GENOME STUDIES OF CEREALS

BY

SONG WEINING

B. Ag. Sc., M. Ag. Sc.

A thesis submitted in fulfilment of the

requirements for the degree of

Doctor of Philosophy

Department of Plant Science

Waite Agricultural Research Institute

The University of Adelaide

August 1992

DECLARATION

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma in any University. To the best of my knowledge and belief, no material described herein has been previously published or written by another person except due reference is made in the text.

If accepted for the award of a Ph.D degree, this thesis will be available for loan and photocopying.

Song Weining

ACKNOWLEDGMENTS

I sincerely thank my supervisor Dr Peter Langridge for his invaluable guidance, support and constructive criticism throughout the progress of the present investigation and in the preparation of this manuscript. I am also indebted to my co-supervisor Dr Ken Shepherd for his support and advice during the course of this project.

I am grateful to Dr Max Tate for his kind encouragement and suggestions.

The following colleagues and friends who contributed to this work and helped me in various ways are also gratefully acknowledged:

P Rogowsky for his provision of the sequence information of R173 λ clones and critical reading of part the manuscript. F Guidet for his encouragement on the seed PCR. D Sparrow and I Dundas for helpful discussions and genetic material. J Nield for her assistance. D Hein and N Shirley for the synthesis of primers. A Dunbar and J for preparing the photographic prints.

Thanks are due to Professor R Wolf (Univ. Maryland Baltimore County) and Professor HC Towles (Univ. Minnesota) for providing DNA probes.

The University of Adelaide Postgraduate Scholarship is gratefully acknowledged.

Finally, and most importantly, I wish to thank my wife Sharon and my parents for their unfailing encouragement and patience.

Publications

Guidet, F., Rogowsky, P., Taylor, C., Song, W. and Langridge P. (1990) Cloning and characterization of a new rye specific repeated sequence. Genome, 34, 81-87.

Weining, S. and Langridge P. (1991) Identification and mapping of polymorphisms in cereals based on the polymerase chain reaction. Theoretical and applied genetics, 82, 209-216

Conference participation

Rogowsky, P., Guidet, F., Song, W., Taylor, C., Karakousis, A., Shepherd, K. and Langridge, P. (1990) Molecular markers for monitoring alien gene transfer to wheat. Australian Gene Mapping Workshop, Sydney, pp26-27.

Weining, S. and Langridge P. (1990) Identification and mapping of polymorphisms in cereals based on PCR. Aust. Soc. Plant Phys., Sydney, p30.

Weining, S., Rogowsky, P., Guidet, F., Taylor, C., Karakousis, A., Shepherd, K. and Langridge, P. (1990) Identification and mapping of polymorphisms in cereals based on DNA probes. Proc. 6th Assembly Aust. Wheat Breeding Soc. pp487-488.

Weining, S. and Langridge P. (1992) Application of PCR to cereal genome mapping and DNA fingerprinting with random, semi-random and specific primers. Third Australasian Gene Mapping Workshop, Brisbane, p67.

ABREVIATION

bp	base pairs
kb	kilobases
Mb	megabase
PCR	polymerase chain reaction
PFGE	pulse-field gel electrophoresis
RFLP	restriction fragment length polymorphism
DGGE	denaturing gradient gel electrophoresis
SSCP	single stranded conformation polymorphism
STS	sequenced tagged sites
YAC	yeast artificial chromosomes
20X SSC	3 M NaCl, 0.3 M Na_3Citrate.
20X SSPE	3.6 M NaCl, 0.2 M NaH_2PO_4, 0.02M Na_2EDTA, pH7.4.
10X TAE	400 mM Tris-HCl, 30 mM sodium acetate, 10mM Na_2EDTA, pH7.8.
10X TBE	1 M Tris base, 10 mM Na_2EDTA, 0.863 M boric acid (pH8.3).
IPTG	Isopropyl β-thiogalactopyranoside
X-gal	Bromo-(5)-4- chloro-3-indolyl-β-galactopyranoside
LB	10g Bacto-Trytone, 5g yeast extract, 10g NaCl, pH 7.0

Table of Contents

Chapter 2

MATERIALS AND METHODS 30

Chapter 3

IDENTIFICATION AND MAPPING OF POLYMORPHISMS THROUGH THE POLYMERASE CHAIN REACTION.

Chapter 4

INTRAVARIETAL AND INTERVARIETAL POLYMORPHISMS REVEALED BY PCR

Summary

This thesis embodies the results of an investigation on the genome analysis of cereals which include wheat, rye and barley. The major objective of the present investigation is to evaluate the feasibility of using the polymerase chain reaction (PCR) to study the cereal genomes.

Firstly, PCR has been adapted to genome mapping in cereals. Known gene sequences can be used to design primers and detect polymorphic PCR products which vary in length. This is demonstrated with primers to the α-amylase gene family. The primers were chosen to hybridise with two conserved sequences that flank a region containing variable sequences. In this way several individual α-amylase genes would be targeted simultaneously and length polymorphisms between the sequences complementary to the primers, could be identified. Polymorphisms were detected between wheat, barley and rye. Using barley addition lines, barley specific bands were mapped to individual barley chromosomes (7H).

A second approach is to use semi-random primers to target diverse regions of the genome. For this purpose the consensus sequences at the intron-exon splice junctions (ISJ) were used. A two-tiered PCR amplification was designed for the ISJ-PCR reaction. The targeting of the intron-exon splice junctions in conjunction with primers of random and defined sequences, such as α-amylase, provides a source of extensive variation in PCR products. These polymorphisms can be mapped as standard genetic markers. The major advantages in using the ISJ sequences to generate primers are that the core of the junctions are highly conserved, the primers matched for the 3 bases at the 3' end was within the highly conserved core of the consensus sequence; that

introns are present in most plant genes and one can, therefore, target an extremely diverse range of genes; that the ISJ approach avoids targeting heterochromatic regions and is likely to get a good chromosomal distribution of polymorphisms; and that the cost of PCR is reduced through the use of the ISJ primers since only one addition primer is required for each reaction and the ISJ primer can be produced in large scale synthesis.

To be useful in linkage mapping programs, markers have to be polymorphic between varieties of a given species. ISJ-PCR was therefore tested on a whole range of barley and wheat varieties. The banding patterns of the PCR products could be used to detect polymorphisms between individual plants and cultivars. This also offers a simple and efficient tool for varietal identification. Identification of cereal varieties has been achieved with the ISJ-PCR based method. Varietal polymorphisms were revealed through the PCR product profiles with ISJ and/or random primers. A maximum of only three primers allows the irrevocable identification of every single one of 30 barley cultivars, which includes all major commercially grown barley varieties in Australia. Nearly all of the individual plants within one barley variety displayed an essentially identical PCR banding pattern.

The results indicate that there is less intervarietal polymorphism in wheat compared that of barley. This phenomenon is considered to be associated with the recent origin of bread wheat. Rye varieties showed the highest intravarietal polymorphism among the cereals tested which is in agreement with the fact that rye is an outbreeder, whereas wheat and barley are inbreeders.

Genomic DNAs isolated from barley seeds were used successfully as templates for PCR reaction. The inhibiting effect from unknown substances present with seed DNA was overcome by the alteration of template and *Taq* polymerase concentrations in the PCR reactions.

The development of techniques for the generation of DNA markers from chromosome segments was described. One of the aims was to identify DNA markers to assist the transferring of CCN resistance from rye to wheat. Cloning of 6-PGD (6-phosphogluconate dehydrogenase) gene in cereals was attempted through PCR with primer derived from DNA and protein sequences of 6-PGD genes from organisms other than cereals. Another aim was to obtain clones from a segment in chromosome 1R of rye. The process of achieving this goal is to clone the flanking regions of rye-specific repetitive sequence in 1R with the combination of PCR and Southern analysis. A number of clones resulted from this approach were used as RFLP probes and mapped to individual rye or barley chromosomes. However, none of the clones tested were mapped to rye chromosome 1R possibly due to the size of the library and the number of clones tested.

Comparison was made between the results obtained from PCR and RFLP. It indicates that PCR can be used to detect and map polymorphisms in cereals. These markers are cheaper and faster to assay than RFLP's. They are technically less demanding to use than RFLP's and, as with standard PCR, do not require radioactivity to detect. The prospect for future work on the analysis of cereal genomes and the application of different techniques, especially PCR, is also discussed.

Chapter 1
LITERATURE REVIEW

1.1. Aspects of molecular genetics in cereals

1.1.1. Introduction

Cereal grains are an important dietary source for the world's population, contributing 70 and 50% of the total calories and protein, respectively (Eskin, 1990). The main cereals are wheat, maize, barley, rye, oats and rice, all members of the Gramineae family. Amongst these cereals, wheat, barley and rye are part of the tribe of Triticeae and, presumably, these now distinct genomes were all derived from a single ancestral species with seven pairs of chromosomes. Their chromosomes have also retained much of the same gene content (see review by Hart, 1987). This grass tribe also contains many wild related species which are potential gene donors to cultivars.

