The granule-bound starch synthase genes of wheat.

Bernadette Bradley

This thesis is presented for the degree of Doctor of Philosophy of Murdoch University 2003

Western Australian State Agricultural Biotechnology Centre
School of Biological Sciences and Biotechnology
Division of Science and Engineering
Murdoch University, Perth, Western Australia
DECLARATION

I declare that this thesis is my own account of my research and contains as its main content work which has not previously been submitted for a degree at any tertiary education institution.

__________________________
Bernadette Bradley
ACKNOWLEDGMENTS

I would like to thank my supervisors Professor Mike Jones and Dr Rob Potter for their training, advice and assistance to complete this research.

I am also grateful for the support of all my colleagues at the SABC, particularly Meredith Carter, Max Paris and Steve Wylie, and Veronika Reck at DAWA.

I gratefully acknowledge financial support from the Grains Research and Development Corporation, the Co-operative Research Centre for Molecular Plant Breeding, and Murdoch University.

Finally I would like to thank my husband Iain for his tireless support during this endeavour.
ABSTRACT

Wheat (*Triticum aestivum*) is the world’s most widely grown and economically important crop. It is both a staple food for humans and a raw material for many industrial processes. World trade in wheat is important for economic stability and an ability to grow wheat is a valuable national resource. Wheat is Australia’s major crop with an annual production of about 23 million tonnes. One-quarter of this is used domestically and meets all of Australia’s requirements; the remaining three-quarters is exported. Therefore, Australia’s wheat industry provides both the national staple food source and the basis of an export industry worth almost 2 billion dollars.

There is great potential for further genetic improvement of wheat, not only by increasing grain yields by improved resistance to pathogens and tolerance to adverse environmental conditions, but also by improving functional quality. For example, one can change the physical properties of the storage components, starch and protein, to increase their usefulness in conventional applications and for novel uses. Some examples of the physical properties of starch affecting its uses are, the large starch granules from wheat that are suitable to make carbonless copy paper (Bligh, 1999), the small starch granules from rice that are used as a fat substitute because the a comparable mouthfeel, the high-amylose starches that have film-forming properties desirable for fried food coating batters and some forms of plastics, and the low-amylose starch that swells more in water and can be used for soft foods such as Asian noodles. Continued improvement of wheat is vital to meet the quantity and quality demands of the local and international wheat markets.

One specialty market for Western Australian wheat, is export to Japan and South Korea for the production of Japanese white salted Udon noodles, an export market worth more than $200M pa (Garlinge, 1996). Udon noodles have specific eating qualities including a light, creamy, uniform colour, a ‘bright’ appearance to the noodle, a soft but elastic texture to the noodle, and a smooth ‘mouthfeel’, all of which result from the quality of the wheat flour starch they are made from.
(Crosbie, 1991; Batey et al., 1997; Zeng et al., 1997). The Australian Standard White Noodle (ASWN) wheat that Australia exports to produce Udon noodles is soft-grained, white coloured, contains between 9.5% and 11.5% protein, and produces flour of fine particle size with little starch damage (V. Reck, DAWA, pers. comm.). The flour also has good starch-swelling characteristics, moderate dough strength and good dough extensibility.

The good starch-swelling characteristics of the flour result, for the most-part, from containing relatively less of the starch amylose than other varieties (22-23% compared to 25%), a property controlled by the GBSS genes (Nelson and Rines, 1962; Garlinge, 1996). When less amylose is present in the starch granule as it is heated in water, the amylopectin matrix inside the granule can swell, causing the finished Udon noodle to be soft. When more amylose is present in the starch granule, the amylopectin matrix cannot swell as much, and the finished noodle is too hard to have the desired ‘mouthfeel’ of an Udon noodle. The amylose fraction of starch is produced by the granule-bound starch synthase (GBSS) enzymes, encoded by the GBSS genes. The overall aim of the research described in this thesis was to investigate the genomic organization of the GBSS genes of wheat. Since the GBSS genes influence wheat starch quality, an understanding of the action of these genes is needed for future improvement of starch quality in noodle-wheats.

There are three loci for GBSS genes in wheat, and these are located on chromosomes 4A, 7A and 7D. Both wild-type alleles and non-functional ‘null’ alleles exist at each locus. At the start of the project, these alleles had not been sequenced and the molecular differences between the alleles were not known. Other GBSS alleles were also thought to exist in Australian varieties that had yet to be identified and characterised. GBSS genes from a selection of wheat varieties, and from all three GBSS loci, were sequenced searching for DNA polymorphisms that were different between the different alleles. If any DNA polymorphisms were found to result in GBSS protein sequence differences, or differences in GBSS enzyme expression, they could influence the functional
characteristics of the starch. Identifying GBSS allelic variants would enable molecular markers to be developed to detect the alleles and investigate their potential effects upon starch quality.

Different PCR-based methods and one non-PCR-based method were used to investigate the genomic organization of the GBSS genes in a selection of genetically diverse wheat varieties. The 31 wheat varieties studied included noodle-wheat varieties from the ASWN classification, varieties with similar genetic background to ASWN wheat varieties but of unsuitable quality for noodle-production, unrelated varieties of Australian Standard White wheat, and were compared with those ‘Chinese Spring’ varieties described in the literature. Most of the varieties are grown in the Western Australian wheatbelt and southern regions, either for export and the production of Asian noodles, or for the production of domestic baked-goods.

A 500bp section from the middle of the GBSS genes was amplified, from a selection of wheat varieties, and sequenced to search for polymorphisms. Twenty-one single nucleotide differences were found between genes at the three loci and two PCR-based tests were designed to validate these differences as Single Nucleotide Polymorphisms (SNPs). A novel microsatellite was also discovered in intron 4 of the GBSS 7A genes. This (TGCCG)n microsatellite was variable between wheat varieties and so defines a novel allele in the Australian germplasm present at a frequency of 40%. A PCR-based test was developed to identify this variable locus. However, the new GBSS allele was not linked to Flour Swelling Volume (FSV) quality properties.

The variable microsatellite locus Xsun1 (Shariflou and Sharp, 1999) in the 3’ untranslated region of the GBSS genes and linked to GBSS allelic variation was used to genotype a wheat breeding population for its GBSS status. The population (n=69) contained combinations of wild-type and null alleles at the 7A and 7D loci. Once genotyped using this marker, the GBSS alleles were assessed for possible linkage to starch variation. Although the trend suggested that the presence of
null alleles increased the FSV, the size of the population tested was too small for the differences in FSV between wild-type and partially-\textit{waxy} wheats to be statistically significant.

The linkage between the \textit{Xsun1} microsatellite variation and the \((\text{TGCCG})_n\) microsatellite variation from intron 4 of the GBSS 7A genes was studied. By combining these two microsatellite loci, which are closely linked to the GBSS coding regions, GBSS genes at the 7A locus could be separated into 12 allelic groups. Although none of these groups could be linked to specific changes in starch qualities, they can be analysed further for functional differences.

In order to access a larger section of the GBSS genes using PCR, new PCR primers were designed and optimized to amplify segments of the GBSS genes. Primers for GBSS genes tend to generate many PCR products, but many of these were shown to be non-specific. These artifacts could be reduced by increasing the annealing temperatures, and non-specific priming was repressed by the presence of the second primer in the PCR reaction. Using one primer set, a nearly 2000bp segment of the GBSS 7A genes from wheat varieties ‘Kulin’ and ‘Eradu’ was amplified and sequenced. These sequences indicated the presence of single nucleotide differences that resulted in changed amino acids in the protein when compared to published GBSS sequences. The sequencing should be repeated to validate this result, which indicates that these are novel alleles, but it does suggest that allelic variation for GBSS exists in Australian wheat varieties and that these alleles are different from those described internationally.

The EcoR1, HindIII and BamH1 restriction enzyme sites surrounding the GBSS genes were identified using Southern hybridisation. This provided the potential to access the entire GBSS gene, including the promoter and untranscribed regions, by restriction enzyme mediated cloning of genomic DNA. However, attempts to clone the genomic GBSS genes into both plasmid and viral vectors were not successful.
The potential existence of pseudogene copies of the GBSS genes in the wheat genome was investigated using both PCR and Southern hybridisation techniques. No evidence of GBSS pseudogenes was found, and this suggests that the wheat genome does not contain them. This result was unexpected since organisms with large genomes, such as wheat, normally contain repeated sequences and pseudogenes. However, the absence of repeated sequences and pseudogenes should be beneficial in molecular wheat breeding because it suggests that there will not be interference from non-coding GBSS sequences in identifying molecular markers to GBSS genes.

The GBSS genes present in Australian wheat varieties were similar enough to those described internationally that Australian breeders can make full use of research and molecular tests for GBSS genes developed elsewhere. However, enough variation exists between overseas and domestic varieties to warrant further investigation of novel GBSS alleles in domestic wheat, which may relate to differences in functionality.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE PAGE</td>
<td>i</td>
</tr>
<tr>
<td>DECLARATION</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td></td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xiv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xxi</td>
</tr>
<tr>
<td>ABBREVIATIONS</td>
<td>xxii</td>
</tr>
<tr>
<td>LIST OF COMPANY ADDRESSES</td>
<td>xxiv</td>
</tr>
</tbody>
</table>

Chapter 1. Introduction and literature review. ........................................ 1
1.1 General introduction to the project. ........................................... 1
1.2 The Australian Standard White Noodle Wheat industry in Western Australia. ........................................ 2
1.3 Starch biosynthesis. ............................................................... 3
1.4 Potential to manipulate starch biosynthesis. ............................ 9
1.5 Control of GBSS enzyme activity in wheat. .................................... 11
1.6 Genomic location of GBSS loci and common nomenclature. ............... 12
1.7 Molecular structure of the GBSS genes. ..................................... 14
1.8 GBSS alleles. ................................................................. 14
1.9 Wild-type alleles. .................................................................... 16
1.10 Null alleles. ....................................................................... 21
1.11 Pseudogene copies of GBSS genes. .......................................... 25
1.12 GBSSII expression in pericarp tissue. ...................................... 26
1.13 The varied activities of the GBSS enzymes from different GBSS loci. ........................................ 27
1.14 Varied activities between enzymes encoded by different wild-type alleles at a single locus. ............ 31
1.15 Effects of amylose levels on starch quality and Flour Swelling Volume (FSV). ........................................ 32
1.16 Molecular markers for noodle quality. ..................................... 33
1.17 Aims of this project. ............................................................ 36

Chapter 2. General Materials and Methods. ........................................ 38
2.0 Description of the Chapter. ..................................................... 38
2.1 Wheat varieties and sources of plant material. ............................ 38
2.2 DNA extraction. ................................................................. 43
   2.2.1 Cetyltrimethylammonium bromide (CTAB) extraction of plant DNA. ........................................ 43
   2.2.2 Extraction of plant DNA based on the method of Dellaporta et al. (Potter, 1991). ........................................ 44
   2.2.3 Quick extraction of plant DNA. ................................................ 45
2.2.4 Plant DNAzol DNA extraction. .................................................. 46
2.2.5 Wheat Ethanol-Perchlorate Method. ........................................ 46

2.3 DNA quantification. ................................................................. 47

2.4 PCR primers and conditions. .................................................... 47

2.5 Gel electrophoresis. ................................................................. 49

2.6 DNA purification by ethanol precipitation. ................................... 49

2.7 PCR fragment isolation by gel purification. .................................. 50

2.7.1 QIAquick Gel Extraction Kit. .................................................. 50
2.7.2 Quick method. ........................................................................ 50

2.8 Ligation of DNA fragments into the pGEM-T vector and transformation through Escherichia coli. .................................................. 50

2.9 Identification of clones containing inserts using PCR. ....................... 51

2.10 Sequencing. ............................................................................ 51

2.10.1 Sequencing the inserts inside plasmid clones. ......................... 51
2.10.2 Direct Sequencing of PCR products. ...................................... 52

2.11 Southern blots. ....................................................................... 52

2.11.1 Transferring DNA from gels to membranes. ............................ 52
2.11.2 Hybridisation of Southern blots of PCR products and Dot Blots. 53
2.11.3 Hybridisation of Southern blots of genomic DNA. ................... 54
2.11.4 Preparation of a radioactive probe. ......................................... 54
2.11.5 Washing the membrane. ......................................................... 55

Chapter 3. A survey of the GBSS alleles present in WA noodle-wheat varieties based on the DNA sequence of a 500bp section of the genes. ................................................................. 56

3.1 Introduction. ............................................................................. 56

3.2 Methods. .................................................................................. 60

3.3 Results. ................................................................................... 61

3.3.1 Comparison of the PCR screening tests used to identify clones with GBSS inserts. ..................................................... 61
3.3.2 Comparison of GBSS alleles at the 4A locus of WA wheat varieties. ............................................................... 65
3.3.3 Design and partial optimisation of a test to validate a SNP at the 4A locus. ............................................................... 67
3.3.4 The GBSS alleles at the 7D locus of WA wheat varieties. .......... 70
3.3.5 Design and partial optimisation of a test to validate a SNP at the 7D locus. ............................................................... 72
3.3.6 Comparison of the GBSS alleles at the 7A locus of WA wheat varieties. ............................................................... 73
3.3.7 The linkage of the (TGCCG)n microsatellite variation and starch quality. ............................................................... 76
3.3.8 Development of a simple PCR test to distinguish the two alleles defined by polymorphism at the (TGCCG)n microsatellite locus. .................................................. 77
3.3.9 Results for screening Australian wheat varieties for the (TGCG)n microsatellite. ................................. 79

3.4 Discussion. .......................... 81
3.4.1 Where are the GBSS gene copies in the genome? .......... 81
3.4.2 How many different GBSS alleles are in the WA wheat germplasm and how does each GBSS allele differ from the others? ......................................................... 82
3.4.3 How can we tag the different GBSS alleles with molecular markers to identify them in a breeding program? .......... 84
3.4.4 Are the GBSS genes present in WA wheat varieties the same alleles described in international publications? 88

Chapter 4. Cloning the genomic GBSS gene sequences. ................. 90
4.1 Introduction. ......................................................... 90
4.2 Methods. ......................................................... 93
4.2.1 DNA extraction methods. ........................................ 93
4.2.2 Digestion of genomic wheat DNA with restriction enzymes and production of Southern blots. ..................... 93
4.2.3 Probing the Southern blots with a GBSS-specific probe. .... 94
4.2.4 Isolation of the 9kbp BamH1 fragments from the wheat genomic DNA. .............................................. 94
4.2.5 Cloning into the pGEM11 plasmid. .............................. 95
4.2.6 Cloning into the EMBL3 viral vector. ......................... 95
4.3 Results. ......................................................... 96
4.3.1 Testing DNA extraction methods to produce high-quality genomic DNA for Southern blotting. ....................... 96
4.3.2 Choice of restriction enzymes. ................................ 99
4.3.3 Identification of the sizes of restriction fragments containing the GBSS genes. ......................................... 100
4.3.4 Choice of restriction fragment to clone (in the initial experiment). ............................................. 103
4.3.5 Isolation of the 9kbp BamH1 restriction fragments from the wheat genome. ........................................ 104
4.3.6 Cloning into the pGEM11 plasmid. ......................... 107
4.3.7 Cloning into the EMBL3 viral vector. ........................ 107
4.4 Discussion. ......................................................... 108
4.4.1 Where are the GBSS gene copies in the genome? .......... 109
4.4.2 How does each GBSS allele differ from the others? .......... 110
4.4.3 How can we tag the different GBSS alleles with molecular markers to identify them in a breeding program? .......... 115
4.4.4 Are the GBSS genes present in WA wheat varieties the same alleles described in international publications? ........ 115

Chapter 5. Characterisation of the GBSS genes using PCR............. 116
5.1 Introduction. ......................................................... 116
5.2 Methods. ........................................................................................................ 119
  5.2.1 PCR. ........................................................................................................ 119
  5.2.2 Southern Blots. ....................................................................................... 120
  5.2.3 Size separation and isolation of GBSS-related PCR fragments from artifacts. ........................................................................................................ 120
  5.2.4 Cloning GBSS1/GBSS2 and GBSS1/GBSS3 PCR fragments. .... 121
  5.2.5 Cloning the GBSS4/GBSS6 PCR fragment. .................................. 121
  5.2.6 Screening clones to identify their inserts using PCR. .................. 122
5.3 Results. .......................................................................................................... 123
  5.3.1 Investigation of non-specific amplification from the wheat genome by the PCR primers GBSS1 to GBSS10……………….. 123
  5.3.2 Amplification using primers GBSS1/GBSS2 and GBSS1/GBSS3. ................................................................. 130
  5.3.3 Identification of GBSS-related PCR bands by hybridization. ... 131
  5.3.4 Isolation and cloning of the GBSS-related PCR products. …… 134
  5.3.5 Design and optimization of novel PCR primers to amplify from the wheat GBSS genes. ......................................................... 138
  5.3.6 Cloning and sequencing the GBSS4/GBSS6 PCR fragment. .... 142
  5.3.7 Sequence of the GBSS4/GBSS6 PCR fragment. .................................. 149
  5.3.8 DNA sequence of the GBSS4/GBSS6 PCR fragment from a second Australian wheat variety ‘Eradu’. ................................. 153
  5.3.9 Use of the GBSS4/GBSS6 PCR primers to access DNA sequence from the null 7A allele. .................................................. 157
5.4 Discussion. .................................................................................................... 160
  5.4.1 Homoproimed bands. ........................................................................ 160
  5.4.2 GBSS1-GBSS3 PCR bands. ................................................................. 161
  5.4.3 GBSS4/GBSS6 PCR bands. ................................................................. 164

Chapter 6. PCR-based methods for accessing novel sections of the GBSS gene. ........................................................................ 167
6.1 Introduction. ................................................................................................ 167
6.2 Methods. ...................................................................................................... 169
  6.2.1 Construction of linker molecules. ...................................................... 169
  6.2.2 Preparation of template DNA. ............................................................. 169
  6.2.3 Testing for non-specific amplification. ............................................ 170
  6.2.4 Restriction site mediated PCR. .......................................................... 170
6.3 Results. ......................................................................................................... 171
  6.3.1 Restriction site mediated PCR on plasmid template. ............... 171
  6.3.2 Non-specific amplification from the wheat template. .......... 174
  6.3.3 Restriction site mediated PCR amplification from wheat. ...... 175
  6.3.4 Amplification using locus-specific primers. ................................ 178
  6.3.5 PCR optimisation for amplifying from wheat. ........................... 181
6.4 Discussion. .................................................................................................. 186
  6.4.1 Primer specificity. ............................................................................. 186
6.4.2 Restriction site prediction based on PCR bands.  ......................... 186
6.4.3 GBSS promoter. ................................................................. 188
6.4.4 Conclusions. ................................................................. 189

Chapter 7. Linking GBSS molecular markers to GBSS alleles and starch quality traits. ......................................................... 191
7.1 Introduction. ................................................................. 191
7.2 Methods. ................................................................. 193
7.3 Results. ................................................................. 195
  7.3.1 Comparing the banding patterns of the Xsun1 marker using two methodologies. ......................................................... 195
  7.3.2 Xsun1 allele variation linked to noodle quality traits. ................. 197
  7.3.3 Linkage between the Xsun1-7A locus and the (TGCCG)₅ microsatellite locus. ......................................................... 199
  7.3.4 Use of the Xsun1 molecular marker to identify the GBSS characteristics of a noodle-wheat breeding population. ........... 201
7.4 Discussion. ................................................................. 206
  7.4.1 Predicted structural differences in the 3’ untranslated region of the GBSS genes based on Xsun1 molecular marker patterns. .... 206
  7.4.2 Allelic groups at the 7A locus support other genetic studies. ...... 209
  7.4.3 Use of the Xsun1 marker to screen two breeding populations and link the GBSS alleles to starch quality ..................... 210
  7.4.4 Using Doubled Haploid technology to produce partially waxy wheats may be hindered by their low survival rate. ............... 212
  7.4.5 Conclusions. ................................................................. 213

Chapter 8. General discussion. ......................................................... 215
8.1 General discussion. ................................................................. 215
  8.1.1 Aims. ................................................................. 215
  8.1.2 Where are the GBSS gene copies in the genome? ................. 215
  8.1.3 How many different GBSS alleles are in the WA wheat breeding germplasm? ......................................................... 218
  8.1.4 How does each GBSS allele differ from the others? ............... 219
  8.1.5 How can we tag the different GBSS alleles with molecular markers to identify them in a breeding program? ............... 222
  8.1.6 Are the GBSS genes present in WA wheat varieties the same alleles described in international publications? ............... 224
8.2 Future directions. ................................................................. 225
8.3 Conclusions. ................................................................. 228

REFERENCES. ................................................................. 229
APPENDIX 1. ................................................................. 233
APPENDIX 2. ................................................................. 235
LIST OF FIGURES

Figure 1.1. Molecular structures of starch polymers. .......................... 5

Figure 1.2. An outline of the metabolic pathway of starch biosynthesis. ...... 6

Figure 1.3. A diagram showing the action of the branching and debranching enzymes on linear amylopectin to produce mature branched amylopectin. ................................................................. 7

Figure 1.4. A drawing representing the structure and organisation of a starch granule. .......................................................... 8

Figure 1.5. Diagrams showing (TOP) the positions of the GBSS (Waxy) loci on chromosomes 7A, 4A and 7D, and (BOTTOM) translocations between chromosomes 4A, 5A and 7B resulting in the structure of the present 4A chromosome. ............................................ 13

Figure 1.6. A diagrammatic representation of the molecular organisation of GBSS genes. ......................................................... 14

Figure 1.7. An alignment of the genomic DNA sequences of the wild-type GBSS genes from nullisomic/tetrasomic ‘Chinese Spring’ wheat, as presented by Murai et al. (1999). ................................................. 18

Figure 1.8. (TOP) A diagrammatic representation of the GBSS genes showing the position of the mutation in the null 7A allele at the junction of exon 1 and intron 1. (BOTTOM) The DNA sequence polymorphism between the wild-type 7A and null 7A alleles in the region of the junction between exon 1 and intron 1. ............ 23

Figure 1.9. (TOP) A diagrammatic representation of the GBSS genes showing the position of the mutation in the null 7D allele in exon 11 and the 3’ untranslated region. (BOTTOM) A diagrammatic comparison of the structures of the wild-type 7D and null 7D alleles in the region of the null 7D deletion mutation in exon 11 and the 3’ untranslated region as shown in Vrinten et al. (1999). ... 24

Figure 1.10. A polyacrylamide gel showing the PCR bands amplified from wheat by the primers GBSS2/GBSS4. ................................. 35

Figure 2.1. Arrangement for Southern Blotting. ................................. 53

Figure 3.1. A diagrammatic representation of the GBSS genes showing the binding positions of the PCR primers GBSS2 and GBSS4 within exons 6 and 4 respectively. ................................. 57

Figure 3.2. DNA sequence alignment of a section from the GBSS genes at all three GBSS loci on chromosomes 4A, 7A and 7D (from Hollingsworth, 1997). ......................................................... 59

Figure 3.3. 1% agarose gel showing clones containing different sized inserts screened in PCR using the T7/Sp6 primers. ........................ 62
Figure 3.4. Positions of PCR primers on the GBSS gene. .......................... 63

Figure 3.5. 1% agarose gel showing use of the PCR primer combination GBSS-Rev/GBSS-4A,7A,7D to screen clones for the inserts they contained. ................................................................. 63

Figure 3.6. 1% agarose gel showing how clones previously identified to contain a GBSS-related insert were screened, using the three locus-specific PCR primer sets, to ascertain which GBSS locus the insert came from. ...................................................... 64

Figure 3.7. DNA sequence alignment of PCR fragments amplified from GBSS genes at the 4A locus of a selection of 4 wheat varieties listed to the left of the alignment. .................................................. 66

Figure 3.8. DNA sequence of the PCR product amplified from the 4A locus of ‘Kulin’ by the primers GBSS2/GBSS4 showing the positions of the new potentially locus-specific primers GBSS4-2 and GBSS-4A-2 designed to amplify a 136bp fragment for SNP analysis........ ................................. 68

Figure 3.9. Three 1% agarose gels showing PCR bands amplified from the genomic ‘Kulin’ DNA over a gradient of annealing temperatures... 69

Figure 3.10. DNA sequence alignment of PCR fragments amplified from GBSS genes at the 7D locus of a selection of wheat varieties listed to the left of the alignment. ................................. 71

Figure 3.11. DNA sequence of the PCR product amplified from the 7D locus of ‘Kulin’ by the primers GBSS2/GBSS4 showing the positions of the locus-specific primer GBSS-7D and new potentially locus-specific primer GBSS-7D-2 designed to amplify a 143bp fragment for SNP analysis. ................................. 73

Figure 3.12. DNA sequence alignment of PCR fragments amplified from GBSS genes at the 7A locus of a selection of wheat varieties listed to the left of the alignment. ................................. 75

Figure 3.13. DNA sequence of the PCR product amplified from the 7A locus of ‘Kulin’ by the primers GBSS2/GBSS4 showing the positions of the locus-specific primer GBSS-7A and new potentially locus-specific primer GBSS-7A-2 designed to amplify a 170bp fragment containing the (TGCCG)n microsatellite region. .......... 78

Figure 3.14. A 10% polyacrylamide gel showing the validation of the GBSS-7A/GBSS-7A-2 PCR test to distinguish PCR bands from GBSS-7A alleles containing either three or four (TGCCG)n microsatellite repeats. .................................................. 79
Figure 3.15. A 10% polyacrylamide gel showing screening of 19 wheat samples for the number of the (TGCCG)_n microsatellite repeats contained in their GBSS-7A alleles using the GBSS-7A/GBSS-7A-2 PCR. ......................................................... 80

Figure 3.16. Diagram of the flow of the GBSS 7A allele marked by the (TGCCG)_4 microsatellite through wheat varieties in the noodle-wheat breeding program........................................... 88

Figure 4.1. 1% agarose gel showing the quality of DNA extracted using two methods, the CTAB and Modified Dellaporta methods. ............ 97

Figure 4.2. 1% agarose gel showing the quality of DNA extracted using the Ethanol-perchlorate method. ......................................................... 98

Figure 4.3. Diagrammatic representation of the positions of the BamH1 restriction enzyme recognition sites in exons 4 and 6 of the GBSS gene and the PCR product used as a probe for Southern Blot analyses. ................................................................. 99

Figure 4.4. An autoradiograph of a Southern blot of wheat genomic DNA digested with EcoR1, BamH1 or HindIII and probed with a section of the GBSS-7A gene. ............................................................. 101

Figure 4.5. An autoradiograph of a Southern blot of genomic wheat DNA digested with the restriction enzyme EcoR1 and probed with a section of the GBSS-7A gene............................................................. 102

Figure 4.6. 1% agarose gels showing the wheat genome digested with BamH1 and the approximately 9kbp fragments extracted from Lane 2. ................................................................. 105

Figure 4.7. 1% agarose gel showing that the 9kbp fragments purified from the genome by gel extraction contained GBSS gene sequences. ... 106

Figure 4.8. Diagrammatic representation of the GBSS gene coding regions of five GBSS alleles shown with the exons marked as red rectangles and numbered at the top, and introns as dotted lines.... 111

Figure 5.1. Positions of PCR primers GBSS1-GBSS4 on the GBSS gene. The GBSS gene is represented in blue. ............................................. 117

Figure 5.2. Positions of PCR primers on the GBSS gene. ......................... 118

Figure 5.3. 1% agarose gels showing PCR bands amplified from the wheat genome (‘Kulin’) in PCR reactions containing single primers only. 123

Figure 5.4. 1% agarose gels (left) showing homoprimed PCR fragments amplified by primers GBSS1-GBSS10 and the autoradiographs (right) of the Southern blots of those PCR fragments probed with a GBSS-specific probe. ......................................................... 125
Figure 5.5. 1% agarose gels showing the homoprimed PCR bands produced by using primers GBSS3-GBSS10 in isolation over an annealing temperature gradient from 58.3°C to 71.9°C. .......................... 127

Figure 5.6. 1% agarose gel showing PCR bands produced in the presence of a single or two primers from the set GBSS7-GBSS10. ............. 129

Figure 5.7. Polyacrylamide gel showing PCR bands amplified using the primer combinations GBSS1/GBSS2 and GBSS1/GBSS3. .......... 131

Figure 5.8. Identification of PCR bands amplified from the GBSS genes by hybridisation to a GBSS-specific probe. ............................. 133

Figure 5.9. 1% agarose gel showing the PCR amplification using primers GBSS1/GBSS2 from a template of a single DNA fragment isolated by capillary electrophoresis. ................................. 135

Figure 5.10. PCR with T7/Sp6 primers showing clones containing different sized inserts. ......................................................... 136

Figure 5.11. Dot Blot of individual clones containing insert PCR fragments amplified by the GBSS1/GBSS2 and GBSS1/GBSS3 primers, probed with a GBSS-specific probe. ................................. 137

Figure 5.12. Sequence of cloned PCR fragments, identified by hybridisation to be related to the GBSS genes, but not similar in sequence. ........ 138

Figure 5.13. 1% agarose gels showing the PCR bands amplified from the wheat genome by combinations of the primers GBSS3-GBSS6. ........ 140

Figure 5.14. 1% agarose gels showing the PCR bands amplified from the wheat genome by combinations of the primers GBSS7-GBSS10. ...... 141

Figure 5.15. 1% agarose gel showing how groups of clones were tested for the inserts they contain using two sets of PCR primers. ............ 143

Figure 5.16. 1% agarose gels showing PCR testing of clones for their inserts. .. 144

Figure 5.17. 1% agarose gel showing the approximately 2kbp PCR fragment amplified from ‘Kulin’ using GBSS4/GBSS6 PCR primers and digested with restriction enzymes. ................................. 145

Figure 5.18. 1% agarose gel comparing a mixed population of plasmids containing a variety of inserts with control plasmids containing inserts of a known size. ............................................. 147

Figure 5.19. 1% agarose gels showing a mixture of linearised plasmids containing variously sized inserts. ................................. 148

Figure 5.20. Comparison of three DNA sequences from GBSS genes. ....... 150

Figure 5.21. A protein sequence alignment in the region of the GBSS-7A protein from exon 4 to the stop codon ending exon 11. ............ 152
Figure 5.22. Comparison of three DNA sequences from GBSS genes. .......... 154

Figure 5.23. A protein sequence alignment in the region of the GBSS-7A protein from exon 4 to the stop codon ending exon 11. ............... 156

Figure 5.24. DNA sequence comparison between a PCR fragment predicted to have been amplified from the null GBSS-7A gene of the wheat variety ‘DHW174’ and the wild-type GBSS-7A gene from ‘Chinese Spring’. ................................................. 158

Figure 5.25. DNA sequence comparison between the first and second PCR fragments predicted to have been amplified from the null GBSS-7A gene of the wheat variety ‘DHW174’ with the wild-type GBSS-7A gene from ‘Chinese Spring’. ................................................. 159

Figure 5.26. A diagram of the potential PCR primer binding positions of primers GBSS1-GBSS3 on the GBSS-7A gene as represented by the black line from the start codon at the 5’ end to the stop codon at the 3’ end. ................................. 162

Figure 6.1. Diagram of restriction site mediated PCR. ...................... 168

Figure 6.2. (Top) A diagram of the plasmid used as the PCR template to optimise the restriction site mediated PCR method. (Bottom) A 1% agarose gel showing the restriction site mediated PCR amplification from the plasmid template. ........................... 167

Figure 6.3. 1% agarose gel showing optimised restriction site mediated PCR amplification from the plasmid template. ......................... 173

Figure 6.4. A 1% agarose gel showing the PCR bands amplified from a template of EcoR1-digested ‘Kulin’ genomic DNA with EcoR1 linkers attached, using only the Gus 3’ primer. ...................... 175

Figure 6.5. 1% agarose gel showing PCR bands amplified across a serial dilution of DNA template concentrations. .......................... 176

Figure 6.6. 1% agarose gel showing PCR bands amplified from the wheat genome using restriction site mediated PCR with primers Gus 3’/GBSS2 and Gus 3’/GBSS4........................................ 177

Figure 6.7. A 1% agarose gel showing PCR bands amplified from a template of Sal1-digested ‘Kulin’ wheat genomic DNA with Sal1 linkers attached, using restriction site mediated PCR with the Gus 3’ primer in combination with one of the locus-specific primers GBSS-Rev, GBSS-4A, GBSS-7A, GBSS-7D. ...................... 179

Figure 6.8. 1% agarose gel showing PCR bands amplified from the wheat genome using restriction site mediated PCR with primers Gus 3’/GBSS2 and Gus 3’/GBSS4........................................ 180
Figure 6.9. A 1% agarose gel showing the different PCR banding patterns after varying the primer concentration and thermal cycling conditions of reactions during restriction site mediated PCR. ........ 182

Figure 6.10. 1% agarose gels showing the PCR bands produced by nested-PCR at various template DNA concentrations. .................................. 184

Figure 6.11. 1% agarose gel showing the PCR bands amplified from a template of ‘Kulin’ DNA with EcoR1 linkers using an Elongase kit and Gus 3'/GBSS-7D primers.................................................. 185

Figure 6.12. Diagrams predicting the positions of the Sal1 restriction sites on the GBSS genes based on the sizes of PCR bands thought to have been amplified from the GBSS 7A and GBSS 7D loci. .......... 187

Figure 6.13. Two possible secondary structures for the GBSS 4A promoter based on the positions of inverted repeats. ................................. 189

Figure 7.1. Two DNA sequences, the genomic and cDNA sequences from ‘Chinese Spring’ wheat (Murai et al., 1999; Clark et al., 1999) showing the positions of the microsatellite marker primers WxD1b1F1, Sun1F and Wx97R2/Sun1R. ................................. 192

Figure 7.2. A 10% polyacrylamide gel showing the XSun1 microsatellite marker bands amplified from a selection of wheat lines by the primers Sun1F/Sun1R. .............................................. 195

Figure 7.3. Three autoradiograph pictures of 6% denaturing polyacrylamide gels displaying the microsatellite marker bands amplified from a selection of wheat varieties by the primers WxD1b1F1/Sun1R. .. 196

Figure 7.4. An autoradiograph picture of a 6% denaturing polyacrylamide gel showing the three differently sized bands amplified from wild-type 7A GBSS genes. .................................................. 197

Figure 7.5. An autoradiograph picture of a 6% denaturing polyacrylamide gel comparing sizes of bands amplified from the 7A locus of wheat varieties with High-FSV and Medium-FSV starch characteristics... 198

Figure 7.6. Every coloured rectangle on the graph represents one wheat variety about which the Xsun1 and (TGCCG)n microsatellite data is known. ................................................................. 200

Figure 7.7. Autoradiograph of a 6% denaturing polyacrylamide gel displaying the XSun1 microsatellite marker bands amplified from a selection of wheat lines containing the four combinations of null and wild-type 7A and 7D alleles by the primers WxD1b1F1/Sun1R. .. 203

Figure 7.8. Two tables listing all the individuals in the Cross 98Y215, their identifying numbers, FSV scores, and GBSS characteristics. ...... 205
**Figure 7.9.** DNA sequences of the 3’ ends of the GBSS cDNA (Clark *et al.*, 1991) and from the wild-type alleles at each locus (Murai *et al.*, 1999).

**Figure 7.10.** Graph of the percentage amylose produced in wheat, per copy the GBSS gene present.

**Figure 7.11** Graph of the FSV values for a selection of wheat varieties containing zero, two, four or six copies of the GBSS genes.
LIST OF TABLES

Table 2.1. Descriptions of the wheat varieties used in the study including pedigrees and GBSS characteristics………………………………. 39

Table 2.2. DNA sequences of PCR primers. ........................................... 48

Table 3.1. Wheat varieties used in the initial survey, their GBSS characteristics and their pedigrees. ................................. 60

Table 3.2. The number of repeats of the microsatellite in wheat varieties with either medium-FSV or high-FSV. ................................. 77

Table 3.3. Wheat varieties and the GBSS 7A allele they contain as differentiated by the number of repeats (3 or 4) of a (TGCCG)_n microsatellite. ............................................................. 80

Table 5.1. Compares the temperature at which no homoprimed bands were amplified, with the melting temperature of the primers. ............ 126

Table 7.1. Table listing the wheat varieties for which molecular marker data is known at both the (TGCCG)_n and Xsun1-7A microsatellite loci. 199

Table 7.2. Table showing molecular data about the Xsun1-7A and (TGCCG)_n microsatellite loci from a selection of wheat varieties that display either high-FSV or Medium-FSV starch characteristics. .............. 201

Table 7.3 comparing the expected and actual genotypes for the 186 doubled-haploid offspring from the first DHWX-174/2* Arrino cross tested………………………………………………………… 202

Table 7.4 Comparing the expected and actual genotypes for the 69 doubled-haploid offspring from the second DHWX-174/2* Arrino cross tested………………………………………………………… 204
ABBREVIATIONS

ANGISAustralian national genome information service
bpbase pair
BSAbovine serum albumin
cDNacomplimentary deoxyribonucleic acid
cmcentimetre
CTABcetyltrimethylammonium bromide
dATP2’deoxyadenosine 5’ triphosphate
dCTP2’deoxyctosine 5’ triphosphate
dGTP2’deoxyguanosine 5’ triphosphate
DNadeoxyribonucleic acid
dNTPdeoxyribonucleotides
dTTP2’deoxythymidine 5’ triphosphate
EDTAethylene diamine tetra-acetic acid (disodium salt)
GBSSgranule-bound starch synthase
ggram
GUSβ-glucuronidase
hrhour
kbpkilobase pair
Llitre
LB Luria-Bertani
m meter
Mmolar
mA milliamps
MALDI TOFmatrix assisted laser desorption/ionisation time of flight
mgmilligram
MgCl₂magnesium chloride
minminute
mlmillilitre
mRNAmessenger ribonucleic acid
MWmolecular weight
ngnanogram
°C degree centigrade
paper annum
PAGEpolyacrylamide gel electrophoresis
PCRpolymerase chain reaction
pmolpicomole
RNAribonucleic acid
RNAselase
rpmrevolutions per minute
secsecond
SDNsodium dodecyl sulphate
ssDNAsingle stranded deoxyribonucleic acid
TAETris-acetate EDTA
TBETris-borate EDTA
TETris-EDTA
TEMED N,N,N’,N’-tetramethylaminomethane
TrisTris-hydroxymethylaminomethane
Uenzyme units
UVultra violet
μμg
μlmicrolitre
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>µM</td>
<td>micromolar</td>
</tr>
<tr>
<td>V</td>
<td>volt</td>
</tr>
<tr>
<td>WADA</td>
<td>Western Australian Department of Agriculture</td>
</tr>
<tr>
<td>w/v</td>
<td>weight:volume ratio</td>
</tr>
</tbody>
</table>
LIST OF COMPANY ADDRESSES

<table>
<thead>
<tr>
<th>Company</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applied Biosystems Inc</td>
<td>Foster City, CA, USA</td>
</tr>
<tr>
<td>BioRad Laboratories</td>
<td>Hercules, CA, USA</td>
</tr>
<tr>
<td>Biotech International</td>
<td>Perth, WA, Australia</td>
</tr>
<tr>
<td>Bresatec</td>
<td>Adelaide, SA, Australia</td>
</tr>
<tr>
<td>Fuji Photo Film</td>
<td>Elmsford, NY, USA</td>
</tr>
<tr>
<td>Geneworks</td>
<td>Adelaide, SA, Australia</td>
</tr>
<tr>
<td>Gibco BRL</td>
<td>Melbourne, VIC, Australia</td>
</tr>
<tr>
<td>Hoefer Scientific Instruments</td>
<td>San Francisco, CA, USA</td>
</tr>
<tr>
<td>Keygene</td>
<td>NL-6700 Wageningen, The Netherlands</td>
</tr>
<tr>
<td>Kodak Australia</td>
<td>Coburg, Victoria, Australia</td>
</tr>
<tr>
<td>New England Biolabs</td>
<td>Beverly, MA, USA</td>
</tr>
<tr>
<td>Perkin-Elmer Corporation</td>
<td>Foster City, CA, USA</td>
</tr>
<tr>
<td>Progen Industries Inc</td>
<td>QLD, Australia</td>
</tr>
<tr>
<td>Promega Corporation</td>
<td>Annandale, NSW, Australia</td>
</tr>
<tr>
<td>Qiagen</td>
<td>VIC, Australia</td>
</tr>
<tr>
<td>Sigma Chemical Co.</td>
<td>Castle Hill, NSW, Australia</td>
</tr>
</tbody>
</table>
Chapter 1. Introduction and literature review.

1.1 General introduction to the project.

Wheat is the most important crop species grown worldwide. More than 200 million hectares of wheat are cultivated annually on which is produced almost 600 million tonnes of grain (FAO; [http://www.fao.org/](http://www.fao.org/)). This provides not only a staple food source in the human diet, but also supplies many industrial applications. The world trade in wheat is important for economic stability, and the ability to grow wheat is a valuable resource. Wheat is a major crop grown in Australia with approximately 23 million tonnes produced annually (FAO; [http://www.fao.org/](http://www.fao.org/)). One-quarter of this grain is used domestically and it fulfills all of the national requirement for wheat, allowing for the remaining three-quarters to be exported. Thus, Australia’s wheat industry supplies both a major domestic food market, and it also supports an export industry worth almost 2 billion dollars.

Over the past 40 years, there has only been a small increase in the area of land on which wheat is cultivated worldwide, but the tonnage of wheat grain produced on this land has tripled as a result of improved farming practices and the development of better wheat varieties (Marshall et al., 2001). There is great potential to improve wheat through breeding, not only to increase grain yields by developing pest resistances and tolerance to poor environments, but also by changing grain starch and protein properties. This could increase its usefulness for both current and novel applications. Such improvements are vital for Australia to continue to meet the quantity and quality demands of an international wheat market.

The overall aim of the research described in this thesis was to investigate the genomic organization of the granule-bound starch synthase (GBSS) (GBSSI, EC 2.4.1.21) genes in wheat (*Triticum aestivum* L.). The GBSS genes play a major part in determining functionality in foods such as Japanese Udon noodles (Oda et al., 1980; Huang and Morrison, 1988; Miura...
and Tanii, 1994; Batey et al., 1997; Zhao et al., 1998). Improving starch quality and functionality depends on understanding the action and effects of these genes. In this Chapter an introduction to the genomic organization of the GBSS genes is presented, together with the role of the GBSS enzymes in starch biosynthesis, and their effects on starch quality.

1.2 The Australian Standard White Noodle Wheat industry in Western Australia.

About one third of Australia’s total wheat production is used overseas to make Asian noodles of various kinds (Crosbie et al., 1990). In Western Australia, the agricultural production of wheat flour from ASWNW (Australian Standard White Noodle Wheat) varieties is a specialty export market worth more than $200M pa (Garlinge, 1996). The grain is exported mainly to Japan and South Korea to make Japanese white salted Udon noodles, and the ASWNW classification was introduced specifically to support this export market for producing Udon noodles.

Udon noodles have specific eating qualities resulting from the quality of the wheat flour starch they are made from (Crosbie, 1991; Batey et al., 1997; Zeng et al., 1997). The flour from ASWNW varieties, such as ‘Arrino’, ‘Cadoux’, ‘Calingiri’, ‘Eradu’ and ‘Gamenya’, has superior starch qualities making it suitable for Udon noodle production (Crosbie, 1991, Zhao et al., 1998). The flour contains between 9.5% and 11.5% protein (V. Reck, DAWA, pers. comm.). The wheat is soft-grained, producing flour of fine particle size and less-damaged starch, which contributes to the smoothness required in the ‘mouthfeel’ test for noodle quality. Japanese cooked noodles are assessed for appearance (brightness, colour and surface), texture and taste. High-quality Udon noodles are soft, elastic in texture, and slightly creamy with no specks. Australian wheat is white-grained, which enables the noodles to be a good, clear, light colour. The flour also has good starch-swelling characteristics, moderate dough strength and good dough extensibility (V. Reck, DAWA, pers. comm.). The good starch-swelling
characteristics of the flour result, for the most-part, from containing relatively less of the starch amylose than other varieties (22-23% compared to 25%), a property controlled by the GBSS genes (Nelson and Rines, 1962; Garlinge, 1996). ASWNW varieties which have these starch characteristics command a higher market price and developing new noodle-wheat varieties is an important aim for wheat breeders.

The specific starch quality required for Udon noodles is an example of the increasing international demand for starches with a specific functionality. Such starches are used both for food and for industrial uses such as pharmaceutical fillers, adhesives and substrates for microbial fermentation (Ball et al., 1996; Bligh, 1999). Different end uses require starch with different physical and chemical properties (Ellis et al., 1998, Graybosch, 1998). For example, flour containing less amylose is required to make some oriental noodles and may extend the shelf-life of bread, whilst flour with more amylose is useful for making batters for fried foods because it has good film-forming properties (Graybosch, 1998, Bligh, 1999). Traditionally, starches have been modified chemically to provide the required starch properties, but development of wheat varieties with improved natural starch qualities is now preferred (Ball et al., 1996; Bligh, 1999). Improved starches eliminate the need for post-mill chemical modification of the starch, reducing both processing costs and possible contamination of the environment. Grain from starch-improved wheat varieties has added-value. One aim of this research was to characterise the various GBSS genes present in Western Australian wheat germplasm and the effects these genes have on starch quality to help wheat breeders develop new varieties with improved starch characteristics.

1.3 Starch biosynthesis.

Wheat endosperm starch consists of two major components, amylose and amylopectin. Since the amylose content of starch affects its physical properties, it is important to consider starch
structure and biosynthesis. Amylopectin makes up approximately 75% of the dry weight of starch, with amylose making up the remaining 25%. The two molecules are similar, both being made of long chains of glucose moieties linked by α-(1,4) linkages (Figure 1.1), but amylopectin is a larger polymer with a molecular weight of approximately $10^9$ and it is highly-branched with α-(1,6) linkages, whilst amylose is a shorter, largely unbranched molecule with a molecular weight of approximately $10^6$ (Davis, 1994; Morell et al., 1995, Ball et al., 1996).

When grain-filling begins, sucrose produced by leaf chloroplasts is transported in phloem to the immature grain. In the endosperm cells the sucrose is degraded to the common metabolic subunit glucose-1-phosphate. This molecule is then transported into the starch storage organelles, the amyloplasts. Starch biosynthesis occurs in the amyloplasts as outlined in Figure 1.2. The first dedicated step is conversion of glucose-1-phosphate to the specific starch precursor ADP-glucose by the enzyme ADP-glucose pyrophosphorylase (Keeling et al., 1988; Preiss et al., 1994). The ADP-glucose units are then polymerized into the α-1,4-glucosyl chains of starch by the two starch synthetic enzymes; soluble starch synthase that synthesizes amylopectin and granule-bound starch synthase that synthesizes amylose (Salehuzzaman et al., 1993; Nakamura et al., 1995). The linear amylopectin chain is then modified by other enzymes including branching enzymes that transfer short α-1,4-glucosyl chains of amylopectin to an α-(1,6) configuration creating the branched amylopectin molecule (Boyer and Preiss, 1978; Guan and Preiss, 1993; Rahman, 1995).
Figure 1.1. Molecular structures of starch polymers. 
(Whistler and Paschall, 1965)
After soluble starch synthase polymerises the linear amylopectin molecule, it is then processed by other enzymes into its mature branched form as shown in Figure 1.3. When the linear amylopectin molecules reach a certain length, the starch branching enzymes cleave the α-1,4-glucosyl chains and re-attach these fragments in an α-(1,6) configuration to create branched chains (Boyer and Preiss, 1978; Guan and Preiss, 1993). Starch branching enzymes require a defined length of amylopectin polymer before the active site can attach to the molecule and cleave off a glucosyl chain (Preiss et al., 1994; Ball et al., 1996). This length of polymer is approximately 6nm or 18 glucose residues (Ball et al., 1996). Thus, when the amylopectin chains reach this threshold length, there is a mass of random branching at that point of the
amylopectin chain. At this stage, the debranching enzymes remove any branches that are not tightly packed together (Ball et al., 1996). These debranching enzymes also have a requirement for space to act and thus only remove widely spaced branches. The soluble starch synthase enzymes then continue to polymerize the amylopectin chains until they are of sufficient length to again be acted upon by branching enzymes.

**Figure 1.3.** A diagram showing the action of the branching and debranching enzymes on linear amylopectin to produce mature branched amylopectin. Enzymes are shown to the right in bold.

The amylopectin molecules grow in alternating layers of tightly packed branch points and long linear sections of polymer (Thompson, 2000). The amylopectin molecules grow outward from a central point, predicted to be a core starch starter molecule amylogenin, like the precursor molecule of glycogen, glycogenin (Lehninger et al., 1993; Keeling et al., 1994), to fill the starch granule as depicted in Figure 1.4.
Figure 1.4. A drawing representing the structure and organisation of a starch granule. The picture shows the amylopectin growing and branching out from a central point. The centre of this drawing has been left blank, but it is thought to be a starch-starter molecule, amylogenin. The amylose is deposited within the gaps of the amylopectin matrix. The long linear molecules of amylose form single helices whilst the linear sections of the amylopectin form double helices. Reprinted from Davis (1994).

The starch granules form inside the amyloplasts. The starch granules are filled with amylopectin to a predetermined maximum size and shape, either a large lenticular A-type granule, a small spherical B-type granule or a very small C-type granule, although the mechanisms that determine the different starch granule types are not known (Parker, 1985).

As the amylopectin chains grow, the granule-bound starch synthase (GBSS) enzymes are attached to the amylopectin matrix and are engulfed within the granule as the amylopectin is deposited around them (Tatge et al., 1999). GBSS bound within the starch granule constitutes the major protein in the starch granule (Rahman et al., 1995). GBSS then polymerises long linear amylose molecules in the spaces between the amylopectin molecules (Jane et al., 1992; Tatge et al., 1999). As GBSS synthesizes amylose within the starch granule after the
amylopectin has been laid down, the growing amylose polymer is spacially and temporally separated from the branching enzymes in the soluble fraction outside the starch granule and thus, amylose remains a linear molecule (Tatge et al., 1999). The long linear amylose molecules form single helices, whereas the linear sections of amylopectin form double helices with adjacent amylopectin molecules (Davis, 1994; Ball et al., 1996).

