	182	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	183	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	184	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	185	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	186	10	rr

Tugela X Tugela-Dn1 ($F_{3/4}$)	187	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	188	9	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	189	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	190	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	191	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	192	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	193	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	194	9	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	195	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	196	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	197	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	198	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	199	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	200	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	201	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	202	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	203	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	204	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	205	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	206	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	207	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	208	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	209	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	210	10	rr

Tugela X Tugela-Dn1 ($F_{3/4}$)	211	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	212	8	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	213	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	214	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	215	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	216	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	217	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	218	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	219	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	220	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	221	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	222	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	223	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	224	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	225	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	226	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	227	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	228	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	229	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	230	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	231	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	232	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	233	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	234	7	rr

Tugela X Tugela-Dn1 ($F_{3/4}$)	235	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	236	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	237	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	238	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	239	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	240	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	241	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	242	8	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	243	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	244	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	245	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	246	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	247	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	248	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	249	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	250	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	251	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	252	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	253	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	254	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	255	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	256	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	257	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	258	10	rr

Tugela X Tugela-Dn1 ($F_{3/4}$)	259	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	260	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	261	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	262	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	263	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	264	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	265	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	266	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	267	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	268	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	269	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	270	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	271	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	272	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	273	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	274	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	275	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	276	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	277	9	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	278	6	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	279	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	280	8	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	281	3	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	282	6	Rr

Tugela X Tugela-Dn1 ($F_{3/4}$)	283	6	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	284	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	285	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	286	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	287	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	288	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	289	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	290	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	291	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	292	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	293	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	294	8	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	295	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	296	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	297	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	298	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	299	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	300	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	301	4	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	302	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	303	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	304	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	305	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	306	6	Rr

Tugela X Tugela-Dn1 ($F_{3/4}$)	307	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	308	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	309	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	310	6	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	311	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	312	9	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	313	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	314	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	315	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	316	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	317	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	318	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	319	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	320	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	321	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	322	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	323	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	324	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	325	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	326	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	327	8	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	328	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	329	4	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	330	2	RR

Tugela X Tugela-Dn1 ($F_{3/4}$)	331	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	332	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	333	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	334	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	335	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	336	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	337	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	338	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	339	6	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	340	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	341	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	342	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	343	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	344	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	345	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	346	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	347	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	348	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	349	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	350	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	351	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	352	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	353	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	354	1	RR

Tugela X Tugela-Dn1 ($F_{3/4}$)	355	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	356	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	357	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	358	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	359	6	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	360	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	361	8	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	362	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	363	3	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	364	6	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	365	6	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	366	8	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	367	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	368	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	369	3	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	370	9	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	371	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	372	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	373	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	374	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	375	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	376	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	377	3	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	378	1	RR

Tugela X Tugela-Dn1 ($F_{3/4}$)	379	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	380	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	381	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	382	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	383	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	384	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	385	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	386	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	387	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	388	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	389	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	390	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	391	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	392	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	393	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	394	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	395	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	396	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	397	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	398	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	399	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	400	8	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	401	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	402	5	Rr

Tugela X Tugela-Dn1 ($F_{3/4}$)	403	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	404	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	405	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	406	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	407	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	408	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	409	6	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	410	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	411	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	412	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	413	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	414	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	415	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	416	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	417	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	418	6	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	419	6	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	420	4	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	421	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	422	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	423	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	424	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	425	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	426	1	RR

Tugela X Tugela-Dn1 ($F_{3/4}$)	427	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	428	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	429	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	430	3	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	431	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	432	7	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	433	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	434	3	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	435	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	436	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	437	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	438	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	439	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	440	4	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	441	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	442	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	443	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	444	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	445	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	446	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	447	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	448	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	449	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	450	2	RR

Tugela X Tugela-Dn1 ($F_{3/4}$)	451	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	452	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	453	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	454	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	455	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	456	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	457	3	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	458	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	459	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	460	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	461	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	462	3	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	463	4	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	464	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	465	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	466	10	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	467	2	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	468	9	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	469	8	rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	470	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	471	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	472	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	473	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	474	1	RR

Tugela X Tugela-Dn1 ($F_{3/4}$)	475	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	476	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	477	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	478	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	479	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	480	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	481	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	482	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	483	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	484	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	485	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	486	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	487	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	488	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	489	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	490	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	491	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	492	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	493	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	494	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	495	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	496	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	497	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	498	1	RR

Tugela X Tugela-Dn1 ($F_{3/4}$)	499	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	500	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	501	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	502	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	503	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	504	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	505	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	506	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	507	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	508	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	509	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	510	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	511	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	512	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	513	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	514	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	515	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	516	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	517	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	518	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	519	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	520	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	521	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	522	1	RR

Tugela X Tugela-Dn1 ($F_{3/4}$)	523	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	524	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	525	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	526	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	527	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	528	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	529	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	530	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	531	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	532	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	533	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	534	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	535	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	536	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	537	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	538	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	539	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	540	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	541	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	542	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	543	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	544	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	545	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	546	1	RR

Tugela X Tugela-Dn1 ($F_{3/4}$)	547	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	548	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	549	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	550	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	551	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	552	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	553	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	554	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	555	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	556	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	557	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	558	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	559	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	560	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	561	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	562	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	563	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	564	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	565	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	566	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	567	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	568	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	569	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	570	1	RR

Tugela X Tugela-Dn1 ($F_{3/4}$)	571	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	572	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	573	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	574	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	575	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	576	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	577	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	578	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	579	1	RR
Tugela X Tugela-Dn1 ($F_{3/4}$)	580	5	Rr
Tugela X Tugela-Dn1 ($F_{3/4}$)	581	6	Rr

Table A.2: PCR primers used for screening in the mapping study.