1.1.2. Wheat

Bread or common wheat, *Triticum aestivum* (2n=6x=42, genomes AABBDD), is an allohexaploid containing three genomes designated A, B and D, each of which consists of seven pairs of chromosomes. The A and D genomes were contributed by two wild grasses, *T. monococuum* and *T. tauschii* (*Aegilops squarrosa*) (for review see Kimber and Sears, 1987). The source of the B genome is still a matter of controversy (Morris and Sears, 1967;

Kimber and Feldman, 1987). It is likely that the original B-genome donor is extinct. Alternatively, the B-genome donor might have been substantially modified through introgressive hybridization with related taxa (Gill and Chen, 1987). As each genome has seven chromosomes, one from each of the homoeologous group and each homoeologous group has three chromosomes, the absence of a particular pair of chromosomes from one genome can be compensated by the corresponding dose increase of chromosomes from the other two genomes, which belong to the same homoeologous group (Sears, 1952; 1966).

Although the nuclear DNA content of bread wheat differs in various reports, according to Bennett (1972) the bread wheat genome has 18.1 pg of DNA per haploid nucleus, which is equivalent to 17.5×10^6 kb (May and Appels, 1987b). The DNA content in the A, B and D genomes were found to be in the ratio of 1.21:1.36:1.00. It has been shown that 22.4% of the total cytosine residues are methylated in wheat DNA. This methylation may reflect regulation of gene expression (for review see May and Appels, 1987b). Over 75% of wheat DNA is highly or moderately repeated (Flavell *et al.*, 1987).

There are different methods that have been used for studies of wheat and its close relatives at the DNA level. Harlan *et al.* (1973) examined the evolution and species relationship on the basis of gene pools as this method provides a good insight into genome organization. It seems that in most cereals evolution has taken place almost entirely within the primary gene pools. One of the widely used techniques is DNA-DNA hybridization in solution followed by a study of deviation in the melting temperature (Tm) in the heteroduplex thus formed. This technique has been utilized to obtain information on the potential donor of B genome in bread wheat (Nath *et al.*, 1983) and tetraploid wheat (Thompson and Nath, 1986), and that of the G genome in *T. timopheevii* (Nath

et al., 1985). It has also been used to study the relationship of four cereals, wheat, rye, barley and oats (Smith and Flavell, 1974). By using the technique of reassociation analysis in the studies of wheat and its relatives, it was shown that about 80% of the DNA is found in the form of repeated sequences; 20% consists of low copy number or unique sequences and only 1% (May and Appels, 1987b) or even less than 0.5% (Flavell *et al.*, 1987) consists of actual coding genes.

In addition to subdividing the genome on the basis of reassociation analysis, buoyant density fractionation in caesium salts has been a valuable analytical tool. In wheat, 5S rRNA genes, rRNA (28S, 18S) genes and satellite DNA have been separated from the bulk DNA through the use of actinomycin-D or Ag^{2+}/Cs_2SO_4 or Hg^{2+}/Cs_2SO_4 density gradients (Gerlach and Bedbrook, 1979; Gerlach and Dyer, 1980; Ranjekar *et al.*, 1978; Appels *et al.*, 1978; Dennis *et al.*, 1980). Restriction endonucleases provide another method of separating and classifying various classes of DNA sequences (Appels *et al.*, 1981).

1.1.3. Barley

Cultivated barley (*Hordeum vulgare* , 2n=2x=14, genome HH) is the fourth most-widely grown cereal on a global scale (FAO, 1982); in many temperate climates it is the most important crop. Barley is also one of the leading experimental organisms in genetic studies of flowering plants. This may be attributed to its diploid nature, low chromosome number, world-wide distribution, predominant self-fertilization yet ease of artificial hybridization, relatively large chromosomes which allow detection of several kinds of chromosome aberrations, and numerous easily classified hereditary characters

(Allard, 1988). The barley genus *Hordeum* includes over 30 species which are spread throughout temperate Eurasia, America, Africa and Australia (Bothmer, 1991). Barley has one near wild relative, *H. vulgare* sp. *spontaneum* (wild barley) which has generally been regarded as the progenitor of cultivated barley (Zohary, 1969). Jorgenson (1986) examined isozyme variation among 32 species of Hordeum for six enzyme systems representing 12 loci. This analysis divided the genus into three groups, one of which (the "vulgare" group) contains *H. vulgare* (both subspecies), *H. bulbosum* and *H. murinum.* The remaining 29 species showed only a distant relationship to cultivated barley.

Bennett and Smith (1976) found that four barley cultivars have a 1C DNA value of 5.5 pg, which equals 5.3×10^9 bp for the haploid barley genome. Sequence data for barley are as yet very limited and provide only a few comparisons of the same gene sequence from different cultivars (Wettstein-Knowles, 1991). One such comparison is that for the Adh1 gene (Trick *et al.*, 1988) as a partial genomic clone from an EMS-mutant of the malting cultivar Proctor, with a cDNA clone of the same gene from the naked cultivar Himalaya (Good *et al.*, 1989). These data point to substantial nucleotide differences between barley cultivars in the coding portion of the genome. In contrast, the base sequence of a cDNA clone for chloroplast glutamine synthetase from cv. Maris Mink (Freeman *et al.*, 1990) is identical with that of a cDNA clone from cv. Carina (Baima *et al.*, 1989).

Molnar *et al* (1989) investigated tandemly repeated DNA sequences containing structural genes encoding ribosomal RNA (rDNA) in 25 species of Hordeum using the wheat rDNA probe pTA71. Intraspecific variation was found in *H. parodii, H. spontaneum* and *H. leporinum,* but not in *H. bulbosum.* Repeat unit length variation was attributed to the intergenic spacer region. The repeat unit of *H. bulbosum* differed from all other species by the

presence of a Hind III site. The closer relationship of *H bulbosum* to *H. leporinum*, *H. murinum* and *H. glaucum* than to *H. vulgare* was indicated by their Bam HI restriction maps.

1.1.4. Rye

Rye, *Secale cereale* (2n=2x=14, genome RR), is the only cultivated species of the genus *Secale*. Its agricultural importance is in part based on the better resistance to winter climatic conditions and other adverse environmental factors, which make it difficult to cultivate other less resistant cereals such as wheat. Furthermore, rye is one of the parentals of the triticales, allopolyploids originated from wheat x rye hybrids. The close genetic relationship between wheat and rye chromosomes has been well established (see Gupta, 1971 for review).

Rye is a diploid species with 7 pairs of chromosomes, numbered 1R to 7R. Extensive studies on DNA content in several genera of the Triticinae by Bennett and Smith (1976) have shown that the DNA content of *Secale* genomes, having the 1C DNA content for 9.5pg, is the same or greater than that of Aegilops and the cultivated wheat. The *S. cereale* genome has about 33% more DNA than the largest of the diploid wheat genomes. The relationship between C-heterochromatin and several families of repeated DNA sequences was examined by *in situ* hybridization. These studies indicated a strong correlation between the location of such families and telomeric C-bands in *Secale* species (Hutchinson *et al.*, 1981). The occurrence and distribution of tandem arrayed DNA in different *Secale* species suggest that *S. silvestre* and its progenitor may be of ancient origin, whereas *S. cereale, S. montanum, and S. vavilovii* may have a more recent origin (Jones and Flavell, 1982b).

Chloroplast (ct) DNA variation in five *Secale* species (*S. cereale, S. africanum, S. montanum, S. silvestre and S. vavilovii),* was studied by using restriction enzymes (Murai *et al.*, 1989). The ct DNA of all the species except *S. silvestre* produced identical patterns with three restriction enzymes applied, also suggesting early separation of *S. silvestre* from the rest of the species.

1.1.5. Repetitive sequences

In contrast to simple organisms like *Escherichia coli*, eukaryote genomes contain large amounts of sequences present in multiple copies. These repetitive sequences are an ubiquitous feature of eukaryotic genomes (for a review see Flavell *et al.*, 1974) and an understanding of their structure and organization plays an important role in the analysis of chromosome structure in general. Repetitive sequences have been supposed to have structural meaning, i.e. in the architecture of the interphase nucleus, meiotic chromatin and/or the metaphase chromosomes (Studer and Epplen, 1990). Changes in repeated sequences are responsible for most of the changes in chromosome size and structure during the evolution and divergence of species (Flavell, 1983).

Cereal genomes contain a very high percentage (>75%) of repeated sequences; this proportion is even higher in *Secale cereale* (Flavell *et al.*, 1974). Different classes of repeated DNA have been observed, from very highly repetitious and palindromic sequences, which can account for 5-10% of the cereal genomes (Appels and McIntyre, 1985), to moderately repeated elements. In oats, barley, wheat, and rye, 26% of the repetitive DNA sequences of rye hybridize to all four species, 26% only to barley, wheat, and rye, 18% only to wheat and rye, and 30% only to rye (Flavell *et al.*, 1977). A relatively small fraction (2-10%) of the genome consists of long stretches (>10 kb) of

single copy DNA (Rimpau et al., 1978, 1980). It has been argued, however, that the amount of DNA in the low copy sequence class is consistently overestimated with a proportion representing diverged sequences rather than coding DNA (Flavell *et al.*, 1980).