To summarise the production of starch, the major structures of the starch granule are formed by biosynthesis of amylopectin by soluble starch synthase enzymes, and the modification of linear amylopectin by branching and debranching enzymes. GBSS enzymes are deposited within the amylopectin matrix and synthesise amylose in the gaps between the amylopectin molecules. Whilst the action of these enzymes account for the polymerization of the major starch structures, it is now predicted that there is also a range of other enzymes that assist in production of starch. One example is the enzyme α-amylase that is usually considered to be involved in starch degradation during germination, but new evidence suggests that it may also be involved in starch biosynthesis, degrading the short amylopectin chains trimmed by debranching enzymes back into ADP-glucose for use by GBSS (Nakamura, 1996). These enzymes will not be discussed further, however once effects of the core starch biosynthetic enzymes are fully understood, the effects of these associated enzymes will need to be investigated.

1.4 Potential to manipulate starch biosynthesis.

The biochemical pathway for starch biosynthesis has several enzymes that could potentially be manipulated to modify starch as reviewed by Bligh (1999). Firstly, the manipulation of the numbers and types of starch granules present in wheat starch could improve starch for specific uses. Native wheat starch consists of both the large lenticular A-type granules and the 10-times smaller spherical B-type granules (Parker, 1985). In developing grain, A-type granules
are initiated 5 days post-anthesis and their final numbers are reached 7 days later when cell division ceases. B-type granules are formed from 14 days post-anthesis and may bud from the A-type granules (Parker, 1985). The two types of starch granules are suitable for distinct and separate uses. The larger wheat starch granules are suitable to make carbonless copy paper, whereas rice starch which contains only small starch granules may be useful as a fat substitute because it gives a comparable mouthfeel (Bligh, 1999). There are also tissue-specific differences between the starch granules found in endosperm and those found elsewhere in the plant (Nakamura et al., 1998). It may be possible to genetically manipulate wheat endosperm to express leaf-specific starch synthases and improve endosperm starch granule characteristics. Thus, controlling the sizes of starch granules would benefit those industries that currently separate the starch granules by size for specific uses. Also, increasing the number of starch granules may increase the yield of starch with obvious economic benefits.

Secondly, there is great potential to manipulate amylopectin structure. Manipulating the activity of starch branching and debranching enzymes will change the branching structure of the amylopectin and its physical properties (Yamamori et al., 2000). Reducing the branching enzyme activity would result in more linear amylopectin which can act more like amylose. High-amylose starches have film-forming properties desirable for fried food coating batters and some forms of plastics (Bligh, 1999). Reducing the debranching enzyme activity would result in the highly branched mutant amylopectin, phytoglycogen, which is similar to animal glycogen. It may also be possible to vary the length of the starch chain required by starch branching enzymes so that the amylopectin branches sooner or later than in the native form. The degree of branching affects the physical and chemical properties of starch, which determine its uses and value.

Thirdly, manipulating the activity of the enzyme GBSS results in variation in the amylose content of the starch (Miura et al., 1994; Fujita et al., 1998; Sasaki et al., 2000). The amylose
content of wheat starch affects the way the starch swells in water and thus the texture of the final food product (Zhao et al., 1998). The maximum percentage of amylose in the starch granule is limited by the space between the amylopectin molecules (Flipse et al., 1996; Tatge et al., 1999). Reducing the amount of amylose allows the starch granule to swell more, giving a suitable texture to foods such as Udon noodles (Zhao et al., 1998). It is because of this relationship between amylose content and starch swelling potential that the manipulation of GBSS enzymes in wheat is of interest.

1.5 Control of GBSS enzyme activity in wheat.

The enzyme GBSS deposits amylose in gaps within the amylopectin matrix of the starch granule. The GBSS enzymes have high activities and amylose production appears to be limited by the space available within the amylopectin matrix (Flipse et al., 1996; Tatge et al., 1999). To reduce the amount of amylose in starch, the amount of GBSS enzyme must be reduced (Yamamori et al., 1992, Yamamori and Quynh, 2000). A reduction in the amount of active GBSS enzyme in the starch matrix does reduce the amount of amylose produced, but the gene dosage effects are not linear because the enzymes are very active and are limited by their environment in vivo. The GBSS enzymes will produce more amylose if given adequate space and resources (Nakamura et al., 1993; Flipse et al., 1996; Yamamori and Quynh, 2000). Sasaki et al. (2000) showed that a single copy of the GBSS gene in the wheat genome can produce starch containing 7.2-7.7% amylose. Two gene copies can produce starch containing 13.5-15.3% amylose and three GBSS gene copies produce starch containing 18.3-20.3% amylose. Four GBSS gene copies produce the 22-23% amylose found in noodle-wheats, whilst five gene copies produce starch containing 25% amylose indistinguishable from the wild-type starch produced by six gene copies (Miura et al., 1994).
The synthesis of GBSS enzymes is controlled by the GBSS genes that encode them. Gene expression is switched on 10 days after anthesis and continues until grain-filling is complete (Vrinten et al., 1999). The GBSS gene expression is not regulated by feedback mechanisms and the activity of GBSS enzyme cannot therefore be controlled in this way. To reduce the amount of amylose in wheat starch the total amount of active GBSS enzyme expressed must be reduced by the presence of null GBSS alleles that do not express an active enzyme.

1.6 Genomic location of GBSS loci and common nomenclature.

The domestic wheat *Triticum aestivum* L., is allohexaploid and contains three genomes A, B and D (AABBDD). There are three loci for the GBSS genes in the hexaploid wheat genome (Nakamura et al., 1993). These loci were originally all on the short arms of the Group 7 chromosomes, but a translocation moved the 7B locus to the long arm of chromosome 4A (Figure 1.5) (Naranjo et al., 1987; Chao et al., 1989; Liu et al., 1992). The section of chromosome translocated from chromosome 7B to 4A contains at least 45 cM of sequence with the GBSS locus at least 15 cM from each end of the translocated fragment (Liu et al., 1992). Thus, the three GBSS loci are now located on chromosomes 4AL, 7AS and 7DS (Chao et al., 1989).

There are several systems of nomenclature for the GBSS genes. GBSS loci, genes and enzymes are also termed *Waxy*, and so the terms GBSS and *waxy* are used interchangeably. The genes at the GBSS loci on chromosomes 7A, 4A and 7D are called *Wx-A1*, *Wx-B1* and *Wx-D1* respectively (Nakamura et al., 1993; Yamamori et al., 1994). Recently this nomenclature has been modified to *Wx-7A*, *Wx-4A* and *Wx-7D* (Murai et al., 1999) to overcome the confusion caused by the translocation of the locus from the B-genome to the 4A chromosome. In this thesis the GBSS loci and genes will be described by the common Australian nomenclature GBSS-7A, GBSS-4A and GBSS-7D.
Current wheat chromosomes containing GBSS genes.

7AS 7AL
XWx-7A

4AL 4AS
XWx-4A

7DS 7DL
XWx-7D

Translocations producing the current wheat 4A chromosome.

Original 5A
Original 4A

Short arm Long arm
translocation

pericentric inversion

translocation

Transient 4A

XWx-7B

Transient 4A
Original 7B

Current 4A

Figure 1.5. Diagrams showing (TOP) the positions of the GBSS (Waxy) loci on chromosomes 7A, 4A and 7D, and (BOTTOM) translocations between chromosomes 4A, 5A and 7B resulting in the structure of the present 4A chromosome. (Liu et al., 1992).
1.7 Molecular structure of the GBSS genes.

The GBSS genes have a highly conserved molecular structure across the three genomes as characterized by Murai et al. (1999). The GBSS gene structure has a 5’ promoter region, 11 exons and 10 introns as depicted in Figure 1.6. Little is known about the structure of the GBSS promoter regions. The lengths of the GBSS genes from start codon to stop codon are between 2781bp and 2862bp. The first 210-213bp of Exon 1 encodes a transit peptide. When the GBSS protein is translated from the mRNA in the cytoplasm, it is a precursor form including a transit peptide that directs the protein to the starch granule where the mature protein is attached to the amylopectin matrix (Murai et al., 1999). Exon 1 also encodes the putative active site of the protein. The remaining exons at the 5’ end of the gene encode the structural part of the protein and possibly the site of attachment to the amylopectin matrix (Vrinten et al., 1999). The dinucleotide sequences at the intron-exon junctions follow the universal GT-AG rule (Murai et al., 1999).

![Figure 1.6](image.png)

**Figure 1.6.** A diagrammatic representation of the molecular organisation of GBSS genes. The 5’ promoter region is shown as a pale blue box while the 3’ untranslated region is shown as a light pink arrow. The 11 exons of the gene are shown as blue boxes, with the 10 introns shown as red lines. (The diagram is to scale for the 7A locus.)

1.8 GBSS alleles.

At each GBSS locus there are several known alleles of the GBSS gene. The alleles can differ by as little as a single base pair, or can differ at many base pairs at different polymorphic loci.
within the gene. If several polymorphic loci exist within the gene, and unless they are all in linkage disequilibrium, then it is appropriate to describe the genes as different gene haplotypes rather than different alleles. However, since little is known about the polymorphic loci within the GBSS genes from each locus, the term ‘allele’ will be used to describe different GBSS genes in this study. It is also understood that genetic polymorphisms may not change the protein sequence or the expression patterns of the gene. However, since the RNA and DNA structural roles, RNA stability role, microRNA signaling role, or epistatic regulatory roles of a polymorphism may not be immediately obvious, no polymorphism can be dismissed as unimportant.

There are both alleles that express an active GBSS enzyme (termed ‘wild-type’) and alleles that do not express an active GBSS enzyme (termed ‘null’). The wild-type GBSS genes contain a functional promoter, a full-length gene copy with a functional active site and express an active GBSS enzyme. Wild-type wheat contains wild-type genes at all three loci and expresses a wild-type complement of GBSS enzymes. The term ‘wild-type’ is often used to describe a single, non-mutated allele, the most common allele in a population of alleles. Since there are potentially many functional alleles of the GBSS genes, and the polymorphisms that distinguish them from each other are not known, the term ‘wild-type’ will be used to describe any and all functioning copies of the GBSS genes. The null GBSS genes do not express an active GBSS enzyme, for a variety of reasons as described later. Though null alleles are very rare, a wheat plant can contain null genes at any of the three loci and a reduced amount of GBSS enzyme is correspondingly produced (Yamamori et al., 1994). Recently wheat varieties have been developed containing null alleles at all three GBSS loci. These plants do not express any active GBSS enzyme and produce only a small amount of amylose, possibly through the action of Soluble Starch Synthase (Nakamura et al., 1995). Starch containing no amylose is obtained from varieties producing no amylose which are termed waxy mutants. There are no known naturally occurring waxy mutants, but waxy wheats have been developed
in various breeding programs (Nakamura et al., 1995). When there exists at least one wild-type and one null gene the wheat is termed a partially waxy mutant.

In the nomenclature of Nakamura et al. (1993) and Yamamori et al. (1994), the wild-type GBSS alleles at each locus present in the majority of their wheat varieties were the first alleles to be described and were termed Wx-A1a, Wx-B1a and Wx-D1a. The first null alleles identified were termed Wx-A1b, Wx-B1b and Wx-D1b, respectively. All alleles discovered since then have been named by other letters progressing through the alphabet in order of discovery (Marcoz-Ragot et al., 2000).

1.9 Wild-type alleles.

The first DNA sequence from a GBSS gene of hexaploid wheat (Japanese cv ‘Chinese Spring’) was a cDNA clone (X57233) described by Clark et al. (1991). This cDNA sequence showed similarities to the GBSS sequences from other species such as rice and barley (X62134, X07932) especially in the coding regions. Clark et al. did not know which locus the cDNA (X57233) had been transcribed from. The work of Hollingsworth (1997) matched the sequence of the cDNA (X57233) to the sequence of the GBSS gene from the 7A locus. This was later confirmed by Murai et al. (1999), but contested by (Shariflou and Sharp, 1999) who believed the cDNA (X57233) was a chimera with a 3’ portion from the 7D locus. One section of the cDNA (X57233) that caused confusion was a 33bp sequence at the junction of the 4th intron. This sequence did not match any of the sequences from the three loci sequenced by Hollingsworth (1997) and later Murai et al. (1999) confirmed this 33bp sequence was not correct. The 33bp sequence reported by Clark et al. (1991) is now accepted as an error and disregarded.
The first GBSS alleles to be described were the wild-type alleles at each locus from ‘Chinese Spring’ wheat (Murai et al., 1999). ‘Chinese Spring’ is distinctly different to the Australian Standard White (ASW) wheat grown in Australia, but is included in the germplasm and is interbred with ASW wheat. Most of the international research into the GBSS genes has been carried out on ‘Chinese Spring’ germplasm that is significantly genetically distant from ASW varieties (Paull et al., 1998). It was not clear how related the wild-type GBSS alleles present in Western Australian varieties were to those described in the literature. For example, is the allele Wx-A1a from the nullisomic-tetrasomic ‘Chinese Spring’ wheat ‘N7D/T7B’ as described by Murai et al. (1999) the same allele that is present in the Australian noodle wheat variety ‘Eradu’? It is clear from personal communications with wheat breeders such as Robin Wilson (WA Department of Agriculture, Perth) and Peter Sharp (University Sydney, Cobbitty) that the genetic distance and genetic relationships between Australian and international wheat varieties is not well understood. One of the aims of this research was to assess the genetic diversity of a sampling of Australian wheat varieties at the GBSS loci to determine the levels of diversity and to compare the GBSS alleles present in Australian germplasm with those reported internationally.

The genomic sequences of the wild-type GBSS alleles (Wx-A1a: AB019622, Wx-B1a: AB019623 and Wx-D1a: AB019624) from all three loci of the nullisomic/tetrasomic variety ‘Chinese Spring’ were described by Murai et al. (1999). The alleles share a similar molecular structure (Figure 1.6). However they differ slightly in length, mostly due to polymorphisms between the three loci within the intron regions. Figure 1.7 compares the DNA sequences of the wild-type alleles from the three loci as described by Murai et al. (1999).
Figure 1.7  An alignment of the genomic DNA sequences of the wild-type GBSS genes from nullisomic/tetrasomic ‘Chinese Spring’ wheat, as presented by Murai et al. (1999). The sequence from each locus is marked at the left with 7A, 4A or 7D. The sequences begin with the start codon and finish at the stop codon highlighted in red. The exon sequences are capitalised in black whilst the intron sequences are in lowercase and dark blue. Dotted lines represent sections of sequence that match that from the 7A locus, while gaps have been made in the sequences to show the positions of insertions or deletions. The sequence encoding N-terminus of the mature GBSS protein is highlighted in pale grey, separating the sequence encoding the transit peptide from that encoding the N-terminus of the mature GBSS protein.
The lengths of the genes from start codon to stop codon were 2781bp in the wild-type 7A allele, 2794bp in the wild-type 4A allele and 2862bp in the wild-type 7D allele. The differences in length were due to deletions and insertions in the intron regions. The only length polymorphism in the exons was a trinucleotide insertion (CAA) in the transit peptide region of the wild-type 4A allele. The DNA sequences from the three alleles were 95.6-96.3% similar in mature protein regions, 88.7-93.0% similar in transit peptide regions and 70.5-
75.2% similar in the introns (Murai et al., 1999). The proteins translated from these sequences were 96.5-97.4% conserved.

As well as these Wx-a alleles, a few other wild-type alleles have also been identified, both in hexaploid bread wheats and in tetraploid durum wheats, which lack the D genome. These have been named alphabetically from ‘c’ onwards. The alleles were identified by differences in protein structure allowing electrophoretic distinction of the novel proteins, or by recognizable differences in GBSS activity. At the GBSS-4A locus there are four other wild-type alleles besides the most common Wx-B1a. Rodriguez-Quijano et al. (1998) identified a novel Wx-B1c’ allele, and Wx-B1d was described by Yamamori et al. (1995) and Urbano et al. (1996). Yamamori and Quynh (2000) described a Wx-B1e allele with a protein mobility similar to the GBSS-7D protein and suggested it was the same allele as Wx-B1c’. Nieto-Taladriz et al. (2000) reported the discovery of Wx-B1f and reported that the presence of the alleles Wx-B1c’ and Wx-B1f resulted in starch containing more amylose than Wx-B1a. At the GBSS-7A locus Wx-A1e was described by Yamamori et al. (1995). The 7D locus seems the least polymorphic with fewer alleles described. It is not certain which of these various alleles are present in Australian wheat varieties as there are no DNA markers to distinguish the alleles and no major screening for these novel GBSS alleles has been done.

To date, there has been little work published on the promoter sequences of the GBSS genes in wheat. Some 5’ untranslated region sequence is available, approximately 10-30bp for most published sequences and 203bp from one GBSS 4A gene (AF286320). Accessing the GBSS promoter sequences was one aim of this research.
1.10 Null alleles.

Null alleles of the GBSS genes are rare and until recently only one null allele from each locus had been described (Vrinten et al., 1999; Marcoz-Ragot, 2000), but a second null 4A allele was discovered recently (M. Shariflou, pers. comm.). Considering the level of redundancy offered in the allohexaploid wheat genome, one might expect more null alleles to have evolved (Vrinten et al., 1999) and their rarity may be testament to the important role GBSS plays in wheat. None of the null GBSS alleles express an active GBSS enzyme. The null GBSS alleles were detected both by their lack of GBSS protein expression as visualized by 2-Dimensional PolyAcrylamide Gel Electrophoresis (2D PAGE) and by the discernable reduction in the level of amylose starch in the endosperm.

The first null allele described was the null 4A allele found homozygously in the ASWN wheat varieties grown in WA for Udon noodles. The null 4A allele does not express either mRNA or protein. The reason for the lack of expression from this allele is that a large deletion of the DNA in the region of the 4A chromosome containing the GBSS locus has removed the entire gene from the chromosome (Vrinten et al., 1999). The lack of the entire gene sequence from the null 4A allele facilitated the design of molecular markers to differentiate the null and wild-type alleles at this locus. The PCR primers that amplified from the wild-type 4A allele couldn’t bind to DNA at the null 4A allele and so did not yield an amplification product (Briney et al., 1998). Even though no DNA sequence encoding a GBSS gene exists at the GBSS 4A locus of the null 4A allele, this allele is treated as any other allele for molecular studies.

The second null 4A allele recently discovered by M. Shariflou (pers. comm.) is different from the first null 4A allele described above since PCR primers bind to and amplify from it. The reason for the lack of GBSS enzyme activity from this second null 4A allele is not yet known,
nor is it known how widely spread this allele is in the Australian wheat germplasm. This allele was not discovered until after the research in this project was completed and it will only be referred to again in the final discussion. All other references to the ‘null 4A allele’ will relate to the first null 4A allele described above.

The null 7A allele was discovered in the variety ‘Kanto 107’ (Yamamori et al., 1992; Nakamura et al., 1993) and characterised by Vrinten et al. (1999) (AF113843). This allele has been introduced into Australian varieties from ‘Kanto 107’ (P. Sharp, Pers. Comm.) and is the null 7A allele investigated in this thesis. The null 7A allele contains a 23bp deletion, with a 4bp insertion of filler DNA at the end of exon 1 and beginning of intron 1 near the 5’ end of the coding region of the gene (Vrinten et al., 1999) as shown in Figure 1.8. When the null 7A allele is transcribed into mRNA the deletion affects the splicing of the RNA transcript such that a cryptic splice site is used and the mature mRNA transcript loses 117bp of sequence. This deletion removes the transit peptide cleavage site and putative ADP-glucose binding site (KTGGL). The transit peptide is affected so the enzyme is not transported to the amyloplast properly and no null 7A GBSS protein is found in the starch granule (Nakamura et al., 1993).

The null 7D allele was found in the variety ‘Bai Huo’ and was also characterised by Vrinten et al. (1999) (AF113844). This allele has been introduced from ‘Bai Huo’ into Australian varieties (P. Sharp, pers. comm.) and is the null 7D allele investigated in this thesis. The null 7D allele has a 588 bp deletion at the 3’ end of the gene, complemented by a novel 12 bp insertion at this site (Vrinten et al., 1999) as shown in Figure 1.9. This forms a premature stop codon such that the resultant protein lacks 30 amino acids (aa) at the C-terminus end, but gains a 5 aa substitution before the stop codon. While the null 7D allele transcribes a mature RNA transcript, the levels of expression are reduced and no GBSS protein resulting from this transcript is found in wheat endosperm, suggesting that the transcript is not translated or that the truncated protein is degraded (Vrinten et al., 1999).
Figure 1.8.  (TOP) A diagrammatic representation of the GBSS genes showing the position of the mutation in the null 7A allele at the junction of exon 1 and intron 1 (boxed in black). The 5’ promoter region is shown as a light blue box while the 3’ untranslated region is a light pink arrow. The 11 exons of the gene are shown as blue boxes, with the 10 introns as red lines. (The diagram is to scale for the 7A locus.)

(BOTTOM) The DNA sequence polymorphism between the wild-type 7A and null 7A alleles in the region of the junction between exon 1 and intron 1 boxed at the top of the Figure. The exons are shown capitalised in blue and the introns are shown in lower case and red. In exon 1 the position of the junction of the transit peptide and the mature GBSS protein is marked with an arrow. The start of the deletion mutation of the null7A allele is also marked with an arrow. The 4 bp insertion is underlined and the deletion is shown with dashes. As the deletion removes the usual splice site for intron 1, an alternate cryptic splice site is used as marked with an arrow. RNA processing at this cryptic splice site interrupts the junction of the transit peptide with the mature protein and truncates the mature protein disrupting the putative active site within exon 1.
Figure 1.9. (TOP) A diagrammatic representation of the GBSS genes showing the position of the mutation in the null 7D allele in exon 11 and the 3’ untranslated region boxed in black. The 5’ promoter region has been shown as a pale blue box while the 3’ untranslated region is shown as a pale pink arrow. The 11 exons of the gene are shown as dark blue boxes, with the 10 introns shown as red lines. (The diagram is to scale for the 7A locus.)

(BOTTOM) A diagrammatic comparison of the structures of the wild-type 7D and null 7D alleles in the region of the null 7D deletion mutation in exon 11 and the 3’ untranslated region as shown in Vrinten et al. (1999). The diagram is not to scale. The structures of the wild-type 7D genomic DNA and cDNA are shown as the upper lines with the lower lines representing the genomic and cDNA structures of the null 7D allele. Intron 10 is coloured red, exons 10 and 11 are shown in dark blue, the 3’ untranslated region is pale pink, and the 12 bp of filler DNA is coloured pale grey. The diagram shows how the null gene lacks 588 bp of exon 11 and the 3’untranslated region, but gains 12 bp of filler DNA. The null cDNA contains 261 bp from the 3’ untranslated region before the poly-A tail.
When the research described here was started, none of the GBSS alleles had been sequenced, thus the DNA polymorphisms between the alleles and the reasons for the lack of expression from the null alleles were not known. An aim of this research was to identify DNA polymorphisms between the different GBSS alleles.

### 1.11 Pseudogene copies of GBSS genes.

In addition to the GBSS genes at the three GBSS loci, the possibility existed for pseudogene copies of the GBSS genes in the wheat genome. Pseudogenes are non-coding copies of genes, derived from ancestrally active genes, that have been inserted into the genome by duplication, recombination or by the action of transposons (Kumar and Bennetzen, 1999; Mighell et al., 2000; Petrov and Hartl, 2000). To date there is no evidence for pseudogenes of GBSS in wheat either from PCR (Murai et al., 1999) or by the hybridisation of GBSS-DNA probes to the wheat genome that anneal only to the coding copies of the GBSS genes (Vrinten et al., 1999). However, these papers had not been published at the start of this study and so the potential existence of GBSS pseudogenes was acknowledged, since large plant genomes often result from retrotransposon action and the genome of wheat is very large. One aim of this research was to search the wheat genome for the presence of pseudogenes and to characterize any if identified.

There were three reasons for interest in GBSS pseudogenes. Firstly, pseudogenes can be a source of novel genetic material. Most pseudogenes are separated from some or all promoter and controlling elements and are therefore extraneous to the wild-type gene expression processes of a plant (Mighell et al., 2000). They are not expressed so pseudogenes have no selection pressure to retain a specific DNA sequence and can mutate faster than coding copies. Such mutation may change the DNA sequence so that it no longer encodes a functional protein, or it may encode a new functional protein with a slightly different activity from the
original. Pseudogenes with mutated DNA sequences could therefore be a resource of mutant genes to be used to genetically modify a plant, relocating the pseudogene copy next to a promoter and expressing it as a novel protein.

Secondly, pseudogenes are useful for studying the evolution of a genome, because they mutate at a faster rate than the coding copies. Thus the level of polymorphism in pseudogenes compared to coding copies gives information about when the pseudogene was created and the evolution of the genome.

Finally, the presence of pseudogene copies of the GBSS genes could interfere with molecular tests for the GBSS genes, since any pseudogenes could also be amplified in these tests. Knowing about the presence of any pseudogenes would allow molecular tests to be designed to avoid their amplification.

1.12 GBSSII expression in pericarp tissue.

As well as the GBSSI enzymes expressed in the endosperm of developing wheat grains, other GBSSII enzymes are expressed in tissues such as pericarp and leaf (Nakamura et al., 1998). The GBSSII enzyme synthesises amylose in pericarp tissue separate from GBSSI amylose synthesis in endosperm. GBSSII enzymes show protein sequence similarity and antisera cross-reactivity with GBSSI, but are expressed in different tissues and are encoded at different loci on group 2 chromosomes (Nakamura et al., 1998; Vrinten and Nakamura, 2000). GBSSI DNA probes do not hybridise to GBSSII gene sequences thus showing significant sequence differences. These GBSS isoforms may provide another source of novel GBSS genes for transgenic improvement of wheat starch quality.
1.13 The varied activities of the GBSS enzymes from different GBSS loci.

The wild-type GBSSI enzymes (60kDa in size) expressed from each of the three GBSS loci differ in protein structure and activity. That the structures of the proteins differed was first evidenced by the fact that the enzymes expressed from different loci have different isoelectric points and could be separated by gel electrophoresis (Crosbie, 1991; Nakamura et al., 1993; Yamamori et al., 1994). In 1999 when Murai et al. published the DNA sequences of the GBSS genes from each locus, the differences in amino acid sequence could be deduced. More recently, the different activities of GBSS enzymes transcribed from each locus were described by Yamamori and Quynh (2000). By developing wheat varieties containing various null alleles, the activities of the wild-type alleles from each locus could be measured. This showed that the activity of the GBSS 4A enzyme was higher than the activities of enzymes from the other two loci, with the GBSS 7A enzymes being the least active. With wild-type alleles at only the 7A locus the starch contained an average of 19.8% amylose. With only the wild-type 7D alleles there was 22.5% amylose, whilst the wild-type 4A alleles alone produced 24.2% amylase, which was 86% of the wild-type amount of amylose in starch. The heightened activity of the GBSS enzymes encoded at the 4A locus may be a result of wild-type 4A alleles expressing more copies of the GBSS enzyme than the genes at other GBSS loci (Miura et al., 1994).

Accurate analysis of GBSS enzyme activities in vivo is difficult, because the GBSS enzymes are highly active but limited by the space inside the starch granule. Any small differences in enzyme activity are masked by the effects of the other enzymes and complicated by the limiting factor of space available in the fixed-sized granule (Miura et al., 1994). A good example of this masking effect is the amylose content of starch from WA noodle-wheat when homozygous null 4A alleles are present. Here the absence of one third of the GBSS enzyme copies, resulting from the lack of expression from the null 4A genes, results in only a 3% reduction in starch amylose content. The lack of GBSS enzymes from the 4A locus is
compensated for by the activity of enzymes encoded by the other two loci. Yet if the GBSS enzymes can so effectively compensate for the absence of some enzyme copies, overproducing amylose up to close to wild-type levels, why is any reduction in amylose seen at all? Why can’t the GBSS enzymes encoded by the 7A and 7D loci produce this final 3% amylose missing from null 4A starch? How do the enzymes encoded by the wild-type 4A alleles produce this additional 3% amylose when the enzymes from the 7A and 7D loci do not?

These questions suggest that the GBSS enzymes encoded by the 4A locus may be temporally or spatially separate from enzymes encoded by 7A and 7D genes. For example, if the GBSS genes at the 4A locus were expressed before genes at the other loci, the 4A enzymes would lay down amylose at the centre of the starch granule before enzymes from the other loci were present. In plants containing the null 4A alleles, no enzyme would be present early in grain-filling to lay down amylose at the centre of the starch granule thus explaining the reduced amylose level even with the compensatory effects of the other GBSS enzymes. Although the expression patterns from the three GBSS loci have not been studied, it is known that A-type starch granules in developing grains are initiated 5 days post-anthesis (Parker, 1985) yet GBSS gene expression from the GBSS 7A or 7D loci has not been detected until 10 days post-anthesis (Vrinten et al., 1999) potentially allowing early GBSS-4A expression to impact amylose deposition.

The genes on the 4A chromosome have been translocated from the 7B chromosome and so early theories suggested that the coding region of the GBSS 4A gene may have been separated from its controlling elements on the 7B chromosome. As discussed below, there may be GBSS controlling elements remaining on the 7B chromosome, however recent work shows that the section of DNA translocated from the 7B chromosome is large – about 60 cM in size, with the GBSS locus in the middle (Liu et al., 1992). Thus, the controlling elements within at
least 15 cM of the coding regions of the gene have been translocated to remain with the GBSS coding region on the 4A chromosome.

The theory of spacial variation suggests that the enzymes encoded by the wild-type 4A gene have transport elements that direct them to different sites within the starch granule that are unavailable to the enzymes encoded by the 7A and 7D genes. Recent sequencing of the GBSS genes showed a trinucleotide insertion in the transit peptide sequence of the wild-type 4A gene (Murai et al., 1999), which may result in enzymes encoded by the 4A genes being transported to unique sites within the starch granule. The absence of the enzymes from the 4A locus would then result in a reduced amylose level, even with the compensatory ability of the other enzymes. The amino acid residues of the GBSS transit peptides are less similar (84.5-94.4%) than the amino acid sequences of the mature GBSS proteins (96.5-97.4%) (Murai et al., 1999). This suggests that some of the differences in starch quality in partially waxy wheat starches lacking various enzymes may be a result of the enzymes being targeted to slightly different sites on the starch granule by the varying transit peptides.

One aim of this research was to identify differences between the GBSS genes at the 4A locus and those at the other two loci to explain the starch qualities seen. In addition, a better understanding of the GBSS promoter elements and transport elements of the genes at each of the three loci was sought.

There may also be other regions of the wheat genome outside the GBSS loci that affect amylose production through epistasis. Epistasis is the interaction between two or more genes to control a single phenotype. Epistasis is probable in a genome as complex as wheat’s, and may lead to unexpected starch quality variations not primarily driven by the GBSS genes. Araki et al. (1999) suggested that there may be allelic variation at a Quantitative Trait Locus.
(QTL) on the short arm of chromosome 4A that can affect the amylose content of wheat starch by 0.3%. Also, Miura et al. (1994) showed that the 7B chromosome results in a nearly 1% reduction in amylose content, suggesting the presence of regulatory elements that suppress GBSS activity. They suggested that the group-7 chromosomes may all carry GBSS regulator genes separate from the structural genes and that the translocation of the GBSS 4A locus has separated the structural component from its regulatory region. Due to the size of the translocation (approximately 60 cM with the Waxy locus at least 15 cM from the site of cross-over (Liu et al., 1992)), it can be deduced that these putative regulatory elements are not positioned adjacent to the structural gene. Thus small differences in amylose content may not result from GBSS gene variation at all. Once the variation in amylose content caused by allelic variation at the GBSS loci is understood, the epistatic effects of other loci in the wheat genome can be investigated.

All these differences between the GBSS enzymes and genes from the three GBSS loci provide potential beneficial mutations for starch improvement. Aside from various combinations of null and wild-type alleles, which result in different starch characteristics based on the varying activities of the wild-type alleles (Yamamori and Quynh, 2000), there is the potential to swap the GBSS alleles trans-genome. Using chromosome substitution varieties, where a chromosome from one wheat genome is replaced by a chromosome from a different genome, such as the nullisomic-tetrasomic variety ‘Chinese Spring’, new combinations of GBSS genes can be produced. For example, wheat containing three GBSS-7A loci may produce different starch characteristics from that containing two GBSS-7D loci and one GBSS-4A locus.
1.14 Varied activities between enzymes encoded by different wild-type alleles at a single locus.

At a single GBSS locus there may be several different wild-type alleles that express proteins with slightly different activities. The effects on starch of different wild-type alleles at a single locus would be difficult to investigate because of the buffering ability of the GBSS enzymes from the other loci. As the amount of amylose in the starch granule has a finite maximum, dictated by the space and raw materials available to GBSS enzymes, most wheat varieties produce starch containing this maximum amount of amylose. Any over-producing GBSS enzymes are difficult to detect in this environment since their action would not be differentiable from the action of other wild-type alleles (Araki et al., 1999). It is only in the reduced-amylose environment of a variety containing null alleles that variations in wild-type enzyme activities are detectable.

There is evidence of multiple wild-type GBSS alleles in ASWN wheat varieties grown in WA. All WA noodle wheats contain the null 4A alleles and less amylose in the starch. This allows the starch granules to swell the desired amount when used to make Udon noodles as discussed below. However, some noodle-wheat varieties have more amylose than predicted and do not swell to the desired level (e.g. ‘Reeves’, ‘Cascades’, ‘Calingiri’, ‘Halberd’ and ‘Spear’). These are not suitable for Udon noodle production and cost wheat breeders time and money to identify. It can be predicted that these varieties contain a different GBSS allele at either the 7A or 7D locus that is responsible for the increased levels of amylose seen. One aim of this research was to identify the presence of putative novel GBSS alleles in the WA wheat varieties unsuitable for classification as ASWNW, and to create molecular markers to identify the novel alleles to allow early detection of the undesirable genotype.

Whilst seen as an impediment to noodle-wheat production, the existence of different wild-type alleles of the GBSS genes is still a resource for wheat breeders developing varieties with novel
starch characteristics for other end-uses. Identifying novel GBSS alleles with different activities will enable novel combinations of GBSS genes to be investigated for their effects on starch quality in wheat. An aim of this project was to identify different wild-type GBSS alleles, and create molecular markers to distinguish the different alleles, toward facilitating tracking GBSS alleles through future wheat breeding programs.

1.15 Effects of amylose levels on starch quality and Flour Swelling Volume (FSV).

For wheat varieties to be classified as ASWNW the flour must contain not only the basic characteristics of clear flour colour, medium protein content and mellow gluten quality (Endo et al., 1988; Garlinge et al., 1996; Konik-Rose et al., 2001), but also the defining starch swelling characteristics. The amount that the flour swells when heated in water is termed the Flour Swelling Volume (FSV), and FSV gives a very accurate indication of the physical cooking properties of the flour (Zhao et al., 1998). Developed by Crosbie (1991), FSV is the benchmark test for starch quality in noodle wheat varieties, and it is used to classify noodle wheat varieties by the extent their flour swells; low-FSV, medium-FSV or high-FSV.

In starch granules, the amylose molecules are deposited within the gaps of the amylopectin matrix and when the starch granule is heated in water, the amylopectin matrix swells outwards (Zeng et al., 1997). The more amylose that is deposited around the amylopectin, the less it can swell. Reducing the amount of amylose surrounding the amylopectin allows the starch granule to swell more. The structure of the amylopectin molecules may also contribute to the FSV (Batey et al., 1997). Wild-type starch contains about 25% amylose, does not swell well and shows a low-FSV value. Wheat with the null 4A allele contains less amylose (22%), the starch granules can swell more and a high-FSV value is seen. Two closely related varieties ‘Kulin’ and ‘Reeves’ differ at the GBSS 4A locus with ‘Kulin’ containing the wild-type 4A
allele and ‘Reeves’ containing the null 4A allele. Their FSV values are 14.8 mL/g for ‘Kulin’ and 17.7 mL/g for ‘Reeves’ thus showing how the presence of the null 4A allele in ‘Reeves’ resulted in an increased FSV (Zhao et al., 1998). Any further reduction in GBSS enzyme reduces the amylose content too far and more waxy characteristics develop.

1.16 Molecular markers for noodle quality.

Wheat breeders in Australia have long been aware of the relationship between the lack of GBSS-4A protein and superior quality starch for Udon noodle production (Oda et al., 1980; Miura and Tanii, 1994; Zhao et al., 1998). Noodle-quality wheat varieties have been tested for their GBSS characteristics for many years, either by extracting the GBSS proteins from the starch granule, separating them electrophoretically by 2D PAGE and viewing the lack of GBSS-4A protein (Crosbie, 1991; Nakamura et al., 1993; Yamamori et al., 1994), or indirectly by milling the flour and performing FSV tests to detect the reduced amylose levels (Konik-Rose et al., 2001). It was because of the linkage between the GBSS null 4A allele and superior starch quality for Udon noodle production that this gene was targeted to develop a molecular marker for wheat starch quality.

When breeding for a particular trait, both the phenotype and the genotype can be examined. A molecular marker is linked to a specific gene of interest, and whereas DNA in plant cells is constant throughout the life of a plant, the expressed gene products can vary with developmental stage and changes in environment. One benefit of using molecular markers in breeding is that DNA can be tested from any part of the plant at any stage of the life-cycle, unlike trait evaluation (e.g. flower colour) that requires a particular developmental stage. The quality of grain can only be tested after the reproductive cycle at the end of the plant’s growing season, and the test is undertaken on the organ that is required for the next generation of plants to be grown (Briney et al., 1997). If the test requires more grain than is produced by a single plant, plants must be bulked to make enough seed to be tested and the seed can be
destroyed by the testing process. In contrast, a DNA marker can be used at a vegetative stage and the selection made immediately. DNA testing can also delay the need for the seed-bulking step necessary in producing flour for quality tests. Another benefit is that recessive or masked genes that are not evident in the phenotype can still be detected as heterozygotes, aiding a breeder in the choice of parents for crosses. The use of molecular markers might therefore reduce the time and cost associated with generating new varieties.

The Polymerase Chain Reaction (PCR) works by enzymatically copying a section of DNA. The section of DNA copied is delimited by two, short (approximately 20bp), single-stranded DNA molecules, called PCR primers, which flank the section of DNA to be copied and allow DNA amplification between them. By amplifying a specific section of DNA, a defined DNA fragment associated with a gene of interest can be isolated from plant genomes.

The first PCR-based molecular marker for Udon-noodle quality starch was developed by Briney et al. (1998) to identify the GBSS null 4A allele. To study the differences between the null and wild-type 4A alleles, two, closely related wheat varieties were developed that segregated for the GBSS-4A protein; ‘Kulin’ which produced a 4A protein and ‘Reeves’ which lacked the 4A protein (Briney et al., 1998; Zhao et al., 1998). As these varieties contained the same genes at most loci, the different alleles at the GBSS 4A locus could be identified more easily. PCR primers were designed to amplify from the coding region of the GBSS genes, because the GBSS genomic genes had not yet been sequenced. The PCR primers were designed to conserved regions of an alignment of the genomic and cDNA sequences of the GBSS genes from barley (X07931, X07932) and rice (X64108, X62134), with the only available GBSS sequence from wheat, the cDNA sequence (X57233) (Clark et al., 1991). One set of the PCR primers (GBSS2/GBSS4) amplified a band from each of the three GBSS loci (Briney et al., 1998). For homozygous null 4A alleles, no band was amplified in this PCR (see Figure 1.10). Ideally, a molecular marker should be co-dominant,
with each allele giving a differentiable band such that the presence of heterozygotes can be detected. However, this marker was not a co-dominant marker so a null allele could not be identified in the presence of the wild-type allele from the same locus, but homozygous null 4A alleles could be identified. This marker has been used to screen the germplasm in the WA ASWNW breeding program to identify potential noodle-wheat genotypes.

**Figure 1.10.** A polyacrylamide gel showing the PCR bands amplified from wheat by the primers GBSS2/GBSS4. The primers amplify one band from each of the GBSS loci on chromosomes 4A, 7A and 7D and thus the PCR bands have been labeled with their locus of origin. The band from the 4A locus is absent from the ‘Reeves’ sample (Lane 2) as this wheat line carries the null 4A allele that does not give a band. Reproduced from Hollingsworth (1997).

Lane 1: 100 bp molecular size marker (BRL). Lane 2: ‘Reeves’. Lane 3: ‘Kulin’.
1.17 Aims of this project.

The overall aim of the research presented in this thesis was to investigate the genomic organization of the GBSS genes to answer the following questions:

1. **Where are the GBSS gene copies in the genome?** There is a single copy of the GBSS gene at each of the three GBSS loci on chromosomes 4A, 7A and 7D, but are there other copies in the genome, such as pseudogene copies? Identifying pseudogenes would eliminate them as a complicating factor in interpreting PCR results and could also give information about the contribution of the GBSS gene sequences to the evolution of the wheat genome. This was a minor aim of the study.

2. **How many different GBSS alleles are in the WA wheat germplasm?** Wheat germplasm in the WA noodle-wheat breeding program contains wild-type and null GBSS alleles at all three loci, but the numbers of alleles and their effects on starch quality are not known. It is aimed to assess the molecular variation present in the noodle-wheat germplasm. This was a major aim of the study.

3. **How does each GBSS allele differ?** At the start of the project, the reasons for the lack of expression from null alleles was not known. Also, any molecular variation between wild-type GBSS alleles had not been identified. It was aimed to identify molecular differences within the population of GBSS alleles in Western Australian germplasm, and identify the reasons for the lack of expression from the null alleles. These differences could then be used to design molecular markers that could be used to assess the possible functional differences those mutations produced. A significant portion of the experimental work in this study had been completed before the genomic GBSS sequences of the null and wild-type alleles were published (Murai et al., 1999; Vrinten et al., 1999). Those published sequences were not from Australian varieties, and so the comparable sequences from Australian varieties were still required. This was a major aim of the study.
4. **How can we tag the different GBSS alleles with molecular markers to identify them in a breeding program?** At the start of this study, the null alleles at the 7A and 7D loci could not be identified using molecular markers and thus could not easily be traced through a breeding program to study their effects on starch quality. Also, the different wild-type alleles predicted to exist at each locus could not be distinguished. Once novel alleles had been identified and tagged with a molecular marker, their effects on starch quality could be investigated and possibly used to develop novel wheat varieties with improved starch qualities. This aim was a direct progression from the results of the second and third aims, so this aim was a secondary aim of the study.

5. **Are the GBSS genes present in WA wheat varieties the same alleles described in international publications?** Most GBSS studies have used ‘Chinese Spring’ germplasm, which is different from the wheat varieties grown in WA. In order to extrapolate the information on GBSS genes in ‘Chinese Spring’, it is important to ascertain how similar the GBSS genes in WA noodle-wheat varieties are to those described in previous studies. This became a major aim of the study when the sequences of GBSS genes from ‘Chinese Spring’ varieties were published.
Chapter 2. General Materials and Methods.

2.1 Wheat varieties and sources of plant material.

A sampling of wheat varieties were used in this study. The selection includes varieties of noodle-wheat currently grown in Western Australia, as well as some of the parental varieties for the noodle-wheats and also some varieties unsuitable for producing noodles. Some of the varieties were chosen because their DNA was easily available - leaf tissue or DNA samples for those varieties were stored at the Western Australian State Agricultural Biotechnology Centre (WA SABC, Murdoch University, Perth), which had been previously sourced from the Department of Agriculture Western Australia (DAWA, Perth). Other varieties, particularly noodle-wheat varieties, were supplied for this study by R. Wilson (DAWA, Perth), whilst the Chinese Spring cultivars were supplied by M. Pallotta (CRC Molecular Plant Breeding, Adelaide). Table 2.1 shows the varieties used and provides a description of their traits. Further information about the percentage of each variety grown and the growing regions, are included as Appendices 1 and 2.

The two varieties used for most of the experiments were the closely-related varieties ‘Reeves’ and ‘Kulin’ (both Bodallin//Gamenya/Inia 66). They were chosen because of their GBSS characteristics: ‘Reeves’ contains the null 4A allele and ‘Kulin’ contains the wild-type 4A allele. These varieties are soft-grained Australian Standard Wheat varieties with good yields and good milling qualities, with white-chaffed and bearded heads (see Table 2.1). ‘Reeves’ was developed to be a noodle-wheat but has unsuitably low FSV.

In addition to the varieties listed above, two breeding populations were used in the study and these were provided as leaf tissue from DAWA. The parents of these crosses were ‘Arrino’ and ‘DHW174’, and the offspring were doubled haploids.
Leaf material from some varieties was taken from stocks provided by R. Wilson and stored at -80°C. Some DNA samples had been previously extracted by A. Briney, M. Carter and H. Li (WA SABC, Murdoch University, Perth). Plants grown from seed at Murdoch University were maintained in a temperature-controlled glasshouse in pots (diameter 15cm) containing washed river sand with Osmocote, a nitrogen-based fertiliser, added 3, 6 and 9 weeks after planting. Young leaves (more than 60% green) were harvested for DNA extraction and stored on ice or at -80°C until required.

2.2 DNA extraction.

Five different DNA extraction methods were used because each gave a slightly different quality and purity of DNA as was required for use in different parts of this study. The CTAB method (2.2.1) was used routinely to extract DNA for PCR, and this DNA was pure enough to be stored for long periods of time. For the PCR screening large numbers of individuals the Quick Extraction method (2.2.3) was used, but samples extracted using this method would degrade after a few days and this DNA could not be stored for later use. The modified method of Dellaporta et al. (2.2.2) had traditionally been used to extract un-sheared DNA for Southern Blotting, and the Plant DNAzol method (2.2.4) was tested for this use, but the Ethanol-perchlorate method (2.2.5) was used for this purpose in this study.

2.2.1 Cetyltrimethylammonium bromide (CTAB) extraction of plant DNA.

DNA was extracted from wheat leaves as described by Rogers and Bendich (1994). Freshly harvested leaf tissue, which had been on ice for less than one hour, or leaves stored at -80°C were ground to a fine powder with a mortar and pestle under liquid nitrogen. The powder was transferred to 1.5 ml microcentrifuge tubes (Sarstedt) to fill the tubes about one third full. To this, 800 µl of hot (65°C) extraction buffer (2% w/v CTAB, 100 mM Tris pH 8.0, 20 mM EDTA pH 8.0, 1.4 M NaCl, 1% w/v polyvinylpyrrolidone (PVP) M, 40 000) was added. After vortexing, tubes were incubated at 65°C for 15 min. One half volume of chloroform/isoamyl
alcohol (24:1 v/v) was added and mixed thoroughly to form an emulsion. Tubes were centrifuged at 11,000 rpm for one min and the aqueous phase was transferred to a new tube. A 1/10 volume of hot (65°C) CTAB solution (10% w/v CTAB, 0.7 M NaCl) was added and mixed by inverting. Another extraction was performed with an equal volume of chloroform/isoamylalcohol and the aqueous phase was again transferred to a new tube. An equal volume of CTAB precipitation buffer (1% w/v CTAB, 50 mM Tris pH 8.0, 10 mM EDTA pH 8.0) was added and mixed gently by inverting. After incubation on ice for between 30 and 180 min, the samples were centrifuged at 14,000 rpm for 2 min. The supernatant was discarded and the pellet was suspended in 50 µl of high salt TE buffer (10 mM Tris pH 8.0, 1 mM EDTA pH 8.0, 1 M NaCl) for 15 min. Two volumes of cold ethanol were added, mixed gently by inverting, and placed on ice for 15 min to allow precipitation of nucleic acids. Samples were centrifuged at 14,000 rpm for 10 min and the supernatant discarded. Fifty µl of cold 80% ethanol was added and the samples were centrifuged at 14,000 rpm for 5 min. The supernatant was discarded and the pellets allowed to air dry inverted on the bench. The pellets were suspended in 50 µl of 0.1 x TE (1 mM Tris pH 8.0, 0.1 mM EDTA pH 8.0) for 15 min at room temperature before being treated with 3 µl of 10 mg/ml RNAase A (Sigma-Aldrich) at 37°C for 30 min. Samples were placed at 4°C overnight to ensure complete resuspension before quantifying the DNA using a Hoefer fluorometer or visually on an agarose gel (see Section 2.3.).

2.2.2 Extraction of plant DNA based on the method of Dellaporta et al. (Potter, 1991).

This method is essentially as described by Potter (1991). Up to 4 g of freshly harvested wheat leaf tissue was ground to a fine powder with a cold mortar and pestle under liquid nitrogen. The powder was transferred to a 50 ml Oakridge tube containing 15 ml of hot (65°C) extraction buffer (100 mM Tris.Cl pH 8.0, 500 mM NaCl, 50 mM EDTA pH 8.0, 0.6% β-mercaptoethanol) (autoclaved before the addition of the β-mercaptoethanol). One ml of 20%
w/v sodium dodecyl sulphate (SDS) was added, mixed by shaking and incubated at 65°C for 10 min. Five ml of 5 M potassium acetate was added, mixed by shaking, and placed on ice for 10 min. Samples were centrifuged at 10 000 rpm for 20 min at 4°C and the supernatant was poured through Miracloth (Calbiochem) into another tube containing 10 ml of cold isopropanol. Tubes were mixed gently by inversion and incubated at -20°C for 30 min before centrifugation at 10 000 rpm for 20 min at 4°C. The supernatant was discarded and the tube allowed to drain inverted for one min before resuspending the pellet in 700 µl of 50TE (50 mM Tris.Cl pH 8.0, 10 mM EDTA pH 8.0). The solution was transferred to a microcentrifuge tube containing 7 µl of 10 mg/ml RNAase A (Sigma-Aldrich) and incubated at 37°C for one hour. The sample was centrifuged at 14 000 rpm for 15 min and the supernatant transferred to a new tube containing 75 µl of 3 M sodium acetate (pH 5.2). The solution was mixed by inversion, 0.5 ml of isopropanol was added and mixed again by inversion before incubation at room temperature for 5 min. DNA was pelleted by centrifugation at 14 000 rpm for 5 min and the supernatant was poured off. The pellet was washed with 75% ethanol and air dried inverted on the bench before resuspending in 200 µl of TE (10 mM Tris.Cl pH 8.0, 1 mM EDTA pH 8.0) overnight at 4°C.