Data on disk: Table A1 MAPPING_PrimerList.xlsx

Table A.3: Expression profile of gene RGA2-29_30 that was significantly up and down regulated between NILs after normalization. Indicated is the GenBank accession number, Affymetrix probe set ID and target description of two replicates. Also indicated is LogFC, average expression, p-value, adjusted p-value (Benjamini *et al.*, 1994) and gene expression. Red = up-regulation; green = down-regulation (adapted from Botha *et al.*, 2014).

Genbank accession	**AJ611054**			**AJ611054**		
Expression Tugela	2.76462081	2.59935715	2.69512947	2.76462081	2.59935715	2.69512947
Expression Tugela-*Dn1*	10.0317978	9.75934196	9.79600467	10.0317978	9.75934196	9.79600467
Expression Tugela-*Dn2*	9.53346407	9.44653453	8.64707066	9.53346407	9.44653453	8.64707066
Expression Tugela-*Dn5*	9.38190809	8.98838039	9.381649804	9.38190809	8.98838039	9.381649804
Affy ID	TaAffx.104444.1.S1_at			TaAffx.104444.1.S1_at		
Gene description	sequence(s) not in UniGene			sequence(s) not in UniGene		
logFC	7.17601			6.56428		
Avergae expression	7.752105			7.752105		
p-value	4.65E-14			1.40E-13		
Adjusted p-value	2.85E-09			3.31E-10		

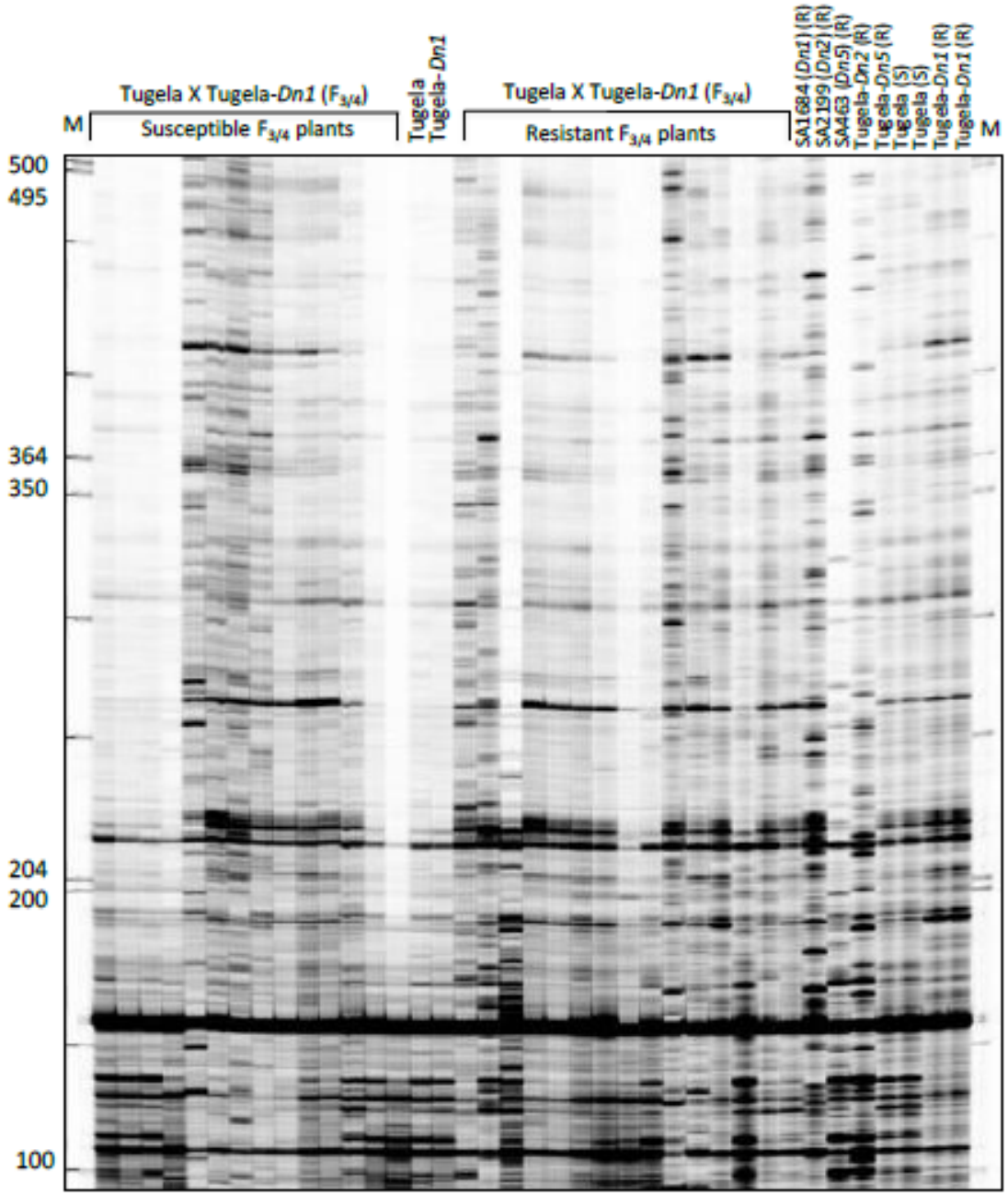


Figure A.1: Comparison of DNA fragments amplified from DNA of the Tugela x Tugela-*Dn1* ($F_{3/4}$) population, *Dn1*, *Dn2* and *Dn5* parental lines, and the different *Dn* progenitors using AFLP primer pair E-AGC/M-CTA. M = Li-COR IR Dye labelled 700 bp ladder.