As in other eukaryotic genomes, repeated sequences in plant genomes are generally divided into tandem arrays and dispersed repeats (see Smyth, 1991 for a review). Unequal crossing-over is believed to be responsible for the generation, maintenance, and variation of the tandemly repeated sequences (Smith, 1976). Whereas the origin of the dispersed sequences could have been transposons and retroelements (Smyth, 1991). The repeated sequences in the first group are simple, tandemly repeated sequence arrays present in centromeric and telomeric heterochromatin. They are normally transcriptionally quiescent. Tandem arrays of repeats appear to be the molecular basis of heterochromatin (John and Miklos, 1979; Flavell, 1980). The major tandem arrays of repeats in rye were shown to be associated with the telomeric blocks of heterochromatin (Bedbrook *et al.*, 1980; Jones and Flavell, 1982a). Tandem arrays displayed frequent variation in the number of repeats between individuals in rye (Jones and Flavell, 1982a, b). The remaining repeated DNA falls into dispersed repeats. In mammals, they are further classified as the short and the long interspersed nucleotide elements (SINEs and LINEs) simply on the basis of their physical lengths (Studer and Epplen, 1990). One dispersed repeat in wheat was found to have a terminal sequence that is identical to the conserved inverted terminal repeat mammalian retroviruses (Flavell *et al.*, 1981). Many of the dispersed repeated sequences present in wheat, barley and rye are much more closely related within a species than between species (Flavell *et al.*, 1977). It appears that the dispersed repeat families evolve "in concert" (Flavell *et al.*, 1986).

A 1.6 kb stem-loop insertion sequence (WIS1) was reported to lie upstream of an α-amylase gene in hexaploid wheat (Martienssen and Baulcombe, 1989). An unusual character of the WIS1 element is a series of short tandem repeats similar to animal minisatellite structures (Jeffreys *et al.*, 1985). The sequence of WIS1 shows that it has many of the features found in transposable elements, including target site duplication and terminal inverted repeats.

Gerlach and Bedbrook (1979) cloned the wheat rDNA units which were shown to be tandemly arranged sequences. The rDNA are mainly located on chromosomes 1B and 6B with a minor site located on chromosome 5D (Appels *et al.*, 1980; Flavell and O'Dell, 1974). The size of the repeat unit in rDNA of wheat, barley and rye, defined by a Eco RI site, is 9 to 10 kb in length (Appels *et al.*, 1980). 5S rRNA genes have been studied in the tribe of Triticeae (Gerlach and Dyer, 1980; Scoles *et al.*, 1988). As in other eukaryotes, they are located separately and not linked with other rRNA genes. The repetitive 5S rRNA genes are arranged as tandem arrays clustered at one or two sites per haploid genome. In Chinese Spring, there are about 10,000 copies of 5S rRNA coding sequences per nucleus (Gerlach and Dyer, 1980) and the repeat unit in both 4x and 6x wheat is small with two size classes arranged in tandem arrays (Appels *et al.*, 1980). The 5s rRNA genes in *Secale* sp. are arranged as tandem arrays of 460- and 480-bp repeating sequences (Peddy and Appels, 1989). The genes assayed in barley were not assigned to particular chromosomes (Appels *et al.*, 1980).

Due to the vast number and variation of the repetitive sequences, there is very little information concerning the molecular structure of the region of the cereal genomes flanking repetitive DNA sequences. According to Rimpau *et al.* (1978), in the case of rye specific sequences, 84% of the flanking sequences

are expected to be interspersed with repetitive sequences common to wheat and rye, whereas 16% should be single or low copy DNA.

1.1.6. Alien genes in wheat improvement

Because wheat is one of the most important crops in the world and the genetic variations have been diminished to an alarming extent within the pure line varieties of cultivated wheat, breeders are turning their attention to the progenitors of wheat and its distant relatives (Sears, 1970). In the genus *Triticum* and in other genera (including *Agropyron, Elymus, Elytrigia, Haynaldia, Hordeum,* and *Secale*) which are more distantly related to wheat, there is a broad range of species that could be used as a supplementary source of useful genes for wheat breeding (see review by Islam and Shepherd, 1991).

Gene transfers from species having at least one genome in common with hexaploid wheat can be achieved by homologous chromosome pairing and recombination. For example, eyespot resistance has been incorporated into the winter wheat Rendezvous from the D genome in *Aegilops ventricosa* Tausch (Day and Lupton, 1987). As the *Ph* locus on wheat chromosome arm 5BL restricts free pairing between wheat and alien chromosomes (Sears and Okamoto, 1958; Riley and Chapman, 1958), wheat breeders have been achieving homoeologous pairing through removal of 5B (nullisomy), utilizing a recessive mutant of the *Ph* locus on 5B, or suppression of the 5B effect by the genome of *Ae. speltoides* / *Ae. multica*. Ionizing radiation has been successfully used to transfer alien chromosomes despite the technical difficulties and low success rates (see reviews of Knott, 1987; Islam and Shepherd, 1991).

Genetic variation from cereal rye is a valuable alien source utilized in wheat breeding. The main objective of wheat breeders in utilizing rye germplasm has

been to breed for high yield, high quality and disease resistance. Whole genomes (Triticales), individual chromosomes (addition or substitution lines), or chromosome arms (translocation lines) have been introduced into wheat (for review see Sharma and Gill, 1983; Gale and Miller, 1987). Rye chromosome arm 1RS is of particular interest as it confers resistance to several foliar diseases (listed in Koebner and Shepherd, 1986) along with wide adaptation and high yield performance (Rajaram *et al.*, 1983). The 1BL.1RS translocation has been included in numerous widely grown cultivars of wheat (Islam and Shepherd, 1991). 1AL. 1RS translocation lines in Chinese Spring wheat background involving (Singh and Shepherd, 1988), a 1BL.1RS translocation in Gabo wheat background (Koebner and Shepherd, 1986) and a 1DL.1RS translocation in Chinese Spring (Shepherd, 1973) have been isolated. However, the wheats carrying these translocations generally produce a flour with serious quality defects in bread making (Dhaliwal *et al.*, 1987). To break the linkage between the desirable and undesirable traits, recombination between 1RS and 1DS was induced using Sears' (1977) ph1b mutant or nullisomic 5B stock, both of which allow rare recombination between wheat and rye chromatin (Koebner and Shepherd, 1986). A number of recombinants have been identified using protein markers.

A rye gene for resistance to cereal cyst nematode (CCN) (*Heterodera avenae* Woll) has been transferred from triticale to wheat through backcrossing (Asiedu, 1986). The resistance is controlled by a single major gene and Asiedu *et al.* (1990) located this gene on the long arm of rye chromosome 6R. Attempts are now being made to transfer this gene to wheat through induced recombination using Sears' *ph1b* mutant, so that this gene may be utilized in the wheat crop. Dundas *et al* (1988) reported several translocations between rye chromosome arm 6RL and wheat chromosome 6D, using dissociation of GOT,

6-PGD and α-amylase isozyme markers located on the long arms of these chromosomes.

1.1.7. RFLP mapping

1.1.7.1. Two types of maps

There are essentially two types of maps for a particular organism. The first, the genetic linkage map, shows the arrangement of genes and markers along the chromosomes as calculated by the frequency with which they are co-inherited. A map unit of linkage maps is a centimorgan (cM) which corresponds to an observation of recombination in 1% of the gametes in the samples (Ott, 1985). The second, termed a physical map, is an actual representation of the chromosomes, providing the physical distance between landmarkers on the chromosome, ideally measured in nucleotide bases (Stephens *et al.*, 1990). The physical map in human is a assembly of different techniques. Sequence-tagged-sites have been proposed to coordinate information from different mapping strategies (Olsen *et al.*, 1989).

DNA polymorphisms can be defined as DNA sequences that vary between homologous chromosomes (Weber, 1990). The primary application of DNA polymorphisms is in the construction of linkage maps and the mapping of genes. Polymorphic DNA markers have become an important tool for linkage studies, and much attention has been given to producing linkage maps of these loci. DNA markers also serve as a bridge between physical and genetic maps. A wide variety of techniques have been developed for the analysis of DNA polymorphisms. The most widely used method, restriction fragment length polymorphism (RFLP) approach, combines restriction endonuclease digestion, agarose gel electrophoresis, blotting to a membrane and hybridization to a DNA

probe (Southern, 1975). The characterization of large numbers of classic RFLPs and other markers has led to the construction of many genetic linkage maps, including RFLP linkage maps for different plants such as *Arabidopsis* (Meyerowitz *et al.*, 1990), cabbage (Slocum *et al.*, 1990), soybean (Keim *et al.*, 1990), lettuce (Kesseli *et al.*, 1990), lentil (Havey and Muehlbauer, 1989), and tomato (Tanksley and Mutschler 1990). Genome projects, combining the two maps, are under way for many organisms with the aim of providing detailed molecular genetic information about fundamental biological processes (see Watson and Cook-Deegan, 1991 for review).

1.1.7.2. Molecular markers

Molecular markers have many advantages over morphological markers: they are generally silent in their effect on the phenotype, more markers and a larger number of alleles can be found, and most molecular markers show a codominant mode of inheritance (Beckmann and Soller, 1986). During the last two decades, isozymes have been used successfully in many aspects of plant breeding and genetics. This technique is both simple and cost-effective. Unfortunately, the number of genetic markers revealed by isozyme assay is insufficient for many breeding programs. Botstein *et al.* (1980) first proposed to use RFLP as genetic markers after their recognition by Kan and Dozy (1978). One of the foremost attributes of RFLP compared to isozymes is that a virtually limitless number of markers can be monitored in a single population. Spectacular success has been achieved by the discovery of RFLPs linked to a number of inherited disorders in humans (see review by Watkins, 1988). This technique has allowed the construction of high-resolution genetic maps for many complex genomes. These maps can be used to analyze quantitative trait loci (QTLs) (Paterson *et al.*, 1988; Keim *et al.*, 1990) and can increase the selection

efficiency for these traits (Stuber 1989; Lander and Thompson, 1990). It was reported that RFLP markers facilitated the transfer of qualitatively expressed genes, such as those conferring pest and disease resistance, from wild species into a cultivated background (Young and Tanksley, 1989). One of the potential applications of RFLP technique could be the analysis of genome organization and implementation of map-based cloning of tightly linked genes (Tanksley *et al.*, 1989).