2.2.3 Quick extraction of plant DNA.

This method was adapted from Edwards et al. (1991). Approximately 2 cm of freshly harvested wheat leaf was macerated in a 1.5 ml microcentrifuge tube using a tissue grinder (Kontes). To this crushed tissue, 400 µl of extraction buffer (200 mM Tris-HCl pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% SDS) was added and mixed vigorously by vortexing. A further 100 µl of 5 M potassium acetate pH 7.5 was added, mixed by shaking, and incubated on ice for 5 min. Samples were centrifuged at 14 000 rpm for one min and 300 µl of the supernatant was transferred to a new tube. To this, 300 µl of cold isopropanol was added, mixed by inverting and incubated at room temperature for 2 min. Samples were then centrifuged at 14 000 rpm for 5 min, the supernatant was poured off, and the pellets allowed to
air dry inverted on the bench for 30 min. Pellets were resuspended in 50 µl of pharmaceutical grade water (Delta West) at 4°C overnight.

2.2.4 **Plant DNAzol DNA extraction.**

DNA was extracted from leaf material using Plant DNAzol Reagent (GibcoBRL, Melbourne Australia) following the manufacturer’s instructions. The extraction was performed both with and without the optional chloroform extraction step.

2.2.5 **Wheat Ethanol-Perchlorate Method.**

This method was essentially as described by Lagudah *et al.* (1991). Leaf material was placed in a pre-cooled mortar with acid-washed sand and liquid nitrogen and ground to a fine powder with a pestle. The powder was transferred to a sterile 10ml centrifuge tube and 7.6ml of the extraction buffer (0.05M Tris, 0.3M EDTA pH 8.0) was added. The tubes were kept on ice until all the samples had been processed. One milligram of Proteinase K (Promega) and 300µl of 20% SDS were added and the tubes were vigorously vortexed. The tubes were incubated for at least one hour at 37°C. Two grams of sodium perchlorate (as 1g/ml Na perchlorate in water) were added to each tube, vortexed and centrifuged at 4000xg for 7 min at 20°C. The supernatant was removed to a new tube, recentrifuged at 14000rpm and again removed to a new 50ml tube. Eight ml of ethanol-perchlorate (120g NaClO4.H2O, 320ml absolute ethanol, 80ml distilled water) was added and the tubes inverted to mix the solutions. The DNA was spooled onto a glass rod and transferred into a 1.5ml tube containing 400µl of TE (10mM Tris, 1mM EDTA, pH 8.4). The 50ml tubes were left to rest on the bench for 5min and then inverted to mix and the resulting DNA precipitate spooled onto a glass rod and removed into TE. The 50ml tubes were then centrifuged at 14000rpm for 5min and the supernatant was removed. The pellet was resuspended in 400µl TE and transferred to a 1.5 ml tube. All 1.5ml tubes were flick-mixed and placed at 4°C overnight to resuspend the DNA. To each tube was added 500µl of phenol and 500µl of chloroform and the tubes agitated at 4°C for 1hr. The
tubes were centrifuged at 12000xg for 15min and the supernatant was transferred to a new tube containing 200µg of RNAaseA (Promega). Tubes were incubated at 37°C for 30min and the phenol:chloroform extraction was repeated. The supernatant was removed to a new tube to which 1/10 volumes of 0.2M Na acetate pH 5.5 and 2 volumes of ethanol were added. Tubes were incubated at -80°C for 30min and centrifuged at 12000xg for 30min at 4°C. The ethanol was removed and the pellet was washed twice with 1ml of cold 70% ethanol. The tubes were centrifuged at 12000xg for 1min and the 70% ethanol removed. The pellet was air dried and resuspended in water overnight at 4°C.

2.3 DNA quantification.

DNA was quantified using a fluorometer (TKO100; Hoefer Scientific Instruments). The instrument was zeroed using 2 ml of fluorescent dye solution (bis benzimidizole also called Hoechst 33258) in a quartz cuvette and calibrated using 2 µl of DNA standard (25 ng/µl calf thymus DNA, Pharmacia Biotech). After rinsing the cuvette with sterile water, 2µl of the DNA sample to be measured was mixed with 2 ml of the fluorescent dye solution and the DNA concentration read from the display.

When DNA had been extracted from a gel, it could not be quantified using a fluorometer and had to be visually quantified on a gel. The DNA concentrations of these samples were estimated by comparison to a molecular marker of known DNA concentration after both had been electrophoresed on a 1% agarose gel.

2.4 PCR primers and conditions.

Typically, PCR reactions were performed in 10 µl volumes containing 67mM Tris-HCl pH 8.8, 16.6mM (NH₄)₂SO₄, 0.45% Triton X-100, 0.2mg/ml gelatin, 0.2mM dNTPs (supplied as 5X polymerisation buffer; Biotech International). Reactions also contained 0.4U Tth⁺ polymerase (Biotech International), 20pmoles of each primer (Table 2.1), 1.5mM MgCl₂ and
20ng of DNA template. Some DNA template solutions were not, or could not, be quantified and thus a standard 2µl of these solutions was included in the PCR. Reactions were made to volume with pharmaceutical grade water (Delta West). Amplification was carried out using Perkin-Elmer thermal cyclers (model GeneAmp PCR System 9600 or 2400) and an Eppendorf Mastercycler gradient thermal cycler (Eppendorf). The DNA sequences of all the PCR primers are given in Table 2.1.

Table 2.2. DNA sequences of PCR primers.

<table>
<thead>
<tr>
<th>PRIMER NAME</th>
<th>DNA SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>GBSS1</td>
<td>5’ GGCGGCATGAACCTCCTGTT 3’</td>
</tr>
<tr>
<td>GBSS2</td>
<td>5’ TTGAGCTGCGCAGAAGTCGTC 3’</td>
</tr>
<tr>
<td>GBSS3</td>
<td>5’ AAGGCTGGAAGCGCTGCTG 3’</td>
</tr>
<tr>
<td>GBSS 4</td>
<td>5’ AACCAGCAGCGCTCAGCCT 3’</td>
</tr>
<tr>
<td>GBSS5</td>
<td>5’ CGTCCAGCTCAGCAACCTCC 3’</td>
</tr>
<tr>
<td>GBSS6</td>
<td>5’ TCAGGGGAGCGCGAGCTTTGCTCAGG 3’</td>
</tr>
<tr>
<td>GBSS7</td>
<td>5’ GTCTCCCACTTCTGGGCGAGGCC 3’</td>
</tr>
<tr>
<td>GBSS8</td>
<td>5’ CCAGGAGAGATCCTGTATCATGC 3’</td>
</tr>
<tr>
<td>GBSS9</td>
<td>5’ GCCGCCAGTCGTCCAGCTCC 3’</td>
</tr>
<tr>
<td>GBSS10</td>
<td>5’ CGTCCAGCGCAGATGCCCCCTGG 3’</td>
</tr>
<tr>
<td>GBSS-Rev</td>
<td>5’ GAGATGTGGTGATGCAGAA 3’</td>
</tr>
<tr>
<td>GBSS-4A</td>
<td>5’ AACAACACCCAGCGCTACTAG 3’</td>
</tr>
<tr>
<td>GBSS-7A</td>
<td>5’ CGGTAAGATCAAGAACAATAGAGT 3’</td>
</tr>
<tr>
<td>GBSS7D</td>
<td>5’ AAGATCAAGCAGCCTACTAGTT 3’</td>
</tr>
<tr>
<td>T7</td>
<td>5’ CGGCCAGTGATTTGTGTAATACGAC 3’</td>
</tr>
<tr>
<td>Sp6</td>
<td>5’ CTATGACCAGTGATTACGGGCAAGC 3’</td>
</tr>
<tr>
<td>Gus 3’</td>
<td>5’ GCAATCTCCACATCCACCGCTT 3’</td>
</tr>
<tr>
<td>Rob26</td>
<td>5’ AGCCTCGTAGACTGCGTACC 3’</td>
</tr>
</tbody>
</table>
The PCR primers GBSS1, GBSS2, GBSS3, GBSS4, T7 and Sp6 were supplied by A. Briney and R. Potter (WA State Agricultural Biotechnology Centre). The primers GBSS-REV, GBSS-4A, GBSS-7A, GBSS-7D, GBSS5, GBSS6, GBSS7, GBSS8, GBSS9 and GBSS10 were designed as part of this project. The GBSS series primers were designed to GBSS DNA sequences whilst the T7/Sp6 primers amplify across the insertion site of the pGEM series plasmids, and the Gus primers amplify from the GUS gene.

2.5 Gel electrophoresis.

The DNA sample to be electrophoresed was combined with 1/5 volumes of loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF, 30% glycerol in water) (Sambrook et al., 1989) before loading onto gels. Two types of gels were used for electrophoresis - polyacrylamide and agarose. Polyacrylamide gels contained 4% 29:1 acrylamide:bisacrylamide (Biorad) in 1x TBE buffer (90 mM Tris-borate, 2 mM EDTA) (Sambrook et al., 1989) and were used as part of a Protean II electrophoresis system (Bio-Rad Laboratories) with 1.5 mm spacers. These gels were run at 110V for 4 hours. Agarose gels contained 1% DNA grade agarose (Progen) in 1x TAE buffer (40 mM Tris-acetate, 1 mM EDTA) (Sambrook et al., 1989) and were electrophoresed for various times as required. All gels were stained with a 0.5\( \mu \)g/ml solution of ethidium bromide for 30min and visualized under UV light (254 nm). Gel images were recorded electronically using a Gel Doc system and Molecular Analyst software (Bio-Rad Laboratories, Los Angeles). A suitable molecular size marker was run on each gel.

2.6 DNA purification by ethanol precipitation.

Purification by ethanol precipitation involved mixing the post-PCR solution with 1/10 volume of 3 M sodium acetate pH 5.2 and 2 vol of cold 95% - 100% ethanol, mixing by inverting and incubating on ice for 30 - 60 min. Samples were then centrifuged at 14 000 rpm for 30 min at 4°C. The supernatant was discarded, the pellets rinsed with 70% - 80% ethanol and the pellets
allowed to air dry inverted on the bench. DNA was resuspended in the desired volume of pharmaceutical grade water (Delta West).

2.7 PCR fragment isolation by gel purification.

To isolate a band of DNA from a gel, the bands were cut from the gel with a scalpel and DNA was extracted from the gel using the following methods.

2.7.1 QIAquick Gel Extraction Kit.

DNA was extracted from gel using a QIAquick Gel Extraction Kit (QIAGEN) as specified by the manufacturers.

2.7.2 Quick method.

Using a Quick gel extraction method, the gel slice was placed in a column containing 3 layers of Watman paper, resting in a 1.5ml centrifuge tube and centrifuged for 1 min at 14000rpm. The resulting supernatant was purified by ethanol precipitation as in Section 2.6.1.

2.8 Ligation of DNA fragments into the pGEM-T vector and transformation through *Escherichia coli*.

DNA fragments were ligated into the pGEM-T vector using a pGEM-T Vector System II kit (Promega) as specified by the manufacturer. Reactions were incubated overnight at 14°C and stored at -20°C until required. Competent cells (JM109) were prepared as outlined by Promega (1991). Ligation reactions were used to transform the cells using the method outlined by Promega (1991). LB media was used throughout this protocol, and cells were allowed to recover on ice for 10min after being subjected to heat shock.
2.9 Identification of clones containing inserts using PCR.

Single bacterial colonies were touched into 500µl of water using sterile pipette tips and the samples were vortexed to disperse the cells through the water. A 2µl sample of this solution was used as PCR template in the screening PCR reactions. Clones were screened for the inserts they contained using two PCR tests, one using the T7/Sp6 primer combination and one using the GBSS-Rev/GBSS-4A,7A,7D primer combination. PCR reactions containing T7/Sp6 primers also contained 1.5mM MgCl₂ and were incubated under the following conditions: 1 cycle of 94°C for 3min, 57°C for 30sec and 72°C for 1min; 35 cycles of 94°C for 30sec, 57°C for 30sec and 72°C for 1min; 1 hold at 72°C for 10min. PCR reactions containing GBSS-Rev, GBSS-4A, GBSS-7A and GBSS-7D primers also contained 2mM MgCl₂ and were incubated under the following conditions: 1 hold at 94°C for 3min; 35 cycles of 94°C for 30sec, 58°C for 30sec, 72°C for 1min; 1 hold at 72°C for 4min. Some PCR reactions contained 20pmol of primer GBSS-Rev and 20pmol of an equimolar mix of the primers GBSS-4A, GBSS-7A and GBSS-7D, while some PCR reactions contained 20pmol of primer GBSS-Rev and 20pmol of one of primers GBSS-4A, GBSS-7A or GBSS-7D.

2.10 Sequencing.

2.10.1 Sequencing the inserts inside plasmid clones.

The desired colony was pricked into 4ml of liquid LB medium containing ampicillin and incubated overnight at 37°C with shaking. Plasmids were extracted from the cells using a QIAprep Spin Plasmid Kit (Quiagen) as specified by the manufacturer, except at step 4 where the 10min centrifugation was increased to 20min at 4°C. Plasmid DNA was eluted from the columns in 100µl of pharmaceutical grade water (Delta West) and was transferred to a new 1.5ml tube and concentrated by ethanol precipitation (Section 2.6.1).

Sequencing reactions were carried out in 10 µl volumes containing 4 µl of ABI PRISM dye terminator cycle sequencing ready reaction mix (A-Dye Terminator, C-Dye Terminator, G-
Dye Terminator, T-Dye Terminator, dITP, dATP, dCTP, dTTP, Tris-HCL pH 9.0, MgCl₂, thermostable pyrophosphatase, AmpliTaq DNA Polymerase, FS) (Perkin Elmer), 100-300ng plasmid DNA template and 1.6 pmol primer (T7 or Sp6). Thermal cycling was carried out on a Perkin-Elmer GeneAmp 2400 thermocycler using 25 cycles of 96°C for 10 sec followed by 60°C for 4 min. Reactions were purified by ethanol precipitation as follows. Sequencing reactions were added to 1 µl of 3 M sodium acetate pH 5.5 and 25 µl of cold 95% ethanol. These were mixed by vortexing and placed on ice for 10 min before centrifugation for 30 min at 14 000 rpm at 4°C. The supernatant was poured off and the pellet washed with 250 µl of 70% ethanol. The pellet was dried in a vacuum centrifuge (Speedvac Concentrator, Savant) and stored at -20°C for less than one week. Samples were sequenced using an ABI 373 sequencer under standard conditions. Sequence data was analysed using Sequence Editor software (ABI).

2.10.2 Direct Sequencing of PCR products.

When PCR products were sequenced directly without being cloned, 40µl of PCR product was purified away from the primers using a Wizard® PCR Preps DNA Purification System (Promega) as specified by the manufacturers and then precipitated with ethanol as in Section 2.6. The PCR product was resuspended in 5µl of water and used directly in a sequencing reaction containing 10pmoles primer as described above. The reactions were incubated for 25 cycles with each cycle consisting of 96°C for 10 sec, 50°C for 5 sec and 60°C for 4 min. The reactions were then cleaned as above.

2.11 Southern blots.

2.11.1 Transferring DNA from gels to membranes.

Gels were soaked for 15 min in 0.4 M NaOH and the blot was set up as shown in Figure 2.1 using Magna charge membrane (Amersham). The blot was left overnight and then the filter
was washed in 5X SSC (43.825 g/L NaCl, 22.05 g/L sodium citrate, pH 7.0) at room
temperature for 10 min before being allowed to dry between paper towels.

---

**Figure 2.1.** Arrangement for Southern Blotting. NaOH from the container is
drawn up by capillary action to pass through the gel, the membrane filter and
finally into the paper towels. The DNA is transferred from the gel with the
NaOH and is retained on the membrane. Plastic film is placed on the filter
paper around the membrane to ensure that the NaOH passes directly upward
from the gel through the membrane and does not bypass it into the paper towels.

2.11.2 Hybridisation of Southern blots of PCR products and Dot Blots.

Southern blots of PCR products were hybridised as follows. The filter was pre-hybridized in
20 ml of hybridization buffer (0.5 M NaH₂PO₄ pH 7.0, 7% SDS, 5 mM EDTA) in a
hybridization bottle and incubated in a hybridisation oven with rotation at 9rpm at 65°C for up
to 1 hr. This solution was discarded and replaced with 15 ml of fresh 65°C hybridization
buffer to which the probe solution had been added. The filter was then incubated at 65°C
overnight rotating at 9 rpm in an oven (XTRON HI 2002, Bartelt instruments).
2.11.3 Hybridisation of Southern blots of genomic DNA.

Southern blots of genomic DNA were hybridised using a method supplied by CRCMPB (Adelaide). The filter was pre-hybridised in 10ml of hybridisation buffer (3ml 25% dextran sulphate, 3ml Denhardt’s III, 3ml 5x HSB and 2.5mg denatured salmon sperm DNA) in a hybridization bottle and incubated in a hybridisation oven with rotation at 9rpm at 65°C for 2 hours. Denhardt’s III was an aqueous solution of 2% BSA Sigma Fraction V, 2% Ficoll 400, 2% Polyvinyl-pyrrolidone 360 and 10% SDS, and 5x HSB was an aqueous solution of 3M NaCl, 100mM PIPES and 25mM Na2EDTA with a pH of 6.8 with 4M NaOH. The hybridisation solution was removed to a bijoux bottle into which was added a denatured probe and an additional 2.5mg salmon sperm DNA. The hybridisation solution was returned to the filter and incubated at 65°C overnight while the bottle rotated at 9 rpm in an oven (XTRON HI 2002, Bartelt instruments).

2.11.4 Preparation of a radioactive probe.

PCR fragments amplified from the GBSS genes at all three loci using GBSS2/GBSS4 primers were cloned and used individually as templates for the radioactive probes. The GBSS fragment was amplified away from the clone using the GBSS2/GBSS4 primers and the PCR product was purified using a Wizard® kit (Section 2.6.2) to remove the PCR primers. Probes were produced in two ways. Firstly, probes were produced in 25 μl reactions containing 50 - 60 ng DNA template, 2.5 μl random nonamers (Biotech International), 2.5 μl of dNTP (2 mM dATP, dGTP, dTTP), 5 units Klenow enzyme (Boehringer Mannheim), 1X Klenow buffer (supplied as 10X buffer containing 500 mM Tris-HCl pH 7.2, 100 mM MgSO4, 1 mM DTT) (Promega) and 50 μCi α-32P dCTP (3 000 mCi/μmol). The DNA template and extra water were heated at 95°C for 10 min in a Hybaid™ Thermal Reactor (AusBiosearch) before the addition of the other reagents. The solution was incubated at 37°C for 30 min followed by 3 min at 95°C. The second method for making probes was using a GIGAPRIME DNA Labelling Kit (Bresatec). The kit was used as specified by the manufacturer for use with α-
32P-dCTP. Steps 1-3 were carried out as specified by the manufacturer and the probe was denatured for 3 min at 95°C before adding it to the Southern blot membrane. The incorporation efficiency of the radioactively labelled probes was not tested, because the required machinery was broken during the course of the study. When blots of PCR products were done, the high copy-numbers of these DNA fragments made high-isotope-incorporation unnecessary.

### 2.11.5 Washing the membrane.

After hybridisation to the probe, the filter was washed to remove unbound or loosely bound probe in a series of solutions of different salt concentrations and at different temperatures. All filters were washed twice with a 50 ml solution of 2x SSC and 0.2% SDS on the bench, and once with a 50 ml solution of 1x SSC and 0.2% SDS at 60°C for 1 hr in a hybridisation oven. Additionally, some filters were washed with solutions containing 0.2% SDS and 0.5x, 0.2x or 0.1x SSC at 65°C for 1 hr in a hybridisation oven. All filters were removed from the hybridisation bottle, wrapped in plastic sheets and placed next to film (Kodak and Fuji) in a lead lined case at -80°C for between 2 hrs and 7 days before the film was processed.
Chapter 3. A survey of the GBSS alleles present in WA noodle-wheat varieties based on the DNA sequence of a 500bp section of the genes.

3.1 Introduction.

One aim of my research was to assess how many different granule-bound starch synthase (GBSS) alleles are present in the wheat breeding populations available to wheat breeders at the Department of Agriculture Western Australia (DAWA), and so a survey of the GBSS genes was undertaken. The approach was to choose a selection of wheat varieties or advanced breeding lines that represented the WA germplasm, sequence the GBSS genes at all three loci and compare those sequences. Polymorphisms could be used to create molecular markers to distinguish the different genes. The novel GBSS alleles might be differentiated by their DNA sequence, even if the differences were in introns and there were no protein differences.

There was good evidence to support the presence of different GBSS alleles in the germplasm of wheat grown in WA, because of the detection of different starch characteristics (Briney et al., 1998; Shariflou and Sharp, 1999). In WA varieties, both wild-type and null GBSS alleles had been identified at all three GBSS loci by examination of their GBSS protein signatures (Crosbie, 1991; Yamamori et al. 1994). It was also predicted that novel GBSS alleles existed in WA wheat varieties that had not yet been described because their enzymes activities were masked by the effects of other alleles. The selection of varieties to be studied included those with null alleles to identify the DNA sequence differences that characterised those alleles. The null allele from the 4A locus resulted from a deletion of the entire gene sequence, therefore this allele was not studied, however the null alleles from the 7A and 7D loci were investigated. Wheat varieties with various starch qualities were included in the study, as well as parents from the noodle-wheat breeding germplasm that displayed high-FSV and medium-FSV starch characteristics. Novel GBSS alleles could also be predicted to occur in genetically distant varieties, because genetic drift causes allelic variation, so wheat varieties with various pedigrees were studied. DNA sequence information from ‘Chinese Spring’ wheats was also included in
the study to investigate how similar their genes are to the alleles present in the WA wheat varieties.

This experiment concentrated on accessing DNA sequence information from the GBSS genes, as opposed to searching for protein or RNA information. DNA polymorphism can mark an allele better than trait polymorphism because it does not vary with the environment that the gene is expressed in. This was particularly important for the GBSS genes where a single mutant copy of a gene is masked by the effects of the other five gene copies, and the effects of a single mutant GBSS enzyme are masked by the compensatory effects of the enzymes from the other five alleles. The DNA mutations that cause allelic variation amongst the GBSS genes can occur anywhere in the gene including both the coding regions and untranscribed promoter regions. To find the DNA sequence variation and to identify novel GBSS alleles, access to the entire genomic DNA sequence of the GBSS gene was needed. At the start of this project, the entire DNA sequence of the GBSS gene was not available; only a cDNA (Clark et al., 1991) had been sequenced. There were no PCR primers to amplify the entire sequence and no genomic libraries of WA wheat varieties existed. However, as discussed in Chapter 1, the PCR primers GBSS2/GBSS4 (Briney et al., 1998) were available and allowed access to an approximately 500bp section of the GBSS genes from all three loci. The PCR primers GBSS2/GBSS4 bound to exons 4 and 6 of the wheat GBSS genes as shown below in Figure 3.1 and amplified bands

![Figure 3.1](image-url). A diagrammatic representation of the GBSS genes showing the binding positions of the PCR primers GBSS2 and GBSS4 within exons 6 and 4 respectively. The 5’ promoter region is shown as a pale blue box whilst the 3’ untranslated region is shown as a light pink arrow. The 11 exons of the gene are shown as blue boxes, with the 10 introns shown as red lines. (The diagram is to scale for the 7A locus.)
from all three GBSS loci as previously shown in Figure 1.8. The primer binding sites in exons 4 and 6 were highly conserved between different species (Briney et al., 1998) and between the three loci (Hollingsworth, 1997), therefore the sequence conservation should extend to all GBSS alleles.

The section of the GBSS genes amplified by this PCR included two introns regions that may vary between different GBSS alleles. The DNA sequences of these introns are not highly conserved between the three genomes of wheat, and they differ between the three GBSS loci (Hollingsworth, 1997). Figure 3.2 shows the DNA sequences of the PCR fragments previously amplified from each of the three GBSS loci by Hollingsworth (1997). The DNA sequence differences between the fragments are coloured to highlight the fact that the intron regions of the genes are more polymorphic than the conserved exon regions. Introns are removed from the mRNA transcribed from the GBSS genes and not translated into protein, therefore the DNA sequences of introns do not affect the protein product of the gene and introns tend to mutate at a faster rate than the exon regions encoding the protein. Thus, two different GBSS alleles might have DNA differences within introns 4 and 5 enabling them to be differentiated even if a mutation affecting the protein was not found.
Figure 3.2. DNA sequence alignment of a section from the GBSS genes at all three GBSS loci on chromosomes 4A, 7A and 7D (from Hollingsworth, 1997). The donor variety was ‘Kulin’ thus the sequences are labeled with the name of the donor variety and with the GBSS locus from which that sequence originated. They are compared to a wheat GBSS cDNA sequence (Ainsworth et al., 1991) which came from ‘Chinese Spring’. The DNA sequence was amplified from the genome by PCR using the primers GBSS2/GBSS4 as labeled on the sequence. The exons and introns within the sequence have also been labeled. DNA sequence polymorphism between the gene fragments from each locus and the cDNA have been labeled beneath with a star and coloured red. Those polymorphisms labeled with ‘1’ show the similarity between the gene from the 7A locus and the cDNA sequence. The polymorphic (TGCCG)n microsatellite has been coloured blue, with ‘Kulin-7A’ only showing three of its four microsatellite repeats.
3.2 Materials.

The wheat varieties chosen to be surveyed for their GBSS alleles in the initial experiment and their pedigrees are shown below in Table 3.1. The varieties represent a selection of Australian Standard White and Australian Hard wheats. The varieties ‘Kulin’ and ‘Reeves’ were genetically similar and have similar backgrounds except at the 4A locus where ‘Reeves’ contains the null 4A allele whilst ‘Kulin’ contains the wild-type 4A allele. The variety ‘DHW174’ is a waxy wheat and supplied the null alleles from the 7A and 7D loci. Although ‘Machete’, ‘Reeves’ and ‘Trident’ contain the null 4A alleles, and were bred to be noodle-wheat varieties, they display medium-FSV starch characteristics that make them unsuitable for Udon noodle production.

Table 3.1. Wheat varieties used in the initial survey, their GBSS characteristics and their pedigrees.

<table>
<thead>
<tr>
<th>WHEAT VARIETY</th>
<th>GBSS</th>
<th>FLOUR SWELLING VOLUME</th>
<th>PEDIGREE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amery</td>
<td>w-type 4A</td>
<td>low</td>
<td>Lr 21 - Sr X/2<em>Shortim//3</em>Bodallin</td>
</tr>
<tr>
<td></td>
<td>w-type 7A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>w-type 7D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DHW174</td>
<td>NULL 4A</td>
<td>very high</td>
<td># sourced from P. Sharp (University of Sydney, NSW). The null 7A allele came from the Japanese wheat ‘Fujimikomugi’ possibly the same allele as in ‘Kanto107’. The null 7D allele came from ‘BaiHou’.</td>
</tr>
<tr>
<td></td>
<td>NULL 7A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NULL 7D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Janz</td>
<td>w-type 4A</td>
<td>low</td>
<td>3Ag3/4*Condor//Cook</td>
</tr>
<tr>
<td></td>
<td>w-type 7A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>w-type 7D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kulin</td>
<td>w-type 4A</td>
<td>low</td>
<td>Bodallin Sib.//(Hyden Sib.)Gamenya/ Inia-66</td>
</tr>
<tr>
<td></td>
<td>w-type 7A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>w-type 7D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Machete</td>
<td>NULL 4A</td>
<td>medium</td>
<td>Mec-3/2*Gabo(RAC177)//Madden</td>
</tr>
<tr>
<td></td>
<td>w-type 7A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>w-type 7D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reeves</td>
<td>NULL 4A</td>
<td>medium</td>
<td>Bodallin//(Hyden Sib.)Gamenya/Inia-66</td>
</tr>
<tr>
<td></td>
<td>w-type 7A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>w-type 7D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tincurrin</td>
<td>w-type 4A</td>
<td>low</td>
<td>Gluclub/3/(AWX10-C-61-1)Chile-1B//Insignia/Falcon.</td>
</tr>
<tr>
<td></td>
<td>w-type 7A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>w-type 7D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trident</td>
<td>NULL 4A</td>
<td>medium</td>
<td>VPM1.5<em>Cook/4</em>Spear</td>
</tr>
<tr>
<td></td>
<td>w-type 7A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>w-type 7D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81W :1137</td>
<td>w-type 4A</td>
<td>low</td>
<td>Bodallin/3/Ciano/Gamenya/2/XBVT223/4/Atlass66/2*Madden</td>
</tr>
<tr>
<td></td>
<td>w-type 7A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>w-type 7D</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In later experiments a larger range of wheat varieties were used and they are described in Section 2.1.

DNA was extracted from the wheat varieties using the CTAB DNA extraction method (Section 2.2.1). The desired PCR fragments were amplified from the GBSS genes in these varieties using the primer combination GBSS2/GBSS4. PCR reactions were prepared in 50µl volumes containing 1.5mM MgCl2 (Section 2.4). The following thermal cycling conditions were used: 1 cycle of 94°C for 3min and 60°C for 2min; 35 cycles of 94°C for 15sec, 60°C for 2min and 72°C for 2min; 1 hold at 72°C for 20min. This PCR amplified a 465bp fragment from the 4A locus, a 492bp fragment from the 7A locus and a 556bp fragment from the 7D locus. Many other bands were also amplified in this PCR and most have previously been identified as being homoprimed PCR artefacts unrelated to the GBSS genes (Hollingsworth, 1997). The PCR products were purified away from the primers using a Wizard® PCR Preps DNA Purification System (Promega) and were cloned into the pGEM-T vector (Section 2.8) using an insert:vector ratio of 8:1. The resulting clones were screened for the inserts they contained using two PCR tests, one using the T7/Sp6 primer combination and one using the GBSS-Rev/GBSS-4A,7A,7D primer combination (Section 2.9). Clones containing inserts were then sequenced (Section 2.10.1). New PCR primers were designed to GBSS sequences and analysed for dimer formation using Amplify software (Engels, 1993). PCR conditions were optimised by varying the annealing temperature.

3.3 Results.

3.3.1 Comparison of the PCR screening tests used to identify clones with GBSS inserts.

The first PCR test used to identify clones with GBSS inserts, used the T7/Sp6 primers to amplify from the vector across the insert, allowing the size of the insert to be seen. As the sizes of the GBSS-related PCR fragments were known to be 465bp, 492bp and 556bp, clones containing these inserts could be identified. The primers amplified from vector sequences so the resulting PCR products were 162bp larger than the
sizes of the inserts themselves. Figure 3.3 shows results of this PCR and the sizes of the inserts within 15 clones. Amplification with the T7/Sp6 primers enabled clones containing inserts of the desired size to be identified, but the inserts had to be sequenced to confirm the presence of GBSS inserts and their originating alleles.

The second PCR test was with primer GBSS-Rev in combination with the primers GBSS-4A, GBSS-7A or GBSS-7D, which amplified from the section of the GBSS genes shown in Figure 3.4. Each set of primers amplified from only one specific GBSS locus, but used in combination, these primers amplified from all GBSS genes. Since these GBSS-Rev/GBSS-4A,7A,7D primers amplified from inside the cloned PCR fragment originally amplified using the GBSS2/GBSS4 primers, they could be used to screen clones containing GBSS2/GBSS4 inserts (Section 2.9).
When a mixture of the primers GBSS-Rev/GBSS-4A,7A,7D was used in PCR to screen clones for the inserts they contained, a single band was amplified from a GBSS insert and no band from an insert unrelated to the GBSS genes (Figure 3.5).

**Figure 3.4.** Positions of PCR primers on the GBSS gene. The GBSS gene is represented by the horizontal line with exons as filled line and introns as dashed line. The positions of primers GBSS4 and GBSS2 in exons 4 and 6 respectively flank the positions of the locus-specific forward primers GBSS-4A, GBSS-7A, GBSS-7D and their shared reverse primer GBSS-Rev.

When a mixture of the primers GBSS-Rev/GBSS-4A,7A,7D was used in PCR to screen clones for the inserts they contained, a single band was amplified from a GBSS insert and no band from an insert unrelated to the GBSS genes (Figure 3.5).

**Figure 3.5.** 1% agarose gel showing use of the PCR primer combination GBSS-Rev/GBSS-4A,7A,7D to screen clones for the inserts they contained. Lane 1: Molecular size marker 100bp DNA ladder; Lanes 2-18: individual clones containing GBSS inserts (lanes 5,11) and non-GBSS inserts (lanes 2-4, 6-10, 12-18) which have been screened using PCR primers GBSS-Rev/GBSS-4A,7A,7D. Each lane contained the PCR test on a different clone. The primers amplified a band from any insert related to the GBSS genes such as those in Lanes 5 and 11. The clones in all other lanes contained non-GBSS inserts and thus no PCR bands were amplified.
Once clones had been identified as containing GBSS-related inserts using the GBSS-Rev/GBSS-4A,7A,7D primer combination, they were screened to find the GBSS locus from which the insert had originated. This was done in three separate PCR reactions using the primer GBSS-Rev with one of each of the primers GBSS-4A, GBSS-7A and GBSS-7D individually (Section 2.9). A band would be amplified in only one of the three reactions - the reaction containing the primer for one specific locus. Figure 3.6 shows PCR results of a screen of five clones that had been previously identified as containing GBSS-related inserts. Using these primers, the inserts could be fully characterised prior to growing up the clones, extracting the plasmids and sequencing the inserts.

![Figure 3.6. 1% agarose gel showing how clones previously identified to contain a GBSS-related insert were screened, using the three locus-specific PCR primer sets, to ascertain which GBSS locus the insert came from. Lane 1: Molecular size marker; Lanes 2-6: five clones identified as containing GBSS-related inserts and screened with three sets of GBSS locus-specific primers to identify which locus 4A, 7A or 7D the insert was derived from. The five clones (Lanes 2-6) were each screened in three locus-specific PCR reactions to identify which locus 4A, 7A or 7D the insert came from. The labels ‘4A’, ‘7A’ and ‘7D’ on the Figure refer to those locus-specific PCR tests. The clone tested in Lane 2 contained an insert from the 4A locus, the insert in the clone in Lane 6 was from the 7A locus and the clones in Lanes 3 and 4 contained inserts from the 7D locus. The insert in the clone in Lane 5 was not identified.](image-url)
When the two PCR tests, T7/Sp6 and GBSS-Rev/GBSS-4A,7A,7D, were compared, this second GBSS-specific method was preferred. This method reduced the time taken to find the GBSS clones and the costs of the experiments because the inserts did not need to be sequenced before identification, and there was no sequencing of the same fragment from multiple clones.

3.3.2 Comparison of GBSS alleles at the 4A locus of WA wheat varieties.

The 465bp PCR fragments amplified from GBSS genes at the 4A locus of a selection of wheat varieties were sequenced and are shown in a DNA sequence alignment in Figure 3.7. Included in the alignment is the wheat cDNA sequence (Clark et al., 1991) commonly used to position the intron regions. This cDNA was from a ‘Chinese Spring’ variety and was thought to have originated from the 7A locus. Also included is the genomic sequence of a GBSS gene from the 4A locus of a ‘Chinese Spring’ variety (Murai et al., 1999). Although these two sequences both originated from ‘Chinese Spring’ varieties, they show DNA sequence differences such as those highlighted by the number ‘5’ (Figure 3.7) that were genome-specific mutations varying between the two loci (Murai et al., 1999). At these nucleotide positions, the Australian varieties were more similar to the ‘Chinese Spring’-4A sample than to the cDNA-7A sample, showing that the mutations were locus-specific markers conserved between wheat varieties. The only exception was the first of these differences, labelled with ‘6’ (Figure 3.7), which fell within a primer sequence and at this nucleotide the Australian varieties were identical to the cDNA from which the primers were designed.

The DNA sequence from ‘Janz’ appeared to be very different from that of the other wheat genotypes. The first section of the sequence from ‘Janz’ comes from the 4A locus as shown by the sequence similarity up to and including the base highlighted with ‘1’ that is a marker for sequences from the 4A locus (Figure 3.2). The second section of the sequence from ‘Janz’ comes from the 7A locus as shown by the 7A-specific marker bases highlighted with ‘2’ (Figure 3.7) and the 7A-specific intron sequence (Figure 3.2). Thus, this PCR product was probably a
GBSS2/GBSS4 PCR fragment was again produced from 'Janz' and sequenced (data not shown). To confirm this hypothesis the GBSS genes at the 4A locus of a selection of 4 varieties listed to the left of the alignment. Included are the genomic 4A gene from 'Chinese Spring’ wheat (Murai et al., 1999) and a wheat cDNA from the 7A locus (Clark et al., 1991) to position the introns. Gaps in the sequence are shown with dashes. DNA polymorphism between the sequences is highlighted with stars or numbers and coloured red.

Figure 3.7. DNA sequence alignment of PCR fragments amplified from GBSS genes at the 4A locus of a selection of 4 varieties listed to the left of the alignment. Included are the genomic 4A gene from ‘Chinese Spring’ wheat (Murai et al., 1999) and a wheat cDNA from the 7A locus (Clark et al., 1991) to position the introns. Gaps in the sequence are shown with dashes. DNA polymorphism between the sequences is highlighted with stars or numbers and coloured red.
shown). This new sequence matched the original sequence from ‘Janz’ up to the point where it began to be amplified from the 7A locus. After this point, the new ‘Janz’ sequence exactly matched the others from other varieties suggesting that the initial sequence had resulted from PCR error. The possibility that the initial sequence had been amplified from a chimeric pseudogene copy of a GBSS gene could not be confirmed as the initial PCR product could not again be amplified.

There were several single base differences between the sequences from the 4A locus of the four varieties tested (Figure 3.7). Some of these may have been PCR errors, such as the two highlighted with ‘7’ (Figure 3.7) which differed between the two replicate sequences from the same variety ‘Tincurrin’. The use of a more specific, high-fidelity DNA polymerase to amplify these sequences in PCR may have reduced these PCR errors, however in this experiment the PCR had previously been optimised to use Tth+ polymerase (Biotech International) and thus the conditions were not modified. Alternatively, since other wheat varieties also showed variation at these bases (‘7’, Figure 3.7) they may be valid single nucleotide polymorphisms (SNPs). One of these bases fell within intron 4 and does not affect the sequence of the protein. The other of these bases was in exon 4, however the nucleotide is the third base in the codon and both the ‘C’ and the ‘T’ code for the same amino acid. Thus neither potential SNP affected the protein sequence of these genes. Nevertheless, if these SNPs were validated and proved to truly exist, molecular markers could be created to identify the different alleles and their possible contribution to starch quality determined.

3.3.3 Design and partial optimisation of a test to validate a SNP at the 4A locus.

Towards validating the ‘C/T’ SNP in exon 4 of the GBSS-4A genes (Figure 3.7), a new PCR test was designed so that the SNP could be investigated using the non-gel-based MALDI TOF (matrix assisted laser desorption/ionization time of flight) mass spectrometry method discussed at the end of the Chapter rather than by sequencing as above. The GBSS2/GBSS4 primers used previously were unsuitable for analysing the SNP using MALDI TOF, because they amplify
many PCR fragments which might confuse the analysis of the SNP and also the PCR fragments were larger than is optimal (150bp) for use in SNP analysis. New PCR primers were designed to only amplify from the 4A locus and to produce an approximately 150bp PCR fragment (Figure 3.8). A sequencing primer was also designed adjacent to the putative SNP.

To optimise the PCR, the new primers GBSS4-2/GBSS-4A-2 were used together, as well as in the primer combination GBSS4/GBSS-4A-2 to amplify from the 4A locus of ‘Kulin’ in PCR with 2mM MgCl₂ (Section 2.4). The reactions were incubated under a gradient of annealing temperatures from 54°C to 62°C as shown in Figure 3.9. Both combinations of primers produced fragments of approximately the correct size to be the expected, a 136bp fragment as circled in the Figure. Both PCRs produced unexpected bands larger than predicted, some of which could be identified as being homoprimed by using the primers singularly in PCR as shown at the bottom of Figure 3.9. The new primers GBSS4-2 and GBSS-4A-2 did not generate homoprimed bands, whilst the primer GBSS4 produced many homoprimed fragments over a range of temperatures. The primer GBSS4-2 may therefore be a candidate to replace

**Figure 3.8.** DNA sequence of the PCR product amplified from the 4A locus of ‘Kulin’ by the primers GBSS2/GBSS4 (underlined) showing the positions of the new potentially locus-specific primers GBSS4-2 and GBSS-4A-2 (red) designed to amplify a 136bp fragment for SNP analysis. The sequencing primer (blue) is shown next to the putative SNP (bold). The exons (capitalised) and introns (lower case) within the sequence have been labeled.
GBSS4 in PCRs to screen for the presence of null 4A alleles, but this was not investigated further. Since the approximately 136bp fragments amplified in both dual-primed PCRs were not homoprimed fragments, they may have been specifically amplified from the 4A GBSS locus.

Figure 3.9. Three 1% agarose gels showing PCR bands amplified from the genomic ‘Kulin’ DNA over a gradient of annealing temperatures. The primers used in each PCR reaction are shown boxed above the gel pictures. The lanes on each of the three gels correspond to each other and are labeled at the top with the annealing temperature used to amplify that reaction. The two bottom gels show PCR reactions containing only a single primer and show multiple homoprimed fragments from the primer GBSS4 but none from the primers GBSS4-2 or GBSS-4A-2. The top gel shows the dual-primed PCR reactions with the expected band circled in white. Some of the bands seen in the top gel can be associated with homoprimed bands seen in the middle gel.
Optimising the PCR to amplify the C/T SNP from the 4A locus was started, but additional work is necessary to either reduce the non-specific amplification or to test if the presence of the extra bands reduces the efficiency of SNP analysis by MALDI TOF MS. The remaining steps towards screening the putative SNP were to clean the PCR product away from primers and single nucleotides, attach the sequencing primer, add a single base of sequence and identify that added base using the MALDI TOF mass spectrometer. Unfortunately, this could not be done due to the time constraints on the project and it is recommended that this part of the project be continued by future researchers.

3.3.4 The GBSS alleles at the 7D locus of WA wheat varieties.

PCR fragments were amplified and sequenced from the GBSS genes at the 7D locus of a selection of wheat varieties and the alignment of those sequences is shown in Figure 3.10. The first four sequences were shorter than the others because they were amplified from the genome using the locus-specific primer pair GBSS-Rev/GBSS-7D (Section 2.9) that bound inside the sequence of the primers GBSS2/GBSS4. These samples were directly sequenced from the PCR product (Section 2.10.2). Included in the alignment (Figure 3.10) was the wheat cDNA thought to be from the 7A locus of ‘Chinese Spring’ wheat (Clark et al., 1991) used to position the intron regions. Also included was the genomic sequence of a GBSS gene from the 7D locus of a ‘Chinese Spring’ variety (Murai et al., 1999). Many polymorphisms could be seen between the ‘Chinese Spring’ genomic sequence from the 7D locus and the cDNA from the 7A locus (Figure 3.10). These differences were genome-specific polymorphisms and it could be seen that the sequences from the 7D locus of the Australian varieties retained those genome-specific bases showing greater similarity to the sequence from the 7D locus of ‘Chinese Spring’ wheat than that from the 7A locus.

The DNA sequence of the null 7D allele from ‘DHW174’ showed little difference from the wild-type 7D alleles of the other samples. The only difference was an ‘A’ to ‘G’ substitution in one of the ‘DHW174’ repeat samples that did not occur when this allele was re-sequenced (data
**Figure 3.10.** DNA sequence alignment of PCR fragments amplified from GBSS genes at the 7D locus of a selection of wheat varieties listed to the left of the alignment. Included are the genomic 7D sequence from ‘Chinese Spring’ wheat (Murai et al., 1991) and a GBSS cDNA sequence from the 7A locus of ‘Chinese Spring’ wheat (Clark et al., 1991). DNA sequence polymorphisms are highlighted below with a star and coloured red. The red star denotes a polymorphic poly-T' region.
not shown). This base change was therefore considered a PCR error and the sequences of the null and wild-type alleles were seen to be identical in this region of the gene. Comparing the sequences from Australian varieties, very high sequence similarity was seen between the genotypes (Figure 3.10). There were 7 positions where single bases differed between the sequences. Some of these may have been PCR error as they were not reproducible across repeated sequencing of the same allele. Again, the use of high-fidelity polymerase may reduce PCR errors when future sequencing projects strive to validate these single nucleotide polymorphisms.

One interesting feature of the sequences shown on Figure 3.10 was the poly ‘T’ motif within the first intron of the sequence. There appeared to be (T)$_n$ polymorphism with some samples having an extra nucleotide. However, poly-‘T’ regions are very difficult for polymerase enzymes and sequencing software to read correctly so the polymorphism may have been a PCR error. More investigation is required to confirm the polymorphism of this motif. As this region was within an intron, its sequence did not affect the resulting protein, however, if this poly-‘T’ was proven to be polymorphic, it might define a novel allele and be useful as a molecular marker for that allele.

### 3.3.5 Design and partial optimisation of a test to validate a SNP at the 7D locus.

Towards characterising the poly-‘T’ region at the 7D locus, a SNP detection assay was designed as described previously in Section 3.8. New PCR primers were designed to amplify an approximately 150bp PCR fragment (Figure 3.11) that could be used as template for a short sequencing reaction, the product of which could be identified by MALDI TOF mass spectrometry. The sequencing primer was designed adjacent to the poly-‘T’ region (Figure 3.11) so that various numbers of dTTPs could be added in the sequencing reaction.
The PCR primer combination GBSS-7D/GBSS-7D-2 was used in PCR with both 2 and 3 mM MgCl₂ (Section 2.4) with the reactions incubated over a range of annealing temperatures between 54°C and 71°C. Only using 57°C annealing temperatures and 2 mM MgCl₂ was a faint band of the expected size seen, with no bands produced at any other temperatures (data not shown). Unfortunately, due to the time constraints of the project, further optimisation of this PCR could not be done.

3.3.6 Comparison of the GBSS alleles at the 7A locus of WA wheat varieties.

Figure 3.12 shows a DNA sequence alignment of the GBSS2/GBSS4 PCR fragments amplified from the 7A locus of a selection of wheat varieties. The first three sequences were shorter than the others because they were amplified from the genome using the locus-specific primer pair GBSS-Rev/GBSS-7A (Section 2.9) that bound inside the sequence of the primers GBSS2/GBSS4. These samples were directly sequenced from the PCR product (Section...
2.10.2). Included in the alignment is the wheat GBSS cDNA thought to be from the 7A locus of ‘Chinese Spring’ wheat (Clark et al., 1991) used to position the intron regions and the genomic sequence from the 7A locus of a ‘Chinese Spring’ variety (Murai et al., 1999). As can be seen, these two genotypes were homologous, the genomic sequence showing greater sequence similarity to the cDNA sequence than sequences from other loci, adding evidence to the argument that the cDNA originated from the 7A locus.

Between the genotypes of Australian wheat there were eight single nucleotide polymorphisms (SNPs) some of which may have been PCR errors because they were not repeatable across multiple sequencing events (Figure 3.12). Some however may prove to be real SNPs upon further study and thus define and mark novel alleles. No DNA polymorphism was seen between the null 7A allele from the ‘DHW174’ sample and the wild-type 7A alleles from the other varieties.

The most polymorphic section of the sequences was a five-base pair motif of ‘TGCCG’ within intron 4 (Figure 3.12). This motif was repeated four times in the varieties ‘Kulin’, ‘Reeves’, ‘Machete’ and ‘Trident’ and three times in all other samples. Repeated elements like this are termed microsatellites. This was the first time this (TGCCG)$_n$ microsatellite had been reported. The reason this microsatellite had not been previously identified could be because it may not be polymorphic between ‘Chinese Spring’ varieties and these are the varieties being studied by researchers outside Australia. However, because this microsatellite is polymorphic within Australian germplasm, it was readily identified here. As this microsatellite polymorphism could be used to separate the 7A alleles in the varieties tested, it defined the genes as separate alleles even though this region was within an intron and thus did not affect the protein sequence. The allele marked by the (TGCCG)$_3$ motif present in the ‘Chinese Spring’ sample had been reported in various literature as $Wx$-$A1a$ so the allele marked by the (TGCCG)$_4$ motif was seen as a novel allele present in Australian varieties.
<table>
<thead>
<tr>
<th>Variety</th>
<th>Sequence 1</th>
<th>Sequence 2</th>
<th>Sequence 3</th>
<th>Sequence 4</th>
<th>Sequence 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reeves_2-7A</strong></td>
<td>GTGCTGCCAT</td>
<td>GCACTGATTT</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
</tr>
<tr>
<td><strong>Kuli in_2-7A</strong></td>
<td>GTGCTGCCAT</td>
<td>GCACTGATTT</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
</tr>
<tr>
<td><strong>DHW_7A</strong></td>
<td>GTGCTGCCAT</td>
<td>GCACTGATTT</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
</tr>
<tr>
<td><strong>Amery_7A</strong></td>
<td>GTGCTGCCAT</td>
<td>GCACTGATTT</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
</tr>
<tr>
<td><strong>Machete_2-7A</strong></td>
<td>GTGCTGCCAT</td>
<td>GCACTGATTT</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
</tr>
<tr>
<td><strong>Machete_1-7A</strong></td>
<td>GTGCTGCCAT</td>
<td>GCACTGATTT</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
</tr>
<tr>
<td><strong>Tincurrin_1-7A</strong></td>
<td>GTGCTGCCAT</td>
<td>GCACTGATTT</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
</tr>
<tr>
<td><strong>Tincurrin_2-7A</strong></td>
<td>GTGCTGCCAT</td>
<td>GCACTGATTT</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
</tr>
<tr>
<td><strong>Ch Spring_7A</strong></td>
<td>GTGCTGCCAT</td>
<td>GCACTGATTT</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
</tr>
<tr>
<td><strong>Wheat_cDNA</strong></td>
<td>GTGCTGCCAT</td>
<td>GCACTGATTT</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
<td>CAGCCGATTC</td>
</tr>
</tbody>
</table>

Figure 3.12. DNA sequence alignment of PCR fragments amplified from GBSS genes at the 7A locus of a selection of wheat varieties listed to the left of the alignment. Included are genomic and cDNA GBSS sequences from the 7A locus of ‘Chinese Spring’ wheat (Murai et al., 1999; Clark et al., 1991). DNA sequence polymorphisms are highlighted with a star and coloured red. The (TGCCG)n microsatellite is coloured blue.
3.3.7 The linkage of the (TGCCG)ₙ microsatellite variation and starch quality.