Appendix B

PALM Microbeam and NGS

Table B.1: Amount and quality of the generated chromosome 7DS dt and 7DL dt sequence data.

	Chromosome or Library		
NGS Sequence data	**Chr 7DS dt**	**Chr 7DL dt**	**Total for Chr 7D**
Total reads	46,439,068	59,096,506	105,535,574
Total bases	4,690,345,868	5,968,747,106	10,659,092,974
Trimmed reads	35,484,806	44,708,526	80,193,332
Trimmed bases	3,583,965,406	4,515,561,126	80,99,526,532
Read count	35,484,806	44,708,526	80,193,32
N (%)[1]	0.0026	0.1704	
CG (%)[2]	42.60	42.13	
Q20 (%)[3]	98.94	98.96	
Q30 (%)[4]	94.86	94.96	

[1] percentage of bases that are unknown or N
[2] percentage bases that are either guanine or cytosine
[3] percentage bases called with a quality score of 20 or more
[4] percentage bases called with a quality score of 30 or more

Table B.2: Contigs, scaffolds and genome coverage of ditelosomic chromosomes 7DS and 7DL after *de novo* sequence assembly using SOAPdenovo.

	Contigs							
Chromosome	**Number of contigs**	**Contigs sum**	**N50**	**Longest contig**	**Shortest contig**	**Average length**	**Genome coverage**	
7DS	38,409	8,000,391	215	2,755 bp	65 bp	208 bp	5.5 x	4.3 x[1]
7DL	78,861	17,376,699	223	3,755 bp	67 bp	220 bp	7.6 x	5.8 x[1]
	Scaffolds							
Chromosome	**Number of scaffolds**	**Scaffolds sum**	**N50**	**Longest scaffold**	**Shortest scaffold**	**Average length**	**Genome coverage (x)**	
7DS	37,093	8,112,289	219	6,700 bp	100 bp	218 bp	5.5 x	4.3 x[1]
7DL	77,068	17,502,816	227	3,755 bp	100 bp	227 bp	7.6 x	5.8 x[1]

[1] trimmed sequence data

Table B.3: GC content of contigs and scaffolds after SOAPdenovo assembly.

Chromosome	**GC content contigs**					
	Number of A[1]	**Number of T**[2]	**Number of G**[3]	**Number of C**[4]	**Number of N** [5]	**GC content**[6]
7DS	2,303,100	2,108,914	1,716,414	1,871,916	47	44.85%
7DL	5,199,043	4,758,523	3,547,700	3,871,373	60	42.70%
Chromosome	**GC content scaffolds**					
	Number of A[1]	**Number of T**[2]	**Number of G**[3]	**Number of C**[4]	**Number of N**[5]	**GC content**[6]
7DS	2,303,100	2,108,914	1,716,414	1,871,916	111,945	44.23%
7DL	5,199,043	4,758,523	3,547,700	3,871,373	126,177	42.39%

[1] the total number of adenine(A)

[2] the total number of Thiamine(T)

[3] the total number of guanine(G)

[4] the total number of cytosine(C)

[5] the total number of ambiguous (N)

[6] the percentage of guanine-cytosine base pairs

Table B.4: Statistics of contigs and scaffolds after assembly using SOAPdenovo.

Chromosome	Contigs				
		Number[1]	Length[2]	Average length[3]	Sum[4]
7DS	N10	1,422	388	562	800,176
	N20	3,877	289	412	1,600,232
	N30	6,874	250	349	2,400,349
	40	10,226	229	312	3,200,227
	N50	13,834	215	289	4,000,240
	N60	17,647	205	272	4,800,355
	N70	21,637	197	258	5,600,321
	N80	25,782	190	248	6,400,432
	N90	31,439	127	229	7,200,369
	N100	38,409	65	208	8,000,391
7DL	N10	3,318	372	523	1,737,892
	N20	8,636	296	402	3,475,474
	N30	14,916	261	349	5,213,254
	N40	21,887	239	317	6,950,859
	N50	29,414	223	295	8,688,492
	N60	37,417	211	278	10,426,175
	N70	45,830	202	265	12,163,785
	N80	54,611	194	254	13,901,531
	N90	64,251	135	243	15,639,127
	N100	78,861	67	220	17,376,699
Chromosome	**Scaffolds**				
		Number[1]	**Length[2]**	**Average length[3]**	**Sum[4]**
7DS	N10	921	555	881	811,773
	N20	2,878	331	563	1,622,661
	N30	5,663	265	429	2,433,838
	N40	8,926	236	363	3,245,146
	N50	12,499	219	324	4,056,361
	N60	16,307	207	298	4,867,578
	N70	20,312	198	279	5,678,697
	N80	24,487	191	265	6,489,956
	N90	30,058	127	242	7,301,154
	N100	37,093	100	218	8,112,289
7DL	N10	2,651	433	660	1,750,464
	N20	7,469	318	468	3,500,846
	N30	13,484	271	389	5,250,936
	N40	20,298	245	344	7,001,364
	N50	27,731	227	315	8,751,626
	N60	35,685	214	294	10,501,858
	N70	44,082	204	277	12,252,099
	N80	52,873	195	264	14,002,412
	N90	62,387	137	252	15,752,652
	N100	77,068	100	227	17,502,816

[1] the number of sequences in the contig/scaffold statistics (N10–N100)
[2] the length of sequence in the contig/scaffold statistics (N10–N100)
[3] the average length in the contig/scaffold statistics (N10–N100)
[4] the sum of the length in the contig/scaffold statistics (N10–N100)

Table B.5: Number of matched reads obtained after mapping against the non-redundant TREP database.