RFLPs detect differences in homologous DNA sequences that alter the length of restriction fragments obtained by digestion with type II restriction endonucleases. These differences result from base changes or other rearrangements (e.g., translocations and inversions) at the recognition site of the restriction enzyme or from internal deletion/insertion events. The restriction fragments are separated according to their size by agarose gel electrophoresis. A Southern blot analysis is subsequently performed: the DNA fragments are transferred to and immobilized on a nylon membrane or a nitrocellulose filter, hybridized to a labelled DNA probe, and visualized by autoradiography (Southern, 1975). Generally, unique or low copy number DNA sequences are preferably used as probes and labelled with ^{32}P. The sources of probes can be cloned DNA of specific genes, cDNA, and genomic DNA. Mapping with RFLPs involves the Southern technique and uses probes to follow the segregation of homologous regions in individuals from segregating populations such as an F2 or backcross.

At present, large scale use of RFLP markers is severely limited by the significant costs, technical complexity, and time required for RFLP assays. Moreover, the multi-step procedure of RFLP detection proves to be very difficult to be automated by the use of machines and robots. As new classes of molecular markers such as oligonucleotide polymorphisms (Beckmann, 1988)

and PCR-derived polymorphisms become available, the future prospects of marker-based techniques for both researchers and breeders are more promising.

1.1.7.3. RFLP mapping in cereals

Genetic manipulation of agronomically important traits in cereals will be greatly enhanced by the development of comparative genetic maps utilizing DNA markers. Such maps will make it possible to evaluate the contribution made to a trait of the same chromosome region in different species and thus to select the best source for gene introgression into cereal cultivars (such as disease resistance, or stress tolerance). The DNA markers will also greatly assist in the introgression of selected genes into cultivars and in the development of new cultivars. Additionally, DNA markers will serve cereal breeders as selection markers for traits with complex inheritance and low heritability (ITMI, 1989). Maps based on DNA markers will greatly facilitate the comparable mapping in Triticeae.

Wheat genetics has lagged behind that of maize, tomato, and barley because the triplication of gene loci greatly complicates isolation of morphological mutants and their genetic analysis. Additionally, the number of polymorphisms for most enzymes is insufficient to construct linkage maps. The use of RFLP markers eliminates these barriers and, if combined with cytogenetic manipulation of wheat genomes, it could greatly improve the status of wheat synteny and linkage maps (ITMI, 1989). However, the development and use of RFLP markers in wheat has been slow because of limited polymorphism (Chao *et al.*, 1989; Kam-Morgan *et al.*, 1989; Liu *et al.*, 1990) probably due to its relatively recent origin (Bell, 1987). An RFLP linkage map

of homoeologous group 7 chromosomes of wheat has been constructed (Chao *et al.*, 1989). This map is based on several intervarietal crosses and contains 18 RFLP, three isozyme and seven phenotypic markers. RFLPs were approximately three times more common in the 7B chromosome than in either the 7A or 7D chromosome. Conserved gene order was observed on the 7A, 7B and 7D chromosomes.

The polyploid nature of bread wheat means that it is able to tolerate aneuloidy, and this has enabled the assemblage over the last 40 years of complete sets of nullisomc, monosomic, tetrasomic, and the telocentric stocks of the standard variety Chinese Spring (Sears, 1954). Wheat can also be crossed to many of its relatives in the Triticeae, and this has resulted in the isolation of sets of alien chromosome addition lines in which single pairs of chromosomes from an alien relative are added to the wheat complement (see review by Gale and Miller, 1987). These aneuploid and addition line stocks allow the DNA probes to be assigned to individual chromosomes or their arms, and hence to a region of the linkage map before detailed linkage analysis is carried out. The close relationship between wheat and other Triticeae means that their chromosomes fall into seven homoeologous groups, each group containing chromosomes of the various species with similar genetic content (Gale and Miller, 1987). The aneupoid stocks of wheat and alien addition lines have been used to locate more than 100 genes encoding enzymes and endosperm storage proteins and a few other molecular markers in chromosome arms (Hart, 1987; Gale and Hart, 1987; Blake *et al.*, 1989). Sharp *et al.* (1989) produced a set of 14 cDNA probes that mark each arm of the basic set of seven chromosomes of the genera in the Triticeae. In another study, eight cDNA clones were located on specific arms (6 on the short arm and 2 on the long arm) of homoeologous group 2 chromosomes with the help of alien addition lines, nullisomic-tetrasomics and the ditelocentrics (Sharp and Soltes-Rak, 1988).

Kam-Morgan *et al.* (1989) initially proposed that linkage mapping may be most efficiently done in *T. tauschii*, the D-genome progenitor of wheat. They located five RFLP loci on chromosome 5D of common wheat which were linked in two blocks in *T. tauschii*. In addition, they found extensive polymorphisms in *T. tauschii*, whereas RFLPs were scarce among wheat cultivars. Gill *et al.* (1991) constructed a detailed linkage map of *T. tauschii* with the majority of loci mapped with an F2 population and the rest with aneuploid stocks. Deletion/insertions and point mutations were found to be equally important for generating RFLPs. Langudah *et al.* (1991) used a similar approach to map the *T. tauschii* genome with genomic DNA clones plus clones from other laboratories containing either anonymous sequences or genes coding for known products.

Genetic mapping in barley was initiated with conventional genetic linkage analysis, and the first linkage map with seven linkage groups was developed by Robertson (1939). Despite the widespread cultivation of barley and its elaborate genetic map, which comprises mostly morphological, isozyme and disease-resistance markers (Sogaard and von Wettstein-Knowles, 1987; O'Brien, 1990), knowledge regarding the application of RFLP markers in this species is limited (Bunce *et al.*, 1986; Blake, 1987; Blake *et al.*, 1989). On the other hand, *in vitro* techniques such as anther or microspore culture are well developed in barley. Their combination with marker-assisted selection might accelerate the breeding process considerably. Graner *et al.* (1990) examined the extent of RFLPs among 48 diverse barley varieties with 23 genomic single-copy probes. The mean diversity index (or average chance that any two of these varieties differ for a random probe) was 0.43. The average degree of polymorphism revealed by a restriction enzyme was positively correlated with average length of fragments produced by the enzyme. This and the failure of other enzymes to uncover additional variation, suggest that most of the RFLPs

were due to insertion/deletion differences. Shin *et al.* (1990) constructed a partial barley map with isozyme, morphological and RFLP markers. Recently, Jahoor *et al.* (1991) isolated more than 100 genomic DNA probes from barley and assigned them to the individual barley chromosomes using wheat/barley addition lines. Heun *et al.* (1991) constructed a RFLP linkage map for barley with 155 RFLPs and two known genes (*hulless* and *Mla 12* powdery mildew resistance). Linkage groups were assigned to chromosomes through disomic wheat/barley addition lines. They found that the number of RFLPs mapped to each chromosome varied considerably, for example, 24.5% to chromosome 1 compared with 5.8% to chromosome 6. There was no difference in the distribution of markers from either cDNA libraries or genomic libraries. The total genetic map is ~~average genetic distance between the mapped markers is~~ 1096 cM, compared with the current tomato maps (1600 cM) and rice maps (2300 cM) (Heun *et al.*, 1991).

Compared to the number of studies that have been carried out on maize, wheat and barley, those providing linkage data in rye are relatively scare (e.g. Lawrence and Appels, 1986). This may be due to the restricted economical importance of this crop as well as to the fact that *Secale cereale* is an outbreeding species with a very effective self-incompability system that may complicate genetic studies. More than 120 morphological and biochemical traits in rye have been localized onto individual chromosomes (for a review see Schlegel *et al.*, 1986). Wehling (1985) and Wehling and Uphoff (1990) presented an isozyme linkage map of genetically analyzed loci which includes all of the seven rye chromosomes and which enabled the marking and chromosomal assignment of the two self-incompatibility loci (Wricke and Wehling, 1985; Gertz and Wricke, 1989). Because of the importance of 1R in wheat-rye introgression, there have been extensive mapping studies on this rye chromosome through the use of conventional and DNA markers (reviewed by Baum and Appels, 1991). For

example, Gustafson *et al.* (1990) physically mapped *Sec-1* and *Sec-3* loci using *in situ* biotin-labelled probes. The *Sec-1* locus was mapped to the satellite region of the short arm of 1R, adjacent to the *Nor* locus. Recently, Wang *et al.* (1991) constructed a linkage map of 1R by using 16 molecular and biochemical loci. They found that loci in the 1R map clustered around the centromere, a phenomenon similar to that reported in wheat homoeologous group 7 chromosomes (Chao *et al.*, 1989).

1.2. Application of the polymerase chain reaction to genome analysis

1.2.1. Introduction

The polymerase chain reaction (PCR) involves the enzymatic amplification of DNA *in vitro* and was originally developed at the Cetus Corporation (Saiki *et al.*, 1985; Mullis and Faloona, 1987). The reaction is based on the enzymatic amplification of a DNA fragment that is flanked by two oligonucleotide primers that hybridize to opposite strands of the target sequence. The primers are oriented with their 3' ends pointing towards each other. Repeated cycles of heat denaturation of the template, annealing of the primers to their complementary sequences and the extension of the annealed primers with a DNA polymerase, result in the amplification of the segment defined by the 5' ends of the PCR primers. Since the extension products of each primer can serve as a template for the other primer, each cycle essentially doubles the DNA fragment produced in the previous cycle. This results in the exponential accumulation of the specific target fragment, up to several millionfold in a few hours.