The (TGCCG)ₙ microsatellite polymorphism at the GBSS 7A locus occurring within intron 4, did not affect the amino acid sequence of the protein product and would not affect the starch quality. However, it defined a novel GBSS allele that might show altered activity due to other mutations within the coding regions and linked to the microsatellite variation. It was predicted that a novel GBSS allele at the 7A or 7D loci was responsible for the variation in FSV amongst noodle-wheat varieties and so the identification of a novel 7A allele marked by the microsatellite polymorphism suggested that this putative novel allele had been discovered. If the microsatellite variation was linked to FSV characteristics it could be used to screen the noodle-wheat breeding stock to eliminate the gene responsible for undesirable starch quality.

The two alleles defined by the microsatellite polymorphism were tested for their linkage to FSV traits.

The previously sequenced varieties ‘Reeves’, ‘Machete’ and ‘Trident’ with the GBSS allele containing the (TGCCG)₄ microsatellite (Figure 3.12) were all null 4A and all showed the medium-FSV values of poor noodle-quality varieties. However, the samples that showed the (TGCCG)₃ allele were all wild-type 4A and thus showed the low-FSV due to the presence of this gene. To test the hypothesis that the microsatellite variation was linked to FSV traits, a selection of wheat varieties with either medium-FSV or high-FVS starch characteristics was collected, as listed below in Table 3.2 and their 7A alleles investigated. As the microsatellite of interest was located between the primer binding sites of the locus-specific PCR primers GBSS-Rev/GBSS-7A, these primers were used as described above to amplify from each of the varieties and directly sequence the microsatellite region in one direction only, using the primer GBSS-7A.

The number of repeats of the microsatellite in each wheat variety were counted and the results are presented in Table 3.2. Whilst most of the varieties with high-FSV characteristics contained the (TGCCG)₃ allele and most of the varieties with medium-FSV characteristics contained the
(TGCCG)$_4$ allele, there were exceptions to both groups. There was little linkage between the number of microsatellite repeats and FSV characteristics. It was therefore concluded that the two alleles characterised by the microsatellite variation were not responsible for the FSV characteristics seen.

<table>
<thead>
<tr>
<th>WHEAT VAR. WITH HIGH-FSV VALUES</th>
<th>NUMBER OF MICROSATELLITE REPEATS</th>
<th>WHEAT VAR. WITH MEDIUM-FSV VALUES</th>
<th>NUMBER OF MICROSATELLITE REPEATS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arrino</td>
<td>3</td>
<td>Aroona</td>
<td>4</td>
</tr>
<tr>
<td>Bodallin</td>
<td>3</td>
<td>Calingiri</td>
<td>4</td>
</tr>
<tr>
<td>Cadoux</td>
<td>4</td>
<td>Carnamah</td>
<td>4</td>
</tr>
<tr>
<td>Eradu</td>
<td>3</td>
<td>Cascades</td>
<td>4</td>
</tr>
<tr>
<td>Express</td>
<td>3</td>
<td>Halberd</td>
<td>3</td>
</tr>
<tr>
<td>Gamenya</td>
<td>4</td>
<td>Machete</td>
<td>4</td>
</tr>
<tr>
<td>Reeves</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Spear</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Tammin</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Trident</td>
<td></td>
<td></td>
<td>4</td>
</tr>
</tbody>
</table>

3.3.8 Development of a simple PCR test to distinguish the two alleles defined by polymorphism at the (TGCCG)$_n$ microsatellite locus.

For the polymorphism at the (TGCCG)$_n$ microsatellite locus to be useful for future genetic research, a simpler method of detection than sequencing was necessary. A PCR test useful for screening many wheat samples was developed by creating a new PCR primer. The new primer GBSS-7A-2 was designed at the beginning of exon 5 (Figure 3.13) and could be used in conjunction with the locus-specific primer GBSS-7A to amplify an approximately 150 bp fragment from the GBSS 7A genes containing the microsatellite locus in intron 4. The PCR reaction contained primers GBSS-7A/GBSS-7A-2 and 2mM MgCl$_2$ (Section 2.4). PCR cycles were 1 hold at 94°C for 3min; 35 cycles of 94°C for 30sec, 61°C for 30sec, 72°C for 1min; 1 hold at 72°C for 3min, though any annealing temperature from 54°C to 62°C amplified only the single specific band.
The PCR fragment was amplified from a selection of wheat varieties with known numbers of microsatellite repeats to optimise a gel system able to distinguish the 5bp size differences between fragments amplified from different alleles. Three gel systems: 4%, 6% and 10% polyacrylamide gels (Section 2.5) were tested (data not shown). The 10% polyacrylamide gel was able to separate the PCR fragments (Figure 3.14). The method was validated on the previously sequenced samples listed above in Table 3.2.
3.3.9 Results for screening Australian varieties for the (TGCCG)n microsatellite.

Having validated the PCR test discussed above, it was then used to investigate the GBSS-7A alleles present in a range of previously unexamined wheat varieties (Figure 3.15). The numbers of microsatellite repeats were identified and are listed in Table 3.3 together with the sequenced data from Table 3.2. The variety ‘Batavia’ was seen to be mixed at this locus as supported by M Shariflou (Pers. Comm.). Of the 41 wheat varieties tested, 60% contained the (TGCCG)₃ allele and 40% contained the (TGCCG)₄ allele with one being heterozygous. This new PCR test was shown to be useful for genotyping GBSS-7A alleles and this is discussed further in Chapter 8.

Figure 3.14. A 10% polyacrylamide gel showing the validation of the GBSS-7A/GBSS-7A-2 PCR test to distinguish PCR bands from GBSS-7A alleles containing either three or four (TGCCG)n microsatellite repeats. The two PCR bands were 165 and 170bp in size and were easily differentiable. (Electrophoresis at 120V for 3 hrs)

Lane 1: Molecular size marker 100 bp ladder (Biotech International); Lane 2: Eradu; Lane 3: Arrino; Lane 4: Gamenya; Lane 5: Cascades; Lane 6: Bodallin; Lane 7: Reeves; Lane 8: Calingiri; Lane 9: Machete; Lane 10: water blank.
Table 3.3. Wheat varieties and the GBSS 7A allele they contain as differentiated by the number of repeats (3 or 4) of a (TGCCG)n microsatellite.

<table>
<thead>
<tr>
<th>Wheat</th>
<th>allele</th>
<th>Wheat</th>
<th>allele</th>
<th>Wheat</th>
<th>allele</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amery</td>
<td>3</td>
<td>Cunningham</td>
<td>3</td>
<td>Molineux</td>
<td>4</td>
</tr>
<tr>
<td>Aroona</td>
<td>4</td>
<td>Dollarbird</td>
<td>3</td>
<td>Pelsart</td>
<td>3</td>
</tr>
<tr>
<td>Arrino</td>
<td>3</td>
<td>DHW174</td>
<td>3</td>
<td>Reeves</td>
<td>4</td>
</tr>
<tr>
<td>Barunga</td>
<td>4</td>
<td>Eradu</td>
<td>3</td>
<td>Rowan</td>
<td>3</td>
</tr>
<tr>
<td>Batavia</td>
<td>mixed</td>
<td>Express</td>
<td>3</td>
<td>Schomburg</td>
<td>4</td>
</tr>
<tr>
<td>Bodallin</td>
<td>3</td>
<td>Excalibur</td>
<td>3</td>
<td>Spear</td>
<td>3</td>
</tr>
<tr>
<td>Cadoux</td>
<td>4</td>
<td>Frame</td>
<td>4</td>
<td>Sunco</td>
<td>3</td>
</tr>
<tr>
<td>Calingiri</td>
<td>4</td>
<td>Gamenya</td>
<td>4</td>
<td>Suneca</td>
<td>3</td>
</tr>
<tr>
<td>Carnamah</td>
<td>4</td>
<td>Halberd</td>
<td>3</td>
<td>Tammin</td>
<td>4</td>
</tr>
<tr>
<td>Cascades</td>
<td>4</td>
<td>Hartog</td>
<td>3</td>
<td>Tasman</td>
<td>3</td>
</tr>
<tr>
<td>Chinese</td>
<td>3</td>
<td>Janz</td>
<td>3</td>
<td>Tincurrin</td>
<td>3</td>
</tr>
<tr>
<td>Spring</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condor</td>
<td>3</td>
<td>Kite</td>
<td>3</td>
<td>Trident</td>
<td>4</td>
</tr>
<tr>
<td>Cook</td>
<td>4</td>
<td>Kulin</td>
<td>4</td>
<td>Yarralinka</td>
<td>3</td>
</tr>
<tr>
<td>Cranbrook</td>
<td>3</td>
<td>Machete</td>
<td>4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.15. A 10% polyacrylamide gel showing screening of 19 wheat samples for the number of the (TGCCG)n microsatellite repeats contained in their GBSS-7A alleles using the GBSS-7A/GBSS-7A-2 PCR. In the sample Batavia (Lane 3) two bands were seen indicating that both alleles were present.

Lane 1: Molecular size marker 100 bp ladder (Biotech International); Lane 2: Condor; Lane 3: Batavia; Lane 4: Cook; Lane 5: Cunningham; Lane 6: Sunco; Lane 7: Pelsart; Lane 8: Schomburg; Lane 9: Molineux; Lane 10: Barunga; Lane 11: Frame; Lane 12: Yarralinka; Lane 13: Hartog; Lane 14: Rowan; Lane 15: Dollarbird; Lane 16: Suneca; Lane 17: Kite; Lane 18: Excalibur; Lane 19: Tasman; Lane 20: Cranbrook.
3.4. Discussion.

The genomic DNA sequences of the GBSS genes of wheat were accessed using PCR to amplify a 500bp region from the middle of the GBSS genes at all three loci. This region of the GBSS genes, from exon 4 to exon 6, was amplified and sequenced from a range of wheat varieties predicted to contain different GBSS alleles. The DNA sequences from each GBSS locus were compared, searching for possible polymorphisms that differentiated varying alleles. The information gained from this experiment is discussed below in relation to the questions posed in the aims of this project.

A search for GBSS protein differences to identify new alleles was not planned since the aim was to produce DNA molecular markers to differentiate the different alleles. It is acknowledged that an analysis of differences in protein expression controlled by GBSS promoters using comparative RNA expression assays should be part of future searches for novel GBSS alleles.

3.4.1 Where are the GBSS gene copies in the genome?

The DNA sequences amplified from the GBSS genes in this experiment were so similar to each other that they were probably amplified from the coding copies of the genes and not from pseudogene copies. Pseudogene copies of a gene would be expected to show more DNA sequence polymorphism due to random mutations without selective pressure (Kumar and Bennetzen, 1999; Mighell et al., 2000; Petrov and Hartl, 2000). Also, some pseudogenes are formed from the reverse-transcription of mRNA and do not contain introns. All sequences identified in the experiments described contained introns. The chimeral sequence amplified from the 4A locus of ‘Janz’ could not be confirmed as a pseudogene and was probably a PCR artefact. The first published cDNA sequence from the GBSS genes (Ainsworth et al., 1991) was also a PCR chimera.

That no identifiable pseudogenes were sequenced suggests one of three possibilities. Firstly, there may be newly-inserted pseudogene copies of the GBSS genes in the genome which do not
show DNA sequence differences to the coding copies. Secondly, there may be distantly related pseudogene copies of these genes with DNA sequences so different to the sequences of the coding copies that the PCR primers could not amplify from them. Thirdly, there may not be any pseudogene copies of the GBSS genes in the wheat genome.

3.4.2 How many different GBSS alleles are in the WA wheat breeding population and how does each GBSS allele differ from the others?

It was thought that a range of different GBSS alleles would exist in the WA noodle-wheat breeding population represented by the wheat varieties examined and that the different alleles would be differentiable by DNA sequence polymorphisms. However, less DNA sequence polymorphism than expected was seen in the region of the GBSS genes studied, from exon 4 to exon 6. In this region of the gene, the coding sequences were highly conserved across the different wheat varieties including both ‘Chinese Spring’ and the Australian varieties. Very little DNA polymorphism was seen even in the third base of the codon that is the ‘wobble’-position for amino acid coding. At this third base many substitutions have no affect on the encoded amino acid and these mutations are under less selective pressure than bases in the first two positions of the codon. Single base substitutions in the ‘wobble’-position are seen between the three genomes of wheat yet they were rare within sequences from a single locus. There were some single base polymorphisms identified between the sequences from different varieties and these will need to be validated before being useful as molecular markers to differentiate alleles. It is possible that this region of the GBSS genes, from exon 4 to exon 6, may be more highly conserved than other regions of the gene, due to encoding a vital region of the GBSS protein. The exact role of this section of the protein is not known, however it does not contain the putative active site (Murai et al., 1999) or the putative amylopectin-binding site (Vrinten et al., 1999) and so probably has a structural function.

The intron regions in this section of the gene, introns 4 and 5, are highly polymorphic between the three loci (Hollingsworth, 1997) and this led to a prediction that polymorphism would also
be seen between the alleles from different varieties at a single locus. However, such a high level of sequence polymorphism was not seen within each locus. At each locus, only a few potential SNPs were discovered between genes from the different wheats and these SNPs have yet to be validated. At the 7A locus a polymorphic (TGCCG)$_n$ microsatellite was discovered in intron 4 allowing the differentiation of the alleles at this locus into two groups based on the number of microsatellite repeats they possessed. The wild-type allele contained the (TGCCG)$_3$ motif and the mutant (TGCCG)$_4$ allele was present in 40% of the Australian wheat varieties tested. This polymorphism seemed to have no functional importance since it was in an intron, nor was it linked to noodle starch quality.

Looking at the sequence conservation of the (TGCCG)$_n$ microsatellite from intron 4 of the 7A genes, two alleles (TGCCG)$_3$ and (TGCCG)$_4$ were discovered. When searching the published genomes of the *Triticaceae* accessible through the NCBI search engine (http://www.ncbi.nlm.nih.gov/) for this microsatellite, all the GBSS genes were found to contain the (TGCCG)$_3$ motif, none contained the (TGCCG)$_4$ motif. This suggests that this microsatellite is not highly variable. Due to the repeating structure of microsatellites they are known to be more variable than other regions of the genome and therefore the lack of variability in this microsatellite might suggest that it has a functional role in the DNA, such as microRNA signalling or DNA tertiary structure. Interestingly, the only other place this microsatellite was found in the published *Triticaceae* genome was a (TGCCG)$_5$ motif 26bp downstream from the stop-codon of the ‘GstA1’ gene for glutathione-S-transferase (X56012) from *Triticum aestivum*. If the microsatellite does have some functional significance, a relationship between the GBSS genes and the GstA1 gene might be seen.

The introns showed similar levels of polymorphism to the exons and this finding was unexpected. It was possible that the selection of varieties tested were so closely related that each contained the same allele. However, the identification of a novel allele at the 7A locus showed that the varieties were not all homozygous at this locus. Also, the ‘Chinese Spring’ variety was more distantly related to the Australian varieties and yet still showed the high
sequence similarity. It may be possible that the intron regions of the GBSS genes are under some kind of selective pressure to maintain their sequence or length. Recent studies suggest that microRNAs cleaved from transcribed introns act as messages in the nucleus, complexing with DNA, RNA and proteins to regulate gene expression (Mattick, 2001; Mattick and Gagen, 2001). Thus the intron sequences may have functional significance and would therefore be under evolutionary pressure to maintain their sequence integrity. However, the differences between the introns from the three different loci seem to contradict this theory.

The lack of DNA sequence polymorphism between the null and wild-type alleles at the 7A and 7D loci showed that the mutations responsible for the lack of expression from the null alleles were not within the DNA region studied. This finding was supported by the findings of Vrinten et al. (1999) that the mutations in the null alleles at the 7A and 7D loci are situated at the 5’ and 3’ ends of the transcribed genes respectively. In order to develop molecular markers to detect the null alleles at these loci, these peripheral regions of the GBSS gene must be accessed.

The high levels of DNA sequence similarity between the GBSS genes in this study made identification of novel alleles difficult and suggested that only a few different alleles exist, and even fewer with differences within their coding regions. It was clear that larger sections of DNA sequence from the GBSS genes and a larger sample of wheat varieties would be required to identify other novel alleles. It is one failing of this study that some ASWN wheat varieties were not included in the initial DNA sequencing study. To truly assess the DNA differences that might be responsible for differences in starch quality between null 4A wheat varieties with different FSV characteristics, some high-FSV noodle-wheat varieties need to be sequenced.

3.4.3 How can we tag the different GBSS alleles with molecular markers to identify them in a breeding program?

The desired outcome was to identify DNA polymorphisms in this region of the GBSS genes and to create molecular markers to screen for such polymorphisms in wheat breeding populations.
When designing a molecular marker to a polymorphism, the characteristics of that polymorphism dictate the type of molecular marker designed. For example, if an insertion/deletion mutation was found that was of sufficient size to allow the PCR products amplified from the wild-type and mutant alleles to be distinguished, then a traditional “PCR and gel” marker viewed on agarose or polyacrylamide gels, could be designed. Equally, if the DNA sequences of two alleles differ enough to disrupt a primer binding to one of the alleles, then the lack of PCR product from this allele could be detected on a gel and distinguished from the wild-type allele. However if the polymorphism is smaller, such as a Single Nucleotide Polymorphism (SNP), then the gels required to distinguish the alleles become less economical and non-gel-based detection methods are preferred. One such method is the use of single-base sequencing of the SNP and detection using a MALDI TOF mass spectrometer to detect the mass difference between the sequences from each allele. The method involves amplifying an approximately 150bp PCR fragment containing the potential SNP, denaturing the PCR product to produce a single strand, annealing a sequencing primer directly adjacent to the SNP and adding a single ddNTP in the SNP position. This short DNA fragment is then analysed by MALDI TOF mass spectrometry to detect the presence of either the wild-type or mutant nucleotide based on their molecular mass. Two molecular markers of this kind were designed in this study to validate SNPs from the 4A and 7D loci. Future work should optimise the conditions needed to use these markers to validate the SNPs.

The new, PCR-based molecular marker developed in this project to distinguish two alleles at the GBSS 7A (TGCCG)n microsatellite locus was validated as useful for genotyping wheat varieties at this locus. Using this molecular marker, it is now possible to investigate the inheritance of these alleles as they are employed in a breeding program. By examining the pedigrees of the wheat varieties genotyped at this locus, it was possible to predict the inheritance of the 7A alleles from particular parents and to trace the progress of a single allele through a breeding program in a way not previously possible. Due to the selfing of advanced breeding lines, these wheat varieties were homozygous and carried only one allele from one parent, having lost the other parental allele. This meant that just because a variety had a certain parent in its pedigree
did not mean it carried the GBSS 7A allele from that parent. Now that a molecular marker is available to differentiate the two 7A alleles, we could identify the parent that contributed the gene to the progeny.

One example of this is the relationship between the variety ‘Eradu’ and its parents ‘Ciano’ and ‘Gamenya’. Eradu’s pedigree was Ciano/Gamenya so Eradu’s GBSS 7A allele could potentially have come from either parent. However, since Eradu showed the (TGCCG)₃ allele while its parent Gamenya showed the (TGCCG)₄ allele, then Eradu’s 7A allele must have originated from its Ciano parent. The loss of the genetic material from one parent could also be seen in the example of the variety ‘Reeves’. ‘Reeves’s pedigree was Bodallin//(Hyden Sib.)Gamenya/Inia-66. Thus ‘Reeves’ could carry the GBSS 7A allele from any of its three listed parents. However, when the microsatellite information was added to the pedigree (in bold type) it read as follows:

Reeves(TGCCG₄) : Bodallin(TGCCG)₃ // (Hyden Sib.)Gamenya(TGCCG)₄ / Inia-66.

It was obvious that the 7A allele in ‘Reeves’ did not come from its parent ‘Bodallin’ but probably came from its parent ‘Gamenya’. Therefore the variety ‘Reeves’ cannot be used to access the GBSS 7A gene from ‘Bodallin’ because it does not contain it.

This kind of genetic analysis will become more common as molecular markers covering every chromosome and linked to many agronomically important or variety-determining genes become available. As quality-control becomes more important, having a unique molecular marker signature for varieties will also become important. The use of molecular markers linked to a selection of genes from each parent in a cross will facilitate the choice of suitable offspring carrying the desired genes to make breeding programs more efficient. When a new molecular marker linked to a trait-of-interest is developed, all the germplasm in the suitable breeding programs will be screened for the marker. Once this is done, and the unsuitable germplasm is removed from the breeding program, the trait will be fixed in that population and should not need to be addressed again until new germplasm is introduced into the breeding program and new crosses are made.
The (TGCCG)$_4$ allele was a new gene different from the previously described allele from ‘Chinese Spring’ wheat, and it could therefore be suggested that this (TGCCG)$_4$ mutant occurred in a particular wheat parent variety that was introduced into the noodle-wheat breeding program. When the pedigrees of the wheat varieties containing the (TGCCG)$_4$ microsatellite were examined (see Table 2.1), patterns of inheritance could be seen that predicted at the inheritance of this gene through the noodle-wheat breeding stock. Figure 3.16 shows a selection of the varieties containing the gene marked by the (TGCCG)$_4$ microsatellite. The number of microsatellite repeats identified in the different varieties have been included with the pedigree to show the inheritance of the 7A genes. As could be seen, the varieties ‘Cadoux’, ‘Tammin’, ‘Reeves’ and ‘Kulin’ all had the variety ‘Gamenya’ in their pedigree and it can be hypothesised that these varieties had inherited their GBSS 7A allele from this parent. It could also be seen that the GBSS 7A allele from the ‘Bodallin’ parent had not been inherited.

The varieties ‘Gamenya’ and ‘Machete’ shared the parent ‘Gabo’ and it was possible that this was the parent from which the (TGCCG)$_4$ gene was inherited. Thus the origin of this particular gene could be seen. The varieties ‘Gabo’ and ‘Gamenya’ were selected and released in New South Wales in 1945 and 1960, respectively, and were widely grown in WA with ‘Gamenya’ the dominant variety in WA until 1985 (Paull et al., 1998). Since Gamenya was such an important variety, it was used as breeding stock in WA and has passed its GBSS 7A (TGCCG)$_4$ allele on to more recently developed varieties in the noodle-wheat breeding program. It will therefore be possible to trace the inheritance of this novel allele through the breeding program to identify the original parent that introduced this gene to WA’s noodle-wheat breeding program. Since this allele was present in so many successful WA wheat varieties, it may in fact have some positive influence on starch quality and should be investigated for its linkage to starch quality.

DNA-based methods of phylogeny can provide additional information about the relationships between parents and offspring. This was demonstrated by Paull et al. (1998), who studied relationships between wheat varieties in a breeding program using molecular markers that
covered the entire genome. Although the phylogenies produced by studying the relationships between the DNA of the plants were similar to those produced by the pedigrees alone, some varieties were more similar to one parent than the other, showing that the DNA of one parent had been transmitted to some progeny and not others.

3.4.4 Are the GBSS genes present in WA wheat varieties the same alleles described in international publications?

The sequences from ‘Chinese Spring’ wheat were very similar and may be identical to those from Australian wheat varieties. This was unexpected because the ‘Chinese Spring’ GBSS
alleles described by Murai et al. (1999) came from nullisomic/tetrasomic varieties distant from those used to breed Australian wheat, except for the newly developed waxy wheat lines such as ‘DHW174’. This suggests that there may be very little genetic diversity amongst the GBSS genes in Australian varieties and that distantly related genetic material may need to be introduced to access variation at these loci. Since Australian wheats may contain the same alleles as those from ‘Chinese Spring’ wheat, information published about the GBSS genes present in ‘Chinese Spring’ wheat should also hold true for Australian varieties.
Chapter 4. Cloning the genomic GBSS gene sequences.

4.1 Introduction.

Since the 500bp section from the middle of the GBSS genes (Chapter 3) had not been highly informative for identifying polymorphic regions of novel alleles and did not contain the major polymorphisms between null and wild-type alleles at the 7A or 7D loci, sections of the GBSS genes outside this region were required.

At the start of this study, the genomic sequences of the GBSS genes had not been published, nor did any wheat contigs containing these sequences exist. Therefore, before the GBSS genomic sequences, and associated PCR primers to access those sequences, were available (Murai et al., 1999), several strategies were used to access new sections of the GBSS genes. Much of the research in this section was completed before the GBSS gene sequences were published. Also, after the GBSS gene sequences were published, they still did not give information about the GBSS promoters.

One approach to sequence more regions of the GBSS genes, was to continue to use PCR to amplify the GBSS genes from genomic DNA for analysis, as discussed in Chapter 3, and another application of this approach is explored in Chapter 5.

A second approach was to access the mRNA transcribed from the GBSS genes by producing cDNA copies. The cDNA would allow identification of any polymorphisms present in the coding regions of the gene, thus showing how the encoded enzymes differ. Whilst useful for studying enzyme function, the study of cDNA has some shortfalls. For example, the DNA sequence of the exon regions of a gene should contain fewer polymorphisms than the intron regions which are not present in the cDNA. More polymorphism should be identified by studying genomic DNA sequence even if the polymorphisms do not affect the encoded protein. Also, GBSS cDNA could not be used to access untranscribed GBSS promoter
regions. These were of great interest to investigate why the effects of the proteins encoded by genes at the 4A locus were much greater than the effects of proteins encoded by genes at the other two loci. The GBSS 4A genes could be expressed at a different time from the genes at the other loci, as controlled by the promoters. Also of interest was whether the 4A promoter sequences were modified when this locus was translocated from the 7B chromosome. For these reasons the cDNA approach was not followed.

A third approach for accessing the GBSS gene sequences was to capture genomic DNA fragments by cloning them into a library, and this is the method followed here. To fully investigate the GBSS genes the sequence of the whole genomic gene, its promoter region and its 5’ and 3’ untranslated regions were desired. Cloning the entire genomic sequences of the GBSS genes would give access to all these sections of the gene. It was aimed to clone and sequence the GBSS genes from a selection of Australian wheat varieties.

Although there were commercially available wheat genomic libraries from which the GBSS genes could have been identified, such libraries were not constructed from local germplasm. A possibility was to have sent DNA from Australian wheat to be cloned into a commercial library. However, because GBSS alleles present in many wheat varieties were of interest, many genomic libraries, from different varieties, would be needed and the cost would have been prohibitive. It was decided to make the genomic libraries to clone the GBSS genes from WA wheat varieties.

To construct a DNA library, the genomic DNA must first be cleaved into small restriction fragments and then cloned. Several different cloning vectors may be needed since many cloning vectors have a restricted insert size range that they can capture. The resulting clones need to be screened to search their inserts for the one containing the GBSS gene. If many wheat varieties were to be studied, this would also have been too large an experiment for the time constraints of the project.
To reduce the experiment to a manageable level, it was decided to reduce the amount of wheat DNA cloned by identifying the specific restriction DNA fragments containing the GBSS genes before cloning. These fragments were cloned into a partial genomic library. This partial genomic library would be greatly enriched for the GBSS genes over the entire genome making those clones easier to identify.

The sizes of the restriction fragments containing the GBSS genes could be identified by Southern DNA hybridisation using a known section of the GBSS gene amplified by PCR (Chapter 3) as a probe. Identification of the restriction fragment containing the GBSS gene by hybridisation would also ensure the ability to detect the resultant clone containing the GBSS gene within the finished library. The clone could be sequenced and the GBSS gene sequence identified. It was aimed to study the sizes of the restriction fragments, from the wheat genome, which contained the GBSS genes. This was done by making genomic southern blots and probing them with GBSS sequences. Towards producing informative genomic southern blots, five DNA extraction methods were tested to produce the large amounts of undamaged DNA required for this study.

When the restriction fragments containing the GBSS genes were identified, they were cloned. The sizes of the restriction fragments dictated the vector system used for cloning. It resulted that a BamHI restriction fragment 9kbp in size was chosen to be cloned (see Section 4.3.4) and therefore two cloning systems, the pGEM11 plasmid and the virus Lambda vector EMBL3, were chosen. They were chosen because they were easily available, contained the correct restriction sites for cloning, would accept inserts of the correct sizes (Sambrook et al., 1989), and because of previous experience with these cloning systems.
4.2 Methods.

4.2.1 DNA extraction methods.

Five DNA extraction methods were tested to produce DNA for Southern blots. The methods were CTAB, Modified Dellaporta, Quick, DNAzol, and Ethanol-perchlorate methods as described in Section 2.2.

4.2.2 Digestion of genomic wheat DNA with restriction enzymes and production of Southern blots.

The sibling wheat advanced breeding lines ‘Kulin’ and ‘Reeves’ were chosen for the Southern blot studies because their genetic similarity should result in the same size restriction fragments from the 7A and 7D loci and yet the differences at the 4A locus would allow identification of the restriction fragments formed from this locus.

DNA was extracted using the Ethanol-perchlorate method (Section 2.2). For each restriction enzyme BamH1, EcoR1 and HindIII, two digests were prepared containing 30µg of DNA from either ‘Kulin’ or ‘Reeves’. Digests also contained 300U restriction enzyme (Promega), 1X enzyme buffer (supplied as 10X buffer), 1/100 BSA, and water to a final volume of 200µl. Samples were mixed by gentle pipette mixing and incubated for two hrs at 37°C. After two hours, a further 5µl of the appropriate restriction enzyme was added to the tubes, mixed and incubated at 37°C for a further two hrs. Samples were concentrated by ethanol precipitation (Section 2.6.1) and resuspended in 30µl water overnight at 4°C.

The restriction fragments were electrophoresed on a 1% agarose gel in 1X TBE (90 mM Tris-borate, 2 mM EDTA) (Sambrook et al., 1989) prepared in a 25cm x 15cm gel tray. The gel was poured thick and the wells of the comb taped together to give six, 1cm-wide wells which could hold a 40µl volume. When the gel was set, it was loaded into the electrophoresis tank.
and covered in 1X TBE. Five µl loading dye (see Section 2.5) was added to each DNA sample and the samples loaded into the wells. A molecular size marker was also loaded. Samples were electrophoresed at a constant 40 mA for 17hrs. The gel was then stained with ethidium bromide and viewed under UV light (see Section 2.5). Excess gel was cut away from the region of the gel containing the samples leaving a small gap around all sides and the gel was photographed to record the positions of the molecular markers on the gel. The gel was exposed to UV light for several minutes to nick the DNA into small fragments that would move more easily from the gel onto the membrane. The DNA was then blotted onto nylon as described in Section 2.11.1. Also, Southern blots of various wheat lines digested with EcoR1 were supplied by J. Kretschmer (Adelaide University, Adelaide).

### 4.2.3 Probing the Southern blots with a GBSS-specific probe.

The probe used to hybridise to the GBSS genes was the 492bp PCR fragment originally amplified from the GBSS 7A locus of ‘Kulin’ using the GBSS2/GBSS4 primers (see Chapter 3). This probe had previously been shown to hybridise to DNA from all three GBSS loci (Hollingsworth, 1997). This PCR product in a pGEM-T vector was re-amplified to produce the radioisotope-labeled probe (Section 2.11.4). The Southern blots were probed with the GBSS-specific probe as in Section 2.11.3.

### 4.2.4 Isolation of the 9kbp BamH1 fragments from the wheat genomic DNA.

To isolate the 9kbp BamH1 fragments, 30µg of DNA from ‘Kulin’ was digested using BamH1. A small sample was electrophoresed on agarose to check the digestion was complete. The remaining sample was concentrated by ethanol precipitation as in Section 2.6 and resuspended in 100µl water overnight at 4°C. The sample was then electrophoresed on a 0.8% low-melting-point agarose gel in 1x TAE at 60V for 1hr. The gel was stained with ethidium-
bromide and viewed under UV light. DNA was removed from the gel using a BRESAclean
DNA Purification Kit (Geneworks, Adelaide).

4.2.5 Cloning into the pGEM11 plasmid.

The plasmid vector pGEM11 was extracted from an overnight culture using a QIAprep Spin
Plasmid Kit (Qiagen) (Section 2.10). Plasmid DNA was digested in a 200µl reaction
containing 5.9µg plasmid DNA, 1x reaction buffer (supplied as 10x buffer), 1/100 volumes
BSA and 60U BamH1 (Promega). The digest was incubated at 37°C for 3hrs. The 5’ ends of
the vector were then dephosphorylated to prevent self-ligation using CIAP enzyme (Promega).

The 9kbp BamH1 restriction fragment insert was ligated into the prepared pGEM11 vector in
ligation reactions as follows using 1:1 and 1:3 insert:vector molar ratios. Reactions were
carried out in 10µl volumes containing 50ng vector DNA, 1x reaction buffer (supplied as 10x
buffer), 10U T4 DNA ligase (Biotech International), water and either 44ng or 133ng insert
DNA. Reactions were incubated at 14°C overnight.

Ligated vectors were used to transform cells as Section 2.8 and also by using 8µl of ligation
mix and allowing the DNA to adsorb to the cells for 120min before the heat shock. The
colonies that resulted were pricked into 500µl water using a sterile pipette tip and screened for
the inserts they contained by PCR using both GBSS-Rev/GBSS-4A,7A,7D and T7/Sp6
primers as described in Section 2.9.

4.2.6 Cloning into the EMBL3 viral vector.

The virus vector chosen to clone the 9kbp BamH1 insert was the Lambda replacement vector
EMBL3 (Promega), which contained a single BamH1 site for inserting the restriction
fragment. The 9kbp BamH1 fragments were ligated into the EMBL3 vector and packaged into phage as directed in the Promega protocols manual (1991). Suitable control ligations were performed using the controls supplied with the vector and the expected results were achieved. The experiment was repeated four times.

4.3 Results.

4.3.1 Testing DNA extraction methods to produce high-quality genomic DNA for Southern blotting.

The sizes of restriction fragments containing the GBSS genes were identified by Southern blotting. Five DNA extraction methods were tested to produce DNA of suitable quality for Southern blots (Section 4.2.1). Using the CTAB, Modified Dellaporta, Quick and DNAzol methods, DNA of suitable quality for Southern blots was not produced. These methods gave DNA that contained an insoluble fraction which was retained in the wells (Figure 4.1). Any DNA bound to the insoluble fraction in the wells might become soluble during electrophoresis, and travel down the gel in a position un-related to its size. Any GBSS-related DNA that did this, would result in giving the blot a smeared appearance. The DNA extracted by the CTAB, DNAzol and Quick methods was also degraded and appeared as a smear of DNA down the gel (Figure 4.1). It is noted however, that J Kritschmer (Adelaide University, Adelaide; Pers Comm.) routinely produced Southern blots from DNA extracted by the CTAB method as seen later in this Chapter.

An attempt was made to solublise the insoluble fraction of the DNA extracted by the Modified Dellaporta method. The DNA sample was diluted 1/5 in water and incubated at 65°C for 30 min with regular flick-mixing, then incubated overnight at 37°C. This treatment did not fully solublise the DNA (data not shown), and probably the samples contained polysaccharide contamination that was insoluble and bound to DNA. To remove the polysaccharide
contamination, the DNA samples were precipitated using 1/10 volumes of NaCl and 2 volumes of ethanol. This reduced the viscosity of the sample but did not reduce the retention of DNA in the gel wells (data not shown). It was decided to reduce the polysaccharide (starch) content of the wheat leaves before DNA extraction. To reduce the polysaccharide

![Image of 1% agarose gel showing the quality of DNA extracted using two methods, CTAB (Lanes 2, 3) and Modified Dellaporta (Lanes 4-6). Each lane contains 5ul of un-quantified, extracted DNA. Note the residue retained in the wells of the gel and the smears of sheared DNA down the gel.](image)

**Figure 4.1.** 1% agarose gel showing the quality of DNA extracted using two methods, CTAB (Lanes 2, 3) and Modified Dellaporta (Lanes 4-6). Each lane contains 5ul of un-quantified, extracted DNA. Note the residue retained in the wells of the gel and the smears of sheared DNA down the gel.

Lane 1: Lambda/HindIII molecular size marker (Biotech International Ltd.); Lane 2: DNA extracted from ‘Kulin’ using the CTAB method; Lane 3: DNA extracted from ‘Reeves’ using the CTAB method; Lanes 4,5: DNA extracted from ‘Kulin’ using the Modified Dellaporta method; Lane 6: DNA extracted from ‘Reeves’ using the Modified Dellaporta method.
content of the leaves, plants were placed in the dark for 48 hrs before being harvested and the DNA extracted. This method did not reduce the retention of DNA in the wells, suggesting that the leaf starch stores had not been depleted or that the insoluble material was not starch.

The ethanol-perchlorate DNA extraction method was tried. This method gave large yields of DNA with little insoluble fraction or DNA degradation as shown in Figure 4.2. The DNA appeared to be of suitable quality for genomic Southern blot production and DNA extracted using this method was used to generate all genomic Southern blots.

Figure 4.2. 1% agarose gel showing the quality of DNA extracted using the Ethanol-perchlorate method. Each lane contains 5ul of unquantified, extracted DNA. Note little residue in the wells and little smearing of degraded DNA down the gel.

Lane 1: Lambda/HindIII molecular size marker (Biotech International Ltd.); Lanes 2-5: DNA extracted from ‘Kulin’; Lanes 6-8: DNA extracted from ‘Reeves’.
4.3.2 Choice of restriction enzymes.

The restriction enzymes used were BamH1, EcoR1 and HindIII. These enzymes were chosen because they have six-base recognition sequences that would cleave the wheat genome into restriction fragments small enough to be sequenced easily but large enough to contain the entire GBSS gene sequence. The restriction enzymes HindIII and EcoR1 were chosen because no recognition sites for these enzymes were found in either the GBSS cDNA sequence (Clark et al., 1991) or the region of the GBSS genes between exons 4 and 6 to which the probe would bind. These enzymes would therefore cut outside the region of the probe, and possibly outside the entire GBSS gene region resulting in restriction fragments containing novel GBSS DNA sequence.

The enzyme BamH1 was chosen because it cleaves within the GBSS gene. From analyzing the GBSS cDNA sequence (Clark et al., 1991), BamH1 restriction sites were known to exist in exons 4 and 6 of the GBSS genes. The restriction site in exon 4 cleaved 50bp inside the probe region, with the restriction site in exon 6 cutting 100bp outside the probe sequence as shown in Figure 4.3. The probe was a 492bp PCR fragment amplified from the GBSS 7A locus of ‘Kulin’ using the GBSS2/GBSS4 primers (see Chapter 3). It was therefore predicted that the BamH1 enzyme would cleave the GBSS genes at these two restriction sites in exons 4 and 6 to produce a 550bp restriction fragment to which the probe would bind. Since 50bp of the probe

![Figure 4.3. Diagrammatic representation of the positions of the BamH1 restriction enzyme recognition sites in exons 4 and 6 of the GBSS gene and the PCR product used as a probe for Southern Blot analyses.](image-url)
would be complimentary to the adjacent restriction fragment containing the region upstream of the restriction site in exon 4, it was also possible that the probe might anneal to this fragment to yield a second band.

4.3.3 Identification of the sizes of restriction fragments containing the GBSS genes.

Genomic Southern blots containing DNA from ‘Kulin’ and ‘Reeves’ digested with EcoR1, BamH1 and HindIII were probed with a GBSS-specific probe. Figures 4.4 and 4.5 show autoradiographs of these hybridisations. On the Southern blot displaying DNA digested with three different restriction enzymes (Figure 4.4) the probe identified a BamH1 fragment approximately 9kbp, an EcoR1 fragment of approximately 8kbp and a HindIII fragment approximately 4kbp in size as containing GBSS gene sequences. On the Southern blot that displayed DNA digested with EcoR1 only (Figure 4.5), either one or two bands approximately 8kbp in size were seen for each sample. The upper band in this Figure was present in those wheat lines with wild-type 4A GBSS characteristics and absent from null 4A GBSS wheat lines. It was predicted that the 8.5kbp band contained the GBSS-4A gene whilst the smaller molecular weight 7.5kbp band contained DNA from the other two loci. Only a single band is visible in the EcoR1 digest on the blot in Figure 4.4, but as this band is very faint, it was concluded that the second band in the ‘Kulin’ samples was too faint to be seen on this blot. Each lane of the blot should have contained 30ug of DNA since the DNA had been quantified and a single stock of DNA was used to make all three digests. However, the faintness of the bands in the outer lanes suggests that less DNA had been loaded into these lanes.

From the DNA sequence of the GBSS cDNA of Clark et al. (1991) BamH1 digested the GBSS gene in exons 4 and 6. The site in exon 6 was 100bp outside the region that the probe bound to, but the site in exon 4 cleaved the DNA 50bp inside the probe region. Thus, it was
Figure 4.4. An autoradiograph of a Southern blot of wheat genomic DNA digested with EcoR1, BamH1 or HindIII and probed with a 492bp section of the GBSS-7A gene. Each lane was loaded with 30ug of DNA. The bands seen on the autoradiograph should contain the GBSS genes. The weak bands in Lanes 2, 3, 6 and 7 have been highlighted with adjacent arrows. The sizes of the molecular marker bands have been included.

Lane 1: Lambda/HindIII molecular size marker (Biotech International Ltd.);
Lanes 2, 4, 6: ‘Kulin’ DNA;
Lanes 3, 5, 7: ‘Reeves’ DNA;
Lanes 2, 3: EcoR1;
Lanes 4, 5: BamH1;
Lanes 6, 7: HindIII.
Figure 4.5. An autoradiograph of a Southern blot of genomic wheat DNA digested with the restriction enzyme EcoR1 and probed with a 492bp section of the GBSS-7A gene. The bands seen on the autoradiograph should contain the GBSS gene. The approximately 21kbp band was undigested genomic DNA, the approximately 8.5kbp band came from the 4A locus as those samples lacking this band are null 4A and the approximately 7.5kbp band was predicted to have come from the other two GBSS loci. The sizes of the molecular size marker have been included (J. Kretschmer, pers comm.).

expected that the probe would bind to the approximately 550bp restriction fragment that it
complimented, and perhaps also to the adjacent restriction fragment to which it matched 50bp
of sequence. Unfortunately in order to separate the high molecular weight restriction
fragments on the gel used to make the Southern blot, the smaller fragments including the
approximately 550bp BamH1 fragment complementing the probe region were electrophoresed
off the end of the gel. Thus this fragment was not seen on the blot in Figure 4.4. However,
the probe also identified an approximately 9kbp restriction fragment (Figure 4.4). This
fragment was hypothesised to be that complimented by 50bp of the probe and should contain
the region of the GBSS genes upstream from exon 4 including exons 1 to 3, the transit peptide
coding region and the promoter elements.

4.3.4 Choice of restriction fragment to clone (in the initial experiment).

Having identified restriction fragments containing GBSS DNA sequences, one of the three
restriction enzymes needed to be chosen for constructing the partial genomic library. The
information gained from the blots (Figures 4.4 and 4.5) and the sequence of the GBSS cDNA
(Clark et al., 1991) indicated that the EcoR1 and HindIII restriction sites were outside the
region of the probe (exons 4 to 6) and that these restriction fragments could potentially contain
the entire GBSS gene sequence. Cloning and sequencing these fragments would provide
novel GBSS sequences.

The BamH1 restriction fragment could not contain the entire GBSS sequence since BamH1
restriction sites digested within the gene sequence at exons 4 and 6. Thus, the entire DNA
sequence could not be accessed using the BamH1 restriction fragments. However, these
fragments did have the benefit of having the GBSS sequences at their peripheries making them
easily accessible for sequencing using primers designed to the vector sequences, whereas the
EcoR1 and HindIII restriction fragments could contain the GBSS sequences anywhere in their
4kbp to 8kbp lengths. Alternately, since the EcoR1 and HindIII fragments contained the
existing GBSS PCR primer sites, the eventual clones containing these fragments could be identified by PCR, whereas the BamH1 fragments could not be identified using existing PCR tests.

It was decided to clone the 9kbp BamH1 restriction fragment in the initial experiment for the following reasons. The 9kbp BamH1 fragment should contain the sequence of the GBSS gene 5’ of exon 4 and was large enough to also contain all the GBSS promoter elements which were of interest in the current study. This clone would contain only half the sequence of the gene, but as the genomic sequences of the GBSS genes had not been published, accessing even part of the GBSS genomic sequence would provide new information. The sequence of the genes 3’ of exon 4 could not be accessed in these clones, but other methods for accessing this region of the gene were being examined as outlined in later Chapters. From the available sequence data, the single BamH1 fragment could potentially have contained sequences from all three GBSS loci, allowing comparison of the gene sequences from the three genomes in a single cloning experiment. Not having the existing GBSS PCR primer sites in the fragment would not allow PCR identification of this fragment, but having it anneal to the GBSS probe in the Southern blots would allow the appropriate clones to be identified by hybridization.

**4.3.5 Isolation of the 9kbp BamH1 restriction fragments from the wheat genome.**

The 9kbp BamH1 restriction fragments were isolated as previously described (Section 4.2.4). Genomic DNA from ‘Kulin’ was digested using BamH1, the restriction fragments were separated by electrophoresis, and the gel containing the 9kbp fragments was isolated and the DNA recovered from the gel. Figure 4.6 shows that the gel extraction procedure effectively separated the 9kbp fragment (Lanes 3, 4) from the other restriction fragments of other sizes (Lane 2).
Although the wheat genomic DNA was well digested by BamH1 (Figure 4.6), there was the potential that the digestion had been incomplete leaving some of the 9kbp BamH1 GBSS fragments still attached to the adjacent 550bp restriction fragment containing the GBSS2/GBSS4 PCR primer binding sites. If this were the case, these fragments could be identified with PCR simplifying the future analysis of clones. To check this theory, the extracted 9kbp fragments were then tested by PCR using the GBSS-Rev/GBSS-4A,7A,7D primers (see Section 2.9) that bound within introns 4 and 5. The PCR tests were positive.

Figure 4.6. 1% agarose gels showing the wheat genome digested with BamH1 (Lane 2) and the approximately 9kbp fragments (Lanes 3, 4) extracted from Lane 2. These fragments have been purified away from fragments of other sizes.

Lane 1: Molecular size marker Lambda/HindIII (Biotech International Ltd.); Lane 2: wheat genomic DNA digested using BamH1; Lanes 3,4: approximately 8kbp DNA fragments isolated from the gel Lane 2.
Figure 4.7 shows the PCR test containing the gel-extracted 9kbp fragments (labeled ‘A’) which were the template in the reaction and the PCR band (labeled ‘B’) amplified from the GBSS genes within the gel-extracted sample. This suggested that either some of the BamH1 fragments contained the region of the GBSS genes across introns 4 and 5, possibly due to an unknown restriction site polymorphism, or that small amounts of the DNA containing these PCR primer sites was contaminating the gel-extracted DNA.

Figure 4.7. 1% agarose gel showing that the 9kbp fragments purified from the genome by gel extraction contained GBSS gene sequences. The 9kbp gel-extracted fragments that were the template in the PCR have been labeled ‘A’. The positive PCR test bands amplified from the GBSS gene sequences within the gel-extracted fragments using the GBSS-Rev/GBSS4A,7A,7D primers have been labeled ‘B’. Each gel extraction can be seen to contain GBSS sequences.

Lane 1: Molecular size marker 100bp DNA Ladder (Promega); Lanes 2-4: PCR reactions of separate gel extraction events; Lane 5: water PCR blank.
4.3.6 Cloning into the pGEM11 plasmid.

To increase the chances of successful cloning, two vector systems were chosen to clone the 9kbp BamH1 restriction fragment, a plasmid and a virus vector. The plasmid vector chosen was pGEM11 that contained a single BamH1 site for inserting the restriction fragment. Following the linearization of the pGEM11 plasmid with BamH1 (Section 4.2.5), the 5’ ends of the vector were dephosphorylated to prevent self-ligation and plasmid concatamers from forming. Control samples of plasmid were used to transform *E.coli* to test the digestion of the plasmid by BamH1 and the dephosphorylation of the plasmid ends. Digested plasmid and dephosphorylated plasmid that had been self-ligated were used. Few colonies were seen proving the validity of these protocols.

The 9kbp BamH1 fragment was then ligated into the prepared pGEM11 vector (Section 4.2.5), transformed into *E.coli* and the colonies that resulted were screened to identify their inserts by PCR using both GBSS-Rev/GBSS-4A,7A,7D and T7/Sp6 primers. None of the clones contained inserts (data not shown). This experiment was repeated three times and no clones containing inserts were produced. It was concluded that there was some barrier to the ligation of the plasmid with the 9kbp insert. Addition of larger quantities of ligase enzyme may have been useful to force the ligation. Possibly the 9kbp insert fragment had been modified by passage through the gel and had not retained the terminal restriction sites allowing it to be cloned. In this case, another method of size-exclusion purification would be required.