Number of matched reads	**Chromosome**							
	7DS dt				**7DL dt**			
Pairwise identity of matched read	100%	90%	80%	70%	100%	90%	80%	70%
Single copy reads	31	2,393	361	20	68	3,211	563	16
All reads	31	2,424	2,785	2,805	68	3,279	3,842	3,858

Table B.6: Summary and numbers of the types of repetitive elements obtained from the sequence data set after mapping against the IWGSC data set.

Data on disk: Table B6_7DL reads matched to TREP.xlsx
Table B6_7DS reads matched to TREP.xlsx

Table B.7: Summary and numbers of mapping data for 7DS and 7DL

Data on disk: Table B7_7DL mapped RANDOMLY.xlsx
Table B7_7DS mapped RANDOMLY.xlsx

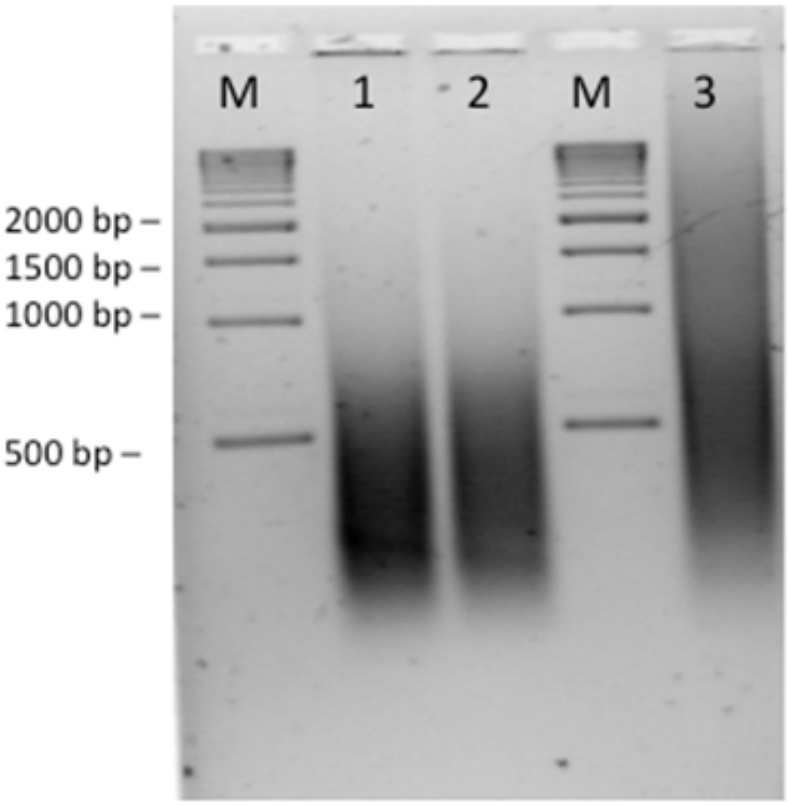
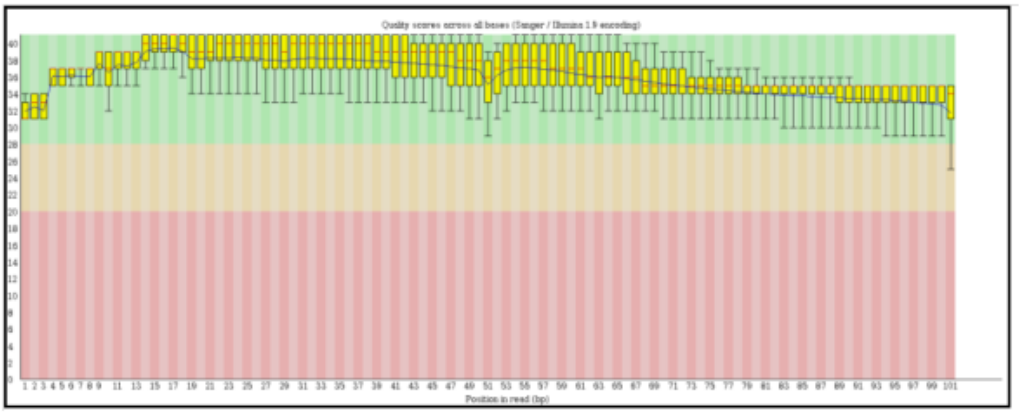


Figure B.1: Whole genome amplification products on 2% agarose gel. M = 500 bp DNA marker; where 1 = 7DS dt; 2 = 7DL dt; and 3 = Control DNA (5 ng/μL).

A.



B.

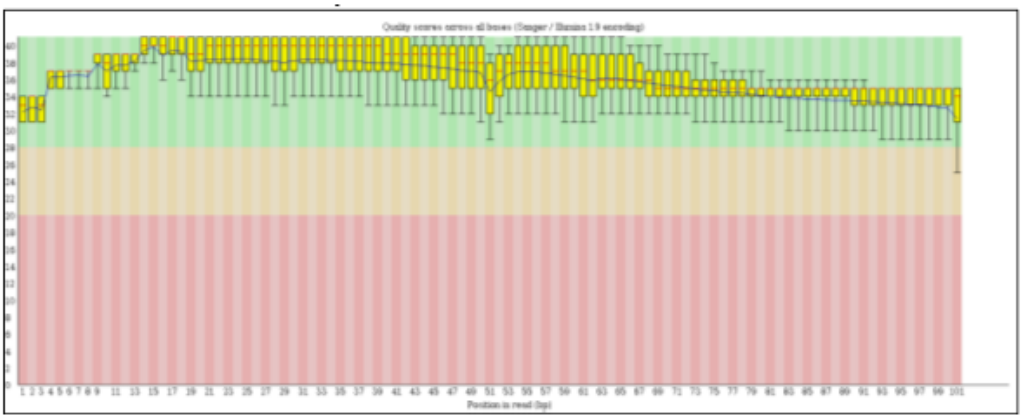


Figure B.2: Quality scores of obtained data after filtering (A) chromosome 7DS dt; and (B) chromosome 7DL dt.