The Klenow fragment of *Escherichia coli* DNA polymerase I was originally used as the DNA polymerase but it has since been replaced through the

use of thermostable *Taq* polymerase, isolated from *Thermus aquaticus*, which can withstand repeated heating up to 94°C (Saiki *et al.*, 1988). The availability of *Taq* polymerase has greatly simplified the automation of the reaction. Currently there are many different commercially available thermocycling instruments suitable for PCR and a number of published descriptions of home-made apparatus (Foulkes *et al.*, 1988). The specificity of the amplification reaction is increased by allowing the use of higher temperatures for primer annealing and extension. In addition to these improvements, the development of automated oligonucleotide synthesis has also promoted wide acceptance and use of PCR (Caruthers *et al.*, 1987). *Taq* polymerase has no 3' to 5' exonuclease activity, but has a 5' to 3' exonuclease activity during polymerization. The initial estimation of the mis-incorporation rate by *Taq* polymerase during PCR was about 10^{-4} nucleotides per cycle (Saiki *et al.*, 1988). Changes in the PCR conditions have since reduced the mis-incorporation rate to less than 10^{-5} nucleotides per cycle (Eckert and Kunkel, 1989; Gelfand and White, 1990). Many of the new thermostable polymerases have additional useful activities. The thermostable Vent™ DNA polymerase, an enzyme isolated from *Thermococcus litoralis*, has the 3' to 5' exonuclease activity, and may therefore have a lower mis-incorporation rate (Cariello *et al.*, 1991). The thermostable DNA polymerase from *Thermus thermophilus (Tth)* can reverse-transcribe RNA efficiently in the presence of $MnCl_2$ at high temperature (Myer and Gelfand, in press). In addition, these new thermostable polymerases may be more resistant to inhibiting components in PCR reaction that hamper the use of *Taq* polymerase (Panaccio and Lew, 1991).

In many circumstances, PCR has advantages over conventional procedures for cloning and analysis because it is faster, simpler, and more flexible (Erlich *et al.*, 1989; Erlich, 1989). One of the virtues of PCR is that it

allows minute amounts of the template DNA to be detected, even down to a single template molecule (Li *et al.*, 1990).

1.2.2. Use of PCR to detect DNA polymorphisms

Skolnic and Wallace (1989) described theoretically three classes of amplified sequence polymorphisms (ASPs) by PCR: restriction site polymorphisms, sequence length polymorphisms, and DNA base pair changes not associated with restriction sites. The simplest use of the PCR in analysing DNA polymorphism is in ascertaining its presence or absence. In human genome analysis, there have been numerous reports of using PCR to replace RFLP analysis when enough sequence information is available (Stephens *et al.*, 1990). The usefulness of this approach was also demonstrated in wheat (D'Ovidio *et al.*, 1990) with primer sequences derived from the sequence of a γ-gliadin gene, and recently in rice (Williams *et al.*, 1991) with primer for the ends of mapped genomic clones. An extension of this approach is the amplification of several DNA regions simultaneously, a so-called mutiplex PCR as demonstrated by Chamberlain *et al.* (1988) in analysis of the dystrophin locus.

To identify new polymorphisms in previously defined genes, Dean *et al.*, (1990) have been using primers 200 to 1000 bp apart to amplify gene sequences. The resulting PCR products are analyzed in two different ways. In one approach, restriction site and insertion/deletion polymorphisms are identified by digesting the amplified DNA with enzymes that cut frequently, and resolving the resulting fragment on a polyacrylamide gel. The second approach involves incorporation of radioactivity into amplified PCR products, followed by

denaturation and separation on polyacrylamide gels to detect single stranded conformation polymorphisms (SSCPs).

Another PCR-based methodology used to detect single base change polymorphisms involves the combined use of PCR and denaturing gradient gel electrophoresis (DGGE) (Myer *et al.*, 1986). This procedure involves the amplification of a target region (200-500 bp) and analyzing the PCR products using DGGE. Single base pair polymorphisms may be detected by the differential migration on this gel system. The use of a GC clamp (a short string of GC nucleotides incorporated into the ends of PCR product), or a modification of DGGE that involved restriction enzyme digestion of the PCR product before gel electrophoresis, enhances the ability to detect single base pair polymorphisms using DGGE (Sheffield *et al.*, 1989).

1.2.3. Amplifying unknown sequences

Although it is normally necessary to know enough sequence information to design the primers for PCR, this method can be used to amplify and analyse unknown intervening sequences. Using evolutionarily conserved portions of a gene for primer construction will allow the amplification of highly variable sequences that lie between the conserved templates for these primers (Kocher *et al.*, 1989). Similarly, from information on amino acid sequences that are conserved among different viruses, a degenerate primer mixture can be constructed that can be used to search for novel viruses that retain significant similarity at the amino acid but not the nucleotide level (Mack and Sninsky, 1988).

One strategy to overcome the limitation of sequence requirement is to create a primer binding site by adding DNA of a known sequence. This was illustrated

initially by the amplification of unknown cDNA sequences cloned into λ gt11 with primers for the vector sequences that flank the insertion site (Saiki *et al.*, 1988). In general, flanking sequences for priming are added by ligation (Kinzler and Vogelstein, 1989) or homopolymer tailing with terminal transferase (Loh *et al.*, 1989).

Another approach for the characterization of fragments adjacent to known sequences is PCR following ligation-mediated circularization (inverse PCR) (Triglia *et al.*, 1988; Ochman *et al.*, 1988; Silver and Keerikatte, 1988). This technique involves the digestion of a DNA fragment by a restriction enzyme and its subsequent ligation at a very low concentration in order to promote the formation of circular products. Judicious choice of restriction enzymes leads to the generation of a circle that contains the previously known sequence on a contiguous circular molecule with the unknown sequence. Primers are synthesized from the known sequence in an orientation that enables amplification around the circle and includes the unknown sequence.

1.2.4. PCR for genome mapping

The versatility and power provided by PCR methodology has encouraged its involvement in almost every aspect of genome projects. The application of PCR has been central in formulating both the conceptual and practical approaches on which the human genome project is based (Rose, 1991). Much of the current emphasis on PCR in genome mapping and sequencing is a direct result of the sequence tagged sites (STS) proposal introduced by Olsen *et al.* (1989). The STS concept requires the use of PCR to generate the 200- to 500-bp sites unique to a given genomic fragment, thus identifying the common element necessary to integrate various genetic and physical maps.

It now appears that the eukaryote genome is densely interspersed with simple tandemly-repeated motifs, termed "microsatellites" that exhibit site-specific variation (Beckmann and Soller, 1990). Poly (G) and poly (A) are the simplest of the microsatellites, while poly (TG) is the most frequent in the mammalian genome (Hamada *et al.*, 1982). Microsatellites are recognized as the preferred sites for performing DNA-based linkage studies because they are twice as informative and can be typed at twice the rate than that of RFLP (Weber, 1990). Analysis of microsatellite polymorphisms involves amplification by PCR of a small fragment of DNA containing a block of repeats followed by electrophoresis of the amplified DNA on denaturing polyacrylamide gel.

PCR has been used to amplify human *Alu* repeat units in isolating human DNA fragments present in the hybrid cell lines or from microdissected or flow-sorted chromosomal regions (Nelson *et al.*, 1989). The applications of *Alu* PCR now reach well beyond the initial aims of detecting human-specific sequences in somatic cell hybrids and the method has now been employed for cloning, mapping and walking strategies (Weber, 1990). An added bonus of the *Alu* PCR is that the ends of the *Alu* elements often exhibit polymorphisms between different individuals and can serve as useful genetic markers (Economou *et al.*, 1990).

One of the most powerful contribution of PCR to genetic mapping has been its use for analysis of DNA sequence polymorphisms in individual gametes, referred to as sperm mapping (Li *et al.*, 1988). This approach offers a significant advantage in that a large number of meiotic products can be rapidly examined from a single individual, allowing determination of recombination frequency between genetic markers that are physically very close.

In a way similar to the RFLP approach, PCR has been used with arbitrarily chosen primers for generating genetic markers, genome mapping and

DNA fingerprinting (Williams *et al.*, 1990; Welsh and McClelland, 1991a, b; Bassam *et al.*, 1991). The approach of Williams *et al.* (1990) and Bassam *et al.* (1991) relies on the amplification of genomic sequences with a short oligonucleotide, typically 8-10 nucleotides in length. The amplified products (RAPDs) segregate in a Mendelian fashion. Whereas Welsh and McClelland's (1991a, b) method involved two cycles of low stringency amplification followed by PCR at high stringency and the amplified products were separated on sequencing gels and visualized by autoradiography. Welsh and McClelland (1991c) also demonstrated that oligonucleotide primers can be used alone or in pairs to generate different PCR products, thus revealing more polymorphisms for DNA fingerprinting and genome mapping. Recently, Martin *et al.* (1991) identified DNA markers linked to a *Pseudomonas* resistance gene in tomato by using the RAPD method and isogentic lines. One of the drawbacks of these markers is that they are scored as dominant markers and heterozygotes cannot be distinguished (Welsh *et al.*, 1991), while RFLP markers are generally codominant and consequently able to identify heterozygotes (Burr *et al.*, 1983).