4.3.7 Cloning into the EMBL3 viral vector.

The virus vector chosen to clone the 9kbp BamH1 insert was the Lambda vector EMBL3 that contained a single suitable BamH1 site for inserting the restriction fragment. There are two ways in which Lambda vectors can incorporate a novel DNA fragment (Sambrook *et al.*, 1989). The first is by inserting a novel insert into the existing Lambda vector sequence.
Insertion-vectors can only receive up to 5kbp of DNA before they cease to replicate. The second form of the Lambda vector is the replacement-vector. A non-essential section of the replacement-vector’s genome is replaced with the insert DNA. These vectors can replace between 7kbp and 20kbp of their sequence with an insert before replication ceases (Sambrook et al., 1989). Thus the 9kbp GBSS fragment was within the size range to be cloned.

The insert fragments were ligated into the EMBL3 vector and packaged into phage as described in Section 4.2.6. Suitable control ligations were performed using the controls supplied with the vector and the expected results were achieved. However, no plaques were seen from reactions containing the GBSS inserts. Analysis of the appropriate control reactions showed that the limiting step was the ligation of the restriction fragments into the vector arms. The experiment was repeated four times with the same results.

Thus, neither the plasmid nor the virus vector could capture the desired restriction fragment insert. The reason for the lack of ligation of inserts into the vectors was not discovered. There were several other cloning systems which could have been used in this study, but time pressures did not allow these to be explored. More recently, a wide range of new cloning systems have also become available, which may provide more positive results in future studies. Although unsuccessful, the results for the pGEM11 and EMBL3 cloning were included in this study to provide information for future research into this area.

4.4 Discussion.

To produce a good Southern blot for wheat, the DNA needs to be of good quality. It should not contain residual polysaccharides, nor should it be nicked or degraded into small, non-specific fragments of DNA which would increase the background signal from the blot. The five DNA extraction methods studied for the quality of DNA they produced (CTAB, Modified
Dellaporta, Quick, DNAzol, and Ethanol-perchlorate methods) were chosen because they largely avoid the use of chloroform or excessive vortexing that might degrade the DNA, yet contain steps to remove contaminating polysaccharides and DNA-degrading enzymes.

Using the ethanol-perchlorate DNA extraction method, Southern blots of wheat genomic DNA were produced and probed with a GBSS-specific probe to show the sizes of restriction fragments containing the GBSS genes, these restriction fragments were isolated from the remaining genomic DNA and attempts were made to clone these fragments, but without success. Although restriction fragments containing the GBSS genes were not cloned, analysis of the sizes of the restriction fragments themselves was informative and are discussed below. The sizes of the restriction fragments of the wheat genome produced by three restriction enzymes and containing the GBSS genes will be useful information for future cloning of the GBSS genes. This should include cloning the EcoRI restriction fragments containing the GBSS genes as these would give access to novel DNA sequence from the GBSS genes.

### 4.4.1 Where are the GBSS gene copies in the genome?

The patterns of bands containing the GBSS gene sequences on the Southern blots gave information about the presence of pseudogene copies of the GBSS genes in the wheat genome. As pseudogene copies of genes are under less selective pressure to maintain their sequence integrity, they mutate at faster rates than coding copies of genes. Thus, the restriction sites within and surrounding a pseudogene should show more polymorphism than sites surrounding coding genes. If pseudogene copies of the GBSS genes were present in the wheat genome, one would predict the detection of multiple bands besides those containing the coding genes. However, for each restriction enzyme used in this study, no more than 2 different sized DNA fragments containing the GBSS genes was observed, suggesting that there are no pseudogene copies of the GBSS genes in the wheat genome. This result was supported by the work of Vrinten *et al.* (1999) who also did not identify any pseudogene copies of the GBSS genes by
Southern blot analyses. Also, Murai et al. (1999) PCR amplified and sequenced 18 fragments from the GBSS gene at each of the three loci using primers in exons 1, 4, 6, 10 and 11 and no pseudogene fragments were identified. Thus no evidence for the existence of wheat GBSS pseudogenes exists.

If any GBSS pseudogenes do exist there are three reasons for their lack of detection, as discussed previously in Chapter 3. Firstly the pseudogenes may be so similar to the coding copies of the gene that their length and associated restriction sites have been conserved forming the same sized restriction fragments as the coding copies. Secondly, the pseudogenes may be so different to the coding copies within the region of the probe that the probe did not recognise them. If this is the case, the pseudogenes would be very difficult to detect without other GBSS sequences as probes and the pseudogene’s sequences would probably not encode an active GBSS protein that could be useful to breeders. Pseudogenes that are very different to the original gene might have diverged to be different genes with different function. Thirdly the pseudogenes may be positioned so close to the coding copies that they fall within the same restriction fragment. The restriction fragments seen were 9kb, 8kb and 4kb in size. As the GBSS genes are at least 3kbp in size, it does not seem likely that these restriction fragments contained pseudogenes adjacent to the coding copies of these genes. One must then conclude that there are no pseudogene copies of the GBSS genes in the wheat genome.

4.4.2 How does each GBSS allele differ from the others?

Five restriction fragments containing DNA from the GBSS genes were identified in this study. These findings, along with others from recently published studies (Murai et al., 1999; Vrinten et al., 1999), enable production of a basic restriction map of the EcoR1, BamH1 and HindIII sites surrounding the GBSS genes, as shown in Figure 4.9. Each of the wild-type alleles from the three loci as well as the null alleles from the 7A and 7D loci contain unique restriction site variations, which can be used to distinguish between the different alleles.
An EcoR1 restriction fragment 8.5kbp in size was identified and which was predicted to contain the gene from the GBSS 4A locus, whilst a second 7.5kbp fragment was identified, which was predicted to contain a sequence from each of the 7A and 7D loci. Comparing these results with the published DNA sequences for null and wild-type GBSS genes (Murai et al.,

Figure 4.8. Diagrammatic representation of the GBSS gene coding regions of five GBSS alleles (left) shown with the exons marked as red rectangles and numbered at the top, and introns as dotted lines. Below each gene are double sided arrows indicating the restriction fragments formed from these alleles by the restriction enzymes EcoR1, HindIII and BamH1 based on both DNA sequence data (Murai et al., 1999; Vrinten et al., 1999), and data from Southern blots performed in this project and by Vrinten et al. (1999). The position of the probe is shown at the top.
1999; Vrinten et al., 1999), confirms that the enzyme EcoRI has no binding sites within the coding regions of the GBSS genes and therefore, these restriction sites must be outside the published sequences (Figure 4.9). The 7.5kbp and 8.5kbp EcoRI restriction fragments should contain the entire sequence of the coding region of these genes and may also contain sequence from the promoter regions. The size difference between the two bands may reflect a polymorphism in the position of EcoRI restriction sites at each locus or DNA sequence length polymorphism between the GBSS genes in the non-coding regions. The sequences of the coding regions of the wild-type GBSS genes from the 4A, 7A and 7D loci were 2794bp, 2781bp and 2862bp, respectively (Murai et al., 1999). Thus, for the fragments from the 7A and 7D loci to be similar in size and the 4A fragment to be 1kbp larger would require some polymorphism between the genomes. Since they are predicted to contain the entire sequence of the GBSS genes, these EcoRI restriction fragments would be useful in characterising the entire GBSS gene sequences.

A HindIII restriction fragment 4kbp in size was identified as containing GBSS DNA sequences. This observation has been supported by Vrinten et al. (1999), who identified a 4.2kbp HindIII restriction fragment from the GBSS wild-type 7A allele. Comparing results with the previously reported GBSS genes sequences (Murai et al., 1999; Vrinten et al., 1999), the wild-type 7A allele contains a HindIII site 3bp from the 5’ end of intron 1. Thus the 4.2kbp HindIII fragment from the 7A locus contained the sequence from intron 1 through the coding region of the gene and into the 3’ untranslated region (Figure 4.9). The hybridization probes used here and by Vrinten et al. were short PCR products that annealed from exons 4 to 6 and exons 5 to 7, respectively, and so they only identified the restriction fragments carrying these regions of the DNA and not the adjacent restriction fragment carrying exon 1 of the GBSS 7A gene. Since the sequences from each locus were polymorphic in this region of intron 1, this restriction site did not exist at the 4A and 7D loci. Neither did this restriction site exist in the null 7A allele. This finding may provide the basis for developing a useful
molecular marker to distinguish null and wild-type alleles at the 7A locus. The null 7A allele produced a 15kbp HindIII restriction fragment (Vrinten et al., 1999).

Vrinten et al. (1999) identified other GBSS HindIII restriction fragments in their study. They identified a 12kbp HindIII fragment from the wild-type 4A allele. Murai et al. (1999) located a HindIII restriction site within intron 7 of this gene. The probe used by Vrinten et al. (1999) was a 434bp PCR fragment amplified from a Barley cDNA clone between exons 5 and 7. It annealed entirely within the 12kbp restriction fragment and did not recognise the adjacent restriction fragment carrying the region downstream of intron 7. Since the probe used in the current study overlapped the region of the probe used by Vrinten et al. (1999), the 12kbp HindIII fragment containing the promoter regions and the coding region up to exon 7 of the GBSS 4A gene should have been recognised in the current study. Though the Southern blots were repeated several times, neither this or other bands were seen. One possible explanation is that this high molecular weight band had been labelled by the probe to the same small extent that the other HindIII fragment had, but it was present in a region of the blot with a higher level of background labelling, or was present in too few copy number, and thus it could not be resolved.

Murai et al., (1999) and Vrinten et al., (1999) demonstrated that neither the null or the wild-type alleles from the 7D locus contained a HindIII site. Vrinten et al. (1999) showed that both these alleles produced an 18kbp HindIII restriction fragment, suggesting the conservation of this restriction site at this locus. This fragment was also not seen in the current experiment, supposedly for the same reasons discussed above, but because it was predicted to contain the entire gene sequence from this locus, it would be a useful fragment to clone in order to obtain a GBSS sequence from the 7D genome without also obtaining sequence information from the other loci.
Two of the HindIII restriction fragments containing the GBSS gene sequences identified by Vrinten *et al.* (1999) were not identified in the current study. The amounts of DNA loaded onto each well had been measured, and using these measurements, equal amounts of DNA had been loaded onto each lane of the gel. However, it appears that this may not have been successful, since the BamH1 digests gave stronger bands, suggesting that more DNA was present in these samples. There may have been only small amounts of DNA present on some lanes of the blots and therefore the bands could not be seen in these samples. The probe was known to bind strongly to DNA sequences from all three loci under the conditions of the experiment (Hollingsworth, 1997), thus limitations due to stringency are also unlikely. It may be that all three GBSS alleles from ‘Kulin’ and ‘Reeves’ have common restriction sites and that the single band contained sequences from all three loci.

The restriction enzyme BamH1 digested the GBSS gene within exons 4 and 6 (Clark *et al.*, 1999). The site in exon 6 was 100bp outside the region that the probe bound to, but the site in exon 4 cleaved the DNA 50bp inside the probe region. It was predicted that the probe would bind to the approximately 550bp restriction fragment that it complemented and also to the adjacent restriction fragment having a 50bp homologous sequence. Unfortunately, in order to separate the high molecular weight restriction fragments on the gel used to make the Southern blot, the smaller fragments including the approximately 550bp BamH1 fragment complementing the probe region were electrophoresed off the end of the gel. The probe identified a 9kbp restriction fragment thought to be one complemented by 50bp of the probe, and containing the region of the GBSS genes 5’ of exon 4, including exons 1 to 3 and promoter elements. However, the recently published GBSS gene sequences indicate that this 9kbp band did not contain sequence from the GBSS 7D locus since these alleles contain a third BamH1 site that cleaves off a 345bp fragment adjacent to the 550bp fragment to which the majority of the probe bound. These two small fragments (which were electrophoresed off the end of the gel) would have bound the probe and no large restriction fragments from the 7D
locus would have been seen. The 9kbp BamH1 restriction fragment could therefore only contain sequences from the 3’ end of the GBSS genes from the 7A and 4A loci.

In general, the Southern blots seemed to have high background levels of signal compared to the specific bands, even though they had been blocked to reduce non-specific binding. Nagano et al. (2000) reported that when doing genomic Southern blots on rice, repetitive sequences from GBSS probes annealed to many places in the genome and gave ‘smeary Southerns’. However, in this study, the sequence of the probe had been explored and no microsatellites were identified.

4.4.3 How can we tag the different GBSS alleles with molecular markers to identify them in a breeding program?

The differences in the restriction enzyme recognition sites outlined above could be used to develop molecular markers to distinguish some alleles. For example, the HindIII restriction fragments formed by the null 7A, wild-type 7A, wild-type 4A and wild-type 7D alleles differed in size such that they could be readily differentiated.

4.4.4 Are the GBSS genes present in WA wheat lines the same alleles described in international publications?

The restriction fragment analysis of the wheat lines ‘Kulin’ and ‘Reeves’ compared to published sequence data supports the hypothesis that the WA wheat lines contained the same GBSS alleles as the ‘Chinese Spring’ wheat lines. The sizes of the restriction fragments predicted by the DNA sequences from ‘Chinese Spring’ wheat lines were the same as those formed from WA wheat lines. However, no ‘Chinese Spring’ samples were included on the blots in this study because a comparison of the GBSS restriction fragment sizes between Australian varieties and ‘Chinese Spring’ was not a major aim of this section of the study.
Chapter 5. Characterisation of the GBSS genes using PCR.

5.1 Introduction.

The central region of the GBSS genes had been PCR-amplified and studied in Chapter 3, but to find additional polymorphisms a larger region of the GBSS genes, outside that already studied, was investigated. The experimental plan was to PCR-amplify previously-unstudied sections of the GBSS genes, and to clone and sequence these regions from a number of wheat varieties, searching for polymorphisms.

Originally, four PCR primers, GBSS1 to GBSS4 (Briney et al., 1998), had been designed to conserved regions from a DNA sequence alignment of the GBSS cDNA sequence of wheat (Clark et al., 1991; X57233) to the GBSS genomic and cDNA sequences from Barley (X07931, X07932) and Rice (X64108, X62134). Since the sequences of the exons from the three GBSS loci are conserved, PCR primers designed to the wheat cDNA sequence from the 7A locus (Clark et al., 1991; X57233) could be used to amplify from all three loci (Hollingsworth, 1997; Briney et al., 1998). The positions of these four original primers, in exons 1, 4 and 6 of the GBSS gene, are shown in Figure 5.1.

The PCR bands amplified by the primers GBSS2/GBSS4 had been studied (Chapter 3). The PCR bands produced by the primer combinations GBSS1/GBSS2 and GBSS1/GBSS3 had not been studied, and so the PCR primers GBSS1-GBSS3 were now used to PCR-amplify the GBSS genes. Also, six new primers, GBSS5 to GBSS10, were designed from the sequence of the wheat GBSS cDNA (Clark et al., 1991). Their binding positions on the GBSS gene are shown in Figure 5.2. They were designed such that, when used in various combinations, they should amplify most sections of the GBSS gene. It was aimed to use combinations of these primers to amplify from the GBSS genes, clone these regions, and sequence them. When this
study was done, the genomic sequences of the GBSS genes had not been published (Murai et al., 1999, Vrinten et al., 1999), and so primers designed in those studies were not used here.

Previously, the use of the original PCR primers, GBSS1 to GBSS4, had been complicated by the presence of non-specific amplification products and homo-primed bands (A. Briney, WA SABC, pers. comm., B. Hollingsworth, 1997). The GBSS2/GBSS4 molecular marker for the null 4A allele is complicated by having an artifactual band of similar size to the marker band (A. Briney, WA SABC, pers. comm.). This band is not always present and this variation was initially thought to be due to variations in the GBSS genes. However, this band was investigated further and found to be homoprimed and unrelated to the GBSS genes (B. Hollingsworth, 1997).

**Figure 5.1.** Positions of PCR primers GBSS1-GBSS4 on the GBSS gene. The GBSS gene is represented in blue. The exons are represented by thick boxes and numbered beneath. The introns are represented by the thin line and numbered above. The lengths of the introns and exons are to scale. The positions of the primers are shown with arrows. The forward primers are coloured pink and shown above their binding positions in exons 1 and 4. The reverse primers are coloured black and shown below their binding positions in exons 4 and 6. The binding sites of the primers GBSS3 and GBSS4 overlap.
Whilst even simple genomes can produce non-specific amplification at low annealing temperatures, the allohexaploid nature of the wheat genome means that potentially each primer binding site occurs in triplicate, and large plant genomes, like that of wheat, can contain many repeated regions, potentially producing other primer-binding sites (Kumar and Bennetzen, 1999). Therefore, PCR primers could anneal to many places in the wheat genome and complicate the amplification of specific PCR products. In this study, the primers designed to the GBSS cDNA sequence (Clark et al., 1991) were being used to amplify bands of unknown length (since the sequences of the GBSS introns were not yet available, Murai et al., 1999). It is normal to amplify many bands, under conditions of low stringency, when optimizing a PCR. Since the sizes of the bands amplified from the GBSS genes would not be known, it was

**Figure 5.2.** Positions of PCR primers on the GBSS gene. The GBSS gene is represented in blue. The eleven exons are represented by thick boxes and numbered beneath. The ten introns are represented by the thin line and numbered above. The lengths of the introns and exons are to scale. The positions of primers are shown with arrows. The forward primers are coloured pink and shown above their binding positions on the gene. The forward primers GBSS9, GBSS5 and GBSS10 are in exon one, the primers GBSS1 and GBSS4 are in exon four, and the primers GBSS-4A, GBSS-7A and GBSS-7D are in intron four. The reverse primers are coloured black and shown below their binding positions on the gene. The reverse primer GBSS-rev is in intron five, the primers GBSS2 and GBSS3 are in exon six, primer GBSS8 is in exon ten, and the primers GBSS7 and GBSS6 are in exon eleven. Only the binding sites of the primers GBSS3 and GBSS4 overlap. The primers GBSS-4A, GBSS-7A and GBSS-7D bind in the same region of the gene, but each only binds to genes at one specific locus.
aimed to identify the sizes of the homoprimed bands produced by the primers, so that they could be dismissed from the analysis of the dual-primed PCR products. Prior to using both the original and new PCR primers to amplify sections of the GBSS genes, a short investigation into the homoprimed bands each primer could amplify, when used singly, was undertaken. It was aimed to identify any artifactual bands, enabling them to be differentiated from the specifically-amplified bands produced in later experiments.

5.2 Methods.

5.2.1 PCR.

To test for homoprimed bands, PCR reactions were prepared in a 20µl volume containing ‘Kulin’ DNA as template, 2mM MgCl₂ and 20pmol of a single primer from the set GBSS1-GBSS10 (see Section 2.4). The PCR was performed on a GeneAmp PCR System 2400 thermal cycler (Perkin Elmer) under the following thermal cycling conditions: 1 hold: 3 min at 94°C; 35 cycles: 30 sec at 94°C, 30 sec at 57/58°C, 3 min at 72°C; 1 hold: 3 min at 72°C. Reactions containing primers GBSS1-GBSS6 were annealed at 58°C, primers GBSS7-GBSS10 were annealed at 57°C. These PCR annealing temperatures were low to enable less-specific binding of the primers to the template and homoprimed fragments to be amplified. Long extension times (3min) were used to allow time for large fragments to be amplified. When the annealing temperatures of the PCRs were raised, the same PCR reactions were amplified over a gradient of annealing temperatures between 58°C and 72°C using a gradient PCR machine.

The dual-primer PCR reactions for primer combinations GBSS1/GBSS2 and GBSS1/GBSS3 were as described (Briney et al., 1998). The primers GBSS3 to GBSS10 were used in all combinations in PCR reactions containing 1.6mM MgCl₂ (Section 2.4), and they were incubated over a gradient of increasing annealing temperatures. Fragments were amplified by
the GBSS4/GBSS6 primers in 40µl reactions containing 1.6mM MgCl₂ (Section 2.4) under the following cycling conditions: 1 hold at 94°C for 4min; 35 cycles of 94°C for 30sec and 72°C for 3min; 1 hold at 72°C for 10min. PCR fragments were electrophoresed on both agarose and polyacrylamide gels (Section 2.5).

5.2.2 Southern Blots.

To determine which homoprimed bands were amplified from the GBSS genes, homoprimed PCR fragments were electrophoresed on agarose gels, blotted onto nylon membranes (Section 2.11.1) and probed with a GBSS-specific DNA probe (see Section 2.11.3). The probe was a cloned, GBSS2/GBSS4 PCR fragment, amplified from the GBSS gene at the 7A locus of the wheat variety ‘Kulin’.

To determine which dual-primed PCR bands were amplified from the GBSS genes, the GBSS1/GBSS2 and GBSS1/GBSS3 PCR fragments were electrophoresed on a polyacrylamide gel (Section 2.5), Southern blotted onto a nylon membrane and probed with a GBSS-specific probe (Section 2.11). Cloned PCR fragments were also screened for their relatedness to GBSS by hybridization. Cell-scrapings of the individual clone colonies were dot-blotted onto nylon membranes as described by the manufacturer (BioRad). The blots were then probed with a GBSS-specific probe made from the GBSS-Rev/GBSS-4A PCR fragment from ‘Kulin’ (Section 2.11).

5.2.3 Size separation and isolation of GBSS-related PCR fragments from artifacts.

Two methods were used to separate the GBSS-related PCR bands from the non-specific fragments before cloning: gel and capillary electrophoresis. GBSS1/GBSS2 PCR fragments
were separated using a BioRad capillary electrophoresis machine. The bands were eluted from the capillary in viscous buffer, which needed to be removed before cloning. The eluted samples were diluted 1 in 3 in water and the bands re-amplified by PCR using the GBSS1/GBSS2 primers. PCR bands were also electrophoresed in a 1% low melting point agarose gel, the desired bands were cut from the gel and the DNA recovered from the gel slice (Section 2.7).

5.2.4 Cloning GBSS1/GBSS2 and GBSS1/GBSS3 PCR fragments.

PCR products were cloned as described in Section 2.8. The total GBSS1/GBSS2 and GBSS1/GBSS3 PCR reactions were first cleaned using the Wizard® PCR Preps DNA Purification System (Promega) before being cloned in reactions containing between 130ng and 190ng of PCR product. When GBSS1/GBSS2 PCR fragments were gel-purified before being cloned, the ligation reactions contained 300ng of PCR product.

5.2.5 Cloning the GBSS4/GBSS6 PCR fragment.

The GBSS4/GBSS6 PCR fragment was cloned whole (Section 2.8) and also cleaved into two fragments using a restriction enzyme before cloning. The whole fragment was cloned after being cleaned using the Wizard® PCR Preps DNA Purification System (Promega). The cleaved fragment was first purified by ethanol precipitation (Section 2.6) and digested in 30μl reactions containing the enzymes BamH1 or Sal1 (Promega) at 37°C for 2hrs. The restriction enzymes were denatured at 80°C for 20min and reactions were again cleaned by ethanol precipitation. To fill in the sticky-ends of the restriction fragments and provide a terminal ‘A’ for cloning, the restriction fragments were resuspended in 40μL PCR reactions (Section 2.4) containing 2mM MgCl₂ but no primers, and incubated at 72°C for 20min. Each sample was then cleaned using a Wizard® PCR Preps DNA Purification System (Promega) and cloned into pGEM-T (Section 2.8).
Sometimes after ligating inserts into the vectors, the percentage of plasmids in the population that contained inserts was concentrated before transforming the \textit{E. coli}. The plasmids were extracted from the mixture of clones, linearised with single-site restriction enzymes SalI or SacI (Promega) and electrophoresed in a low melting point agarose gel with control plasmids containing either no insert or inserts of known size. The bands containing the desired plasmids were cut individually from the gel, and the DNA extracted from the gel. The gel-purified plasmid samples were recircularised using a large reaction volume to reduce concatamer formation and transformed into \textit{E. coli} (Section 2.8).

\textbf{5.2.6 Screening clones to identify their inserts using PCR.}

Clones were screened to identify their insert fragments using the T7/Sp6 primer set and the GBSS-Rev/GBSS-4A,7A,7D primer combination (Section 2.9). To determine the size and orientation of cleaved insert fragments, clones were examined using PCR with various combinations of the primers T7, Sp6, GBSS4 and GBSS6. Since the original PCR product amplified by the GBSS4/GBSS6 primers had been digested with a restriction enzyme and then cloned, the insert contained only one of its original primer sites and therefore amplifying the insert using one of GBSS4 or GBSS6 required the use of the vector-based primers T7 or Sp6.

To screen many clones at once, several colonies were pricked into the same water sample and this mixed template was tested using a single PCR. When a group of clones was found to contain a suitable insert, the members of the group were then screened individually. Plasmids were sequenced as described in Section 2.10.1.
5.3 Results.

5.3.1 Investigation of non-specific amplification from the wheat genome by the PCR primers GBSS1 to GBSS 10.

To identify homoprimed bands produced by the primers GBSS1 to GBSS 10, these primers were used in reactions singularly, and the bands they produced are shown in Figure 5.3. The annealing temperatures (57/58°C) were lower than the melting temperature calculated for the

![Figure 5.3](image.png)

**Figure 5.3.** 1% agarose gels showing PCR bands amplified from the wheat genome (‘Kulin’) in PCR reactions containing single primers only. The single primers tested were GBSS1-GBSS10. Lanes 1, 8: 1Kb Plus DNA Ladder molecular size marker (Gibco BRL); Lane 2: primer GBSS1; Lane 3: primer GBSS2; Lane 4: GBSS3; Lane 5: GBSS4; Lane 6: GBSS5; Lane 7: GBSS6; Lane 9: GBSS7; Lane 10: GBSS8; Lane 11: GBSS9; Lane 12: GBSS10; Lane 13: water PCR blank.
primers (see Table 5.1) for all except GBSS 8 (56°C). The melting temperatures for the other primers varied between 60°C and 72°C. The primers amplified many bands ranging in size from 200bp to 3000bp (the maximum expected size when using a 3min extension time). The primer GBSS2 (Lane 3) amplified fewer bands than the other primers, and GBSS8 (Lane 10) did not amplify, probably due to using a too-high annealing temperature. Primer GBSS2 was therefore considered to be good candidates for generating GBSS-related bands, since the interference from homoprimed fragments amplified by it was less than for other primers.

To assess if any of these bands had been amplified from GBSS sequences, they were blotted onto nylon and probed with a GBSS sequence. Figure 5.4 compares the DNA banding patterns on the gels from Figure 5.3 with autoradiographs of the blotted DNA probed with the GBSS probe. The probe hybridized to the GBSS plasmid positive control (Figure 5.4, Lane P) as was expected (the plasmid contained an insert of DNA from a GBSS allele from the 7D locus). Since the probe, which was from the 7A locus, hybridized to the control fragment from the 7D locus shows the high level of DNA sequence similarity between the loci and the ability of a probe from one locus to detect DNA sequences from all three loci. Unexpectedly, the probe also hybridized to one of the bands in the molecular size marker (Figure 5.4, Lane M) which was not related to the wheat GBSS genes. The DNA sequence of this marker band was compared to the sequence of the probe and no similarities could be seen (data not shown). The GBSS probe did not hybridise to any of the homoprimed PCR fragments (Figure 5.4, Lanes 1-10), demonstrating that none of the bands were amplified from the GBSS genes within the region of the probe.
Figure 5.4. 1% agarose gels (left) showing homoprimed PCR fragments amplified by primers GBSS1-GBSS10 and the autoradiographs (right) of the Southern blots of those PCR fragments probed with a GBSS-specific probe. The probe was the GBSS2/GBSS4 PCR fragment from the 7A locus. The probe hybridised to the unlinearised positive control plasmid (Lane ‘P’) and one band from the molecular marker (Lane ‘M’). None of the homoprimed bands hybridised with the probe thus none were amplified from the GBSS genes. The Southern Blots were washed with 0.5xSSc/0.2% SDS solution at 65°C for 1hr. The blots were exposed to film for 3hrs and 19hrs respectively.

Lane M: Molecular size marker 1 Kb Plus DNA Ladder (Gibco BRL): Lanes 1-10: fragments amplified in PCR containing the single primers GBSS1-GBSS10 respectively; Lane P: plasmid positive control containing the GBSS2/GBSS4 PCR fragment from the 7D locus.
Having concluded that none of the homoprimered PCR fragments were amplified from the GBSS genes by these primers, methods to reduce the generation of the homoprimered bands were investigated. It was aimed to reduce the production of homoprimered fragments, by changing the PCR reaction conditions, so that they would not complicate the future use of these primers. It was important to reduce the non-specific amplification, since the sizes of the specific bands were not known when this experiment was done, and any homoprimered bands produced under more stringent conditions might be confused for specific bands.

The annealing temperature of the PCR conditions was increased to see at which temperatures these homoprimered bands were no longer amplified. The PCR reactions, each containing a single GBSS-series primer, were incubated under similar conditions as described previously, but over a gradient of annealing temperatures from 58.3°C to 71.9°C. The homoprimered bands are shown in Figure 5.5. Table 5.1 compares the temperature at which no homoprimered bands were amplified, with the melting temperature of the primers, calculated using the website http://www.promega.com/biomath/default.htm.

**Table 5.1.** Compares the temperature at which no homoprimered bands were amplified, with the melting temperature of the primers.

<table>
<thead>
<tr>
<th>primer</th>
<th>melting temperature (°C)</th>
<th>temperature at which no homoprimered bands were amplified</th>
</tr>
</thead>
<tbody>
<tr>
<td>GBSS1</td>
<td>60</td>
<td>58</td>
</tr>
<tr>
<td>GBSS2</td>
<td>60</td>
<td>58</td>
</tr>
<tr>
<td>GBSS3</td>
<td>61</td>
<td>66</td>
</tr>
<tr>
<td>GBSS4</td>
<td>61</td>
<td>64</td>
</tr>
<tr>
<td>GBSS5</td>
<td>64</td>
<td>66</td>
</tr>
<tr>
<td>GBSS6</td>
<td>66</td>
<td>64</td>
</tr>
<tr>
<td>GBSS7</td>
<td>65</td>
<td>68</td>
</tr>
<tr>
<td>GBSS8</td>
<td>56</td>
<td>58</td>
</tr>
<tr>
<td>GBSS9</td>
<td>64</td>
<td>58</td>
</tr>
<tr>
<td>GBSS10</td>
<td>72</td>
<td>68</td>
</tr>
</tbody>
</table>
Figure 5.5. 1% agarose gels showing the homoprimed PCR bands produced by using primers GBSS3-GBSS10 in isolation over an annealing temperature gradient from 58.3°C to 71.9°C. The name of the primer used in the PCR reactions is shown on the Figure under the lanes containing those PCR reactions. The numbers of bands decreased as the temperature increased. Some reactions with higher annealing temperatures have not been shown due to space restrictions, and no bands were seen in those samples. The DNA template was 'Kulin' DNA.

Lane 1: 1 Kb Plus DNA Ladder molecular size marker (Gibco BRL); Lane 2: annealing temperature of 58.3°C; Lane 3: 58.9°C; Lane 4: 60.2°C; Lane 5: 61.8°C; Lane 6: 63.7°C; Lane 7: 65.8°C; Lane 8: 67.7°C; Lane 9: 69.5°C; Lane 10: 70.8°C; Lane 11: 71.9°C.
The primers GBSS1 and GBSS2 did not produce any bands under these conditions (data not shown), nor did GBSS8 and GBSS9 (Figure 5.5). Since primers GBSS1 and GBSS2 had previously produced homoprimed bands under similar conditions but cycled in a different PCR thermal cycler, this suggested that the intrinsic properties of different PCR machines may affect the PCR bands produced. All PCR machines cycle under slightly different conditions due to the ramping speeds of each machine, the temperature of the room at the time of cycling (which may vary during a day) and the number of tubes contained in the machine. Caution must be taken when optimizing a PCR, to ensure that the reaction conditions are robust such that different machines do not produce different banding patterns. For a molecular marker, such as that for the GBSS 4A allele (Briney et al., 1998), which is not co-dominant (producing either a band, or no band, to distinguish the alleles) optimization must ensure that all PCR machines do amplify the band from the allele, or false-negatives will occur. The reproducibility of amplifying these non-specific bands was not studied, because this section of the study was only a small investigation.

The primer GBSS8 did not generate homoprimed bands under any conditions, probably due to using too high an annealing temperature, but GBSS9 had produced homoprimed bands in the previous experiment using an annealing temperature of 57°C. Increasing the annealing temperature to 58.3°C in this experiment reduced the bands produced (Table 5.1). These temperatures (Table 5.1) formed a limit for each primer above which no homoprimed bands were formed. Any bands amplified in PCR containing a combination of two of these primers, at annealing temperatures above these temperature thresholds, should produce specifically amplified, GBSS sequences.

As a final experiment, the primers were tested to see if they acted in the same manner when used in isolation as when used in combination with another primer. PCR reactions were performed using PCR primers GBSS7-GBSS10 singly and in combination. PCR reactions
were prepared containing either 20pmol of one primer or 20pmol of each of two primers.

Reactions were incubated under the PCR conditions outlined previously, with a 57°C
annealing temperature. Figure 5.6 shows the PCR bands formed. As expected, the lanes
containing PCR products from reactions with the single primers GBSS7, GBSS9 and GBSS10
(Figure 5.6, Lanes 7, 9, 5), showed bright smears of homoprimed artifact bands. The PCR
reactions containing two PCR primers (Figure 5.6, Lanes 2, 4, 6, 8) did not produce the smears
of homoprimed bands seen in the reactions containing only one primer. This suggested that
PCR reactions containing two primers may have repressed the production of homoprimed
artifacts. If this was the case, then the amplification of the dual-primed PCR bands themselves

<table>
<thead>
<tr>
<th>Lanes</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular size marker (bp)</td>
<td>2000</td>
<td>1650</td>
<td>1000</td>
<td>650</td>
<td>400</td>
<td>2000</td>
<td>1650</td>
<td>1000</td>
<td>650</td>
<td>400</td>
</tr>
</tbody>
</table>

**Figure 5.6.** 1% agarose gel showing PCR bands produced in the presence of a
single or two primers from the set GBSS7-GBSS10. The homoprimed bands
were repressed in reactions containing two different primers. The DNA
template for the PCRs was from ‘Kulin’.

Lane 1: molecular size marker; Lane 2: primers GBSS8 and GBSS9; Lane 3:
primer GBSS8; Lane 4: primers GBSS8 and GBSS10; Lane 5: primer
GBSS10; Lane 6: primers GBSS7 and GBSS10; Lane 7: primer GBSS7; Lane
8: primers GBSS7 and GBSS9; Lane 9: primer GBSS9; Lane 10: water PCR
blank.
might limit the contaminating effects of the homoprimed fragments. The bands amplified in these reactions were not blotted and probed to identify GBSS-related fragments, since this was investigated in other experiments discussed later in this Chapter.

5.3.2 Amplification using primers GBSS1/GBSS2 and GBSS1/GBSS3

Two templates ‘Kulin’ and ‘Reeves’ were amplified using the primer combinations GBSS1/GBSS2 and GBSS1/GBSS3 (Section 5.2.1). These two varieties were chosen because they carried different GBSS 4A alleles, and thus any bands amplified from this locus could be identified. The bands amplified by these primers are shown in Figure 5.7. Both primer combinations amplified many bands with sizes ranging from 150bp to larger than 1500bp in size. The bands amplified specifically from the GBSS genes were not easily identified. At this annealing temperature (60°C), the primer GBSS3 had produced homoprimed fragments (Section 6.3.1). The homoprimed bands amplified by this primer were compared to the bands produced in the dual-primed PCR reactions and some of the fragments shown in Figure 5.7 were identified as being homoprimed fragments (data not shown).

From the length of the cDNA sequence, the bands amplified from GBSS by primers GBSS1/GBSS3 were calculated to be larger than 370bp, while the bands amplified from GBSS using GBSS1/GBSS2 would be larger than 700bp. Bands smaller than these minimum sizes were rejected as not specific to the GBSS genes. However, for each primer combination, some banding differences could be seen between the two wheat samples as highlighted on the Figure. The two wheat varieties are genetically similar except for at the GBSS 4A locus and so any differing PCR bands should be amplified from differing GBSS alleles at the 4A locus.
Due to the large numbers of PCR bands produced by primers GBSS1/GBSS2 and GBSS1/GBSS3 and the significant proportion of these bands that were predicted to be

**Figure 5.7.** Polyacrylamide gel showing PCR bands amplified using the primer combinations GBSS1/GBSS2 (lanes 2, 3) and GBSS1/GBSS3 (lanes 4, 5). Bands that differ between the two DNA samples ‘Kulin’ (lanes 2, 4) and ‘Reeves’ (lanes 3, 5) are highlighted with double-sided arrows.

Lane 1: Molecular size marker 100bp DNA Ladder (Promega); Lane 2: ‘Kulin’ DNA amplified using GBSS1/GBSS2 primers; Lane 3: ‘Reeves’ DNA amplified using GBSS1/GBSS2 primers; Lane 4: ‘Kulin’ DNA amplified using GBSS1/GBSS3 primers; Lane 5: ‘Reeves’ DNA amplified using GBSS1/GBSS3 primers.

### 5.3.3 Identification of GBSS-related PCR bands by hybridization.

Due to the large numbers of PCR bands produced by primers GBSS1/GBSS2 and GBSS1/GBSS3 and the significant proportion of these bands that were predicted to be
homoprimed PCR artifacts, GBSS-related bands were difficult to identify. Although many non-specific PCR bands could be eliminated by increasing the annealing temperature (Section 5.3.1), the sizes of the GBSS-specific PCR bands were not known so the PCR conditions could not be optimized to favor their amplification. The PCR bands were investigated to find GBSS-related bands.

Since the fragments amplified from the GBSS genes by the primers GBSS1/GBSS2 would encompass the section of DNA amplified by the GBSS-specific primers GBSS-Rev/GBSS-4A, this smaller PCR fragment could be used as a probe to hybridise to and identify GBSS-related bands amplified by the GBSS1/GBSS2 primers. The probe region was outside the primer binding sites of the GBSS1/GBSS3 primers and so no GBSS-related bands would be seen in these samples. However, the sizes of GBSS1/GBSS3 bands could be calculated once the sizes of the encompassing GBSS1/GBSS2 PCR fragments were found.

The PCR bands shown in Figure 5.7 were blotted onto a nylon membrane and probed with the GBSS-specific probe (Section 6.2.2). Figure 5.8 shows the comparison of the PCR bands on the gel with those identified with sequence similarity to the GBSS probe. For the primer combination GBSS1/GBSS2, many bands larger than 1kbp were identified as GBSS-related. From the size of the GBSS gene from Barley (X07931), the sizes of the GBSS1/GBSS2 PCR bands from wheat were predicted to have been approximately 1130bp. For the primer combination GBSS1/GBSS3, no bands were expected because the region of the GBSS gene that these primers amplified was outside the region to which the probe bound. However, for these primers, a discrete band of approximately 600bp and a band cluster between 1200-1400bp was identified (Figure 5.8). The only difference seen between the GBSS-related bands amplified from ‘Kulin’ and ‘Reeves’ was an approximately 100bp band amplified from ‘Reeves’ by the GBSS1/GBSS2 primers but absent from the ‘Kulin’ sample.
Figure 5.8. Identification of PCR bands amplified from the GBSS genes by hybridisation to a GBSS-specific probe. The polyacrylamide gel (left) shows the PCR bands amplified using the primer combinations GBSS1/GBSS2 (lanes 2, 3) and GBSS1/GBSS3 (lanes 4, 5). The two autoradiographs (middle, right) show bands related to the GBSS genes by DNA sequence hybridisation. The Southern blot membranes were probed with a GBSS-Rev/GBSS-4A PCR fragment amplified from the 4A locus of ‘Kulin’, washed to a stringency of 1xSSc at 60°C and exposed to film for 24hrs (middle) and 72hrs (right). For the primer combination GBSS1/GBSS2 (lanes 2, 3), all bands larger than 1kbp in size hybridised to the GBSS-specific probe. For primer combination GBSS1/GBSS3 (lanes 4, 5) a discrete band approximately 600bp in size and the band cluster between 1200-1400bp in size were identified by hybridisation. Other bands from both primer combinations were more faintly labeled.

Lane 1: Molecular size marker 100bp DNA Ladder (Promega); Lane 2: ‘Kulin’ DNA amplified using GBSS1/GBSS2 primers; Lane 3: ‘Reeves’ DNA amplified using GBSS1/GBSS2 primers; Lane 4: ‘Kulin’ DNA amplified using GBSS1/GBSS3 primers; Lane 5: ‘Reeves’ DNA amplified using GBSS1/GBSS3 primers.
5.3.4 Isolation and cloning of the GBSS-related PCR products.

After identifying the GBSS-related PCR products by sequence hybridisation, the next steps were cloning and sequencing them. The PCR was not further optimized to reduce non-specific amplification. Since the sizes of the desired bands had been identified, those bands could be separated from the bands unrelated to the GBSS genes before cloning. GBSS-related PCR bands over 1000bp amplified with the GBSS1/GBSS2 primers were purified from those bands smaller than 1000bp.

PCR fragments were separated using gel and capillary electrophoresis. Using capillary electrophoresis (Section 5.2.3), four PCR bands larger than 1000bp were detected. The bands were eluted from the capillary in viscous buffer, diluted 1 in 3 in water and re-amplified using the GBSS1/GBSS2 primers (Section 5.2.1). Figure 5.9 shows the PCR products amplified from the isolated bands. Many bands of different sizes were amplified.

The GBSS-related GBSS1/GBSS2 PCR bands were also separated by gel electrophoresis, the bands were cut from the gel and the DNA recovered (Section 5.2.3). An aliquot of the PCR fragments isolated from the gel was electrophoresed on agarose and the preferred PCR fragments were successfully purified away from smaller undesirable fragments (data not shown). The gel-purified PCR fragments were cloned (Section 5.2.4). Only eight transformants were obtained and none was found to contain inserts when screened with T7/Sp6 primers (Section 5.2.6). Analysis of suitable control ligation reactions showed that the gel-purified inserts had not ligated into the pGEM-T vector efficiently. Thus the gel-purified GBSS-related PCR products could not be cloned.
These methods to isolate the GBSS-related fragments had not been successful and so the experiment was modified to clone the entire range of bands. The fragments amplified from ‘Kulin’ by both GBSS1/GBSS2 and GBSS1/GBSS3 were cloned (Section 5.2.4). The sizes of the GBSS-related bands had been identified and the sizes of the inserts were screened using T7/Sp6 PCR primers. These primers amplified from the vector across the insert showing the size of the insert and an example of the sizes of the inserts is shown in Figure 5.10. Clones with a range of inserts were generated, showing that all the different sized PCR fragments had been cloned.

**Figure 5.9.** 1% agarose gel showing the PCR amplification using primers GBSS1/GBSS2 from a template of a single DNA fragment isolated by capillary electrophoresis.

Lane 1: Molecular size marker 100bp DNA Ladder (Promega); Lanes 2-5: PCR using a single band isolated by capillary electrophoresis as template; Lane 6: PCR using a mixture of four bands isolated by capillary electrophoresis as template; Lane 7: electrophoresis elution buffer control.
To identify the inserts, cell-scrapings of the individual clone colonies were dot-blotted onto nylon membranes and probed with a GBSS-specific probe (Section 5.2.2). Figure 5.11 shows the clones identified by hybridization to a GBSS-specific probe. The positive control, labeled ‘P’, hybridised to the probe and two clones (Clone 10 and Clone 19) were also identified as containing GBSS sequences.

The inserts in Clones 10 and 19 were originally produced by the GBSS1/GBSS2 primers. The GBSS-related PCR fragments amplified by these primers had been identified by hybridization as being larger than 1kbp (Section 5.3.2). The inserts in Clones 10 and 19 were expected to have been larger than 1000bp. However, the T7/Sp6 PCR screen of these clones indicated that they were about 600bp in size (data not shown), i.e. shorter than the predicted value. The inserts in Clones 10 and 19 were sequenced (Figure 5.12). They were similar to each other suggesting they were from the same fragment. However, the sequences could not be aligned with the sequence of the GBSS-specific probe. When the cloned sequence was used to

**Figure 5.10.** PCR with T7/Sp6 primers showing clones containing different sized inserts. Inserts in Lanes 2-8 were GBSS1/GBSS3 PCR fragments, inserts in Lanes 9-13 were GBSS1/GBSS2 PCR fragments. The 2% agarose gel was electrophoresed at 50V for 2hrs.

Lane 1: 100bp DNA Ladder molecular size marker (Promega); Lanes 2-13: insert sizes of pGEM-T clones.
Figure 5.11. Dot Blot of individual clones containing insert PCR fragments amplified by the GBSS1/GBSS2 and GBSS1/GBSS3 primers, probed with a GBSS-specific probe. The two images are the same blot exposed to film for 17hrs (top) and 42hrs (bottom). The blot was washed with 1xSSC at 60°C. The positive hybridisation control hybridised strongly to the probe (Column 1, Row 1). Two clones also strongly hybridised to the probe, Clone 10 and Clone 19 (Column 12, Row 1; Column 9, Row 2). Other clones hybridised more weakly to the probe for example Clone 44 (Column 10, Row 4) and Clone 72 (Column 2, Row 7).

(Column 1, Row 1): Positive control plasmid containing a GBSS insert; (Column 2, Row 1): Negative control E.coli cell sample; All other dots contain individual clones.
BLAST the genome databases (www.ncbi.nlm.nih.gov/Blast/), it was significantly similar to a sequence from the diploid wheat *Triticum monococcum* from a non-coding region containing microsatellites and inverted repeats (AF326781). It is not clear how the GBSS-specific probe hybridized with these cloned sequences.

5.3.5 Design and optimization of novel PCR primers to amplify from the wheat GBSS genes.

The GBSS1–GBSS3 PCR primers had amplified from many places in the wheat genome, producing non-specific amplification. Six new PCR primers (GBSS5–GBSS10) were designed from the sequence of the wheat GBSS cDNA (Clark et al., 1991). They were designed such that, used in various combinations, they should amplify most sections of the

---

**Figure 5.12.** Sequence of cloned PCR fragments, identified by hybridisation to be related to the GBSS genes, but not similar in sequence. The sequence reads continuously from left to right, but has been displayed to show its highly repetitive structure with similar sequences directly below each other and dissimilar sequences separated by a line space. The fragment was homoprimed using the GBSS1 primer as shown in bold.
GBSS gene (Figure 5.2). The primers were used in all logical combinations to amplify from the wheat genome (Section 5.2.1) and were run over a gradient of increasing annealing temperatures to reduce the production of homoprimed bands (Figures 5.13 and 5.14). At lower annealing temperatures, a range of bands were produced by most primer sets. Most of these bands were not amplified at higher temperatures suggesting imperfect binding of the primers to the template. The PCR reactions with annealing temperatures high enough to repress the production of homoprimed bands have been highlighted on the Figures. Any PCR bands amplified above these temperatures should have been dual-primed bands specific to the GBSS genes.

Primer combinations GBSS5/GBSS6, GBSS4/GBSS6, GBSS7/GBSS10 and GBSS8/GBSS10 all generated PCR bands at higher annealing temperatures than the threshold temperature predicted to repress homoprimed fragments. The bands produced by primer combinations GBSS4/GBSS6 and GBSS7/GBSS10 both produced a single band up to an annealing temperature of 72°C. These bands were of an appropriate size compared to the predicted size of the fragment from the GBSS genes and appeared to be good candidates for having been amplified from the GBSS genes. If the primer binding sites were conserved between the three GBSS loci, the primers should have amplified from all three loci. Since the GBSS genes from each locus were predicted to be different sizes due to intron length differences (Chapter 3), it was expected that up to three bands would be amplified from the GBSS genes. That only one band was present suggested that either the GBSS genes from each locus were the same size or that the PCR had only amplified from a single locus.
Figure 5.13. 1% agarose gels showing the PCR bands amplified from the wheat genome by combinations of the primers GBSS3-GBSS6. The primer combinations used in each set of reactions are shown under the lanes. Reactions were incubated over a gradient of annealing temperatures. The reaction incubated at the lowest annealing temperature at which no homoprimed bands should have been amplified has been highlighted with a white arrow. Bands in reactions incubated at this temperature or higher should be dual-primed from the GBSS genes. Thermal cycling conditions were 1 hold at 94°C for 4min; 10 cycles of 94°C for 30sec, annealing for 30sec and 72°C for 2.5min; 25 cycles of 94°C for 30sec, annealing for 30sec and 72°C for 2.5min (+5sec/cycle); 1 hold at 72°C for 4min.

Lane 1: Molecular size marker mix containing Lambda/HindIII (Biotech International Ltd) and 100bp DNA Ladder (Promega); Lane 2: annealing temperature of 57.5°C; Lane 3: 58.5°C; Lane 4: 60.3°C; Lane 5: 62.7°C; Lane 6: 65.4°C; Lane 7: 68.3°C; Lane 8: 71.1°C; Lanes 9-12: 72°C.
Figure 5.14. 1% agarose gels showing the PCR bands amplified from the wheat genome by combinations of the primers GBSS7-GBSS10. The primer combinations used in each set of reactions are shown under the lanes. The reaction incubated at the lowest annealing temperature at which no homoprime bands should have been amplified has been highlighted with a white arrow. Bands in reactions incubated at this temperature or higher should be dual-primed from the GBSS genes. Thermal cycling conditions were as outlined previously in Chapter 5.