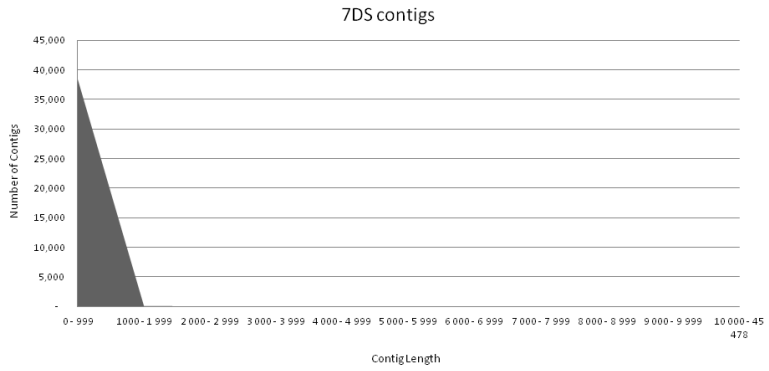


Figure B.3: Graph representing the contig size distribution after SOAPdenovo assembly of the chromosomes 7DS dt.

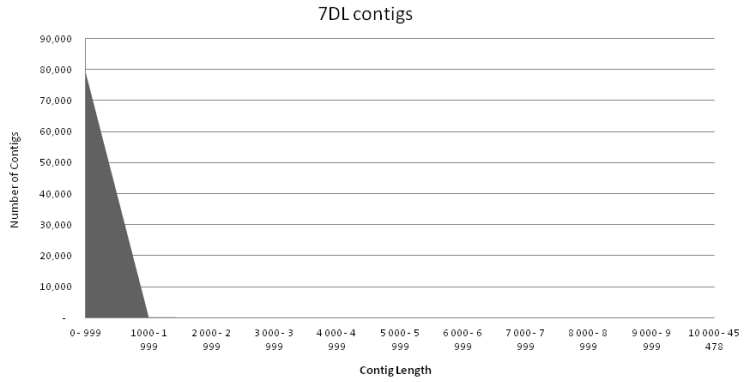


Figure B.4: Graph representing the contig size distribution after SOAPdenovo assembly of the chromosomes 7DL dt.

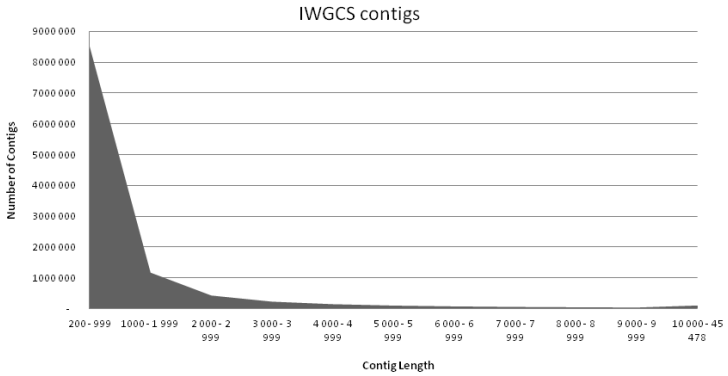


Figure B.5: Graph of the contig size distribution of the IWGSC dataset.

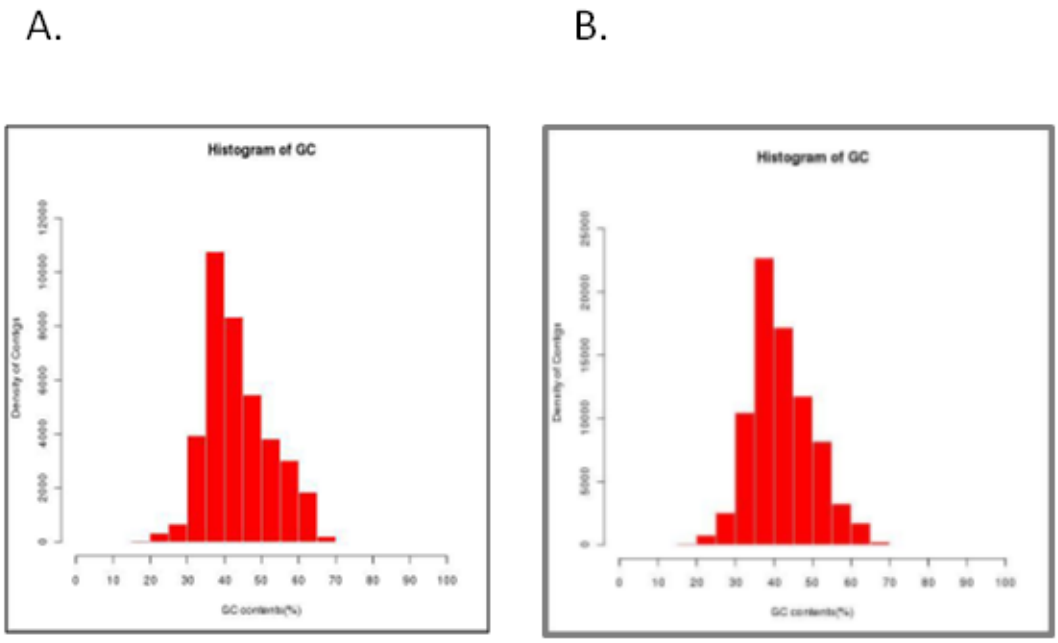


Figure B.6: Histogram representing GC content of the contigs assembled from chromosomes 7DS dt (A) and 7DL dt (B) sequences using SOAPdenovo.