1.2.5. PCR-based sequencing

PCR can obviate the need for many of the laborious subcloning and preparative scale procedures currently required to obtain sufficient DNA for sequence analysis because DNA fragments of a length suitable for sequencing can be readily obtained by amplification from genomic or large insert cloned DNA. Several procedures have been developed for the direct sequencing of double-stranded PCR products (Egelke *et al.*, 1988; Innis *et al.*, 1988). PCR can be used to generate single-stranded DNA products to serve as sequencing templates by minor modifications of the amplification protocol (see review by Gyllensten, 1989). Similarly, conventional PCR has been performed with one

of the primers chemically modified with biotin. The resulting PCR products have one of the two strands specifically end-labelled with biotin that can be immobilized on an avidin-coated solid matrix for strand separation (Mitchell and Merril, 1989). Enzymatic sequencing reactions can be thermally cycled with a thermostable DNA polymerase similar to PCR amplification. This approach utilizes the dideoxy-terminator method of Sanger (1977) and a sequencing primer either specific to the fragment under study, or the universal sequencing primer utilizing primer binding sites introduced by cloning or PCR (Gyllensten and Erlich, 1988; Innis *et al.*, 1988). As for PCR products, there are three most common ways to sequence them: subcloning, asymmetric PCR, and direct sequencing.

A contig is a set of partially overlapping cloned DNA fragments that cover a stretch of the genome without interuption. Sequence contigs can be ordered by performing the PCR reaction with all or a subset of the pairwise combinations of primers generated from the end sequences of the contigs. Once the contigs are ordered and the gaps between the contigs are determined, the appropriate primer pairs can be used to generate the physical DNA necessary to fill in the gap sequences (Hunkapiller *et al.*, 1991).

In attempts to increase both the throughput and consistency of sequencing efforts, as well as decrease the overall cost, several automated DNA sequencers have been developed (for review see Hunkapiller *et al.*, 1991). The fluorescent sequencing is well established in the Applied Biosystems Inc. (ABI) machine that allows for all four DNA sequencing reactions to be electrophoresed together in a single gel lane (Connell *et al.*, 1987). Based on the same principal, ABI also developed a fluorescent DNA fragment analyzer which can accommodate four DNA samples in a single gel lane by the use of four different colour fluorescent dyes.

1.2.6. Contamination and its prevention

The extreme sensitivity of PCR makes the procedure very vulnerable to contamination due to minute amounts of material which could be products of a previous PCR reaction (carry over), exogenous DNA, or other biological material. One approach to inactivate the contaminating template involves short-wavelength ultraviolet irradiation of the reaction mixture prior to amplification (Sarkar and Sommer, 1990). An alternative inactivating method involves the photochemical modification of the amplified DNA, thereby blocking the *Taq* polymerase from further extension after it encounters a modified base in the template strand (Isaacs *et al.*, 1991).

1.3. Aims and scope of this thesis

Molecular markers, especially RFLPs, have become important tools for genetic research and practical breeding in the last decade. A large number of molecular markers are needed for genetic mapping and other analyses. As cereals are the most important food crops in terms of production and trade, research groups around the world have been generating and mapping DNA probes and other biochemical markers to the cereal genomes. RFLP, the most widely used DNA marker in humans, is also the main technique employed in cereal mapping programs. However, the high-cost and technical complexity associated with RFLP analysis have hampered its application. For the same reasons, plant breeders have been reluctant to accept and apply this technique. It is not surprising that there was little information available for RFLP analysis in cereals when the present study was initiated.

In the last few years, PCR has emerged as a promising technique in human genetic studies. In the area of plant molecular biology, PCR could serve a similar role but had not been utilized for mapping in plants before the commencement of this study. To realize its full potential in the genomic analysis of cereal genomes, a comprehensive study was needed for the evaluation of the usefulness of this technique. The objectives set out for this study are as follows:

1. The primary goal was to adapt PCR for genetic mapping in cereals. The new method should offer improved sensitivity and speed over RFLP. This technique is also expected to be less complex and does not require the use of radioactivity. PCR is a powerful technique for amplifying specific sequences from extremely small amounts of DNA. But the ideal technique for genetic mapping should be able to provide a near

infinite number of markers which map to all parts of the genome. Two types of primers have been chosen for the target region of PCR to generate genetic markers, those that target specific, known sequences and those that target intron-exon splice junctions.

2. The application of PCR for the investigation of intravarietal and intervarietal polymorphisms in cereals. Another aim was to develop a cheap, fast and sensitive method for cultivar identification with the basic criteria of distinguishable variation, reproducibility, minimal intravarietal polymorphism, and environmental stability.

3. The development of techniques for the generation of DNA markers from chromosome segments. One of the aims was to identify DNA markers, e.g. 6-PGD gene, to assist the transferring of CCN resistance from rye to wheat. Cloning of 6-PGD (6-phosphogluconate dehydrogenase) gene in cereals was attempted through PCR with primer derived from DNA and protein sequences of 6-PGD genes from organisms other than cereals. Another aim was to obtain clones from a segment in chromosome 1R of rye. The process of achieving this goal is to clone the flanking regions of rye-specific repetitive sequence in 1R with the combination of PCR and Southern analysis.

4. Improving the methods of PCR analysis. Automation of the detection of PCR products offers the prospects of improved sensitivity and speed. The Applied Biosystem's DNA sequencer and fluorescent dyes were chosen for the demonstration of this technique.

5. DNA isolated from a single or half seed is not suitable for normal Southerns. It seems that there are certain substances in the seeds also hampering the PCR amplification. The development of a method for the amplification of seed DNA was included in the present study. This

method would be particularily suitable for cereal breeders.

Chapter 2
MATERIALS AND METHODS

2.1. Materials

2.1.1. Chemicals

All the chemicals were of the highest purity available. The suppliers of special chemicals are indicated in the section describing their use.

2.1.2 Genetic materials

Bulk seeds of hexaploid wheat (*Triticum aestivum*) cv. Chinese Spring, the other wheat varieties and breeding lines, rye (*Secale cereale*) cv. Imperial and cv. South Australia, the seven wheat-rye addition lines (1R-7R), barley (*Hordeum vulgare)* cv. Betzes, and the six wheat-barley addition lines described by Islam *et al.* (1981) were obtained from Dr. K. W. Shepherd (Waite Institute). The triple translocation line (TTL) of wheat cv. Chinese Spring (C.S. 18II + 1AL-1RS +1BL-1RS + 1DL-1RS), in which the short arm of chromosome 1A, 1B, and 1D are all substituted by the short arm of chromosome 1R from rye cv. Imperial, was also provided by Dr. K. W. Shepherd. The barley 1H* addition line (Islam and Shepherd, unpublished) was very generously provided by Dr A.K.R.M. Islam (Waite Institute). The barley varieties and breeding lines were provided by Dr. D. Sparrow (Waite Institute). The nomenclature for the barley chromosomes is based upon the equivalent wheat homoeologous group (traditional barley chromosomes numbers are given in parentheses): 1H* (5), 2H (2), 3H (3), 4H (4), 5H (7),

6H (6), and 7H (1). It should be noted that the 1H* addition line is a double monosomic addition containing one copy each of 1H and 6H.

2.1.3. Sources and description of the probes

RW229/pMN1 was kindly supplied by Professor R Wolf, Dept. Biol. Science, Univ. Maryland Baltimore County, USA. 6-PGD-1 was a generous gift from Professor HC Towles, Dept. Biochem.,Univ. Minnesota, USA.

2.1.4. PCR machines

The PCR machines used in this study were DNA Thermal Cycler (Perkin Elmer Cetus); Intelligent Heating Block, Thermal Reactor, Combi Thermal Reactor TR2(Hybaid); FTS-1 Fast Thermal Sequencer (Corbett Research); DNA Amplifier GM-10 (Bartelt Instrument); Programmable Thermal Controller (MJ Research).

2.2. Laboratory Methods

2.2.1. Transformation of *E.coli* by plasmid DNA

Transformation of the *E. coli* strain JM109 was based on the method of Hanahan (1985). 500ml of JM109 cells at log phase (A600=0.45-0.55) were collected by centrifugation at 2500xg for 15 min at 4°C. The cells were resuspended in 17 ml of TFB buffer and kept on ice for 10 min. The cells were recentrifuged and resuspended in 4 ml of TFB. 140 ul of DMSO was added to the suspension and 315 ul of 1M DTT added 5 min later. After 10 min, 150 ul DMSO was added to the suspension was kept on ice for further 5 min. 10 ng DNA was mixed with 200 ul cell suspension and incubated on ice for 30

minutes. The cell suspension was 'heat shocked' at 42°C for 2 min. 800 ul of SOC was then added to the suspension and the cells were incubated in a 37°C shaker for 45 min. The cells were plated onto LB-agar plates with appropriate antibiotic, IPTG and X-gal. The plates were incubated overnight to establish colonies.