Lane 1: 1 Kb Plus DNA Ladder molecular size marker (Gibco BRL); Lane 2: 58.3°C; Lane 3: 58.9°C; Lane 4: 60.2°C; Lane 5: 61.8°C; Lane 6: 63.7°C; Lane 7: 65.8°C; Lane 8: 67.7°C; Lane 9: 69.5°C; Lane 10: 70.8°C; Lane 11: 71.9°C.
5.3.6 Cloning and sequencing the GBSS4/GBSS6 PCR fragment.

The PCR band amplified by the GBSS4/GBSS6 primers was chosen to be cloned and sequenced. It contained sequence 3’ of the known region from exons 4 to 6, providing new GBSS DNA sequence as well as the known region which would allow the locus of origin to be identified. Many clones were produced from this transformation event and screened for GBSS-related inserts using the GBSS-Rev/GBSS-4A,7A,7D primer combination (Section 2.9) to amplify from within the GBSS4/GBSS6 PCR fragment insert. Initially, clones were screened individually, but this proved inefficient because of the small percentage of clones that contained the GBSS4/GBSS6 fragment (approximately 0.3%). The method was optimized to screen many clones at once by sampling several colonies into the same water sample and testing this mixed template using a single PCR (Section 5.2.6).

Figure 5.15 shows results for PCR tests to screen the groups of clones for their inserts. The upper set of reactions are using the GBSS-Rev/GBSS-4A,7A,7D primer combination to identify GBSS-related sequences. This PCR identified the group of clones tested in Lane 8 (Figure 5.15) that contained a GBSS sequence. At least one of the clones in this set must contain a GBSS-related insert. The lower set of PCR reactions contained the T7/Sp6 primers that amplified across the insert in the vector to size the insert. This PCR did not indicate that the clone group in Lane 8 contained the large-sized GBSS-related insert identified by the other PCR primer set. It was possible that the T7/Sp6 PCR test was not as sensitive as the GBSS-Rev/GBSS-4A,7A,7D primer test or that when the T7/Sp6 primers amplified from the mixed template containing different sized inserts, the smaller inserts were preferentially amplified before the larger inserts.

Each clone in the group was then screened individually using the GBSS-Rev/GBSS-4A,7A,7D primer combination (Figure 5.16), and one clone had a GBSS insert (Lane 14). The insert fragment from this clone was smaller than the insert fragment from the positive control (Lane
The positive control fragment was from the GBSS 7D locus, the locus that produced the largest sized PCR fragment. The GBSS fragment in the new clone was predicted to have come from the GBSS 7A or 4A loci, because these loci gave a smaller PCR fragment than the 7D locus. The clone was tested with the three locus-specific PCR primer sets individually.

**Figure 5.15.** 1% agarose gel showing how groups of clones were tested for the inserts they contain using two sets of PCR primers. The top gel shows the PCR test containing primers GBSS-Rev/GBSS-4A,7A,7D. These primers amplify from GBSS sequences and amplified from both the positive control (Lane 18) and the set of clones tested in Lane 8. Thus one of the clones in this set contains a GBSS insert. The lower gel shows the PCR testing of the samples using the primers T7/Sp6. The wells of the two gels align and the same samples were tested in aligned wells. The T7/Sp6 primers did not identify any of the approximately 2kbp GBSS-related inserts.

Lane 1: Molecular size marker (upper gel) 100bp DNA Ladder (Promega) and (lower gel) λ/HindIII (Biotech International Ltd.); Lanes 2-17: groups of clones; Lane 18: GBSS positive control plasmid containing a GBSS insert; Lane 19: water blank.
Only the GBSS-Rev/GBSS-7A primers amplified, suggesting that the GBSS4/GBSS6 insert in the new plasmid originated from the GBSS 7A locus.

**Figure 5.16.** 1% agarose gels showing PCR testing of clones for their inserts. The top gel shows the screening of 14 individual clones using the GBSS-Rev/GBSS-4A,7A,7D primers to identify which clones contained a GBSS-related insert. The lower gel shows the assessment of a desired clone with locus-specific primers to determine at which locus the insert originated. The band in Lane 19 was produced by primers GBSS-Rev/GBSS-7A and so the insert originates from the 7A locus.

Lanes 1,17: 100bp DNA Ladder molecular size marker (Promega); Lanes 2-14: individual clones tested with the GBSS-Rev/GBSS-4A,7A,7D primers; Lane 15: GBSS positive control PCR of a GBSS 7D fragment; Lanes 16, 21: water blanks; Lanes 18-20: PCR tests on the clone from Lane 14 using the GBSS-Rev/GBSS-4A, GBSS-Rev/GBSS-7A and GBSS-Rev/GBSS-7D primer sets respectively.
The plasmid containing the GBSS4/GBSS6 insert was extracted and sequenced from both ends. The insert was about 2000bp in size and too long to be sequenced as a single fragment. Initially 600bp was sequenced from each end of the fragment confirming that it originated from GBSS genes. The full sequence was obtained by first digesting the PCR fragment with restriction enzymes BamH1, EcoR1, HindIII and Sal1 (Section 5.2.5) - only BamH1 and Sal1 cut it, into two pieces approximately 1400bp and 550bp in size (Figure 5.17). The pieces were then cloned.

**Figure 5.17.** 1% agarose gel showing the approximately 2kbp PCR fragment amplified from ‘Kulin’ using GBSS4/GBSS6 PCR primers and digested with restriction enzymes. The enzymes EcoR1 and HindIII (Lanes 3, 4) did not digest the fragment, but the enzymes BamH1 and Sal1 (Lanes 2, 5) did.

Lane 1: Molecular size marker mix containing 100bp DNA Ladder (Promega) and Lambda/HindIII (Biotech International Ltd.); Lane 2: BamH1; Lane 3: EcoR1; Lane 4: HindIII; Lane 5: Sal1.
None of the first 23 clones screened contained an insert. To concentrate the percentage of plasmids in the population that contained inserts, the remaining clones were pooled and all their plasmids were extracted. This plasmid mix was then linearised with single-site restriction enzymes and compared to control plasmids containing either no insert or inserts of known size (Section 5.2.5). The majority of the plasmids in the population contained no inserts (Figure 5.18). A small percentage of plasmids did contain inserts. One of these was about the same size as a plasmid containing a 500bp insert whilst another was a larger insert.

Plasmids containing inserts were purified away from those not containing inserts by cutting those bands from the gel and extracting the DNA. Plasmids containing both the smaller and larger inserts were recovered (Figure 5.19). These samples still retained a small number of contaminating plasmids without inserts, but the number was greatly reduced.

Gel-purified plasmids were recircularised and transformed into E. coli (Section 5.2.5). The resulting clones were screened for inserts using T7/Sp6 primers and the inserts were examined using PCR with various combinations of the primers T7, Sp6, GBSS4 and GBSS6. Since the original PCR product amplified by the GBSS4/GBSS6 primers had been digested with restriction enzyme and then cloned, the insert contained only one of its original primer sites and therefore amplifying the insert using one of GBSS4 or GBSS6 required the use of the vector-based primers T7 or Sp6. The orientation and length of the inserts was checked before plasmids were extracted from the clones and the inserts sequenced.
Figure 5.18. 1% agarose gel comparing a mixed population of plasmids containing a variety of inserts (Lanes 2, 3) with control plasmids containing inserts of a known size (Lanes 4-7). The control plasmids in Lanes 4 and 5 did not contain an insert and appeared as a single band 3kbp in size. The major band in the mixed-plasmid samples (Lanes 2, 3) corresponded to this 3kbp band indicating that the majority of plasmids in the mixed population did not contain an insert. The control plasmids in Lanes 6 and 7 contained 500bp inserts. In the mixed-plasmid samples (Lanes 2, 3) there was a minor band that corresponded to this 500bp-insert band. There was also a larger band in the mixed-plasmid samples that corresponded to a plasmid containing a 1500bp insert. The remaining lanes on the gel were restriction digest controls.

Lane 1: Molecular size marker (Promega); Lane 2: mixed plasmid sample linearised using Sac1; Lane 3: mixed plasmid sample linearised using Sal1; Lane 4: plasmid containing no insert linearised using Sac1; Lane 5: plasmid containing no insert linearised using Sal1; Lane 6: plasmid containing a 500bp insert linearised using Sac1; Lane 7: plasmid containing a 500bp insert linearised using Sal1; Lane 8: non-linearised mixed plasmid sample; Lane 9: non-linearised mixed plasmid sample; Lane 10: non-linearised plasmid containing a 500bp insert.
Figure 5.19. 1% agarose gels showing a mixture of linearised plasmids containing variously sized inserts. The top gel shows the mixed population of plasmids resulting from a ligation event. The major band (labeled ‘A’) corresponded to plasmids containing no inserts, while the minor bands (labeled ‘B’ and ‘C’) corresponded to plasmids containing inserts. As only plasmids containing inserts were of interest, the plasmid population was enriched for these plasmids by gel-extraction. The plasmids in bands ‘B’ and ‘C’ in the top gel were excised, gel-purified and electrophoresed on the bottom gel. Comparing the two gels, the amount of the insert-less plasmid ‘A’ had been greatly reduced in the bottom gel by the gel extraction procedure.

Lanes 1, 4: Lambda/HindIII molecular size marker (Biotech International Ltd.); Lane 2, 3: mixed plasmid samples containing variously sized inserts; Lane 5: gel purified plasmid sample cut from Lane 2 band ‘C’; Lane 6: gel purified plasmid sample cut from Lane 2 band ‘B’; Lane 7: gel purified plasmid sample cut from Lane 3 band ‘C’; Lane 8: gel purified plasmid sample cut from Lane 3 band ‘B’.
5.3.7 Sequence of the GBSS4/GBSS6 PCR fragment.

The DNA sequences of the two sections of the GBSS4/GBSS6 PCR fragment were joined into one sequence 1951bp in length. Comparing this sequence to the GBSS locus-specific sequences from Chapter 3 confirmed that the cloned sequence was from the GBSS 7A locus and the sequence was therefore termed ‘Kulin_7A’. The sequence extended from exon four to the end of the coding region in exon 11 and had not previously been sequenced from Australian wheat varieties. The GBSS-7A allele from ‘Kulin’ was sequenced only once and its sequence needs to be confirmed.

The ‘Kulin_7A’ sequence was compared to the published wheat genomic and cDNA GBSS sequences from the 7A locus of Chinese Spring wheat (Murai et al., 1999, Clark et al., 1991). An alignment of these sequences is shown in Figure 5.20. There were 14 single base-pair substitutions between the three DNA samples. Thirteen of these differences occurred between the sequence from ‘Kulin’ and the sequences from ‘Chinese Spring’, suggesting that these bases were variety-specific between the two wheat genotypes. There were also two insertions in the ‘Kulin’ sequence that did not occur in ‘Chinese Spring’, one of which was the (TGCCG)_n microsatellite identified in Chapter 3.
Figure 5.20. (Two pages)
The DNA sequences from both ‘Chinese Spring’ and ‘Kulin’ had their intron regions removed, were translated into amino acid sequences and aligned to each other using ANGIS software (www.angis.org.au). This alignment is presented as Figure 5.21. Of the 12 single nucleotide polymorphisms that occurred between the two DNA sequences within the exon regions, eight resulted in amino acid variation between the proteins. The two alleles encoded slightly different proteins. The allele from ‘Kulin’ was novel and not before reported. The

**Figure 5.20.** (Continued from the previous page) Comparison of three DNA sequences from GBSS genes. The first sequence (ChSpr_7A) is a genomic DNA sequence from the GBSS gene at the 7A locus of ‘Chinese Spring’ wheat (Murai et al., 1999). The second sequence (Kulin_7A) is the partial genomic DNA sequence from the GBSS gene at the 7A locus of ‘Kulin’ wheat sequenced in the current project. The third sequence (ChSpr_cDNA) is a cDNA sequence from the 7A locus of ‘Chinese Spring’ wheat (Clark et al., 1991). The positions of the PCR primers GBSS4 and GBSS6 that were used to amplify the fragment ‘Kulin_7A’ have been coloured blue. Sequence differences between the three DNA sequences have been coloured red and highlighted with a star. There are 14 single nucleotide polymorphisms between the three sequences, 13 of which occur between the sequence from the wheat variety ‘Kulin’ and those from ‘Chinese Spring’ varieties. These differences may be accession-specific and result from the genetic distance between these types of wheat. In addition there are two insertions in the sequence ‘Kulin_7A’ that do not occur in the sample from ‘Chinese Spring’ wheat. One of these insertions is a microsatellite polymorphic in WA noodle wheat varieties. As the insertions occur within intron regions, they do not affect the resulting protein sequence.
GBSS alleles have traditionally been named alphabetically, but the number of GBSS alleles is not well defined, and thus, assigning names can be ambiguous. It is proposed that this new GBSS allele be named WaxyA1k in reference to the wheat variety ‘Kulin’ in which it was found.

**Figure 5.21.** A protein sequence alignment in the region of the GBSS-7A protein from exon 4 (top of alignment) to the stop codon ending exon 11. The two protein sequences come from ‘Chinese Spring’ wheat (Murai et al., 1999) and ‘Kulin’. There are 8 single-amino acid differences between the two protein sequences, highlighted with stars and coloured red.
5.3.8 DNA sequence of the GBSS4/GBSS6 PCR fragment from a second Australian wheat variety ‘Eradu’.

The GBSS alleles \textit{WaxyA1a} from ‘Chinese Spring’ wheat and \textit{WaxyA1k} from ‘Kulin’ differed in their (TGCCG)$_3$ microsatellite configuration, but it was not known whether all varieties that contained the (TGCCG)$_3$ motif seen in \textit{WaxyA1a} also contained the amino acid differences seen between \textit{WaxyA1a} and \textit{WaxyA1k}. The GBSS-7A allele from a second Australian wheat variety ‘Eradu’ shared the same (TGCCG)$_3$ microsatellite allele as the \textit{WaxyA1a} allele. The GBSS4/GBSS6 PCR product from ‘Eradu’ was amplified, cloned and sequenced. It was sequenced only once and needs to be confirmed.

The DNA sequence of the GBSS4/GBSS6 PCR fragment from the GBSS-7A gene of ‘Eradu’ is given in Figure 5.22 compared to the sequences from ‘Kulin’ and ‘Chinese Spring’. Many of the DNA polymorphisms which differed between the ‘Chinese Spring’ sequence and that from ‘Kulin’ did not differ between ‘Chinese Spring’ and ‘Eradu’. The sequences from ‘Kulin’ and ‘Eradu’ differed at several points suggesting that these were different alleles with the ‘Eradu’ allele being more similar to the ‘Chinese Spring’ allele. However, there were also DNA polymorphisms within exon 6 of the ‘Eradu’ sample that differed from both the other sequences. All three sequences were in fact different.

When the protein sequences were compared (Figure 5.23) the predicted amino acid sequence from the GBSS-7A gene of ‘Eradu’ was similar to that from ‘Chinese Spring’ differing at only a single residue, and was less similar to the ‘Kulin’ sample differing at 9 amino acids.
Figure 5.22. (Two pages)
Figure 5.22. (Continued from the previous page). Comparison of three DNA sequences from GBSS genes. The first sequence (ChSpr_7A) is a genomic DNA sequence from the GBSS gene at the 7A locus of ‘Chinese Spring’ wheat (Murai et al., 1999). The second sequence (Kulin_7A) is the partial genomic DNA sequence from the GBSS gene at the 7A locus of ‘Kulin’ wheat sequenced in the current project. The third sequence (Eradu_7A) is the partial genomic DNA sequence from the GBSS gene at the 7A locus of ‘Eradu’ wheat sequenced in the current project. The positions of the PCR primers GBSS4 and GBSS6 that were used to amplify the fragment ‘Kulin_7A’ have been coloured blue. Sequence differences between the three DNA sequences have been coloured red and highlighted with a star.

Many of the DNA polymorphisms which differed between the ‘Chinese Spring’ sequence and that from ‘Kulin’ were not different between ‘Chinese Spring’ and ‘Eradu’. Thus the sequences from ‘Kulin’ and ‘Eradu’ differed at several points. There were also DNA polymorphisms within exon 6 of the ‘Eradu’ sample that were different to both the other sequences. It could thus be seen that all three sequences differed.
**Figure 5.23.** A protein sequence alignment in the region of the GBSS-7A protein from exon 4 (top of alignment) to the stop codon ending exon 11. The three protein sequences come from ‘Chinese Spring’ wheat (Murai et al., 1999), ‘Kulin’ and ‘Eradu’. The 9 single-amino acid differences between the three protein sequences are highlighted with stars and coloured red. The protein sequence from ‘Kulin’ differed from the other two sequences at 8 or 9 aa, but the ‘Eradu’ sequence differed from the ‘Chinese Spring’ sequence at only a single aa (marked with #). Those residues assigned an ‘X’ were not identified due to unavailable sequence data.
5.3.9 Use of the GBSS4/GBSS6 PCR primers to access DNA sequence from the null 7A allele.

The primers GBSS4/GBSS6 were used to amplify the null 7A allele from ‘DHW174’ wheat to look for polymorphisms in the 3’ region of this gene. The PCR band was produced as described for ‘Kulin_7A’. The 5’ end of the PCR product had previously been sequenced (Chapter 3) and so the 3’ end of the PCR product was directly sequenced using primer GBSS6 (Section 2.10.2).

Figure 5.24 shows the DNA sequence of this null 7A PCR product compared to the sequence of the wild-type 7A gene from ‘Chinese Spring’ published by Murai et al. (1999). The PCR product showed high sequence similarity to the GBSS-7A gene from ‘Chinese Spring’. The null 7A sequence had a premature stop-codon at the start of exon 11, 105bp before the stop-codon of the wild-type 7A gene. This result suggested that the null 7A allele is caused by the insertion of a premature stop-codon resulting in a truncated and inactive protein.

In 1999, Vrinten et al. published the sequence of the null 7A gene from ‘Chinese Spring’ wheat and this sequence showed homology with the wild-type 7A gene (Murai et al., 1999) at the 3’ end of the gene. The null 7A allele from ‘Chinese Spring’ wheat therefore had a different DNA sequence to the null 7A PCR product sequenced here. It was therefore possible that the null 7A gene from ‘DHW174’ wheat was a previously-undescribed allele truncated at the 3’ terminus. However, the pedigree of ‘DHW174’ wheat suggested that it shared the same allele as in ‘Chinese Spring’ wheat. The null 7A allele from ‘DHW174’ was re-sequenced to confirm the polymorphism seen in the initial sequence.

The second time the null 7A gene from ‘DHW174’ was sequenced, it showed DNA sequence homology to the null 7A and wild-type 7A alleles from ‘Chinese Spring’ wheat (Figure 5.25).
The premature stop-codon seen in the first sequence was not seen in the second sequencing event.

---

**Figure 5.24.** DNA sequence comparison between a PCR fragment predicted to have been amplified from the null GBSS-7A gene of the wheat variety ‘DHW174’ (DHW174_7A) and the wild-type GBSS-7A gene from ‘Chinese Spring’ (ChSpr_7A). The DNA sequence of the PCR product covered the region from intron 8 to exon 11. The PCR product showed high sequence similarity to the ‘Chinese Spring’ sample suggesting it was amplified from the 7A locus, not from the other two loci. The PCR fragment from ‘DHW174’ displayed a premature ‘TAA’ stop codon (coloured red) at the beginning of exon 11. After this point its sequence deviated from that of the ‘Chinese Spring’ sample.
Although this region of the null 7A allele was re-amplified and re-sequenced twice more, it was not possible to again amplify the truncated null 7A sequence from the genome and could not be studied further.

**Figure 5.25.** DNA sequence comparison between the first and second PCR fragments predicted to have been amplified from the null GBSS-7A gene of the wheat variety ‘DHW174’ (DHW174_7A, DHW174_7A/2) with the wild-type GBSS-7A gene from ‘Chinese Spring’ (ChSpr_7A). The high level of DNA sequence similarity between the three sequences suggested they were all amplified from the 7A locus, not from the other two loci. The two samples from ‘DHW174’ wheat vary at the 3’ end with the second sample (DHW174_7A/2) exactly matching the sequence from ‘Chinese Spring’ wheat (ChSpr_7A).
5.4 Discussion.

5.4.1 Homoprimed bands.

The primers GBSS1 to GBSS10 had produced multiple homoprimed bands from the wheat genome, supporting the theory that primers bind to many places in this genome, and that PCR from wheat is complicated by non-specific amplification. It was predicted that the primers may have bound to an intron region of the GBSS genes, the sequence of which was not known at the time this study was done, giving a GBSS-related homoprimed band. Whilst it is not common for GBSS introns to have similar sequences to exons, until these regions had been characterized, it could not be assumed that all the non-specific bands were amplified from outside the GBSS genes. Homoprimed bands might also have been amplified from a region containing DNA sequence from pseudogene copies of the GBSS gene where duplication, repetition and recombination had resulted in tandem or reversed copies of the primer binding site. However the homoprimed bands were found to be unrelated to the GBSS gene by hybridization. This suggested that there were no GBSS pseudogene complexes in the genome. It also showed that the intron sequences were not similar to the exons, a result confirmed when sequence data from the introns became available (Murai et al., 1999).

When the genomic sequences of the GBSS genes became available (Murai et al., 1999), the ten PCR primers were compared to the GBSS sequences, using the software Amplify (Engels, 1993). At lower levels of stringency, all the primers, excepting GBSS6 and GBSS7, were predicted to bind more than once within the GBSS genes and produce homoprimed fragments. This did not occur, indicating that the PCR conditions used were too stringent to allow GBSS-specific homoprimed bands to be amplified and that the homoprimed bands produced were amplified from strong primer-binding events with highly complementary sequence outside the GBSS genes.
The 10 PCR primers designed to the GBSS cDNA (Clark et al., 1991) all gave homoprimed PCR fragments when used alone (though GBSS8 only produced faint bands). This result supports findings from Murai et al. (1999) who used 34 PCR primers designed from wheat cDNA in various combinations to amplify GBSS genes. They reported that each set of PCR primers amplified at least five different PCR fragments and some produced ten fragments. PCR from the GBSS loci should generate a maximum of three bands. When sequenced, the remaining bands were not from GBSS genes (Murai et al., 1999). It seems that primers designed to this cDNA sequence anneal at many places in the wheat genome and this may be a general feature of primers designed to GBSS coding sequences.

Although PCR primers designed to the cDNA sequence seemed to bind to many places in the wheat genome and amplify non-specific bands, the primers designed to the intron regions of the GBSS genes, such as GBSS-Rev/4A/7A/7D, were more specific and did not produce homoprimed bands or non-specific dual-primed bands (Chapter 3). It is possible that sequences from the exons of the GBSS genes are replicated more than intron sequences in the wheat genome. Maybe there is conservation of sequences encoding certain structural elements of proteins. For example, if a protein structure such as a loop or a helix is encoded in one gene by a specific sequence, the same sequence could also be used to encode loops or helixes in other genes. When designing future molecular markers to the GBSS genes, primers that target the more-specific intron sequences should reduce non-specific amplification.

### 5.4.2 GBSS1-GBSS3 PCR bands.

Many PCR bands were amplified using the primers GBSS1/GBSS2 and those related to the GBSS genes were identified by hybridization. The genes at the three loci are different sizes and so three bands were expected to have hybridized to the probe. However, more than three
bands hybridized to the probe suggesting multiple binding sites for one or both primers, flanking the region of the probe.

Primers GBSS1/GBSS3 also amplified many bands, but since the primers bound outside the probe region, none were expected to hybridise to the probe. However bands did anneal to the probe, including one about 600bp that was the same size as this section of the GBSS gene from Barley (X07931). A 1200-1400bp band cluster was also identified by the probe, suggesting that one or both of the primers had alternative binding sites resulting in PCR fragments including the probe region.

Computer analysis (‘Amplify’ Engels, 1993) of the GBSS genomic DNA sequences (Murai et al., 1999) showed potential GBSS1-GBSS3 primer binding sites that, under conditions of low stringency, would produce the kinds of PCR bands seen in this study. An example of this is presented in Figure 5.26. It was likely that the many bands identified as being GBSS-specific

![Figure 5.26](image_url)

**Figure 5.26.** A diagram of the potential PCR primer binding positions of primers GBSS1-GBSS3 (arrows with the number of the primer above) on the GBSS-7A gene as represented by the black line from the start codon at the 5’ end to the stop codon at the 3’ end. The primers would only bind at all these positions under conditions of low stringency such as low annealing temperature. The position of the gene fragment used as a probe is shown as a black rectangle on the gene. Thus many PCR fragments amplified by combinations of these primers would anneal to the probe.
after hybridisation were truly amplified from the GBSS genes. Thus, sequencing those fragments would give access to GBSS DNA sequence information.

Two varieties ‘Kulin’ and ‘Reeves’ were used as template because they contained different 4A alleles and should give different banding patterns. Since ‘Reeves’ contained the null 4A allele, it was expected that one of the bands amplified in the ‘Kulin’ sample would be missing from the ‘Reeves’ sample. The amplification of a 100bp GBSS-related band from ‘Reeves’, that was not amplified from ‘Kulin’, suggests that its genome contains GBSS sequences outside the GBSS loci. This was not supported by the genomic Southern blot hybridization findings (Chapter 4) or by other similar work by Vrinten et al. (1999).

The GBSS-related GBSS1/GBSS2 PCR bands were separated from unrelated bands using both gel and capillary electrophoresis. Capillary electrophoresis was carried out in a viscous buffer that was predicted to inhibit downstream reactions, so the eluted bands were diluted in water and the DNA re-amplified. However, a range of unexpected bands were amplified from the template. The capillary electrophoresis procedure had separated the different sized fragments enough to be detected as individual bands, but the full range of fragment sizes may still have been present in small amounts in the elute. The small-sized fragments were then re-amplified in the PCR. A second possibility is that the large PCR fragments contained internal primer binding sites that resulted in many PCR bands being amplified using the purified large PCR band as template. Since smaller fragments are amplified preferentially by PCR, the requirement for a PCR step to remove the large-sized GBSS-related fragments from the viscous electrophoresis buffer limited the usefulness of this method.

Gel electrophoretic purification of the GBSS-related PCR bands was also unsuccessful because the recovered DNA could not be cloned. Difficulty inserting DNA fragments that
have been passed through a gel into cloning vectors was encountered in other parts of the project (Chapter 4). Possibly passage through the gel or the gel purification procedure damaged the DNA (by removing its single-stranded overhangs or changing its chemistry) reducing its ability to be efficiently cloned into the pGEM-T vector.

Although no GBSS-related GBSS1-GBSS3 PCR fragments were sequenced, they were identified by hybridization amongst the mixture of unrelated fragments amplified by these primers. This provides a point from which future work can continue, including optimizing reaction conditions (temperature and MgCl₂ concentration) to reduce non-specific PCR amplification.

5.4.3 GBSS4/GBSS6 PCR bands.

New primers were also designed to amplify the GBSS genes, and the primers GBSS4/GBSS6 were used to clone and sequence the 3’ section of the GBSS 7A genes from three varieties ‘Kulin’, ‘Eradu’ and ‘DHW174’. Although these primers were designed to amplify from all three loci, they only amplified from the 7A locus. The reason for this was revealed by comparing the DNA sequences from the three GBSS loci (Murai et al., 1999), which revealed that the DNA base in the position that complemented the 3’ end of the GBSS6 primer was polymorphic between the three loci. The 7A locus has a cytosine residue while the other two loci have adenosine residues. Since the cDNA sequence from which the primers were designed had originated from the 7A locus, the primer matched only this sequence. The GBSS6 primer is locus-specific for the 7A locus. While this allowed easy sequencing of the GBSS 7A gene, it denied access to the other loci and new primers with complementary 3’ ends would need to be developed to amplify from these loci.
Three 7A-alleles from ‘Kulin’, ‘Eradu’ and null 7A were sequenced using GBSS4/GBSS6 primers and each was found to be different, with polymorphisms in the coding and intron regions. Since the 7A-alleles from ‘Kulin’ and ‘Eradu’ contain different (TGCCG)$_n$ microsatellites from intron 4 and their coding regions also differ, the microsatellite is a molecular marker linked to coding sequence differences. However, there was DNA sequence variation between the two different alleles from ‘Eradu’ and ‘DHW174’ that both contain the (TGCCG)$_3$ microsatellite and so the linkage of the microsatellite to the alleles needs further investigation and validation before it is useful for genetic studies.

Two novel wild-type 7A alleles differing from that described by Murai et al. (1999) were identified in the Australian varieties ‘Kulin’ and ‘Eradu’. These novel alleles differed from the published sequence in both DNA and protein sequences showing that there is allelic variation in the Australian wheat germplasm separate from that reported in for ‘Chinese Spring’. The Australian wheat varieties need to be studied in more depth to identify possible different alleles. Though distinctly different from the previously described alleles, the Australian GBSS alleles were shown to be similar enough so that the PCR primers designed to published sequences remain effective tools for characterising Australian wheat varieties. The GBSS4/GBSS6 primers developed in this study represent a molecular tool for characterising alleles in Australian germplasm and should be used to screen additional genotypes for the 7A-alleles.

The first time the null 7A allele was sequenced from ‘DHW174’, it contained a premature stop codon and a different sequence from that of Vrinten et al. (1999). All subsequent sequencing from this allele has matched the published sequence as was expected, since the gene present in Australian varieties originated from parents studied by Vrinten et al. (1999). There are two possible reasons for finding this sequence difference. It may have been a PCR artifact formed by the hybridization of a GBSS sequence with an unrelated sequence. If this were the case,
then the correct PCR fragment should also have been amplified during subsequent PCR cycles. Since all the PCR products were directly sequenced without being cloned, this mixture of correct and hybrid PCR fragments would have presented the signature ‘mixed-template’ DNA sequence with both sequences being present. This was not the case. The DNA sequence was clear and had no contaminating peaks. A second explanation for the truncated sequence was that it was the correct sequence of a pseudogene copy of a GBSS-7A gene. However, no other evidence of GBSS pseudogenes has previously been found, and the truncated sequence has never again been amplified from the genome, but neither has variety ‘DHW174’ been characterized for the presence of pseudogenes.
Chapter 6. PCR-based methods for accessing novel sections of the GBSS gene.

6.1 Introduction.

In Chapter 6, PCR primers designed to GBSS cDNA sequences (Clark et al., 1991) were used to amplify unknown intron regions. However, this approach could not be used to access the promoter and 5’ untranscribed regions of the GBSS genes because no DNA sequence data was available upstream from the promoter. Two methods that could be used to amplify into unknown regions adjacent to the GBSS coding regions are ‘restriction site mediated’ PCR and inverse PCR. These methods combine the use of restriction enzymes to create large random fragments of DNA with specific PCR to amplify those fragments containing the GBSS genes.

Inverse PCR involves cleaving genomic DNA using restriction enzymes and self-ligating these restriction fragments into loops of DNA. The PCR primers are designed to known sequences to amplify outward into unknown neighbouring sequences. When the restriction fragments are circularised, it is possible to amplify the flanking sequences in a loop from primers within the known regions.

Restriction site mediated PCR (Figure 6.1) involves digesting the genome using restriction enzymes, attaching linkers at the restriction sites and allowing PCR amplification between the linkers at the ends of the DNA fragment and a specific PCR primer in the known region. Both of these methods rely on the specificity of the PCR primers to amplify only those fragments containing the gene of interest. As discussed in Chapter 5, PCR primers designed to wheat GBSS genes are not highly specific and readily form homoprimed artifacts. Nevertheless, it was decided to attempt to access the GBSS promoter regions using restriction site mediated PCR.
The ‘known’ primers GBSS2 and GBSS4 were chosen and they amplify fragments between 465bp and 556bp from the GBSS genes. Four restriction enzymes (EcoR1, Xba1, Sal1 and
HindIII) were chosen, which do not digest within this PCR fragment, so the GBSS restriction fragments should be larger than 465bp to 556bp. Based on the known sizes of the EcoR1 restriction fragments containing the GBSS genes (7.5kbp and 8.5kbp, Chapter 4), the restriction fragments could be kilobases in size, and since the positions of the primer binding sites within the restriction fragments were not known, kilobase sized PCR fragments could be amplified. However, any PCR fragments smaller in size than 465bp to 556bp would not be specifically amplified.

6.2 Methods

6.2.1 Construction of linker molecules

Four types of linker molecules were constructed using EcoR1, Xba1, Sal1 and HindIII linkers from an AFLP kit (Perkin-Elmer). A long single-stranded molecule was constructed with the Gus 3’ primer sequence attached adjacent to the sequence of one of the linker strands (Section 2.4). The original linker strand was replaced by the long overhang fragment in 10µl reactions containing 5ng/µl of each molecule and incubating at 94°C for 1min and 50°C for 10min. The resultant molecule had both the long overhanging Gus 3’ sequence and a sticky ended linker to match the restriction site on the DNA template. The remaining single strand from the original linker was not removed.

6.2.2 Preparation of template DNA.

The plasmid template was a pGEM-T vector containing an insert of the GBSS2/GBSS4 PCR product from the GBSS-7A gene of ‘Tincurrin’. The plasmid was attached to the Sal1 linker molecule in 20µl digestion/ligation reactions containing 300ng plasmid template, 100mM NaCl, 1x Ligase Buffer (supplied as 10x concentrate, ), 1/100 BSA (Promega), 6pmol linker molecule, 1mM ATP, 10U Sal1 enzyme (Promega), and 10U T4 DNA Ligase (Biotech International). Reactions were incubated at 37°C for 2hr.
DNA from ‘Kulin’ was digested in a 20μl reaction containing 1μg DNA, 12U restriction enzyme (Promega), 1x Buffer (supplied as 10x buffer, ) and 1/100BSA. The reaction was incubated at 37°C for 2hr and then cleaned by ethanol precipitation as described in Section 2.6. Linkers were attached to the DNA in a 10μl ligation reaction containing 150ng DNA, 6pmol linker, 0.5U T4 DNA Ligase (Biotech International Ltd) and 1x Ligase Buffer (supplied as 10x buffer, Biotech International Ltd). Ligase reactions were incubated overnight at 14°C.

6.2.3 Testing for non-specific amplification.

To test that the linkers had ligated to the plasmid, PCR amplification was performed using the rob26 primer, which bound to the double-stranded section of the linker molecule. The 10μl PCR reactions contained between 30ng and 0.03ng plasmid template, 2mM MgCl₂ and were incubated under the following conditions: 1 hold at 94°C for 2min, 2 cycles each of 94°C for 30sec, 65°C-57°C (-2°C/2cycles) for 30sec and 72°C for 2min, 30 cycles of 94°C for 30sec, 55°C for 30sec and 72°C for 2min, followed by a final hold at 72°C for 4min.

To test for auto-amplification of the wheat template, 10μl PCR reactions (Section 2.4) were prepared containing 20ng of DNA with no primer included and incubated in PCR at a variety of different annealing temperatures. The Gus 3’ primer was investigated for its ability to amplify directly from the wheat template. Ten microliter reactions were prepared containing either 0.6ng or 0.04ng DNA template, 2Mm MgCl₂ and the Gus 3’ primer singly (Section 2.4) and were incubated as follows: 1 hold at 94°C for 3min; 2 cycles of 94°C for 30sec, 58°C for 45sec and 72°C for 1min; 35 cycles of 94°C for 30sec, 55°C for 45sec and 72°C for 1min; and 1 hold at 72°C for 4min.

6.2.4 Restriction site mediated PCR.

Ten microliter reactions were prepared containing 20ng template, 2mM MgCl₂, the Gus 3’ primer (that bound to the linker overhang sequence) and one of GBSS2 or GBSS4. Initially
the reactions were incubated for 1 hold at 94°C for 3min; 35 cycles of 94°C for 30sec, 63°C or 65°C, respectively, for 30sec and 72°C for 3min; and 1 hold at 72°C for 3min.

6.3 Results

6.3.1 Restriction site mediated PCR on plasmid template.

The restriction site mediated PCR technique was initially tested and optimised on a simple DNA template, a pGEM-T vector containing a GBSS insert. The modified plasmid was first tested to see if linkers had successfully been attached, using the rob26 primer, which bound to the double-stranded section of the linker molecule but not within the plasmid sequence. The plasmid was amplified (data not shown) so it must have had linkers attached to each end. Then restriction site mediated PCR was tested using the Gus 3’ primer (that bound to the linker overhang sequence) and one of the specific-primers GBSS2 or GBSS4. It was expected that the two primer combinations would amplify bands of approximately 600bp and 3.6kbp respectively (Figure 6.2). The Gus 3’/GBSS2 primers amplified the expected 600bp fragment, however the 3.6kbp band from the Gus 3’/GBSS4 primers was not seen (Figure 6.2). If the Gus 3’ primers or the contaminating single strand from the linker had amplified by themselves, a band of approximately 3.6kbp would have been in every sample, but this was not seen.

The PCR reactions were optimised further by reducing the DNA template concentration in a 0.1x serial dilution, varying the annealing temperature between 65°C and 55°C and increasing the extension time to 3min. The expected bands were amplified under these conditions (Figure 6.3).
Figure 6.2. (Top) A diagram of the plasmid used as the PCR template to optimise the restriction site mediated PCR method. The plasmid (black line) contained a section of GBSS gene as an insert (black box). The plasmid was linearised with SalI and linkers (red lines) were attached to the restriction sites. The diagram also shows the binding positions of the PCR primers Gus 3’, GBSS4 and GBSS2 (arrows) as well as the predicted sizes of the PCR fragments expected to be amplified by those primers (dashed lines).

(Bottom) A 1% agarose gel showing the restriction site mediated PCR amplification from the plasmid template. The 600bp fragment expected from the Gus 3’/GBSS2 primers is present (Lanes 2, 3), but the 3600bp fragment expected from the Gus 3’/GBSS4 primers is not (Lanes 4, 5). Other unexpected bands were amplified by both primer combinations.

Lane 1: Molecular size marker combination Lambda/HimdIII (Biotech International) and 100bp DNA Ladder; Lanes 2,3: primers Gus/GBSS2; Lanes 4,5: primers Gus/GBSS4, Lane 6: water PCR blank containing all primers.
Reducing the amount of template DNA 300-fold also reduced the number of bands amplified (data not shown). The restriction site mediated PCR strategy worked on the simple template of the plasmid.

**Figure 6.3.** 1% agarose gel showing optimised restriction site mediated PCR amplification from the plasmid template. Both the 600bp fragment expected using the Gus 3'/GBSS2 primers (Lane 6) and the 3600bp fragment expected from the Gus 3'/GBSS4 primers (Lane 4) are present. Other unexpected bands were also seen for both primer combinations. The annealing temperature of the PCR reactions in Lanes 2-5 was 65°C, Lanes 6-7 was 63°C.

Lane 1: Molecular size marker combination Lambda/HindIII (Biotech International) and 100bp DNA Ladder; Lanes 2,6: primers Gus 3'/GBSS2; Lanes 3,7: water control reactions containing primers Gus 3'/GBSS2; Lane 4: primers Gus 3’s/GBSS4; Lane 5: water control reaction containing primers Gus/GBSS4.
6.3.2 Non-specific amplification from the wheat template.

Before attempting to specifically amplify the GBSS sequences from the wheat template, amplification of non-specific PCR fragments was investigated. The modified DNA template itself contained contaminating DNA fragments from the construction of the linker molecule that could potentially act as primers to allow non-specific PCR amplification. The major contaminant was the strand of the original linker molecule that had been replaced by the single stranded overhang. This fragment had not been removed, and since it complemented the linker sequence that was on both ends of every DNA template molecule, there was the potential for it to act as a primer to amplify from every template molecule. Although this had not happened for the plasmid template, it was still examined for the wheat template (Section 6.2.3). No bands were amplified under any conditions (data not shown), so the contaminating linker molecules did not act as primers.

Next the Gus 3’ primer was investigated for its ability to amplify directly from the wheat DNA template (Section 6.2.3). The Gus 3’ primer did not amplify from unmodified wheat DNA (data not shown), and was subsequently tested to see if it would amplify from the DNA with linkers. Since the Gus 3’ primer sequence exactly matched the sequence of the linker overhang, it should not amplify from the DNA template until another primer had produced the complementary strand. However, the Gus 3’ primer did amplify from the modified template (Figure 6.4). This unexpected interaction of the Gus 3’ primer with the template, and possibly the contaminating linker molecules, was not seen using the plasmid template and may therefore have resulted from the relative complexity of the wheat template.
Although the Gus 3’ primer had amplified non-specific bands from the wheat template when used in isolation, the dynamics of a single primer changed in the presence of other primer species (Chapter 5), and so the PCR bands produced using primers GBSS2 or GBSS4 in combination with Gus 3’ were examined. The PCR reactions containing the modified wheat template.

**Figure 6.4.** A 1% agarose gel showing the PCR bands amplified from a template of EcoR1-digested ‘Kulin’ genomic DNA with EcoR1 linkers attached, using only the Gus 3’ primer. As the sequence of the primer was homologous to the sequence of the linker, no amplification should have occurred. The amplification seen showed an unpredicted primer interaction.

Lane 1: Molecular size marker combination Lambda/HindIII (Biotech International Ltd.) and 100bp Ladder; Lane 2: PCR reaction containing 0.6ng template DNA; Lane 3: PCR reaction containing 0.04ng template DNA; Lane 4: water PCR blank.

### 6.3.3 Restriction site mediated PCR amplification from wheat.

Although the Gus 3’ primer had amplified non-specific bands from the wheat template when used in isolation, the dynamics of a single primer changed in the presence of other primer species (Chapter 5), and so the PCR bands produced using primers GBSS2 or GBSS4 in combination with Gus 3’ were examined. The PCR reactions containing the modified wheat template...
template were prepared (Section 6.2.4) using information from the plasmid template model and, since DNA template concentration had affected the banding patterns amplified from the plasmid template, a serial dilution of wheat DNA template concentrations was used (Figure 6.5). Smear PCR bands similar to those produced by the Gus 3’ primer alone were seen in the dual-primed reactions, however reducing the amount of DNA template in the 10µl PCR reactions to below 0.3ng reduced the smearing and allowed amplification of discrete bands. More template DNA resulted in the amplification of many smeared bands, while less template resulted in few or no bands amplified. Many of the bands were smaller than the 465bp to 556bp fragments expected to be the smallest GBSS-specific fragments. The smaller PCR fragments were probably non-specifically amplified. It is also important to note that PCR reactions containing different primers amplified different bands, suggesting that the presence of the specific-primers had changed the conditions to preferentially amplify using those

![Figure 6.5](image.png)

**Figure 6.5.** 1% agarose gel showing PCR bands amplified across a serial dilution of DNA template concentrations. The PCR primers were Gus 3’/GBSS2 (left set) and Gus 3’/GBSS4 (right set). Amounts of DNA template between 0.3ng and 0.03ng per 10µl PCR reaction (Lanes 3, 4) resulted in discrete PCR bands.

Lane 1: Molecular size marker combination Lambda/HindIII (Biotech International Ltd) and 100bp DNA Ladder; Lane 2: 3ng template DNA; Lane 3: 0.3ng template; Lane 4: 30pg template; Lane 5: 3pg template; Lane 6: 0.3pg template; Lane 7: 30fpg template; Lane 8: 3fpg template.
primers. Whilst the banding pattern differences could have been due to the different homoprimed bands from each primer, it is also possible that the interaction of the specific-primer had repressed the non-specific amplification (Chapter 5).

Using the DNA template concentrations that promoted amplification of discrete bands and decreasing the annealing temperature of the PCR cycles to 60°C, the PCR bands shown in Figure 6.6 were produced. Some of the bands were discrete and potentially suitable to clone.

<table>
<thead>
<tr>
<th>Lanes</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular size marker (kbp)</td>
<td>4.4</td>
<td>2.3</td>
<td>2</td>
<td>1</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

**Figure 6.6.** 1% agarose gel showing PCR bands amplified from the wheat genome using restriction site mediated PCR with primers Gus 3’/GBSS2 (Lanes 2-4) and Gus 3’/GBSS4 (Lanes 5-7). Lanes 2/3 and 5/6 show that reducing the amount of template DNA reduced the numbers of PCR fragments amplified. The contamination in the PCR controls (Lanes 4, 7) was subsequently eliminated.

Lane1: Molecular size marker combination Lamda/HindIII (Biotech International Ltd) and 100bp DNA Ladder; Lanes 2, 5: PCR containing 0.3ng template DNA; Lanes 3, 6: PCR containing 0.03ng template DNA; Lanes 4, 7: water PCR blanks.
However, it was not known whether these bands were amplified from the GBSS genes, homoprimed or were from other places in the genome where primers GBSS2 and GBSS4 were bound.

6.3.4 Amplification using locus-specific primers.

To reduce non-specific amplification, the PCR reactions were repeated using the locus-specific primers GBSS-Rev, GBSS-4A, GBSS-7A and GBSS-7D, which bound specifically within the GBSS genes and did not amplify non-specific PCR bands (Figure 6.7). The resulting PCR products should therefore have been amplified specifically from the GBSS genes. Since the GBSS-Rev, 4A, 7A, 7D primers bound approximately 40bp and 90bp away from the GBSS2 and GBSS4 primers respectively, the new PCR products should be smaller than the GBSS-specific GBSS2/GBSS4 bands by these amounts. It was expected that the primer GBSS-Rev would amplify three fragments, one from each locus. However, it only amplified two fragments (Lane 3) and the smaller band was too small to have been amplified from the GBSS genes since it placed a restriction site where none existed. The other three primers were expected to amplify one fragment each for their specific locus. While primer GBSS-7D did amplify one band (Lane 9), the primer GBSS-7A amplified two bands (Lane 7) and primer GBSS-4A (Lane 4) did not amplify at all, possibly because the annealing temperature was too high.

Based on the information gained from the locus-specific primers, the sizes of the corresponding bands amplified by the GBSS2 or GBSS4 primers were predicted and are shown on Figure 6.8. No bands exactly matched the sizes of those predicted. This suggested that most or all the bands seen in Figure 6.8 were not amplified from the GBSS genes, but were either homoprimed by GBSS2 or GBSS4, or were amplified by the primers in combination with Gus 3’ from other parts of the genome.
Figure 6.7. A 1% agarose gel showing PCR bands amplified from a template of SalI-digested ‘Kulin’ wheat genomic DNA with SalI linkers attached, using restriction site mediated PCR with the Gus 3’ primer in combination with one of the locus-specific primers GBSS-Rev, GBSS-4A, GBSS-7A, GBSS-7D. The primer combination Gus 3’/GBSS-Rev (Lane 3) amplified two bands approximately 1.5kbp and 200bp in size. The primer combination Gus 3’/GBSS-4A (Lane 5) did not amplify from the wheat template. The primer combination Gus 3’/GBSS-7A (Lane 7) amplified two bands approximately 1.6kbp and 1.3kbp in size. The primer combination Gus 3’/GBSS-7D (Lane 9) amplified one band approximately 500bp in size.

Lane 1: Molecular size marker combination Lambda/HindIII (Biotech International Ltd.) and 100bp DNA Ladder; Lane 2: water PCR blank with Gus 3’/GBSS-Rev primers; Lane 3: ‘Kulin’ DNA template with Gus 3’/GBSS-Rev primers; Lane 4: water PCR blank with Gus 3’/GBSS-4A primers; Lane 5: ‘Kulin’ DNA template with Gus 3’/GBSS-4A primers; Lane 6: water PCR blank with Gus 3’/GBSS-7A primers; Lane 7: ‘Kulin’ DNA template with Gus 3’/GBSS-7A primers; Lane 8: water PCR blank with Gus 3’/GBSS-7D primers; Lane 9: ‘Kulin’ DNA template with Gus 3’/GBSS-7D primers.
To validate that the bands seen in Figures 6.6 and 6.7 had been amplified from the GBSS genes, the PCR products were cloned into pGEM-T (Section 2.8) and clones were screened for the inserts they contained using PCR with the GBSS-Rev,4A,7A,7D primer combination (Section 2.9). None of the inserts were identified by PCR as coming from the GBSS genes (data not shown).

**Figure 6.8.** 1% agarose gel showing PCR bands amplified from the wheat genome using restriction site mediated PCR with primers Gus 3’/GBSS2 (Lanes 2-4) and Gus 3’/GBSS4 (Lanes 5-7). The positions of the PCR bands predicted by the sizes of the PCR bands amplified by the locus-specific primers are shown with white arrows. Included next to some arrows are the loci from which the band is predicted to have been amplified.

Lane1: Molecular size marker combination Lamda/HindIII (Biotech International Ltd) and 100bp DNA Ladder; Lanes 2, 5: PCR containing 0.3ng template DNA; Lanes 3, 6: PCR containing 0.03ng template DNA; Lanes 4, 7: water PCR blanks.
6.3.5 PCR optimisation for amplifying from wheat.

The PCR bands amplified from the wheat genome using restriction site mediated PCR were not easily reproducible, sometimes being smeared or failing to be amplified at all. This suggested that restriction site mediated PCR of the complex wheat genome was affected by small variations in the PCR conditions. To investigate the effects of various PCR conditions and to optimise the methodology, different PCR conditions were tested.

The plasmid model had shown that reducing the amount of template DNA to between 0.3ng and 0.03ng per 10µl PCR reaction resulted in the amplification of discrete PCR bands and so the amount of wheat template used was reduced to these levels. Also, amplifying the plasmid template using the primers GBSS2 and GBSS4 and annealing temperatures of 63°C and 65°C respectively gave discrete bands. Addressing this second variable using wheat DNA template, the annealing temperature of the PCR cycles was varied between 67°C and 55°C over a range of cycling conditions. As expected, more bands amplified at lower annealing temperatures.

Various ‘touchdown’ PCR protocols were also used, with the first PCR cycles using a high annealing temperature and reducing the annealing temperature by varying degrees every two cycles to 55°C was achieved and annealing the remainder of the cycles at this temperature. Touchdown PCR is used to increase the copy numbers of the fragments the primers anneal to at high temperatures such that the amplification of these more-specific fragments is promoted. Even though the majority of cycles are carried out at the lower annealing temperatures, the higher temperatures used for the first cycles effect the bands amplified as shown in Figure 6.9.