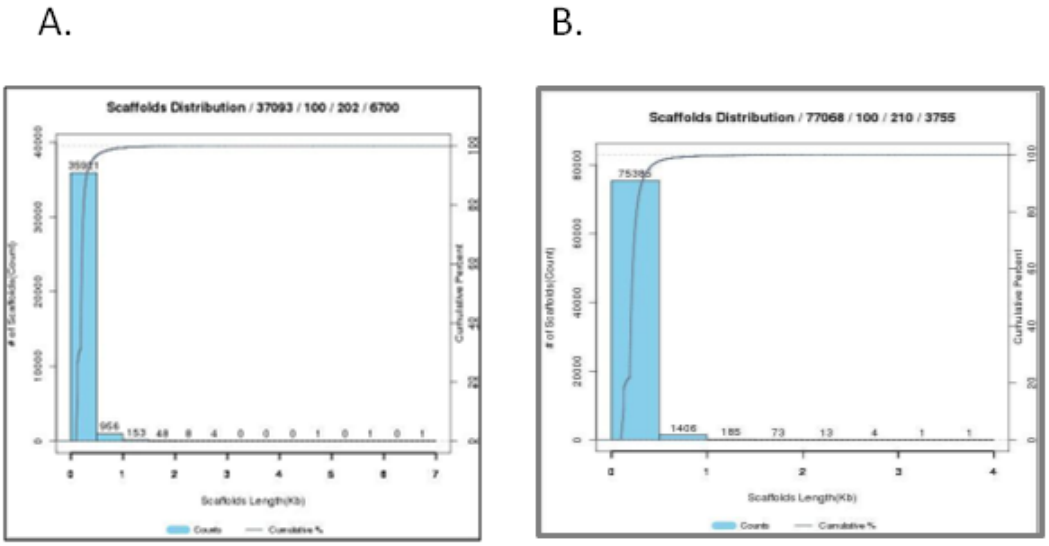


Figure B.7: Graph representing the length of the obtained scaffolds from the assembled chromosomes 7DS dt (A) and 7DL dt (B) sequence using SOAPdenovo.

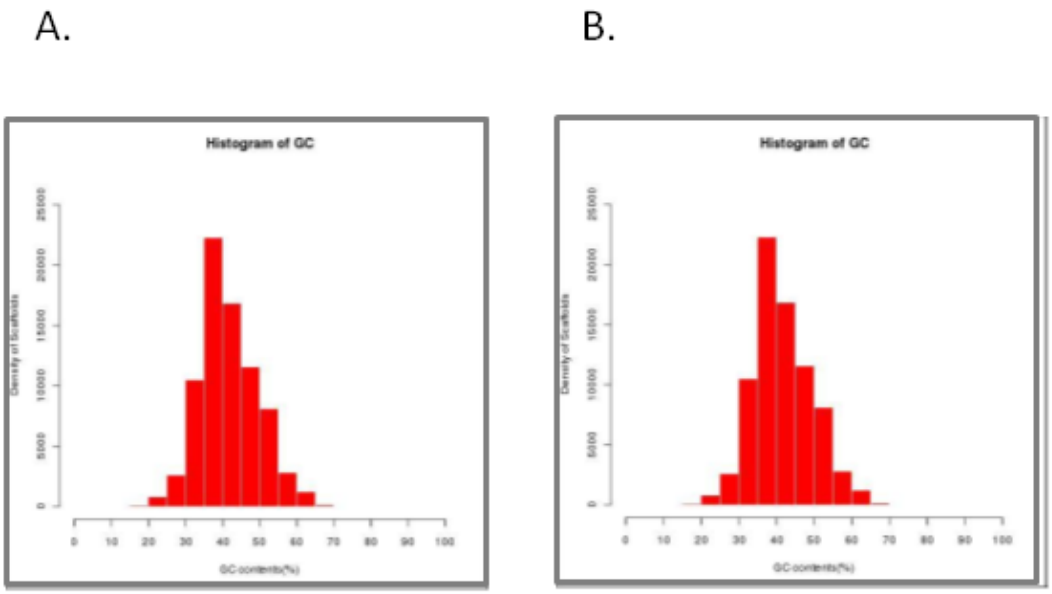


Figure B.8: Histogram representing the GC content of the obtained scaffolds build from the contigs obtained from chromosomes 7DS dt (A) and 7DL dt (B) sequence after SOAPdenovo assembly.