2.2.2. Minipreparation of the plasmid probes

A single *E. coli* colony containing a recombinant plasmid was used to inoculate 2ml of LB containing 5ul of antibiotics (10ug/ml). Cells were grown overnight at 37°C with vigorous shaking. The culture was transferred to an Eppendorf tube and spun for one min in a microfuge. The pellet was resuspended in buffer 1 (50mM glucose, 25mM Tris-HCl (pH8), 10mM EDTA), to which 10ul lysosyme (40ug/ml) was added, and incubated on ice for 20 min. Then 200ul of buffer 2 (0.2M NaOH, 0.1% SDS) was added, the tube was gently inverted twice and left on ice for 5 min. 150ul of buffer 3 (3M NaOH pH 4.8) was added, gently mixed and the solution was left to stand on ice. After 10 min the mixture was centrifuged for 5min; around 400ul supernatant was taken into a new tube and extracted with equal volumes of phenol/chloroform. After centrifugation for 2 min, an equal volume of isopropanol was added to the supernatant. The tube was inverted a few times and stood on ice for 15 min. The precipitate was collected by centrifugation for 5 min and the pellet was washed twice with 75% ethanol and dried under vacuum. The pellet was resuspended in 50ul TE containing RNase (final concentration 10ug/ml). If necessary, the solution was extracted once more with phenol/chloroform. The supernatant was precipitated by adding acid salt and isopropanol, freezing, spinning and washing in 75% ethanol as described above.

2.2.3. Large-scale DNA extraction from plant tissue

Large-scale DNA isolations were made from approximately 5 g fresh leaves. The leaves were ground to a fine powder under liquid nitrogen and then suspended in 100 ml extraction buffer (4% sarkosyl, 0.1 M Tris-HCl, 10 mM EDTA, pH 8.0). An equal volume of phenol/chloroform/isoamylalcohol (25:24:1) was added and the slurry mixed for 90 minutes at 4°C. After separation of the phases by centrifugation, the upper aqueous phase was poured off and filtered through fine nylon mesh. The solution was mixed with 10 ml sodium acetate (3 M, pH 4.8) and 250 ml ice cold ethanol. The DNA was scooped out with a spatula, washed twice with 70% ethanol and dissolved in 7 ml TE (5 mM Tris-HCl, 0.2 mM EDTA, pH 7.5). The DNA was further purified via CsCl equilibrium centrifugation (Maniatis et al. 1982).

2.2.4. Small-scale DNA extraction from plant tissue

Small-scale DNA isolations used 50-200 mg of young leaves. These were ground to a powder in 2 ml Eppendorf tubes under liquid nitrogen. The powder was then mixed with 1 ml extraction buffer and, subsequently, with 0.75 ml phenol/chloroform/isoamylalcohol . The whole mixture was shaken for 20 to 30 seconds and the aqueous phase recovered after centrifugation. The phenol/chloroform/isoamylalcohol extraction was repeated and the DNA precipitated by ethanol precipitation.

2.2.5. Isolation of bacteriophage λ DNA from plate lysates

Phage DNA was isolated by the standard method described by Maniatis *et al.* (1982).

2.2.6. Restriction digests, fragment isolation and ligation

Restriction enzymes were purchased from either Boehringer Mannheim, Pharmacia or New England Biolabs. The enzyme buffers were obtained from the manufacturer or made up at ten times the final concentration as recommended by them. DNA fragments were isolated with a GENE CLEAN (BIO 101) kit following supplier's instructions. Dephosphorylation of fragment ends was achieved by incubation of DNA with 1 unit of calf-intestinal alkaline phosphatase (Boehringer-Mannheim) for 30 minutes at 37°C in buffer as recommended by supplier. DNA ligations were carried out using T4 DNA ligase (New England Biolab) in the recommended buffer.

2.2.7. Gel electrophoresis

Electrophoresis separation of DNA fragment was usually in 0.8-1.5% agarose gels with TAE buffer. After electrophoresis, the gel was immersed for 15 min in the buffer containing 1 mg/l ethidium bromide and photographed by Polaroid camera fitted with a red filter under UV light (254nm).

After electrophoresis, the gel was immersed for 15 min in the buffer containing 1 mg/l ethidium bromide and photographed by Polaroid camera fitted with a red filter under UV light (254nm).

Some of the PCR products were fractionated by 10% polyacrylamide gels Polyacrylamide gels were prepared according to Maniatis *et al.* (1982) in vertical electrophoresis tanks made at the Waite Institute. The gel was prepared from 33 ml of acrylamide-bisacrylamide (30:1, w/w), 10 ml of 10xTBE buffer, 2.1 ml of 3% ammonium persulphate, 30 ul of TEMED (N,N,N',N'-

tetramethylethylene diamine), and 55 ml of H_2O. The gel was allowed to polymerize for 2 hour then pre-run at 35 mA for 15 min. Lanes were washed with buffer before loading the samples. The running buffer was also TBE and the current 35 mA. The gel was stained with ethidium bromide and visualized as described above.

2.2.8. Southern transfer of DNA fragment

Approximately 5 μg of total plant DNA was digested to completion with restriction endonucleases as recommended by the manufacturer, separated on 0.8-1.5% agarose gels in TAE buffer (0.04 Tris-acetate, 0.001 M EDTA). Following eletrophoresis, the DNA was partially hydrolyzed by soaking in 0.25M HCl for 20 min prior to alkali denaturation (0.4 N NaOH for 30 min) and high-salt neutralization (1.5 M NaCl, 0.5 M Tris-HCl, pH 7.5 for 30 min) and then transferred (overnight) with 10x SSC (1.5 M sodium chloride-0.15M sodium citrate) from the gel to the surface of Hybond N^+ membrane (Amersham) by the method described by the manufacturer, which was a modification of that originally described by Southern (1975).

2.2.9. Dot blot

DNA transfer was performed in a dot blotting apparatus made by Schleicher &Schuell following the manufacturer's recommendation. The DNA to be blotted was diluted to a final volume of 100ul with TE and denatured by adding 900ul of 0.4M NaOH and incubating at room temperature for 5 min. A piece of nylon membrane (Hybond N^+, Amersham) and a piece of filter paper (Whatman No.1) were cut to required size and soaked for 5 min in 0.4M NaOH. The filter paper was placed on the bottom section of the dot blot manifold and overlaid with the

nylon membrane. The top of the manifold was clamped into place and the device connected to a vacuum line. The denatured DNA was applied to the membrane via the wells in the top of the dot blot apparatus and washed twice with 500ul of 0.4M NaOH. The membrane was removed from the apparatus and rinsed twice in 10xSSC. The subsequent hybridization was carried out as described in section 2.2.11 (Hybridization and washing of membranes).

2.2.10. Labelling of DNA with ^{32}P

(A) Probes were radioactively labelled with [α-^{32}P]dCTP according to the oligolabelling procedure of Feinberg and Vogelstein (1983). Relevant DNA fragments were either isolated from preparative agarose gels and labelled with the "Oligo Priming Labelling Kit" from Amersham or they were cloned into pUC 19 and specifically labelled by using the forward and reverse sequencing primers instead of the mix of random primers in the labelling reaction. The labelled products were separated from unincorporated nucleotides on G75 (Pharmacia) minicolumns prepared in Pasteur pipettes.

(B) Cereal genomic DNAs were labelled with [α-^{32}P]dCTP by a "Nick-translation" kit from Bresatec (Rigby *et al.*, 1977). The labelled products were separated on the minicolumns described above.

2.2.11. Hybridization and washing of membranes

Hybridizations were carried out overnight at 65^{o}C in 0.6 M NaCl, 20 mM PIPES, 5 mM EDTA, 1% SDS, 0.2% gelatin, 0.2% ficoll, 0.2% polyvinylpyrroidone, 0.5% tetrasodium pyrophosphate, 500 ug/ml of carrier DNA, pH 6.8. The membranes were washed 20 min at room temperature in 2x SSC, 0.1% SDS, 20 min at 65^{o}C in 0.5x SSC, 0.1% SDS; and 20 min at

65°C in 0.2x SSC, 0.1% SDS. The membranes were then dried on blotting paper (Whatman No.1) and wrapped in Glad Wrap for autoradiography.

2.2.12. Autoradiography

The mounted membranes were placed with a sheet of X-ray films (Fuji, NIF) in a cassette fitted with intensifying screen. The cassette was kept at -70°C for 1-7 days and the film was subsequently developed manually or in a Agfa Curix 60 film processor.

2.2.13. Programs for PCR machines

The standard reaction consisted of 45 cycles each of 1.5 min at 95°C, 2 min at 55°C and 2 min at 72°C. With the intron splice junction primers and/or random primers the first 6 cycles were at 94°C for 1 min, 40°C 1.8 min, 72°C 2 min. This was followed by a further 28 cycles of 94°C 1 min, 58°C 1.5 min and 72°C 2 min. Alteration to these conditions are described in the text.

2.2.14. PCR reaction conditions

The polymerase chain reactions were carried out in a 25 or 100µl volume containing 0.2 - 0.5 µg of genomic DNA template, 0.2 µM of each primer, 200 µM each of dATP, dCTP, dGTP and dTTP, 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, 0.01% gelatin and 0.2 - 0.5 units of Taq polymerase (Perkin Elmer Cetus or Promega). The PCR was performed in one of the PCR machines described in 2.1.7. If not specified, Programmable Thermal Controller (from MJ Research) was used for the PCR amplification.

2.2.15. Method of inverse-PCR

This method is based on the inverse-PCR procedure by Ochman *et al.*(1988). Restriction digests were carried out using 5 ug of TTL genomic DNA treated with 10 units of Hind III (section 2.2.6). The digested DNA was extracted twice with phenol/chloroform/isoamylalcohol, and the DNA precipitated by ethanol precipitation.