The annealing times of the PCR cycles were varied between 30sec and 45sec to allow more time for the primers to bind to the template, and the extension times of the PCR cycles were increased from 1min up to 3min to allow time for large (>1kbp) fragments to be amplified.
Since the restriction fragments carrying the GBSS genes were large, the PCR fragments amplified from them were also predicted to be at least 1kbp in size.

**Figure 6.9.** A 1% agarose gel showing the different PCR banding patterns after varying the primer concentration and thermal cycling conditions of reactions during restriction site mediated PCR. The DNA template was ‘Kulin’ DNA with EcoR1 linkers and the primers were Gus 3’/GBSS-7D. The gel shows a set of PCR reactions containing various primer concentrations (Lanes 2-6) repeated at different thermal cycling conditions. The set on the left was a touchdown PCR using the annealing temperatures 60°C, 58°C and 55°C, while the set on the right used only the annealing temperatures 58°C and 55°C. There are clear banding pattern differences between the two sets, showing the effects of varying thermal cycling conditions. Lane 2 showed the bands formed at the usual primer concentration of 20pmoles of each primer per 10µl PCR reaction. The effects of increasing the amount of primer two-fold and ten-fold are shown in Lanes 3 and 4, while reducing the amount of primer by these same factors gave no amplification (Lanes 5, 6).

Lane 1: Molecular size marker combination Lamda/HindIII (Biotech International Ltd) and 100bp DNA Ladder; Lane 2: 20pmoles of each primer; Lane 2: 20pmoles of each primer; Lane 3: 40pmoles of each primer; Lane 4: 200pmoles of each primer; Lane 5: 10pmoles of each primer; Lane 6: 2pmoles of each primer; Lane 7: water PCR blank.
The amount of PCR primer was varied between 200pmoles and 2pmoles per 10µl PCR reaction. As shown in Figure 6.9, increasing the amount of each primer above 20pmoles per 10µl PCR reaction increased both the numbers of bands amplified and the primer-primer interactions, while reducing the amount of primer below 20pmoles per reaction resulted in no amplification.

To increase the specificity of the PCR amplification, ‘nested’ PCR was undertaken. This involves enriching the DNA template by amplifying using one set of PCR primers and then using this PCR product as a template for a second round of PCR using a second set of primers that bind within the first PCR product. The first round of PCR would increase the percentage of DNA fragments containing the GBSS genes so that amplification from this GBSS-enriched template would be promoted in the second round of PCR. First the wheat DNA was amplified using either primer GBSS2 or GBSS4 with Gus 3’. This PCR product was then used as template for the second round of PCR amplification using Gus 3’ and the appropriate primer consisting of GBSS-Rev, GBSS-7A or GBSS-7D. The PCR products from the first round of PCR were diluted serially until they could no longer be seen on the gel, reducing the un-amplified ‘Kulin’ DNA template such that the second round of PCR amplification was encouraged to be from the first-round PCR product template.

At most template concentrations (Figure 6.10, Lanes 1-6) a smear of PCR products was seen. When discrete bands were seen (Lanes 4-8) they were usually less than 1kbp in size. The template for the original round of PCR had been EcoR1 restriction fragments about 7.5kbp in size, so the bands amplified by the two sets of primers should have added to 8kbp (based on data from Chapter 4). The small discrete bands seen were probably not amplified from the GBSS genes. The nested PCR method was therefore not effective in increasing the specificity of amplification from the GBSS genes.
Figure 6.10. 1% agarose gels showing the PCR bands produced by nested-PCR at various template DNA concentrations. The original template was ‘Kulin’ DNA with EcoR1 linkers attached. The top gel shows PCR bands amplified using primers Gus 3’/GBSS-Rev at various concentrations of the secondary template DNA that was PCR product amplified using Gus 3’/GBSS2 primers. The bottom gel shows PCR bands amplified using primers Gus 3’/GBSS-7D at various concentrations of the secondary template DNA that was PCR product amplified using Gus 3’/GBSS4 primers.

Lane 1: Molecular size marker combination Lambda/HindIII (Biotech International Ltd) and 100bp DNA Ladder; Lane 2: 30ng template DNA; Lane 3: 6ng template DNA; Lane 4: 3ng template DNA; Lane 5: 0.3ng template DNA; Lane 6: 30pg template DNA; Lane 7: 3pg template DNA; Lane 8: 0.3pg template DNA; Lane 9: 30f template DNA; Lane 10: 3f template DNA; Lane 11: 0.3f template DNA; Lane 12: 0.03f template DNA.
Since PCR fragments kilobases in length were expected, a PCR system Elongase (Gibco) designed to promote amplification of long PCR products was used. Figure 6.11 shows how this PCR system did promote the amplification of many large PCR fragments between 1kbp and 7kbp in size. Most of the PCR fragments were larger than the theoretical minimum size.

**Figure 6.11.** 1% agarose gel showing the PCR bands amplified from a template of ‘Kulin’ DNA with EcoR1 linkers using an Elongase kit (Gibco) and Gus 3’/GBSS-7D primers. Lanes 2 and 3 were duplicates, as were Lanes 4 and 5, and so PCR band differences seen between the repeated samples suggested non-specific amplification. There was 10x as much template DNA in the PCR reactions shown in Lanes 4 and 5 than in Lanes 2 and 3. Reducing the template concentration generated more high molecular weight bands.

Lane 1: Molecular size marker combination Lambda/HindIII (Biotech International Ltd) and 100bp DNA Ladder; Lanes 2, 3: PCR containing 0.1ng DNA template; Lanes 4, 5: PCR containing 1.5ng DNA template; Lane 6: water PCR blank.
for a GBSS-specific PCR fragment (465bp-556bp) suggesting some inhibition of non-specific amplification. However, discrete bands were not seen and there were banding pattern differences between PCR repeats, suggesting that this was non-specific amplification.

6.4 Discussion.

6.4.1 Primer specificity.

The specificity of restriction site mediated PCR depends on the specificity of a single primer, not the interaction of two specific primers as for other types of PCR. This does not matter for simple templates such as the plasmid where the primers bind at only one site. However for complex templates, such as the wheat genome where the primers bind to many sites, it creates non-specific amplification. Here a selection of primers that anneal to the GBSS genes, including GBSS2, GBSS4, GBSS-Rev, GBSS-4A, GBSS-7A and GBSS-7D, were used in combination with the Gus 3’ primer. The specificity of the primers seemed directly related to the number of bands produced in restriction site mediated PCR. The primers GBSS2 and GBSS4 that bind to many sites on the wheat genome (Chapter 5) amplified many more bands than the more-specific primers GBSS-Rev, GBSS-4A, GBSS-7A and GBSS-7D, even under thermal cycling conditions that do not usually produce homoprime bands.

6.4.2 Restriction site prediction based on PCR bands.

Using the GBSS-7A primer, two bands were amplified from the wheat Sal1 restriction fragment template. This suggests two Sal1 restriction sites near each other and an incomplete digestion (Figure 6.12), or that the primers amplified from another part of the wheat genome, either another GBSS locus or a site on another chromosome with no corresponding GBSS-Rev site.
Figure 6.12. Diagrams predicting the positions of the Sal1 restriction sites on the GBSS genes based on the sizes of PCR bands thought to have been amplified from the GBSS 7A and GBSS 7D loci. The top diagram predicts the positions of Sal1 sites on the GBSS 7D locus, while the bottom diagram predicts the positions of Sal1 sites at the GBSS 7A locus. The GBSS gene is shown as a black line with the positions of the Sal1 and primer binding sites marked. The predicted base-pair distances between the sites are shown in blue. The sizes of the PCR bands (Figure 6.7) amplified by one of the locus-specific primers GBSS-Rev, GBSS-7D and GBSS-7A in combination with the Gus 3’ primer that bound to the Sal1 linker are shown in grey. The predicted sizes of the PCR fragments amplified by the primers GBSS2 and GBSS4 in combination with Gus 3’ are shown in red. The sizes of the two restriction fragments were predicted to differ by approximately 102bp.
Based on the sizes of PCR fragments expected to have been amplified from the GBSS 7A and GBSS 7D loci by their locus-specific primers, it was possible to predict the positions of the Sal1 sites within the GBSS genes. Figure 6.12 shows that the size of the Sal1 restriction fragment from the 7D loci was predicted to be approximately 1680bp in size. Since there were two PCR fragments amplified with the GBSS-7A there were two sizes of Sal1 restriction fragment predicted from the GBSS 7A locus; 2442bp and 2742bp in size. These restriction fragment sizes were not confirmed by genomic Southern blot, nor were any of the known restriction fragments described in Chapter 4 able to be amplified by PCR in this way.

6.4.3 GBSS promoter.

To date, there has been little work published on the promoter sequences of the GBSS genes in wheat. Some 5’ untranslated region sequence is available from the Genbank database, approximately 10-30bp for most sequences and 203bp from one GBSS 4A gene (AF286320). The promoter sequence AF286320 was analysed using Angis software (www.angis.org.au) and found to contained short (4 or 5 bp) inverted repeats as well as two microsatellites; an ‘AGG’ motif with four repeats and a ‘CACT’ motif with three repeats. Using the positions of the inverted repeats, predictions about possible secondary structures for this region were made (Figure 13). When comparing this limited amount of sequence from all the published wheat GBSS promoters (data not shown), only one polymorphism seems significant; at position -1 from the start of translation the GBSS 4A genes have a mutant ‘C’>’G’ substitution that may hold transcriptional or translational significance. More 5’ untranslated sequence is available from the GBSS genes of rice and barley, but they show little sequence similarity to each other or to the wheat promoter sequences.
6.4.4 Conclusions.

Having investigated various elements of the restriction site mediated PCR technique on both a simple plasmid template and the more complex template of the wheat genome, it was decided that while this technique was effective on the simpler plasmid template, it was not robust
enough to allow specific amplification from more complex templates like the wheat genome.

It was more sensitive to variations in the PCR conditions than traditional PCR. Both these statements were supported by our colleagues who had also used similar methods (J. Kretschmer, WAITE Adelaide, pers. comm.).

Since the restriction site mediated PCR method had not succeeded in amplifying the GBSS promoter sequences from the wheat genome, it was decided not to attempt the similar method, inverse PCR. It may not be effective to attempt to use these methods to amplify from the wheat genome, unless very highly specific PCR primers are used.
Chapter 7. Linking GBSS molecular markers to GBSS alleles and starch quality traits.

7.1 Introduction.

All the PCR primers developed in this project were positioned within the coding regions of the GBSS gene with the most 3’ primer, GBSS6, positioned at the stop codon. The cDNA (Clark et al., 1991) to which the primers were designed contained an additional 257bp of 3’ untranslated region before the poly-A tail and Shariflou and Sharp (1999) developed a molecular marker in this region. The primers amplify from the end of exon ten, across intron ten, through exon eleven and into the 3’ untranslated region containing an imperfect (AT)$_n$ microsatellite locus, Xsun1 (Figure 7.1). Microsatellite variation at the Xsun1 locus is linked, by proximity, to GBSS allele variation, providing a potential molecular marker to the GBSS genes.

The molecular marker can be used to distinguish null and wild-type GBSS alleles at all three loci and is therefore valuable for use in the DAWA waxy wheat breeding projects aiming to produce novel wheats containing various combinations of null and wild-type GBSS genes. As part of this project, the Xsun1 molecular marker was obtained, optimized for use in the WA SABC wheat genotyping laboratories, and used to identify the GBSS status of a breeding population of GBSS-variable, partially-waxy, doubled-haploid wheat individuals. This was done to gain access to this important new molecular marker for use in Western Australia’s waxy-wheat breeding program, to support the collaborative relationship between breeders at DAWA and the genetics lab at SABC, to try to link the XSun1 variation to starch quality in noodle wheat varieties, and also to link the XSun1 marker to the (TGCCG)$_n$ marker that had been discovered in this project.
The nature of the collaboration between wheat breeders at DAWA and the DNA-testing service at SABC was that wheat breeders from DAWA could provide leaf tissue samples from their populations, these would be genotyped at the SABC, and results sent back to DAWA for analysis. Due to the service-provider role of the SABC, very little information about the parentage, growing conditions or phenotypes of the plants was available to this study. Choices about wheat crosses were made by wheat breeders at DAWA without the input of staff at the SABC, and FSV testing was carried out at DAWA laboratories, using their own techniques.

![Figure 7.1](data:image/png;base64,iVBORw0KGgoAAAANSUhEUgAAAAEAAAABAQMAAABJO7e3AAAABGdBTUEAALGPC/xhBQAAAABlBMVEX///8AAABVwtJUwADJQQAAAABJRU5ErkJggg==)

**Figure 7.1.** Two DNA sequences, the genomic (ChSpr-7A) and cDNA (ChSpr-cDNA) sequences from ‘Chinese Spring’ wheat (Murai et al., 1999; Clark et al., 1999) showing the positions of the microsatellite marker primers WxD1b1F1, Sun1F and Wx97R2/Sun1R (red). The primer WxD1b1F1 binds at the end of intron 10 adjacent to the intron, Sun1F binds at the stop-codon, while the primer Wx97R2 or Sun1R binds within the 3’ untranslated region 63bp upstream from the poly ‘A’ tail of the cDNA. The primers were designed around the imperfect (AT)n microsatellite Xsun1 shown in dark blue.
Unfortunately this resulted in the data set for this current study being incomplete. However, this study aimed to show the potential uses of molecular markers to assess genotypes in a wheat breeding program, and compare them to very preliminary phenotypic data, rather than to be a comprehensive study.

7.2 Methods.

The $\text{Xsun1}$ marker was analysed using two different sets of conditions. The first set of conditions were as described by Shariflou and Sharp (1999) using the Sun1F ($5'\text{-CGCTCCCTGAAGAGAGAAAGAA-3'}$) and Sun1R ($5\text{-ATAGGCACAACCCCTAAC-3'}$) primers and Biotech International reagents, with 15ul of PCR product being electrophoresed on 10% polyacrylamide in TBE gels for 1280Vhrs staining with ethidium bromide (Sections 2.4, 2.5).

The second method for accessing the $\text{Xsun1}$ marker was suggested by M. Shariflou (pers. comm.) and then further optimized. The forward primer was WxD1b1F1 ($5'\text{-ACAGGATCTCTCCTGGAAG-3'}$) used with the Sun1R primer (also called Wx97R2 by Shariflou and Sharp). The original PCR reaction conditions supplied by M. Shariflou contained 5µM each primer, 200µM dNTPs, 1.5µM MgCl$_2$, 1x AB buffer, 2U Taq polymerase and 100ng DNA template in a 10µl volume. These conditions were optimised here to reduce the amounts of reagents required and were radioisotope-labelled to allow visualisation on sequencing gels. The optimised reaction contained 1x polymerisation buffer (supplied as 5X polymerisation buffer containing 67mM Tris-HCl pH 8.8, 16.6mM (NH$_4$)$_2$SO$_4$, 0.45% Triton X-100, 0.2mg/ml gelatin, 0.2mM dNTPs, Biotech International), 50µM $\alpha$-P$^{33}$ dCTP, 0.8U Tth$^+$ polymerase (Biotech International), 40pmoles of each primer, 1.5mM MgCl$_2$ and 4µl DNA template in a 20µl volume.
Reactions were incubated under the following conditions: 1 hold at 95°C for 3min; 14 cycles of 94°C for 1min, 64°C to 58°C at -1°C/2 cycles for 1min, 72°C for 1min; 25 cycles of 94°C for 1min, 57°C for 1min, 72°C for 30sec; 1 hold at 72°C for 5min. An equal volume of denaturing loading buffer (Sambrook et al., 1989) was added to each sample which was then denatured at 95°C for 5min and immediately transferred to ice before loading onto the gels.

Several types of gel were tested to obtain the necessary resolution for distinguishing the microsatellite bands, including 1% agarose in 1x TAE, 6% polyacrylamide in 1x TBE and denaturing polyacrylamide containing 8M urea (Section 2.5). The gels that resolved the microsatellite bands most clearly were sequencing polyacrylamide gels containing 6% 29:1 acrylamide:bisacrylamide (Biorad), 8M urea, 1/1000 APS, TEMED and 1x TBE buffer (90 mM Tris-borate, 2 mM EDTA) (Sambrook et al., 1989). Gels were electrophoresed in Protean II electrophoresis equipment (Bio-Rad Laboratories) with 0.4 mm spacers. Gels were electrophoresed at 100V for 4hrs, cut to size using a scalpel, blotted onto Watman paper and dried in a gel drier. Dried gels were placed next to film (Fuji) for 72hrs and then developed.

For the initial parts of this study, a selection of wheat varieties were tested and these varieties were described in Chapter 3. Also, the three nullisomic/tetrasomic ‘Chinese Spring’ varieties lacking the 4A, 7A and 7D chromosomes, respectively, were used. In addition, offspring from two, doubled-haploid breeding crosses were tested and leaf tissue from these plants was supplied by R. Wilson (WADA). The crosses were the waxy wheat ‘DHWX-174’ crossed twice with ‘Arrino’ which contains the null 4A allele and wild-type alleles at the 7A and 7D loci. The offspring were doubled haploids. No further information was supplied about the methods used to produce the cross or grow the offspring, nor was an analysis of the potential grain and quality yields for the growing season provided. DNA was extracted from these plants in collaboration with H. Li (Western Australian State Agricultural Biotechnology Centre, SABC). DNA was extracted from leaf tissue using the ‘Quick’ DNA extraction
protocol as described in Section 2.2.3. The DNA samples were diluted 1/100 before use in the PCR reactions. The PCR reactions were as described above using the WxD1b1F1/ Sun1R primers. Flour Swelling Volume data was supplied by A. Briney and M. Carter (SABC) and R. Wilson (WADA). The FSV testing was carried out at DAWA using their own methods and those methods were not supplied.

7.3 Results.

7.3.1 Comparing the banding patterns of the Xsun1 marker using two methodologies.

The two methods for accessing the XSun1 locus were used on a selection of wheat varieties containing different combinations of GBSS alleles. One of the methods was the one described in the original report for the XSun1 marker (Shariflou and Sharp, 1999), and the other method

![Figure 7.2](image.png)

**Figure 7.2.** A 10% polyacrylamide gel showing the Xsun1 microsatellite marker bands amplified from samples of wheat DNA by the primers Sun1F/Sun1R. Different sized bands were amplified from the wild-type 7A, null 7A and wild-type 7D GBSS alleles as marked to the right of the gel. The presence of the null 7D allele did not produce a band. The primers did not amplify from the 4A locus.

Lane 1: Molecular size marker (100bp DNA Ladder, Promega); Lanes 2-6: five samples of mixed wheat DNA containing various combinations of GBSS alleles. The DNA from the null 7A band had come from ‘DHWX-174’, the DNA from the wild-type bands had come from ‘Kulin’, and the wild-type 7D bands in lanes 3 and 6 are from nullisomic 7A/tetrasomic 7D. These DNA samples were mixed to give the combinations of bands.
was provided directly to us from the Authors. Figure 7.2 shows the PCR bands amplified using the Sun1F/Sun1R primers, while Figure 7.3 shows examples of the microsatellite bands.

![Figure 7.3](image)

**Figure 7.3.** Three autoradiograph pictures of 6% denaturing polyacrylamide gels displaying the microsatellite marker bands amplified from a selection of wheat varieties by the primers WxD1b1F1/ Sun1R. Different sized bands were amplified from the three GBSS loci as marked to the left of the gels. The band amplified from the GBSS 7A locus was approximately 475bp in size (e.g. Lanes 4, 8, 11, 14) and displayed the typical ‘stuttering’ shown by microsatellite bands. The bands from the 4A and 7D loci were similar in size at approximately 450bp. The 4A band was slightly larger in size (e.g. Lanes 4, 5, 7) while the 7D band was the smallest (e.g. Lanes 2, 3, 13). The 4A and 7D bands were typically double bands (Lanes 2-5) though sometimes a triple band was amplified from a single locus (Lane 6).

When null GBSS alleles were present in the varieties (listed with the Lane guide), no bands were amplified by the microsatellite primers. The presence of the null 4A allele in the variety resulted in the lack of the 4A band (Lanes 2, 3, 10), just as the lack of the 7D band (Lanes 4, 11, 14) denoted the null 7D status of these varieties. The null 7A allele resulted in a lack of amplification from that locus (Lanes 3, 5, 12, 13), however the amplification of the 7A band was less reliable than the amplification of the other two bands so the primers did not amplify from some wild-type 7A alleles (Lanes 2, 6, 7, 8, 9).

Lane 1: Molecular size marker (100bp DNA Ladder, Promega); Lane 2: ‘Arrino’ (null 4A); Lane 3: ‘Sharp 22’ (null 4A, null 7A); Lane 4: Chinese Spring nullisomic 7D tetrasomic 7A; Lane 5: Chinese Spring nullisomic 7A tetrasomic 7D; Lane 6: Chinese Spring nullisomic 4A tetrasomic 4D; Lane 7: ‘Amery’; Lane 8: ‘Trident’ (null 4A); Lane 9: 81W:1138 (null 4A); Lanes 10-15: doubled-haploid progeny from the DHWX-174//2*Arrino cross (all null 4A, with various null and normal alleles at 7A and 7D..
amplified by the WxD1b1F1/Sun1R primers. The nullisomic/tetrasomic lines have been included on Figure 7.3 to identify which locus each band had been amplified from, and it would have been ideal to have included these samples on all the gels shown in this Chapter to confirm the assignments of these bands to the GBSS loci. Both methods were successfully used to analyse the GBSS alleles present in a range of wheat varieties, though the Sun1F/Sun1R primers could not be used to access alleles at the GBSS 4A locus.

7.3.2 Xsun1 allele variation linked to noodle quality traits.

Shariflou and Sharp (1999) had reported size variation of the marker bands from the Xsun1-7A locus suggesting variations in the numbers of Xsun1 microsatellite repeats present in the different wheat varieties. This was also noticed in the current project. Figure 7.4 shows an example of three different sized marker bands amplified from 7A loci of a selection of wheat varieties containing wild-type 7A genes using the WxD1b1F1/Sun1R primers.

![Figure 7.4](image)

**Figure 7.4.** An autoradiograph picture of a 6% denaturing polyacrylamide gel showing the three differently sized bands amplified from wild-type 7A GBSS genes. The upper bands were amplified from the 7A locus and the lower band was amplified from the 7D locus as labeled. Each Lane contains bands amplified from a different wheat variety, from left to right, two repeats each of: Eradu (null 4A), Cadoux (null 4A), Reeves (null 4A), Halberd (null 4A), Machete (null 4A, one repeat only), Tincurrin (one repeat only), Trident (null 4A), Kulin, Janz, Corrigin, and Amery. The Lanes have been numbered based on which of the three differently sized 7A bands they contained.
Since the $X_{sun1}$ microsatellite locus is located close to the GBSS coding regions, the two are closely linked and allelic differences at the GBSS-7A locus may be able to be associated with size polymorphism at the $X_{sun1}$-7A locus. Variations in FSV starch quality between WA noodle wheat varieties possibly result from allelic variation at the 7A locus and so the $X_{sun1}$-7A microsatellite bands from a small selection of null 4A wheat varieties with both High-FSV and Medium-FSV were compared (Figure 7.5). There did not seem to be linkage between the band sizes and the FSV values for the starch from those varieties. The $X_{sun1}$ microsatellite marker was found not to be a marker for the High/Medium FSV characteristics studied.

![Figure 7.5](image)

**Figure 7.5.** An autoradiograph picture of a 6% denaturing polyacrylamide gel comparing sizes of bands amplified from the 7A locus of wheat varieties with High-FSV (H) and Medium-FSV (M) starch characteristics. Each Lane contains the microsatellite bands amplified from a different variety. Though all the varieties with High-FSV values showed the same 7A microsatellite band, the varieties with Medium-FSV values showed both this band and a second, larger band. The 7A microsatellite bands were not therefore linked to the wheat FSV values.

Lanes H: High FSV varieties; Lanes M: Medium FSV varieties. Lanes contain (from left to right): Eradu (two repeats), Cadoux (two repeats), Reeves (two repeats), Halberd (two repeats), Trident, 81W:1138, Machete. All varieties are null 4A.
7.3.3 Linkage between the Xsun1-7A locus and the (TGCCG)n microsatellite locus.

The different Xsun1-7A alleles were also analysed for their linkage to the (TGCCG)n microsatellite alleles from intron 4 of the GBSS-7A genes (Chapter 3). Since the (TGCCG)n microsatellite is located within intron 4 of the GBSS gene and the Xsun1 locus is located 120bp 3’ of the GBSS stop-codon, both are closely linked to the GBSS coding regions and therefore to each other. Table 7.1 lists the wheat varieties about which the molecular

<table>
<thead>
<tr>
<th>Wheat name</th>
<th>(TGCCG)n allele repeat number</th>
<th>Xsun1 allele size (bp)</th>
<th>Wheat name</th>
<th>(TGCCG)n allele repeat number</th>
<th>Xsun1 allele size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rowan</td>
<td>3</td>
<td>260</td>
<td>Frame</td>
<td>4</td>
<td>268</td>
</tr>
<tr>
<td>Pelsart</td>
<td>3</td>
<td>265</td>
<td>Aroona</td>
<td>4</td>
<td>270</td>
</tr>
<tr>
<td>Tincurrun</td>
<td>3</td>
<td>270</td>
<td>Schomburg</td>
<td>4</td>
<td>270</td>
</tr>
<tr>
<td>Spear</td>
<td>3</td>
<td>272</td>
<td>Barunga</td>
<td>4</td>
<td>272</td>
</tr>
<tr>
<td>Condor</td>
<td>3</td>
<td>275</td>
<td>Molineux</td>
<td>4</td>
<td>272</td>
</tr>
<tr>
<td>Cunningham</td>
<td>3</td>
<td>275</td>
<td>Cook</td>
<td>4</td>
<td>275</td>
</tr>
<tr>
<td>Dollarbird</td>
<td>3</td>
<td>275</td>
<td>Gamenya</td>
<td>4</td>
<td>275</td>
</tr>
<tr>
<td>Eradu</td>
<td>3</td>
<td>275</td>
<td>Kulin</td>
<td>4</td>
<td>275</td>
</tr>
<tr>
<td>Halberd</td>
<td>3</td>
<td>275</td>
<td>Machete</td>
<td>4</td>
<td>275</td>
</tr>
<tr>
<td>Hartog</td>
<td>3</td>
<td>275</td>
<td>Reeves</td>
<td>4</td>
<td>275</td>
</tr>
<tr>
<td>Janz</td>
<td>3</td>
<td>275</td>
<td>Trident</td>
<td>4</td>
<td>275</td>
</tr>
<tr>
<td>Sunco</td>
<td>3</td>
<td>275</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suneca</td>
<td>3</td>
<td>275</td>
<td>Batavia</td>
<td>mixed 3/4</td>
<td>mixed 270/275</td>
</tr>
<tr>
<td>Tasman</td>
<td>3</td>
<td>275</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kite</td>
<td>3</td>
<td>278</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cranbrook</td>
<td>3</td>
<td>280</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excalibur</td>
<td>3</td>
<td>285</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.1. Table listing the wheat varieties for which molecular marker data is known at both the (TGCCG)n and Xsun1-7A microsatellite loci. Using the Xsun1-7A data, the (TGCCG)n wheat groupings can be further separated.
information from both the (TGCCG)$_n$ and \textbf{Xsun1-7A} loci has been gathered. About half of the varieties are null 4A and were developed in the noodle-wheat breeding program with ‘Gamenya’ as a parent. Most of these varieties are unsuitable for Udon noodle production due to medium-FSV. The selection of varieties also includes non-noodle wheat varieties, and non-Western Australian varieties. The (TGCCG)$_n$ locus data was collected in this project, and the \textbf{Xsun1-7A} data was supplied by M. Shariflou (pers. comm.).

Table 7.1 shows how the \textbf{Xsun1-7A} data was used to further classify GBSS-7A alleles that had been previously separated on the basis of the number of (TGCCG)$_n$ microsatellite repeats they possessed. The reverse was also true. Some groups of wheat varieties classified into a single allele grouping by the size of the \textbf{Xsun1-7A} band they produced could now be separated by the number of (TGCCG)$_n$ microsatellite repeats they possessed (Figure 7.6). Using these two molecular marker loci, the wheat varieties tested could be separated into 12 allelic groups.

**Figure 7.6.** Every coloured rectangle on the graph represents one wheat variety about which the \textbf{Xsun1} and (TGCCG)$_n$ microsatellite data is known. Varieties have been categorised first by the size of the \textbf{Xsun1} allele and then into groups with either 3 or 4 (TGCCG)$_n$ microsatellite repeats. Three of the \textbf{Xsun1} allele groups could be further sectioned by the differing numbers of (TGCCG)$_n$ microsatellite repeats they contained.
The GBSS-7A alleles classified by the (TGCCG)$_n$ and Xsun1-7A data were also compared to Single Nucleotide Polymorphism data from the centre of the GBSS-7A genes discussed in Chapter 3. From the wheat varieties tested, no SNPs were seen to be linked to the GBSS-7A alleles classified.

Using molecular data from both the Xsun1-7A and (TGCCG)$_n$ loci, the linkage between the molecular markers and High-FSV or medium-FSV starch quality traits shown by ASWN wheat varieties was investigated. Table 7.2 shows that no combination of markers could be linked to the different starch qualities.

<table>
<thead>
<tr>
<th>Wheat variety displaying high-FSV starch characteristics.</th>
<th>Molecular data about the (TGCCG)$_n$/Xsun1-7A loci.</th>
<th>Wheat variety displaying medium-FSV starch characteristics.</th>
<th>Molecular data about the (TGCCG)$_n$/Xsun1-7A loci.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eradu</td>
<td>3/275</td>
<td>Halberd</td>
<td>3/275</td>
</tr>
<tr>
<td>Gamenya</td>
<td>4/275</td>
<td>Machete</td>
<td>4/275</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reeves</td>
<td>4/275</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trident</td>
<td>4/275</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aroona</td>
<td>4/270</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spear</td>
<td>3/272</td>
</tr>
</tbody>
</table>

7.3.4 Use of the Xsun1 molecular marker to identify the GBSS characteristics of a noodle-wheat breeding population.

Towards understanding the differences in wheat starch quality associated with various combinations of GBSS null and wild-type genes, the Xsun1 marker was used to genotype two of the most recent, partially-waxy, wheat breeding populations produced by DAWA. Both
populations were progeny from ‘DHW174’ (triple null) crossed twice to ‘Arrino’ (null 4A),
and then passed through the doubled-haploid procedure to be made homozygous at all loci.
The progeny would all be null 4A, but have combinations of null and wild-type alleles at the
7A and 7D loci. This study was a collaboration between wheat breeders at DAWA and the
genotyping facility at the WA SABC. The progeny from this cross were produced at DAWA,
were to be genotyped at the SABC and have FSV testing done by DAWA. It was hoped to
find distinct starch quality differences between wheats containing the four combinations of
null and wild-type GBSS 7A and 7D alleles.

The first cross yielded a population of 186 individuals. The members of the population were
genotyped and individuals containing all 4 combinations of GBSS-7A and GBSS-7D null and
wild-type alleles were seen (Figure 7.7). The numbers of each genotype are listed in Table
7.3. About double the expected numbers of null 4A, null 7A, wild-type 7D individuals were
seen, but only 80% of the expected numbers of null 4A, wild-type 7A, wild-type 7D offspring
were seen. During the breeding process, some offspring are lost because they fail to sprout or
die in the field. Also, some individuals are lost in the process of producing doubled-haploids.

<table>
<thead>
<tr>
<th>4A</th>
<th>7A</th>
<th>7D</th>
<th>% expected offspring</th>
<th>% actual offspring</th>
</tr>
</thead>
<tbody>
<tr>
<td>null</td>
<td>null</td>
<td>null</td>
<td>12.5</td>
<td>13</td>
</tr>
<tr>
<td>null</td>
<td>null</td>
<td>wild-type</td>
<td>12.5</td>
<td>25</td>
</tr>
<tr>
<td>null</td>
<td>wild-type</td>
<td>null</td>
<td>12.5</td>
<td>13</td>
</tr>
<tr>
<td>null</td>
<td>wild-type</td>
<td>wild-type</td>
<td>62.5</td>
<td>49</td>
</tr>
</tbody>
</table>

Unfortunately the starch quality testing of individuals in this first cross had not been
completed by our collaborators at DAWA before the end of this project, and the FSV data for
these individuals was not available at the time of publication of this report.
The progeny from the second wheat cross (WADA Cross code 98Y215, DHWX-174/2*Arrino) was also tested for their GBSS alleles, this time using the Sun1F/Sun1R primers. This cross yielded fewer offspring, 69 individuals, only one of which was a waxy genotype. The remaining individuals all contained the wild-type 7D allele and approximately 50% of these contained the wild-type 7A allele, see Table 7.4. The numbers of individuals of each genotype was very different from that expected. As mentioned previously, some progeny from a cross are lost during the breeding process. However, the ratios of surviving progeny from this cross seem especially different from that expected, and seem to suggest a strong

![Figure 7.7. Autoradiograph of a 6% denaturing polyacrylamide gel displaying the Xsun1 microsatellite marker bands amplified from a selection of doubled-haploid wheat progeny from the cross DHWX-174//2*Arrino, containing the four combinations of null and wild-type 7A and 7D alleles by the primers WxD1b1F1/ Sun1R. The band amplified from each allele and its approximate size is marked to the left of the autoradiograph.](image)

Lane 1: plant contains wild-type 7A and wild-type 7D alleles; Lane 2: plant contains wild-type 7A and null 7D alleles; Lane 3: plant contains null 7A and wild-type 7D alleles; Lane 4: plant contains null 7A and null 7D alleles; Lane 5: water PCR blank.
selective pressure to contain the wild-type 7D allele, but little to contain the wild-type 7A allele. Perhaps the environmental conditions during this growing season were not ideal, resulting in many progeny dying, or not all the progeny were taken through to doubled-haploids.

<table>
<thead>
<tr>
<th>4A</th>
<th>7A</th>
<th>7D</th>
<th>% expected offspring</th>
<th>% actual offspring</th>
</tr>
</thead>
<tbody>
<tr>
<td>null</td>
<td>null</td>
<td>null</td>
<td>12.5</td>
<td>1.5</td>
</tr>
<tr>
<td>null</td>
<td>null</td>
<td>wild-type</td>
<td>12.5</td>
<td>55.2</td>
</tr>
<tr>
<td>null</td>
<td>wild-type</td>
<td>null</td>
<td>12.5</td>
<td>0</td>
</tr>
<tr>
<td>null</td>
<td>wild-type</td>
<td>wild-type</td>
<td>62.5</td>
<td>43.3</td>
</tr>
</tbody>
</table>

FSV starch quality information was available for some of these plants (Figure 7.8). The average FSV for a plant containing wild-type alleles at both 7A and 7D loci was 18.2 (+/- 2.9) ml/g and the average for a plant containing the wild-type 7D and null 7A alleles was 20.8 (+/- 4.0) ml/g. Although the difference between these two FSV was not statistically significant due to the small number of individuals in the study, the trend for the presence of the null 7A allele to produce a higher FSV is an exciting preliminary result and is just what was expected.

A wheat variety with no null alleles, such as ‘Kulin’, has a low FSV of about 13.7 ml/g (A. Briney, WA SABC, pers. comm.). A null 4A variety with medium FSV, such as ‘Reeves’, has an FSV of about 17.5 ml/g. A null 4A, noodle-wheat variety with high-FSV, such as ‘Arrino’ has an FSV of 18.3 ml/g (Figure 7.8). The doubled-haploid progeny with the same GBSS alleles as ‘Arrino’ had an average FSV of 18.2 ml/g. For the doubled haploid progeny with only 2 functional copies of the GBSS genes, at the 7D locus, the average FSV was 20.8 ml/g. The waxy parent ‘DHWX-174’ had an FSV of 43.2 ml/g. Therefore the 2.6ml/g difference in FSV between the progeny with wild-type 7A alleles and with null 7A alleles
<table>
<thead>
<tr>
<th>Cross/ Seln Number</th>
<th>FSV Mge</th>
<th>7D allele</th>
<th>7A allele</th>
</tr>
</thead>
<tbody>
<tr>
<td>98Y215-D2-001</td>
<td>19.2</td>
<td>wild-type</td>
<td>wild-type</td>
</tr>
<tr>
<td>98Y215-D3-002</td>
<td>22.1</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D2-003</td>
<td></td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D2-004</td>
<td>21.5</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D2-005</td>
<td>21.8</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D2-006</td>
<td>20.8</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-007</td>
<td>17.8</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-008</td>
<td>19.8</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-010</td>
<td>21.5</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-012</td>
<td></td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-014</td>
<td>21.1</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-015</td>
<td>17.8</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>DHWX-2-001</td>
<td>42.8</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>Arrino</td>
<td>18.3</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-016</td>
<td></td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-017</td>
<td></td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-018</td>
<td>21.5</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-020</td>
<td></td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-021</td>
<td>21.8</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-022</td>
<td></td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-024</td>
<td></td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-025</td>
<td></td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-026</td>
<td>18.2</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-027</td>
<td>20.5</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-029</td>
<td></td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-030</td>
<td></td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-031</td>
<td>20.8</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-032</td>
<td>21.5</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-034</td>
<td>19.8</td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-035</td>
<td></td>
<td>wild-type</td>
<td>null</td>
</tr>
<tr>
<td>98Y215-D3-036</td>
<td>17.7</td>
<td>type</td>
<td>null</td>
</tr>
</tbody>
</table>

**Figure 7.8.** Table listing all the offspring in the Cross 98Y215, their identifying numbers, FSV scores, and GBSS characteristics. For some individuals there was no FSV data available.
shows a trend in the expected direction and may indicate a real relationship between the GBSS alleles present and the FSV they produce. Larger crosses and crosses that take into account different genetic environments, different growing seasons and different field locations will be needed to fully examine the phenotypic effects of these combinations of GBSS genes.

7.4 Discussion.

7.4.1 Predicted structural differences in the 3’ untranslated region of the GBSS genes based on Xsun1 molecular marker patterns.

Using the Xsun1 molecular marker in this project and the data reported by Shariflou and Sharp (1999), it was possible to predict the molecular structure of the 3’ untranslated region of the GBSS genes. These predictions could have been tested by cloning and sequencing the Xsun1 microsatellite bands. However since this sequencing was being carried out by our collaborator M. Reza Shariflou (pers. comm.), who had requested that we not study this region of the GBSS genes, to support our collaborative relationship, this sequencing was not done in this project. Therefore, only predictions about the structure of the 3’ untranslated regions of the GBSS genes can be made until these regions are actually sequenced.

The primers Sun1F/Sun1R amplified a 204bp PCR fragment from the wild-type 7D allele, as reported by Shariflou and Sharp (1999) and confirmed in this project. This sized fragment exactly matched the size of the 3’ untranslated region of the GBSS cDNA (Clark et al., 1991) which was from the 7A locus (Figure 7.9). Since the cDNA had originated from the 7A locus, it is predicted that the 3’ untranslated region of the wild-type 7D allele closely matches that found in the ‘Chinese Spring’ wild-type 7A allele, from which the cDNA was copied. No fragment was amplified from the null 7D allele in this study, because of the 3’-truncation of this allele and the deletion of the primer binding sites.
Figure 7.9. DNA sequences of the 3’ ends of the GBSS cDNA (Clark et al., 1991) and from the wild-type alleles at each locus (Murai et al., 1999). Positions of primers have been highlighted in red, the stop-codon in grey and the Xsun1 microsatellite in blue. Each line of sequence contains 50 nucleotides in 10-nucleotide blocks.
From the sequence of the GBSS cDNA, originally transcribed from a ‘Chinese Spring’ wild-type 7A allele, the Sun1F/Sun1R primers were predicted to amplify a 204bp PCR fragment from 7A alleles. However both we, and Shariflou and Sharp (1999), reported a 265bp fragment amplified from the wild-type 7A alleles of Australian varieties and ‘Chinese Spring’. This suggests the presence of a 61bp insert in the 3’ untranslated region of these varieties. Unless the cDNA sequenced by Clark et al. (1991) had been post-transcriptionally modified, this 61bp size difference is a marker for the GBSS 7A allele present in the ‘Chinese Spring’ variety used by Clark et al. (1991). The insertion of 61bp into the Australian varieties may reflect microsatellite polymorphism at the Xsun1 locus. However it is also possible that the 3’ region of the sequenced cDNA was chimerical, and had originated from the 7D locus, as proposed by Shariflou and Sharp (1999). The Sun1F/Sun1R primers amplified a smaller band of approximately 240bp from the null 7A allele. This suggests fewer Xsun1 microsatellite repeats compared to the wild-type 7A alleles.

No band was amplified from the wild-type 4A allele by the Sun1F/Sun1R primers due to DNA sequence polymorphism at the 4A locus within the binding site of the Sun1F primer at the stop-codon (Figure 7.9). The polymorphism occurs after the stop-codon and thus does not affect the protein, but interrupts primer-binding preventing amplification from this locus.

Shariflou and Sharp (1999) reported that the Sun1EF/Sun1R primers amplified 362bp fragments from the 4A and 7D loci, and a 420bp fragment from the 7A locus. This suggested that the length of the 3’ untranslated region of the GBSS 4A allele was similar to that of the 7D allele, with the 7A allele being larger. Comparing the GBSS sequences presented in Figure 7.9, it would seem impossible for those primers to amplify the sized bands reported. However, removing the length of intron 10 from the sequences (i.e. using the cDNA sequence) gives the sizes of bands reported and therefore it is suggested that the Shariflou and Sharp (1999) report contains an error.
In this project the WxD1b1F1/Sun1R primers amplified three differently-sized bands from the three GBSS loci. Since this PCR fragment spanned intron 10, some of the size variation was due to polymorphism in this intron. The PCR fragments amplified from the wild-type alleles at the 4A and 7D loci were very close in size, at approximately 450bp. If the 3’ untranslated regions of these two alleles are similar to that reported for the cDNA (Clark et al., 1991), then the PCR fragment from the 7D allele would be 446bp in size, with the band from the 4A allele being 445bp in size. These predicted sizes are similar to the actual sizes of PCR bands amplified from these loci, except that the band amplified from the 4A locus was larger than that from the 7D locus suggesting more microsatellite repeats at the Xsun1 locus of the 4A gene.

From the wild-type 7A allele, the WxD1b1F1/Sun1R primers amplified a band approximately 475bp in size. The 7A allele has 61bp more 3’ untranslated region (Shariflou and Sharp, 1999) and 22bp less of intron 10 (Murai et al., 1999) than the 7D allele, and so the WxD1b1F1/Sun1R PCR fragment from the 7A allele is approximately 40bp larger than that from the 7D allele.

The WxD1b1F1/Sun1R primers did not amplify from any of the three null GBSS alleles. The null 4A allele results from a deletion and the null 7D allele is truncated at the 3’ end, so the primers could not bind to these alleles. However, since both the WxD1b1F1 and Sun1R sites exist in the null 7A allele, the reason for the lack of amplification from this allele was not discovered.

### 7.4.2 Allelic groups at the 7A locus support other genetic studies.

The 12 allelic groups into which the two molecular markers Xsun1 and (TGCCG)n separated the wheat varieties supported the genetic relationships between these varieties as described by Paull et al. (1998). They had investigated the genotypes of 124 Australian wheat accessions.
using RFLPs selected to give uniform genome coverage. The 124 accessions had been divided into 4 Groups. Group 1 was a closely linked group containing 12 members. Four of these wheats (Reeves, Kulin, Gamenya and Machete) had also been studied in the current project and all contained the ‘4/275’ GBSS-7A allele. Group 3 contained many wheat accessions that had been separated into two main phylogenetic branches. One of the branches contained wheats with ‘3/275’ GBSS-7A alleles (Hartog, Dollarbird, Suneca and Eradu) while the other branch contained the accessions Molineux and Barunga with their ‘4/272’ GBSS-7A alleles. The classification of the GBSS-7A alleles based on the \textit{Xsun1} and (TGCCG)\textsubscript{n} molecular data therefore seems to provide an accurate indication of the genetic distance of those GBSS-7A alleles. Future work could include multiplexing these two molecular marker tests to create a more efficient test to distinguish GBSS-7A alleles.

### 7.4.3 Use of the \textit{XSun1} marker to screen two breeding populations and link the GBSS alleles to starch quality.

The use of the \textit{Xsun1} molecular marker to screen the GBSS characteristics of the partially-waxy wheat breeding populations was successful. The populations were useful for optimizing the use of the \textit{XSun1} marker for genotyping, but their doubled-haploid status meant that no heterozygotes were able to be screened, and therefore PCR effects, such as differential amplification of one allele over another, could not be studied. Genotypes were obtained for all the progeny of these crosses, and the preliminary linkage of those GBSS alleles with starch quality was examined. All the progeny were null 4A, but those that were also null 7A showed an average higher FSV than those that were wild-type 7A. As Figures 7.10 and 7.11 show, this linkage fits the non-linear trend for FSV versus GBSS copy number, and reflects the relationship between GBSS copy number and percentage amylose in the starch. These non-
Figure 7.10. Graph of the percentage amylose produced in wheat, per copy of the GBSS gene present. Graphed from data from Sasaki et al. (2000).

Figure 7.11. Graph of the FSV values for a selection of wheat varieties containing zero, two, four or six copies of the GBSS genes. The varieties were ‘DHWX-174’, the average for the 2-copy progeny from the DHWX-174/2*Arrino cross, ‘Arrino’, and ‘Kulin’, respectively.
linear trends are due to the high activity of the GBSS enzymes to compensate for the loss of some functional copies.

7.4.4 Using Doubled Haploid technology to produce partially waxy wheats may be hindered by their low survival rate.

In the second ‘DHWX-174/2*Arrino’, doubled-haploid breeding cross, different numbers of progeny with each combination of alleles were seen, than expected. There was only waxy progeny seen and no null 4A, wild-type 7A, null 7D progeny seen. During the breeding process, some individuals are inevitably lost at each step from plants dying in the field or not producing viable seed. Since all genotypes were seen in the first cross, perhaps the second cross was grown under less than ideal conditions, adversely affecting survival of these progeny. It is interesting to note that, in the second cross, plants containing wild-type GBSS alleles at both 7A and 7D loci had a much higher survival rate than those containing the null alleles. This suggested that higher levels of amylose in the starch had direct benefits for the survival of the plants. This was supported by the fact that those plants possessing the wild-type 7D allele (the higher starch producer of the two alleles 7A and 7D) and the null 7A allele had higher levels of survival than those containing the wild-type 7A and null 7D alleles.

While the data cannot be used to empirically measure the true survival rate of the wheat plants, it does seem to suggest that the presence of wild-type GBSS alleles confers some survival benefit over the null alleles. Most probably this benefit is due to the higher amylose levels in the starch encouraging seed viability. Perhaps in the course of breeding the partially-waxy wheats it may be seen that development and continued production of low-amylose wheat varieties is hindered by their lower survival rate compared to the higher-amylose wheats.
7.4.5 Conclusions.

The Xsun1-7A marker is a very useful tool for identifying wild-type alleles at the GBSS 7A locus. Shariflou and Sharp (1999) had identified 8 alleles in their selection of wheat varieties and a further 5 alleles since that publication (pers. comm.). When used in combination with the (TGCCG)$_n$ marker, more than 17 alleles can be distinguished. It is recommended that the WA germplasm be surveyed using these markers to identify the GBSS-7A alleles contained within. The alleles lacking from some breeding programs can then be introduced to increase variation. Whilst genetic variation does not automatically provide phenotypic variety, and having genetic variation is not necessarily desirous in a breeding program, genetic variation is vital for introducing new traits into a breeding program. Even polymorphisms that have been tested in a segregating population, and do not seem to effect the phenotype, may effect the phenotype in a different genetic environment due to epistatic effects or in the presence of different populations of transcription factors. This means that the effects of a polymorphism may need to be tested in several crosses. It also means that having polymorphism present in the germplasm, which is presumed non-functional can, in certain genetic environments, produce unexpected and undesired phenotypes. Even if a polymorphism is neutral, it may be linked to a functional mutation and could be used as a molecular marker for that trait. Without genetic variation there is no potential for new phenotypes and so discovering, and testing the effects of, polymorphisms is the only non-transgenic way to introduce phenotypic variation into wheat varieties.

Wheat varieties containing different combinations of GBSS-7A alleles can be tested for the starch characteristics they contain to try to link the Xsun1 marker variation to starch quality. However, in this study, the FSV variation in noodle-wheats could not be linked to the microsatellite markers. This may suggest that 7A allele variation may not be the reason for these starch quality differences, or that the epistatic factors are too complex to allow the true linkage between the 7A alleles and starch quality to be seen.
The GBSS genes are highly conserved and stable, with mutated genes being very rare, suggesting that most mutations cause an evolutionary disadvantage. Since the $X_{sun1}$ and $(TGCCG)_n$ microsatellite loci are both located close to the coding regions of the GBSS-7A gene, they should also be highly conserved and tightly linked to each other creating stable haplotypes. However, three of the $X_{sun1}$ allele groups contained members with both $(TGCCG)_n$ alleles suggesting that either these two microsatellite loci are currently mutating or that there is significant recombination in the DNA between the two loci. Either of these possibilities would mean that the linkages of the microsatellite alleles to the coding regions of the GBSS gene may not be as simple or as stable as we would prefer. Tight linkage between the microsatellite loci and the coding regions of the GBSS gene needs to be confirmed before these markers can be effectively used in wheat breeding efforts. While the linkage between the null alleles and the markers is easy to study because of protein differences, the multi-generational crosses required for these studies have not been produced and the null alleles are rare in existing varieties. Ideally, new markers need to be found in the coding regions of the gene, which are linked to the microsatellite loci. Using modified GBSS4/GBSS6 primers to amplify the region between the microsatellite loci to find markers in this region might reveal the true haplotype structure of the GBSS genes and identify potential ‘hotspots’ of recombination, which might confound the use of the microsatellite markers in breeding programs.
Chapter 8. General discussion.