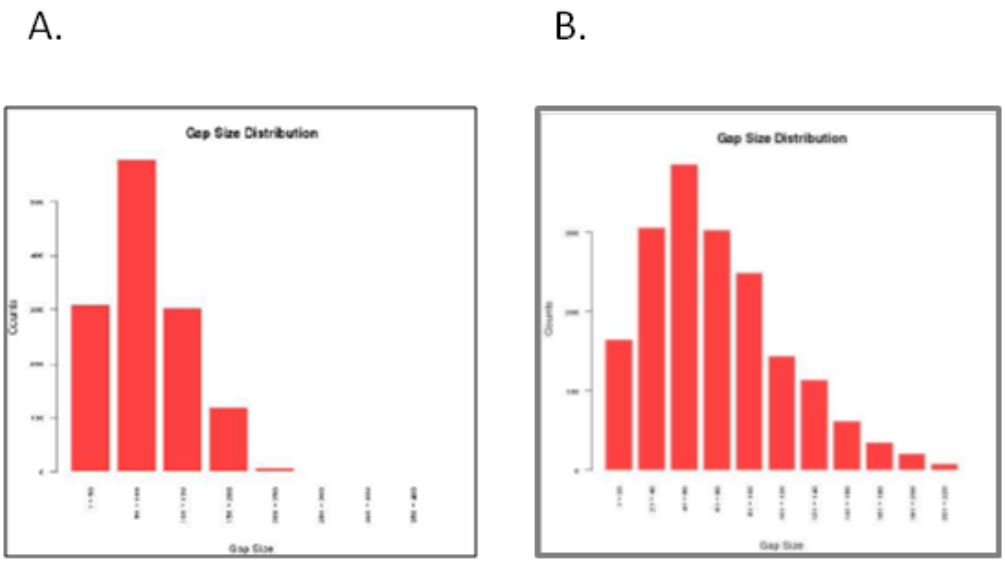


Figure B.9: Gap size distribution in the obtained scaffolds from the chromosomes 7DS dt (A) and 7DL dt (B) sequence data sets.

Figure B.10: Contig order in built scaffolds obtained from chromosomes 7DS dt (A) and 7DL dt (B) sequences.

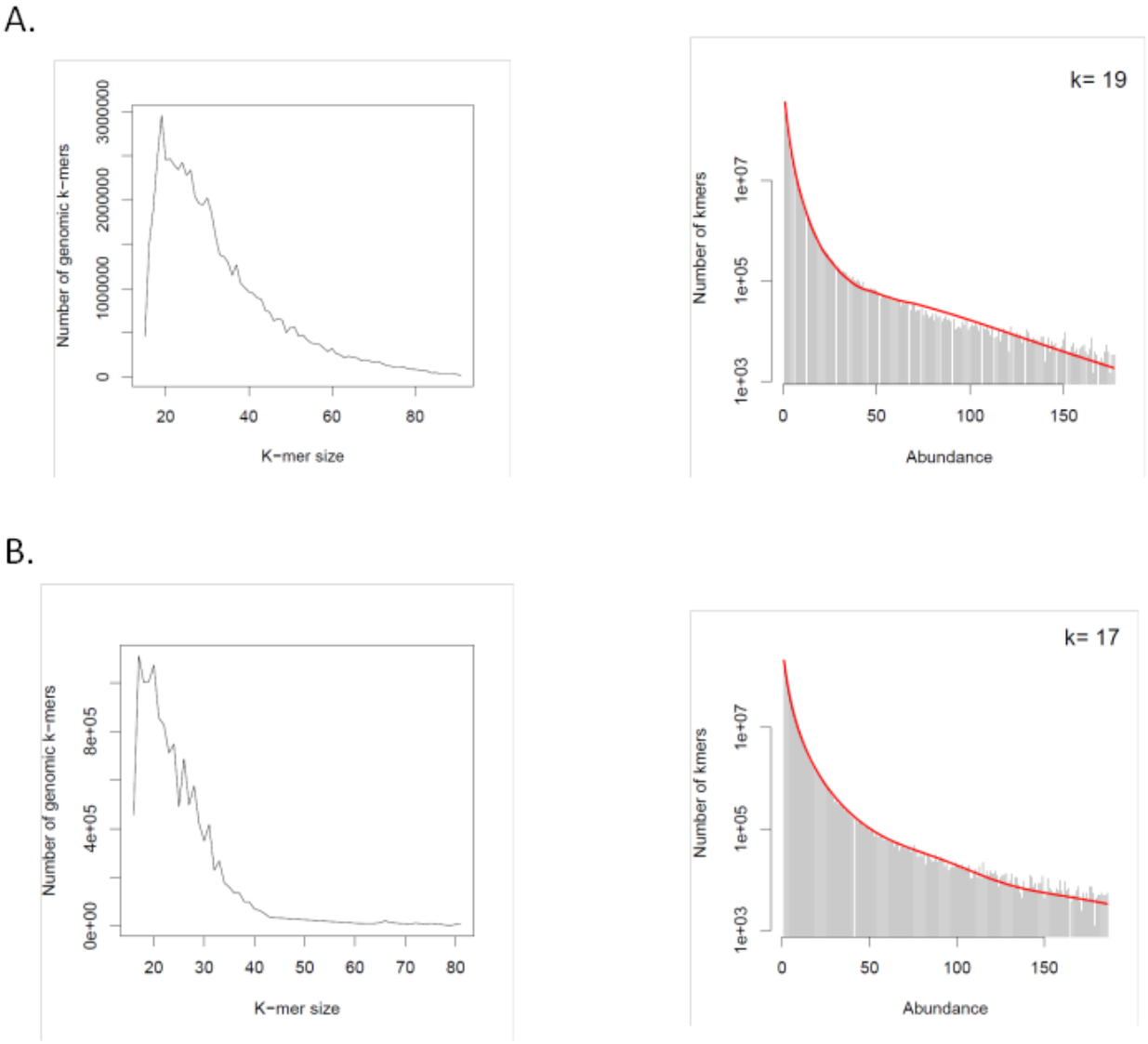


Figure B.11: Graph representing the number of genomic k-mers of chromosomes 7DS dt (k=19) (A) and 7DL dt (k=17) (B) data sets after k-mer analysis.