8.1 General discussion.

8.1.1 Aims.

The GBSS genes from Australian wheats were studied because of their suspected impact on starch quality. The small differences in starch amylose content resulting from mutations at the GBSS loci influence the utility of different varieties in food products, such as Udon noodles. Wheat suitable for Udon noodles must contain null 4A alleles, which have a lower flour amylose content and increased Flour Swelling Volume (FSV). There are also other contributing factors to noodle-quality since some of the noodle-wheat varieties have low FSV, possibly due to allelic variation at the 7A or 7D loci, or to epistatic interactions beyond the scope of this study. The overall aim was to investigate the genomic organisation of the GBSS genes, searching for different alleles and their effects on starch quality. By developing molecular markers to detect variation among GBSS alleles, the aim was to improve the efficiency of breeding for starch quality.

In Chapter 1 the specific aims of the project were posed as five questions. These questions have been posed again here and their answers discussed in the context of the research.

8.1.2 Where are the GBSS gene copies in the genome?

There is a single copy of the GBSS gene at each of the three GBSS loci on chromosomes 4A, 7A and 7D, but are there other copies in the genome such as pseudogene copies? Pseudogene copies of the GBSS genes could be amplified in PCR tests for the GBSS genes and interfere with our analysis of the coding copies of the GBSS genes in a wheat variety. Identifying pseudogenes will allow their contribution to PCR results to be disregarded, but will also give information about the contribution of the GBSS gene sequences to the evolution of the wheat genome.
GBSS pseudogenes were sought using both PCR and hybridization methods. Using PCR, sections of GBSS genes from exon 4 to exon 11 were amplified. The DNA sequences of most of these PCR fragments were so similar to each other and to published GBSS sequences, that they were considered to have been amplified from the coding copies of the GBSS genes and not from pseudogene copies, which were expected to show DNA sequence polymorphism. Of the many PCR fragments sequenced, there were two that did not match the accepted GBSS sequences. One was amplified from the region between exons 4 and 6 of the GBSS-4A gene of ‘Janz’ and seemed to be a chimera constructed of sequences from both the 4A and 7A loci. The other was amplified from the region near the stop-codon from the null 7A allele of ‘DHW-174’ and showed a premature stop-codon. While it was possible that these fragments were amplified from pseudogene copies of the GBSS genes, it was also possible that they resulted from PCR errors since they could not be re-amplified from those wheat varieties.

Six new PCR primers were designed to amplify from various sections of the GBSS genes. Many of these primers produced non-specific and homoprimed PCR bands. The PCR products were hybridized to known GBSS sequences to discover if the primers were amplifying from pseudogene sequences. No evidence was found that the primers were amplifying from GBSS pseudogenes. This was supported recently by Murai et al. (1999) who amplified and sequenced 18 fragments from the GBSS genes at all three loci using primers in exons 1, 4, 6, 10 and 11. They also detected no pseudogene fragments.

The entire wheat genome was examined for pseudogene copies of the GBSS genes by hybridization of a GBSS sequence to genomic Southern blots of wheat DNA. Since pseudogene copies of genes are under less selective pressure to maintain their sequence integrity, they mutate at faster rates than coding copies of genes and the restriction sites within and surrounding a pseudogene is expected to show more polymorphism than sites surrounding coding genes. If pseudogene copies of the GBSS genes were present in the wheat genome, one should see evidence for their existence as multiple bands in addition to those containing
the coding genes. However, the RFLPs that would have characterized pseudogene copies of the GBSS genes were not detected. This result was also supported by the work of Vrinten et al. (1999), who did not identify any pseudogene copies of the GBSS genes using Southern blot analyses.

Since no pseudogenes in wheat were identified during the course of this project, and none are reported in the literature, there likely are no GBSS pseudogene copies in the wheat genome. This result was unexpected since there is an expectation that large plant genomes are comprised of repeated regions and pseudogenes. It is now thought that non-coding regions of the genome can be involved in small-mRNA signalling and gene regulation processes (Mattick, 2001; Mattick and Gagen, 2001).

From the large numbers of non-specific PCR products found in this study and from literature reports (Murai et al., 1999), it is clear that primers designed to the coding regions of the GBSS genes can bind to many places in the genome, possibly because of conservation of sequences encoding structural protein elements. However, the four primers designed to introns were specific, and this information should be utilised when designing new primers to the GBSS genes.

These findings all suggest that amplification of molecular markers from GBSS genes should not be confounded by interference from pseudogene sequences, but that non-specific amplification will be a problem and high-fidelity polymerases might need to be used for reducing PCR sequence errors.
8.1.3 How many different GBSS alleles are in the WA wheat breeding germplasm?

*Wheat germplasm in the WA noodle-wheat breeding program contains wild-type and null GBSS alleles at all three loci, but the numbers of alleles and their effects on starch quality are not known. It is aimed to assess the molecular variation present in the noodle-wheat germplasm.*

Three regions of the GBSS genes were examined from a selection of wheat varieties. In the first region, between exons 4 and 6, both the coding and non-coding regions were highly conserved across the different wheat varieties. A total of only 22 sequence differences were seen and some of these may have been sequencing errors since they were not repeatable. The single base polymorphisms will need to be confirmed before being useful as molecular markers to differentiate alleles. Two PCR tests were designed, but their optimisation was not completed and should be included in future work.

One of the polymorphisms was confirmed; the (TGCCG)$_n$ microsatellite from intron 4 of the 7A locus. This (TGCCG)$_n$ polymorphic marker allowed the differentiation of the alleles at this locus into two groups based upon the number of microsatellite repeats they possessed (three or four) and allowed the differentiation of a novel 7A allele from that described by Murai *et al.* (1999). This novel (TGCCG)$_4$ allele was present in 40% of the Australian varieties studied.

The second region of the GBSS genes to be studied was between exon 4 and the stop-codon of the GBSS-7A alleles. Three distinct GBSS-7A alleles were identified; the alleles present in the Australian varieties ‘Kulin’ and ‘Eradu’ were different to each other and to the published ‘Chinese Spring’ sequence (Murai *et al.*, 1999) both at the DNA and protein sequence levels.
Thus, three wild-type 7A alleles were identified in the WA wheat breeding germplasm and a PCR tool now exists for more alleles to be identified.

The third region of the GBSS genes to be studied was the region surrounding the Xsun1 locus in the 3’ untranslated region of the GBSS genes. This molecular marker was linked to the GBSS genes and could distinguish between null and wild-type alleles at the 7A and 7D loci thereby allowing the identification of these alleles. Due to variation at the Xsun1-7A locus, the GBSS-7A alleles could be separated into 9 allelic groups. Using the (TGCCG)n microsatellite marker from intron 4 of the GBSS-7A genes, three of these allelic groups could be further subdivided. Thus, the GBSS-7A alleles in a selection of WA wheat varieties could be separated into 12 groups as identified by molecular markers.

Using information from all three of these regions, the selection of wheat varieties examined contained one wild-type 4A allele, two null 4A alleles (M. Shariflou, pers. comm.), one wild-type 7D allele, one null 7D allele, one null 7A allele, and 12 wild-type 7A alleles. This group of 18 differentiable alleles provides a basis for producing a wide range of genetically different varieties, possibly with different starch qualities. This breeding work has begun at WADA with partially-waxy, doubled haploid populations being produced. The GBSS alleles from two of these populations were characterized here using these molecular markers, to test how efficient the marker can be. The starch qualities produced by the different partially-waxy individuals trended with the more-waxy individuals having higher FSV, but the small population size meant that the FSV differences were not significantly different.

8.1.4 How does each GBSS allele differ from the others?

At the start of the project, the reasons for the lack of expression from null alleles was not known. Also, any molecular variation between wild-type GBSS alleles had not been identified.
It was aimed to identify molecular differences within the population of GBSS alleles in Western Australian germplasm, and identify the reasons for the lack of expression from the null alleles. These differences could then be used to design molecular markers that could be used to assess the possible functional differences those mutations produced. A significant portion of the experimental work in this study had been completed before the genomic GBSS sequences of the null and wild-type alleles were published (Murai et al., 1999; Vrinten et al., 1999). Those published sequences were not from Australian varieties, and so the comparable sequences from Australian varieties were still required.

Many of these questions were answered when the DNA sequences of the wild-type and null GBSS alleles at each locus from ‘Chinese Spring’ wheat were published in 1999 (Murai et al., 1999; Vrinten et al., 1999). Differences in predicted protein sequence between the three wild-type alleles were revealed, as were the reasons for the lack of expression from the null alleles - the 5’-truncation of the null 7A allele and the 3’-truncation of the null 7D allele.

In this project, investigating how the GBSS genes present in Australian wheat varieties differed was also of interest. Using sequence information from the GBSS genes between exons four and six, some differences between the genes present in different wheat varieties were seen. Several putative single nucleotide polymorphisms were seen and when these are validated they will mark novel alleles. Any SNP that falls within a coding region and changes an amino-acid residue will also change the protein expressed from that allele and will need to be confirmed as having an effect on starch production.

The 7A alleles showed more sequence differences compared with those from the other loci. GBSS 7A alleles show variation at the \((\text{TGCCG})_n\) microsatellite in intron 4 with two alleles described. They also show variation at the \(X_{sun1}\) microsatellite locus in their 3’ untranslated regions, with at least 13 alleles discovered so far. The GBSS 7A alleles also differ in their
coding regions as was demonstrated by sequencing two wild-type 7A alleles from Australian wheat varieties ‘Kulin’ and ‘Eradu’. Since the proteins expressed from these alleles differ, their effects on starch quality must also be investigated. Since ‘Kulin’ contains the wild-type 4A allele, it has low FSV and the effect of the 7A allele is masked by the 4A gene. ‘Kulin’ does have other closely related varieties, including ‘Reeves’, which are null 4A, however they are members of the group of noodle-wheat varieties with inexplicably low FSV, so again the effects of the 7A genes are masked. Ideally, future work should identify other copies of the 7A allele from ‘Kulin’ present in noodle-wheat varieties with wild-type FSV characteristics such that their starch quality could be compared to that of high-quality ‘Eradu’.

Some of the GBSS alleles contain restriction site polymorphisms within and surrounding them. The polymorphism within the genes should be sequenced to see if it changes the expressed protein, while the polymorphism surrounding the gene could alter its expression. As more of the DNA in the region of the GBSS genes is sequenced, PCR primers can be developed to amplify the regions containing the polymorphism and then the PCR/RFLPs can be used as markers. Alternatively, the restriction sites can be used as starting points for directed AFLP amplification of the DNA surrounding the GBSS genes. This would allow access to unknown regions, such as the promoter. In this study, the restriction sites were used as basis for restriction site mediated PCR, but this proved too complex since it relied on the use of specific primers that have proven difficult to find in wheat, possibly to repetitive DNA in the species. However, knowing the sizes of the restriction fragments and having alternate alleles that give different bands should allow the AFLP amplification and identification of other GBSS fragments.

There is little sequence available from the GBSS promoters since it comes from the ends of cDNA copies of the genes and is not very long (10 to 30bp). However, one polymorphism seems significant. In position -1 from the start of translation, the GBSS 4A genes have a mutant ‘C’>’G’ substitution that may hold transcriptional or translational significance. Future
work should include designing a primer in this 30bp region of the promoter sequence to amplify and validate this polymorphism.

The one longer promoter sequence (203bp from one GBSS 4A gene AF286320) was found to contain short (4 or 5 bp) inverted repeats as well as two microsatellites; an ‘AGG’ motif with four repeats and a ‘CACT’ motif with three repeats. To study these elements, primers could be designed from the sequence that might be conserved across the three loci, enabling an examination of the microsatellite differences in the promoters from the 7A and 7D loci.

8.1.5 How can we tag the different GBSS alleles with molecular markers to identify them in a breeding program?

At the commencement of this project, the null alleles at the 7A and 7D loci could not be identified using molecular markers and thus could not easily be traced through a breeding program to study their effects on starch quality. Also, the different wild-type alleles predicted to exist at each locus could not be distinguished. Once novel alleles had been identified and tagged with a molecular marker, their effects on starch quality could be investigated towards breeding novel wheat varieties with improved starch qualities. Tagging different GBSS alleles with molecular markers was a major aim of this project.

Most useful molecular markers in breeding are PCR-based, so this project devoted significant investigation into examining how PCR primers designed to the GBSS sequences amplify from the wheat genome. The PCR primers designed to the GBSS cDNA (Clark et al., 1991) in this and previous projects were largely seen to produce homoprimed PCR fragments when used singularly. A recently published report (Murai et al., 1999) demonstrated that primers designed to this cDNA sequence annealed to many places in the wheat genome, and this may be a feature of primers designed to GBSS coding sequence. However, primers designed to intron sequences may be more unique to the GBSS gene regions. Previously, use of molecular
markers to the GBSS genes had been complicated by non-specific PCR amplification (Briney et al., 1998). When designing future molecular markers to the GBSS genes, it may be prudent to design primers to more-specific intron sequences.

The homoprimed bands amplified could be reduced and repressed by increasing the annealing temperatures of the PCR reactions. Reducing the annealing temperature of the PCR increased the stringency so non-specific primer binding was thermodynamically unfavourable. Taking advantage of the annealing temperatures at which non-specific primer binding was discouraged, reducing the confounding effect of non-specific amplification. Future work could also include changing the MgCl₂ concentration and the use of additives such as betaine or DMSO to destabilise weak primer binding and reduce non-specific bands. The current work demonstrated that the presence of a second primer species in the PCR may repress the formation of homoprimed artifacts. Perhaps the combination of the two primers reduced the less-specific amplification that occurred in the single-primed reactions.

A (TGCCG)n microsatellite was discovered in intron 4 of the GBSS 7A genes. A novel PCR test was developed to distinguish the two alleles at this locus and its subsequent use on Australian wheat varieties showed that this locus is not linked to starch quality. This study showed the usefulness of developing molecular markers to the GBSS genes. Also, 12 GBSS-7A alleles were described by a specific combination of marker bands from the Xsun1 and (TGCCG)n loci. The two simple PCR tests can now be used to identify each of these markers to genotype a wheat plant at these loci. Future work could include multiplexing these two tests into a single test since the PCR products produced are sufficiently different, and will enable detection on a single gel lane.

Differences in the restriction enzyme recognition sites were discovered, and these differences could be used to develop molecular markers to distinguish some alleles. For example, the
HindIII restriction fragments formed by the null 7A, wild-type 7A, wild-type 4A and wild-type 7D alleles differed in size such that they could be identified and differentiated.

By combining the PCR primers designed here with those designed in other projects (Murai et al., 1999; Shariflou and Sharp, 1999; Vrinten et al., 1999), the entire DNA sequence of the GBSS genes can be amplified, with the exception of the promoter regions. Additional research might include the use of these tools to screen the GBSS genes present in Australian wheat varieties to screen novel polymorphisms.

8.1.6 Are the GBSS genes present in WA wheat varieties the same alleles described in international publications?

Internationally, most work on the GBSS genes of wheat has been carried out in varieties of ‘Chinese Spring’ wheat which are different from the wheat varieties grown in WA, but present in our wheat breeding programs. In order to use the information gathered about the GBSS genes in ‘Chinese Spring’ wheat by international researchers, it is important to ascertain how similar the GBSS genes in WA noodle-wheat varieties are to those previously described from ‘Chinese Spring’ wheat.

Generally, there is GBSS gene sequence conservation between ‘Chinese Spring’ and the Australian wheat varieties examined. The PCR primers designed to ‘Chinese Spring’ sequences should also amplify from Australian wheat and serve as markers for Australian molecular-breeding programs. There are some differences between wild-type 7A alleles in Australian wheat and ‘Chinese Spring’, such as the SNPs which change amino acids in ‘Kulin’ and ‘Eradu’, the intron 4 microsatellite differences, and the 3’-untranslated region Xsun1 difference. However, the published null alleles at each locus (Vrinten et al., 1999) are exactly the same alleles as are present in Australian wheats.
The only published sequence from the 3’ untranslated region of the GBSS genes is the GBSS cDNA (Clark et al., 1991), originally transcribed from the 7A allele of a ‘Chinese Spring’ variety. Using the Xsun1 molecular marker PCR test, a PCR band corresponding to the size of the GBSS-7A 3’ untranslated region as displayed in the cDNA could not be amplified. Thus the ‘Chinese Spring’ variety used by Clark et al. (1991) seemed to lack 61bp of GBSS-7A 3’ untranslated region present in Australian wheat varieties. This suggests that ‘Chinese Spring’ wheat’s genome may be significantly different compared with those of Australian wheat varieties, particularly in non-coding regions. This will become clearer when the 3’ untranslated region of the GBSS gene is sequenced by our collaborator M. Sharp (University of Sydney, Sydney). If sequence from the 3’ untranslated region of the GBSS genes is not forthcoming from other sources, future work should include amplifying and sequencing these regions of the genes using the available primers.

8.2 Future directions.

The main problem to be solved towards advancing the breeding of improved Udon noodle wheat varieties, is that a molecular genetic link has not been established with starch functionality. Whilst it was predicted that the differing FSV characteristics would result from allelic variation at the 7A or 7D GBSS loci, this was not proven. Genes at the 7D locus were studied within the region between exons four and six, but no polymorphism was linked to functional starch quality. Genes at the 7A locus were more extensively studied both in the corresponding region between exons four and six where a polymorphic (TGCCG)ₙ microsatellite was discovered, as well as at the polymorphic Xsun1 locus in the 3’ untranslated region. Although 12 GBSS-7A alleles were described using these loci, none was linked to functional starch quality. If the different FSV characteristics do, in fact, result from allelic variation at the 7A or 7D GBSS loci, perhaps a causal polymorphism does not occur in the regions examined. The entire GBSS gene sequences from all three loci, including the promoter regions, will need to be examined and any polymorphisms found within these
regions examined for their linkage to starch quality. Only after this has been completed, can the effects of epistasis from other regions of the genome be investigated. There is also a chance that the predicted protein difference between the 7A proteins from ‘Kulin’ and ‘Eradu’ affects starch quality and this will need to be investigated.

The other regions of the GBSS genes, such as exons one to four and the promoter region, need to be examined for possible polymorphisms linked to starch quality variation. PCR primers designed in the current project and by other researchers (Murai et al., 1999; Vrinten et al., 1999) can now be used to amplify from the coding regions of the genes to search for polymorphisms in these regions. Sequencing the promoter and 5’ controlling element regions of the GBSS genes may be necessary to fully understand the expression of these genes. The sizes of some restriction fragments containing the GBSS promoters have been identified for cloning and sequencing these regions, and detecting possible polymorphisms.

It is important to noodle wheat breeding programs that the causes of FSV differences be understood. The hypothesis explaining FSV variation due to allelic variation at the GBSS loci is still relevant, however variation at other loci cannot be discounted. Preliminary findings from this research did not link GBSS-variation to starch quality and suggest that other loci may be influencing starch quality through epistasis and contributing to FSV variation.

Breeding partially waxy wheat populations containing individuals with various combinations of null and wild-type GBSS genes has been initiated using an Xsun1 molecular marker. Whilst preliminary results showed little significant variation in FSV between the wheat genotypes, the numbers of genotypes examined should be expanded. The difficulty is in the testing of the starch characteristics. Thus the use of the molecular marker will be important in association with quality testing. Future work will include the use of the Xsun1 molecular marker to genotype the partially-waxy genotypes and possibly link starch quality to GBSS allele variation.
An assessment of the genetic variation at the GBSS locus, and indeed at all loci, is important to fully understand the genetic component of the phenotype. Shariflou and Sharp (1999) studied the genetic diversity of the GBSS 7A alleles present in 135 varieties, primarily of Australian origin, using the closely-linked *Xsun1* marker. They have expanded the wheat genotypes studied (pers. comm.) and other varieties were examined in this study. However, a more extensive survey of the GBSS 7A alleles present among the Western Australian genotypes would provide valuable information about the different types of GBSS 7A alleles present. When other research groups screen their GBSS 7A alleles using this same molecular marker, it will become possible to identify and import novel alleles to increase the genetic diversity of Australian wheat. Thus assessing the GBSS 7A alleles present in Australian germplasm is recommended.

Specific future directions should include the following:

- Using the primer combination GBSS4/GBSS6 to amplify and sequence the 7A alleles from a selection of genetically diverse wheat varieties to validate polymorphisms and possibly identify novel polymorphisms. GBSS6 should be altered slightly to allow amplification from the 4A and 7D loci so that these genes can also be examined. Once alleles that produce different proteins are found, the effects of the different isoforms on starch quality can be studied by breeding partially waxy isolines containing each allele.

- The primers used by Murai *et al.* (1999) and Vrinten *et al.* (1999) to amplify the entire coding region of the GBSS genes should be sourced and used to access the 5’ region of the GBSS genes. Again the genotypes should be searched for novel alleles and their effects on starch quality studied.

- Primers should be designed to the available promoter sequence (AF286320) and used to study variation in this region.
- The existing molecular markers to the (TGCCG)$_n$ and $Xsun1$ loci should be used to screen germplasm, but their linkage to the GBSS coding regions needs to be established to map the haplotypes and characterise recombination.

- A wheat genomic library should be sourced and the clones containing the GBSS genes identified and sequenced. This will allow PCR primers to be designed to amplify the entire promoter region of GBSS genes from Australian varieties. This might enable one to find promoter polymorphisms and to detect functional differences by inserting the promoters into expression vectors and driving the expression of a reporter gene.

**8.3 Conclusions.**

This study successfully collected the current information about GBSS genes and used it to investigate the genomic organisation of these genes. Both existing and novel PCR primers were used to amplify the GBSS genes and compare the alleles from various wheat varieties in the Australian breeding programs. Non-PCR and various PCR methods were used to characterise unsequenced regions of the genes. The tools exist to fully sequence the GBSS genes from wheat and examine them for their impact on starch qualities. Perhaps this information can be used to improve the efficiency of developing Australian varieties with desirable starch qualities.
REFERENCES


**Table 2.1.** Descriptions of the wheat varieties used in the study including pedigrees and GBSS characteristics.

<table>
<thead>
<tr>
<th>VARIETY</th>
<th>NULL GBSS</th>
<th>FSV</th>
<th>ASWN</th>
<th>PEDIGREE</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>81W:1138</td>
<td>null 4A</td>
<td>medium</td>
<td>Bodallin/Ciano/Gamenya/Atlas/Madden</td>
<td>A breeding line from DAWA. It is a sister line to Tammin.</td>
<td></td>
</tr>
<tr>
<td>Amery</td>
<td>none</td>
<td>low</td>
<td>Lr21-SrX/~2<em>Shortim/~3</em>Bodallin</td>
<td>An early-maturing, Australian Premium Hard or Australian Hard wheat. It has good flour colour for noodles and is used as a colour control for Rammen noodle testing. Previously grown extensively in the wheatbelt, this variety was superseded by Carnamah due to leaf rust susceptibility.</td>
<td></td>
</tr>
<tr>
<td>Aroona</td>
<td>null 4A</td>
<td>medium</td>
<td>Lerma Rojo-64//Norin-10/Brevor-14/3/3*Andes(WW15)/4/Raven</td>
<td>An early-maturing, Australian Premium White wheat that was recommended in the past but has been superseded by newer varieties. Poorly adapted for soils of the eastern wheatbelt (deep, sandy, acidic). Grown in South Australia. It is semi-dwarf, hard-grained, has high milling extraction, but weak dough properties, is very extensible and has yellow flour. Suitable as a noodle-blend with Gamenya and Eradu.</td>
<td></td>
</tr>
<tr>
<td>Arrino</td>
<td>null 4A</td>
<td>high</td>
<td>yes</td>
<td>Complex pedigree(77W:660)/Eradu</td>
<td>A mid-maturing Australian Standard White Noodle wheat, released in 1997 in Western Australia. It has good yields, tolerates low rainfall, and has white bearded heads. Suitable for producing Udon noodles.</td>
</tr>
<tr>
<td>Barunga</td>
<td>null 4A</td>
<td>medium</td>
<td>Halberd/Aroona//3<em>Schomburgk/3/2</em>Molineux</td>
<td>An early-maturing, Australian Hard wheat that is the preferred AH in South Australia. It has medium height, small grains, white awns, and is Boron tolerant.</td>
<td></td>
</tr>
<tr>
<td>Batavia</td>
<td>null 4A</td>
<td></td>
<td></td>
<td>Brochis ’S’/Banks;Ciano-67/Bluebird//Cardinal/4/~SietteCerros/3/~Lermarojo 64/~Inia 66//~Inia 66/~Bluebird/5/~Banks</td>
<td>A late-maturing, Australian Premium Hard wheat. This variety is from Queensland and one of its parents is from Mexico.</td>
</tr>
<tr>
<td>Bodallin</td>
<td>null 4A</td>
<td>high</td>
<td>yes</td>
<td>Bokal/~Siete Cerros</td>
<td>This variety was released in 1981.</td>
</tr>
<tr>
<td>Batavia</td>
<td>null 4A</td>
<td></td>
<td></td>
<td>Brochis ’S’/Banks;Ciano-67/Bluebird//Cardinal/4/~SietteCerros/3/~Lermarojo 64/~Inia 66//~Inia 66/~Bluebird/5/~Banks</td>
<td>A late-maturing variety Australian Premium Hard variety from Queensland, it was once the preferred APH in this region. Derived from a cross between Banks and a high-yielding Mexican variety, this wheat is resistant to Septoria but susceptible to Crown Rot.</td>
</tr>
<tr>
<td>Cadoux</td>
<td>null 4A</td>
<td>high</td>
<td>yes</td>
<td>Centrifen/Gamenya (F3)//Gamenya/3/Jacup</td>
<td>A late-maturing, Australian Standard White Noodle wheat. Grown in the wheatbelt, Cadoux generally outyields Eradu and Gamenya, but is lower yielding than Spear and Amery. A soft-grained wheat with good milling quality similar to Eradu. It has good quality for white salted noodles (Udon) and is best suited for this use. Marketed into Japan for Udon noodles and South Korean dry noodles.</td>
</tr>
<tr>
<td>Calingiri</td>
<td>null 4A</td>
<td>medium</td>
<td>yes (no) Chino/Kulin//Reeves</td>
<td>A mid-maturing wheat classified as Australian Standard White Noodle but with unsuitable FSV unless blended. High yields, needs good rainfall but sensitive to waterlogging, not tolerant of alkaline soils or of SU herbicides.</td>
<td></td>
</tr>
<tr>
<td>Carnimah</td>
<td>null 4A</td>
<td>medium</td>
<td>(RAC529:911)/(77W:660)</td>
<td>A mid-maturing, Australian Hard wheat. It has very high yields, good baking quality and extensibility, has brown bearded heads, and is susceptible to fungal staining.</td>
<td></td>
</tr>
<tr>
<td>Cascades</td>
<td>null 4A</td>
<td>medium</td>
<td>Aroona*3/(AusenVII-95)Tadorna/Inia66</td>
<td>A mid-maturing, marginal Australian Hard wheat, registered in 1994 in Western Australia. It has excellent flour yield, yellow flour, good dough strength and extensibility, and a fully awned head.</td>
<td></td>
</tr>
<tr>
<td>Chinese Spring nullisomic 7D 7D alleles</td>
<td>null 4A</td>
<td>medium</td>
<td></td>
<td>A chromosome substitution variety that does not contain the 7D chromosome.</td>
<td></td>
</tr>
<tr>
<td>Chinese Spring nullisomic 7A 7A alleles</td>
<td>null 4A</td>
<td>medium</td>
<td></td>
<td>A chromosome substitution variety that does not contain the 7A chromosome.</td>
<td></td>
</tr>
<tr>
<td>Variety</td>
<td>Parental Background</td>
<td>Characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>---------------------</td>
<td>-----------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chinese Spring nullisomic 4A tetrasomic 4D</td>
<td>lacks GBSS 4A alleles</td>
<td>A chromosome substitution variety that does not contain the 4A chromosome.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cook</td>
<td>Timgales/Condor sib/Condar</td>
<td>An early-maturing variety from Queensland, released in 1977, but has low yield.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrigin</td>
<td>none low yes (no) Tincerrin*2//Gamenya//Jassul</td>
<td>A mid-maturing, Australian Standard White Noodle wheat. A soft-grained wheat with marginal quality for the A. Soft grade. Dough extensibility is better than Tincerrin, but water absorption of the flour is slightly higher which is a disadvantage. Its baking quality is very poor. Very marketable at protein levels below 9.0 per cent but difficult to sell at higher levels. Protein levels are often higher than Tincerrin, which can be a disadvantage. Has a tendency to be susceptible to fungal staining, making the wheat unacceptable in many markets due to visual appearance and associated milling problems. Preferred markets - South Korea, Thailand, Singapore, Malaysia, Taiwan, Philippines, domestic markets. Ideal end use Sweet biscuits, cakes, confectionery, snack food products and fermented beverages.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cranbrook</td>
<td>Wren-Mex/Ciano’S’/Noroeste-66/3/Zambezi</td>
<td>A mid-maturing, Australian Standard White wheat, released in 1984 in Western Australia. It is semi-dwarf and intolerant to Aluminum, it has moderately hard grains but is out-yielded by other varieties. It suffers from pre-harvest sprouting and is used to assess pre-harvest sprouting in trials.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cunningham</td>
<td>3Ag3/4*Condor//Cook</td>
<td>A mid-maturing, Premium Hard wheat from Queensland. It is a derivative of Cook and Condor, with stem and leaf Rust resistance but susceptibility to Crown Rot.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DHW 174</td>
<td>null 4A null 7A null 7D very high</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dollarbird</td>
<td>Wren/Gaboto/Kalyansona/Bluebird</td>
<td>An Australian Hard wheat, grown in southern New South Wales and tolerant to acid soils.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eradu</td>
<td>null 4A high yes Ciano/Gamenya</td>
<td>A mid-maturing Australian Standard White Noodle wheat, released in 1981 in Western Australia. A good-quality, soft-grained, semi-dwarf wheat. It is suitable for producing Udon noodles, but has lower yields than other varieties and a paler white flour colour compared to the ideal Cadoux. It has good textural properties and is being used as breeding stock in the ASWN program.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Express</td>
<td>null 4A high yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frame</td>
<td></td>
<td>An early to mid-maturing, Australina Premium White wheat grown in Victoria and South Australia. It is white awned with boron tolerance.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gluchub</td>
<td>Gluyas Early//Clubhead</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goroke</td>
<td>none low</td>
<td>TM56<em>2//Ausen 4-2/1--77-702D (3AG3</em>4*Condor)</td>
<td>Grown as an Australian Premium White wheat in Victoria.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Halberd</td>
<td>null 4A medium</td>
<td>Scimitar//Kenya C6042/ -Bobin/2/</td>
<td>A mid-maturing Australian Premium White wheat. Released in 1969, this variety is grown in the south-eastern wheatbelt because it has good boron toxicity tolerance. An awnless, free-milling wheat,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variety</td>
<td>Father/Mother/Parent(s)</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insignia 49</td>
<td>~Ciato/'S'/Siete Cerros/3/Kalyansona/Bluebird</td>
<td>but too hard-grained for the protein levels it achieves. It has poor dough strength and its extensibility is low. Not ideally suited for pan bread production, but has good starch quality. Preferred markets - Middle East Gulf markets and as part of the noodle blend for Japan and South Korea. Its good starch quality makes it acceptable in blends of wheat for noodle production, but it can only be used in limited proportions due to hard-grained characteristics. Ideal end use - Domestic flour, Arabic breads such as pocket and flat breads, and as a noodle blending wheat.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hartog</td>
<td>Vicam-71/Ciano/'S'/Siete Cerros/3/Kalyansona/Bluebird</td>
<td>An early-maturing, Prime Hard released in Queensland in 1982, and adaptable to many environments. It was selected from a cross with germplasm introduced from Mexico. It has good yields, high dough strength, is low tillering and is tolerant to root lesion nematodes, but has marginal colour stability.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inia</td>
<td>Napo Calidad Ciano</td>
<td>Released in 1978 by DAWA.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jacup</td>
<td>Bencubbin Charter Sword Kenya Mexico Gamernya</td>
<td>An early-maturing, Prime Hard released in Queensland in 1982, and adaptable to many environments. It was selected from a cross with germplasm introduced from Mexico. It has good yields, high dough strength, is low tillering and is tolerant to root lesion nematodes, but has marginal colour stability.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Janz</td>
<td>3AG/4<em>Cordor/4</em>Cook</td>
<td>A mid-maturing, Australian Premium White or Australia Hard wheat. Registered in 1989 in Queensland. It has been extensively grown during recent seasons because it has good Stripe rust resistance, and is grown with high quality in all areas of Australia. In Western Australia it produced over 11.5% protein. It is a short, semi-dwarf variety, with small grains and an intolerance to Aluminium in acid soils.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kulin</td>
<td>Bodallin/2/-'Gamenya/-Inia 66</td>
<td>Old variety released by Ag, WA in 1985. It is heavily awned, has a square head taper, medium head density and cream coloured head colour.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Machete</td>
<td>Sonora 64/2/~Tezanos Pintos Precoz/~Yaqui54/3/~Gabo/4/~Madden</td>
<td>A mid-maturing, Australian Hard wheat grown in various areas of WA and SA. Outclassed for yield by Carnamah. A hard-grained variety with average milling quality. Good dough strength and extensibility but has a poor flour colour. Protein must be more than 11.5 per cent for A. Hard. Preferred markets - Middle Eastern markets, although these markets principally buy ASW wheats from Western Australia and not A. Hard. Ideal end use - Arabic breads, pan breads and as a blend wheat. Western Australian Hard wheat is bought mainly by South Korea, Singapore and Malaysia and is used in noodle and bread grists. Machete is unsuitable for noodle production. It also has insufficient dough strength for bread production in these markets.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Molineux</td>
<td>Pitic-62/Festiguay/2*Warigal Seln.-21/7/16</td>
<td>A mid-maturing, Australian Hard wheat released in 1988. It is of medium height, with a white awned head and small grains.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pelsart</td>
<td>Potam-70/4*Cook</td>
<td>A mid-maturing, Australian Premium Hard variety, released in Queensland in 1996. It is a derivative of Cook, and has good yields, but can produce dark brown or black heads.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reeves</td>
<td>Bodallin/-'Gamenya/-Inia 66</td>
<td>Released in 1989 by AgWA. This variety is unsuitable for Udon noodle production due to unsuitable FSV. It has cream coloured head colour, medium head density, square head taper and is heavily awned.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rowan</td>
<td>Vicam-71/Ciano/'S'/Siete Cerros/3/Kalyansona/Bluebird</td>
<td>An early-maturing, good quality, Australian Hard wheat released in Queensland in 1995. It is similar to Hartog but awnless, and is suitable for feed or baling if affected by drought or frost after head emergence.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schomburg</td>
<td>Joral-66/Gamut(QT2338)//4*Hartog</td>
<td>A mid-maturing, Australian Premium Hard or Australian Hard wheat, similar to Aroona with yellow flour colour. It mills well, produces strong and extensible dough, and is most suitable blended for Middle-Eastern pan breads. Not suitable for normal bread or noodles. The variety is semi-dwarf, and is intolerant of Boron and Aluminium.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spear</td>
<td>Sabre/~Mec 3/2/~Insignia</td>
<td>An Australian Premium White or Australian Standard White wheat with variable maturity. Grown widely, its yield is superceded in most areas of Western Australia by more recent varieties. Medium</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variety</td>
<td>Height</td>
<td>Protein Tolerance</td>
<td>Extensibility</td>
<td>Market Preferences</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>----------------------------</td>
<td>---------------------------------</td>
<td>-----------------</td>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>Stretton</td>
<td>semi-dwarf</td>
<td>Good tolerance of boron toxicity</td>
<td>Better dough extensibility and strength than Halberd. Average starch quality and is hard-grained which precludes the variety being suitable for soft white noodle production. Preferred markets - Good for the Middle Eastern markets.</td>
<td>Good for the Middle Eastern markets.</td>
<td></td>
</tr>
<tr>
<td>Sunco</td>
<td>mid-maturing</td>
<td>Australian Premium Hard or Australian Hard variety, that has excellent quality in Queensland and New South Wales. It is low yielding in Western Australia, but produces more than 13% protein in this state. Released in 1986, this semi-dwarf variety has relatively small grains and Crown Rot tolerance. It is the national reference standard for alkaline noodles.</td>
<td>A mid-maturing, hard-grained, Australian Premium White wheat. Similar yield to Spear. Registered by DAWA in 1993, it is outyielded by new varieties. It has a tendency to high screenings, which may necessitate grading before delivery. It has large grain size, but with some small grains. A Stripe rust resistant wheat suitable for the medium rainfall area of the wheatbelt.</td>
<td>Good for the Middle Eastern markets.</td>
<td></td>
</tr>
<tr>
<td>Suneca</td>
<td>Ciano//Spice/Amber Mutant Sonora-64</td>
<td>unknown</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tammin</td>
<td>null 4A</td>
<td>Bodallin/-(74W07-47)/Eradu-sib- XBUT223/ Atlas 66/-*Madden</td>
<td>A mid-maturing, feed-quality wheat. It has high yields in Western Australia and broad adaptation to a wide range of sowing times. It is a soft grained wheat with plump grain, fair milling quality, and similar strength to Gamenya but lower extensibility. Registered in 1994 by DAWA, it has a white fully-awned head, but occasional taller, bold and brown off-types are present. It is not liable to develop black point and has low pre-harvest sprouting.</td>
<td>A mid-maturing, Australian Premium White wheat. Similar yield to Datatine and lower than Corrigin. A medium height, semi-dwarf variety, it is soft wheat ideal for biscuit and cake production. It has low water absorption, has weak extensible dough properties and its baking quality is very poor. Very marketable at protein levels between 8.0 and 9.0 per cent for end users of soft wheat, but difficult to sell with higher protein levels. Preferred markets – South Korea, Thailand, Singapore, Malaysia, Taiwan, Philippines, domestic market. Ideal end use - Sweet biscuits, cakes, snack foods, fermented beverages.</td>
<td>An early-maturing Australian Hard wheat released in Queensland in 1993.</td>
</tr>
<tr>
<td>Tasman</td>
<td></td>
<td>Torres/3/Gaboto/Siette Cerros/Bluebird/Ciano(11th IBWSN 45/Torres)</td>
<td>A mid-maturing, AGP1 wheat. Registered in 1993, it is outyielded by many varieties. It has poor milling yield, is weak and lacks extensibility. Moderately tolerant to soil boron, similar to Spear. VPM, B tolerant, CCN resist.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tuncurin</td>
<td>none</td>
<td>Gluculb/3/~Chile 1B/2/~Insignia/~Falcon</td>
<td>A mid-maturing, Australian Standard White wheat. Registered in 1977 by DAWA. Similar yield to Datatine and lower than Corrigin. A medium height, semi-dwarf variety, it is soft wheat ideal for biscuit and cake production. It has low water absorption, has weak extensible dough properties and its baking quality is very poor. Very marketable at protein levels between 8.0 and 9.0 per cent for end users of soft wheat, but difficult to sell with higher protein levels. Preferred markets – South Korea, Thailand, Singapore, Malaysia, Taiwan, Philippines, domestic market. Ideal end use - Sweet biscuits, cakes, snack foods, fermented beverages.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trident</td>
<td>null 4A</td>
<td>VPMI/~5<em>Cook/~4</em>Spear</td>
<td>A mid-maturing, Australian Premium White wheat. Similar yield to Halberd, Gutha, Nyabing and Westonia. A minimum 9.5 per cent protein requirement exists, with a seasonal upper limit of 11.5 per cent protein if needed in certain seasons.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yarralinka</td>
<td></td>
<td>Mengavi/Siete Cerros/3/Mengavi/Siete Cerros/Crim/4/Combination III/2/Warigal</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ASW Noodle (ASWN): Segregation limited to the varieties Arrino, Cadoux, Calingiri, Eradu and Gamenya at protein 9.5 per cent to 11.5 per cent.
Australian Soft (ASFT): Specific soft grain varieties Corrigin, Datatine, Harrismith and Tuncurin at protein less than 9.5 per cent. May be segregated at lower levels for certain markets.
ASWT: Market specific segregation for the varieties Cascades, Halberd, Gutha, Nyabing and Westonia. A minimum 9.5 per cent protein requirement exists, with a seasonal upper limit of 11.5 per cent protein if needed in certain seasons.
Australian Hard (AH1 and AH13): Specific varieties at protein above 11.5 per cent protein (AH1) and above 13 per cent protein (AH13).
Australian Premium White (APW): Specific varieties at protein above 10 per cent.
Australian Standard White (ASW): A grade with mixed grain hardness and no specific protein requirements. Eligible varieties include AH, APW and ASWN varieties.
Australian Durum (ADR1, ADR2 and ADR3): Limited to Durum varieties Kimaloari, Tamaroi, Wallaro and Yallaro with good protein levels. Durum failing to meet receival standards is downgraded to Feed.

In addition to premium milling wheats, the Australian Wheat Board segregates General Purpose and Feed grades and will introduce specific segregation’s based on seasonal conditions.

All information was provided by V. Reck, DAWA, Esperence, Western Australia. Where information is missing for a variety, no information was available from the DAWA intranet information resource.
Appendix 1. Information about the percentages of area sown to crop varieties 1995/96 to 1999/00 seasons

_data from Co-operative Bulk Handling Ltd_

Varieties listed 0.02% or greater of crop area. (* Year variety registered)

<table>
<thead>
<tr>
<th></th>
<th>95/96</th>
<th>96/97</th>
<th>97/98</th>
<th>98/99</th>
<th>99/00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spear</td>
<td>17.86</td>
<td>16.33</td>
<td>12.95</td>
<td>12.56</td>
<td>11.57</td>
</tr>
<tr>
<td>Stiletto</td>
<td>0.66</td>
<td>5.35</td>
<td>8.01</td>
<td>10.3</td>
<td>10.58</td>
</tr>
<tr>
<td>Amery</td>
<td>0.28</td>
<td>4.89</td>
<td>11.99</td>
<td>12.54</td>
<td>10.23</td>
</tr>
<tr>
<td>Eradu</td>
<td>11.82</td>
<td>11.66</td>
<td>11.59</td>
<td>10.82</td>
<td>9.48</td>
</tr>
<tr>
<td>Halberd</td>
<td>13.54</td>
<td>11.79</td>
<td>11.29</td>
<td>10.09</td>
<td>7.49</td>
</tr>
<tr>
<td>Machete</td>
<td>5.63</td>
<td>6.28</td>
<td>6.09</td>
<td>6.56</td>
<td>6.47</td>
</tr>
<tr>
<td>Carnamah</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0.23</td>
<td>5.21</td>
</tr>
<tr>
<td>Cadox</td>
<td>16.2</td>
<td>16.06</td>
<td>8.25</td>
<td>7.25</td>
<td>5.2</td>
</tr>
<tr>
<td>Cascades</td>
<td>*</td>
<td>0.2</td>
<td>2.26</td>
<td>4.47</td>
<td>4.82</td>
</tr>
<tr>
<td>Cunderdin</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0.5</td>
<td>4.38</td>
</tr>
<tr>
<td>Perenjori</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0.36</td>
<td>3.22</td>
</tr>
<tr>
<td>Blade</td>
<td>2.05</td>
<td>2.62</td>
<td>3.24</td>
<td>3.22</td>
<td>2.91</td>
</tr>
<tr>
<td>Tincurry</td>
<td>3.05</td>
<td>3.64</td>
<td>3.25</td>
<td>2.6</td>
<td>2.43</td>
</tr>
<tr>
<td>Wilgoyne</td>
<td>6.68</td>
<td>5.28</td>
<td>5.26</td>
<td>3.33</td>
<td>2</td>
</tr>
<tr>
<td>Kalannie</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0.14</td>
<td>1.91</td>
</tr>
<tr>
<td>Dagger</td>
<td>1.24</td>
<td>1.16</td>
<td>1.21</td>
<td>1.53</td>
<td>1.24</td>
</tr>
<tr>
<td>Aroona</td>
<td>3.92</td>
<td>3.43</td>
<td>2.39</td>
<td>1.65</td>
<td>1.22</td>
</tr>
<tr>
<td>Datatine</td>
<td>*</td>
<td>0.08</td>
<td>0.53</td>
<td>1.18</td>
<td>1.2</td>
</tr>
<tr>
<td>Janz</td>
<td>1.9</td>
<td>2.45</td>
<td>1.92</td>
<td>1.68</td>
<td>1.11</td>
</tr>
<tr>
<td>Brookton</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-*</td>
<td>0.84</td>
</tr>
<tr>
<td>Sunelg</td>
<td>1.44</td>
<td>0.9</td>
<td>1.13</td>
<td>0.81</td>
<td>0.62</td>
</tr>
<tr>
<td>Gutha</td>
<td>2.44</td>
<td>1.26</td>
<td>1.38</td>
<td>0.8</td>
<td>0.54</td>
</tr>
<tr>
<td>Krichauff</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0.07</td>
<td>0.53</td>
</tr>
<tr>
<td>Stretton</td>
<td>0.15</td>
<td>1.06</td>
<td>0.85</td>
<td>0.79</td>
<td>0.5</td>
</tr>
<tr>
<td>Corrigin</td>
<td>1.03</td>
<td>1.13</td>
<td>0.84</td>
<td>0.62</td>
<td>0.5</td>
</tr>
<tr>
<td>Westonia</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-*</td>
<td>0.49</td>
</tr>
<tr>
<td>Arrino</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-*</td>
<td>0.45</td>
</tr>
<tr>
<td>Gamenya</td>
<td>1.84</td>
<td>0.98</td>
<td>0.77</td>
<td>0.51</td>
<td>0.4</td>
</tr>
<tr>
<td>Tammin</td>
<td>*</td>
<td>0.44</td>
<td>2.46</td>
<td>1.59</td>
<td>0.27</td>
</tr>
<tr>
<td>BT-</td>
<td>0.55</td>
<td>0.18</td>
<td>0.35</td>
<td>0.34</td>
<td>0.26</td>
</tr>
<tr>
<td>Schomburgk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schomburgk</td>
<td>0.14</td>
<td>0.47</td>
<td>0.26</td>
<td>0.21</td>
<td>0.22</td>
</tr>
<tr>
<td>Calingiri</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-*</td>
<td>0.19</td>
</tr>
<tr>
<td>Nyahing</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-*</td>
<td>0.17</td>
</tr>
<tr>
<td>Canna</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>0.17</td>
</tr>
<tr>
<td>Sunco</td>
<td>0.38</td>
<td>0.27</td>
<td>0.11</td>
<td>0.13</td>
<td>0.15</td>
</tr>
<tr>
<td>Wialki</td>
<td>0.2</td>
<td>0.17</td>
<td>0.2</td>
<td>0.17</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>95/96</td>
<td>96/97</td>
<td>97/98</td>
<td>98/99</td>
<td>99/00</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>Frame</td>
<td>-</td>
<td>-</td>
<td>0.01</td>
<td>0.07</td>
<td>0.11</td>
</tr>
<tr>
<td>Madden</td>
<td>0.3</td>
<td>0.16</td>
<td>0.18</td>
<td>0.15</td>
<td>0.1</td>
</tr>
<tr>
<td>Silverstar</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.09</td>
</tr>
<tr>
<td>Trident</td>
<td>0.03</td>
<td>0.27</td>
<td>0.36</td>
<td>1.37</td>
<td>0.08</td>
</tr>
<tr>
<td>Kite</td>
<td>0.17</td>
<td>0.1</td>
<td>0.12</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>Cunningham</td>
<td>0.11</td>
<td>0.13</td>
<td>0.04</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>Suneca</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.04</td>
</tr>
<tr>
<td>Dollarbird</td>
<td>0.45</td>
<td>0.26</td>
<td>0.12</td>
<td>0.06</td>
<td>0.03</td>
</tr>
<tr>
<td>Katunga</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>Cranbrook</td>
<td>0.12</td>
<td>0.08</td>
<td>0.06</td>
<td>0.04</td>
<td>0.02</td>
</tr>
<tr>
<td>Millewa</td>
<td>0.22</td>
<td>0.08</td>
<td>0.07</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Goldmark</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Appendix 2. Regions of Western Australia that the wheats are grown in.

All information was provided by V. Reck, DAWA, Esperence, Western Australia, from the intranet information pages of DAWA.

<table>
<thead>
<tr>
<th>Region Zone</th>
<th>High rain</th>
<th>Medium rain</th>
<th>Low rain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>APW</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amery A</td>
<td>A</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Brookton A</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Carnamah A</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Cascades A</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td>Cunderdin A</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Janz</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kalannie A</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Karlgarin A</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Perenjori A</td>
<td>S</td>
<td>S</td>
<td>A</td>
</tr>
<tr>
<td>Spear</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td>Stiletto A</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td>Westonia A</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>ASWT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nyahing A</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>ASW</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ajana A</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Noodles</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arrino A</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Cadoux</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td>Calingiri A</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Eradu</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td>A. Hard</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amery A</td>
<td>S</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Carnamah A</td>
<td>S</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kalannie A</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A. Soft</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrigin</td>
<td>-</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Datatine A</td>
<td>-</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Tincurren</td>
<td>-</td>
<td>-</td>
<td>S</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S = Suggested</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A = Alternative</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

H5W is the area west of the Chester Pass Road.
H5C is the area between Chester Pass Road and the Fitzgerald National Park.
H5E is the area east of the Fitzgerald National Park.
M5W is the area west of Ongerup.
M5C is the area between Ongerup and Ravensthorpe.
M5E is the area east of Ravensthorpe.