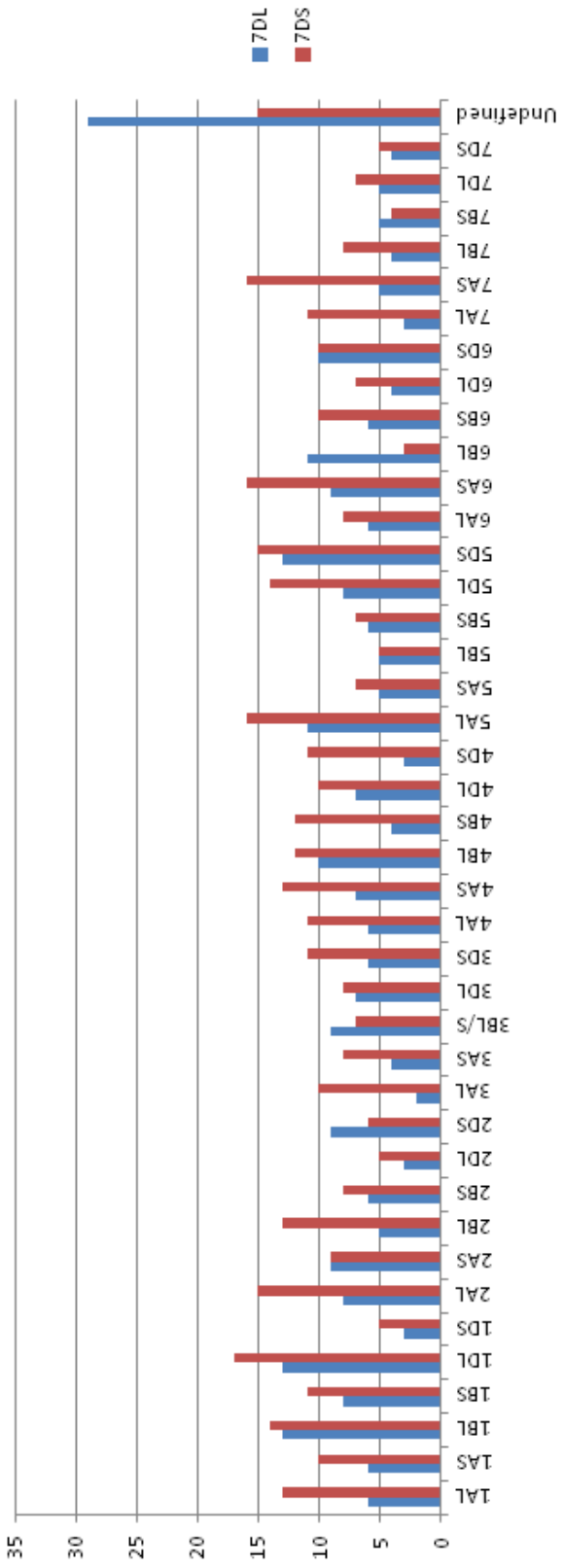


Figure B.12: Predicted location of the obtained PCGs after mapping against the IWGSC scaffolds available on ENSEMBL (http://www.ensembl.org).

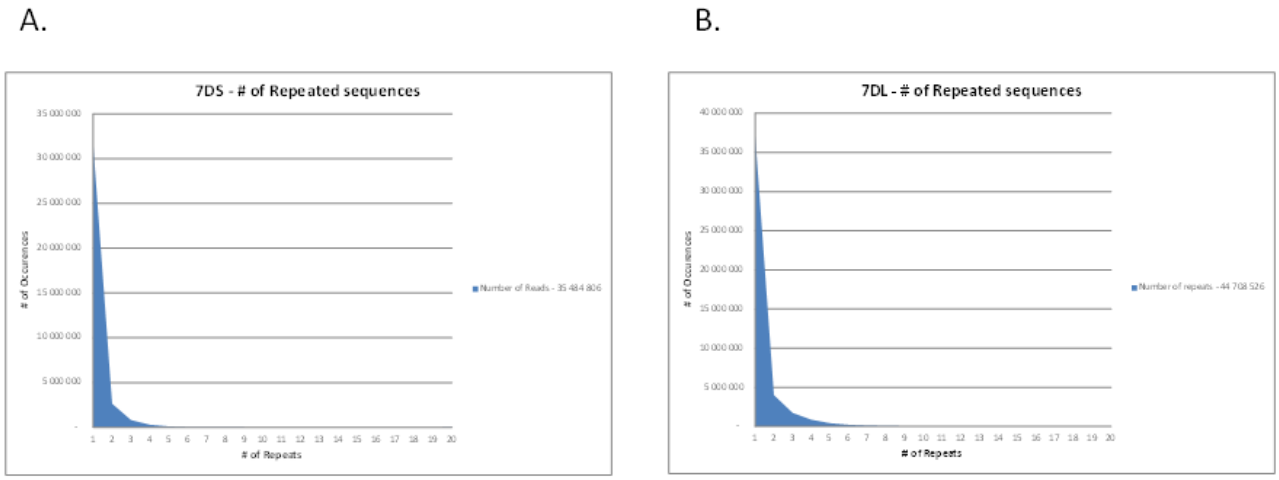


Figure B.13: Number of repeated sequences after duplicated reads were collapsed on the NGS data sets from chromosomes 7DS dt (A) and 7DL dt (B).