For circularization, 0.1 ug of the Hind III restriction fragment was diluted to a concentration of 0.5 ug/ml in ligation buffer (section 2.2.6). The ligation reaction was initiated by the addition of T4 DNA ligase (New England Biolab) to a concentration of 1 unit/ul and the reaction was allowed to proceed for 16 hr at 12^{o}C. The ligated sample was then treated with phenol/chloroform/isoamylalcohol and precipitated as described above.

The PCR was performed in reactions containing 0.1 ug of circularized DNA (section 2.2.13). The PCR machine used was Programmable Thermal Controller (MJ Research) with 35 cycles each of 1.5 min at 95^{o}C, 3 min at 55^{o}C and 2 min at 72^{o}C. The amplified products were extracted twice with phenol/chloroform/isoamylalcohol, and the DNA precipitated as before, and then digested with Pst I and ligated into pUC 19 cleaved with Pst I.

2.2.16. Nucleotide sequence analysis

The sequencing reactions were performed on double stranded plasmid DNA using dye labelled primers and following the cycle sequencing protocol (ABI). The procedures used are as follows:

The following reagent were aliquoted into four 0.5 microcentrifuge tubes:

Reaget	A	C	G	T
d/ddNtP MIx	1ul	1ul	2ul	2ul
Dye primer	1ul	1ul	2ul	2ul
5x Buffer	1ul	1ul	2ul	2ul
DNA template	1ul	1ul	2ul	2ul
Taq (1unit/ul)	1ul	1ul	2ul	2ul
Total Volume	5ul	5ul	10ul	10ul

Each of the tubes were overlayed with 20 ul mineral oil and placed in a PCR machine.

A DNA Thermal Cycler (Perkin Elmer Cetus) was used for the sequencing reaction with 15 cycles each of 0.5 min at 95°C, 0.5 min at 55°C and 1 min at 70°C, and another 15 cycles each of 0.5 min at 95°C and 1 min at 70°C.

After sequencing reaction, the four separate reactions were pooled, ethanol precipitated and dried in a vacuum. The pellet was then resuspended in 6 ul of deionized formamide/50mM EDTA (pH8.0) 5/1 (v/v). The sample was heated at 90°C for 2 min and loaded onto a preelectrophoresed acrylamide gel in a 373A DNA Sequencing System (ABI).

Fluorescence analyses and base calling were performed with the ABI software version 1.30.

2.3. Statistical Methods

Parsimony analysis of the PCR data was carried out using the computer programs of PAUP developed by Dr. D. L. Swofford (Illinois Natural History Survey) and Phylip developed by Felsenstein (1987).

Chapter 3

Identification and mapping of polymorphisms through the polymerase chain reaction.

3.1. Introduction

Genetic markers are of great value in practical breeding programs and genetic research. Traditionally, markers based on morphological differences between individuals have been used. The subsequent development of isozyme and other biochemical markers represented a significant improvement since they offered greater diversity (for review see Tanksley and Orton, 1983). However, markers based upon DNA probes have introduced a new dimension to the development of genetic maps and the mapping of agronomically and physiologically important characters. The major strength of DNA probes, so far mainly RFLPs, is that they have the potential to reveal an almost unlimited number of polymorphisms (Kan and Dozy, 1978; Botstein *et al.*, 1980; Wyman and White, 1980).

RFLPs depend on the use of probes to identify single or low copy sequences in DNA. They are very powerful and have been used to construct detailed linkage maps of several crop species including tomato, potato and maize (reviewed by Tanksley *et al.*, 1989). However, they are expensive and time consuming and technically difficult to use in some species with large and complex genomes, particularly wheat. This complexity of the wheat and related genomes

has delayed the development of such maps for these plants even though recent world-wide efforts are likely soon to remedy this situation (ITMI, 1989).

An alternative method, with the potential to overcome some of the current limitations in cereal genome mapping, is based on PCR (Saiki *et al.*, 1985; Mullis and Faloona, 1987). In this case, polymorphisms are sought in the distance between two short target sequences rather than the presence or absence of restriction endonuclease sites as is the case for standard RFLP's (Skolnic and Wallace, 1989). Oligonucleotides that anneal to the target sequences are used to prime the polymerase reactions. Careful selection of the primers will allow many polymorphisms to be detected and mapped as standard genetic markers. D'Ovidio *et al.* (1990) reported that PCR can be used to detect genetic polymorphism in wheat with primer sequences derived from the sequence of a γ-gliadin gene.

It is necessary for the bases at the 3' end of each primer to provide a perfect match to the target, whereas extensive mismatching in the remainder of the primer/target can be tolerated (Sommer and Tautz, 1989). This gives the option of using sequences with only partial homology to the target sequence to prime the PCR. An extension of this principle has been developed by Williams *et al.* (1990) who have used entirely random primers to generate polymorphic PCR bands.

PCR has provided an alternative approach to many procedures in molecular biology and is replacing many standard techniques (Erlich, 1989). It can be used to amplify specific target sequences for subsequent cloning and it provides an extremely sensitive method for the detection of specific RNA and DNA sequences. This has led to the use of PCR to reveal variability of simple sequences in eukaryotic genomes (Litt and Luty, 1989; Tautz, 1989; Weber and May, 1989). PCR methods based on families of repeated sequences, such as *Alu*-directed PCR

in human genome analysis, provides the prospect of a rapid genome mapping technique (Nelson *et al.*, 1989).

This chapter describes the use of PCR to reveal and map polymorphisms in cereals. Two types of primers have been tested: those that target specific, known sequences (α-amylase) and those that target intron-exon splice junctions. Both types offer valuable new sources of variability that can be more rapidly and cheaply exploited than RFLP markers.

3.2. PCR with primers derived from defined gene sequences (α-amylase)

α-amylases (EC 3.2.1.1) is an important enzyme in biology and industry because of its central role in the hydrolysis of starch (for a review, see Robyt and Whelan, 1968). The α-amylase genes comprise multigene families in wheat and barley. As many as 12-14 *α-amylase1* and 10-11 *α-amylase2* genes were found in wheat by hybridization to α-amylase cDNA probes (Martienssen, 1986). In Southern hybridizations a large number of bands appear and these have been correlated to genes on group 5, 6 and 7 chromosomes in wheat and 6H and 7H in barley (Lazarus *et al.*, 1985; Muthukrishnan *et al.*, 1984; Baulcombe *et al.*, 1987). There are various degrees of homology between these gene sequences (Baulcombe *et al.*, 1987; Chandler *et al.*, 1984; Huttly *et al.*, 1988; Knox *et al* ., 1987; Lazarus *et al.*, 1985).

3.2.1. Comparisons of α-amylase gene sequences and the designing of primers for PCR

The published nucleotide sequences of three wheat and two barley α-amylases were compared by a software package: "DNA Inspector IIe" (Textco). The names of the α-amylase clones where the sequences are from and their chromosomal locations are listed in Table 3.1.

Table 3.1 α-amylase clones and their chromosomal location

α-amylase clones	sources	chromosomal location
p141*	barley	6H
p155*	barley	7H
AMY3/33**	wheat	group 5
AMY2/34***	wheat	7D
AMY2/54***	wheat	7A

* Knox *et al.* (1987).

** Baulcombe *et al.* (1987).

*** Huttly *et al.* (1988).

The comparisons between barley α-amylase p141 and other α-amylases were displayed in Fig. 3.1. The results of the comparisons are generally in agreement with those available in the original publications (Baulcombe *et al.*, 1987; Chandler *et al.*, 1984; Huttly *et al.*, 1988; Knox *et al.*, 1987) although one should be aware that different computer programs were used. The mapping of the PCR products was attempted using wheat-barley addition lines as this approach is simpler than using wheat aneuploid stocks. The primers were chosen to hybridize with two conserved sequences that flank a region containing variable sequences. In this way several individual α-amylase genes would be targeted simultaneously and length polymorphisms between the sequences complementary to the primers, could be identified (Fig. 3.2). A Nae I restriction site was incorporated into each of the primers to allow restriction digestion and cloning of the amplified DNA if required. Both of the 3' ends of the A1 and A2 primers were perfectly matched to the target sequences.

3.2.2. Optimisation of reaction conditions for PCR

Fig. 3.3 shows the amplification products generated when the two primers, A1 and A2, were used with DNA from wheat (cv. Chinese Spring), barley (cv. Betzes) and rye (cv. Imperial). The predicted length of the reaction product for barley α-amylase (α-Amy 1) is 163 bp. A band of approximately 163 bp is the most prominent band in wheat and rye, and was also present in barley. At least two polymorphisms were observed between wheat and barley genomes on agarose gels (compare lanes 1 and 3, Fig. 3.3). The wheat bands at 350 and 450 bp are replaced by bands at 400 and 440 bp in barley. Two polymorphic bands could also be detected between wheat and rye genomes (compare lanes 1 and 2, Fig. 3.3).

Fig. 3.1. Comparison of α-amylase sequences.

The α-amylase sequences were compared by a software package: "DNA Inspector IIe" (Textco). Individual dots on the diagram indicate homology between short stretches of nucleotide sequence (stretches of 12 nucleotides in this study). Regions which are highly conserved between sequences appear as many dots in close proximity, thus forming lines. The sources of the sequences are indicated in Table 3.1.

A. Comparison of barley α-amylase sequences p141 and p155.

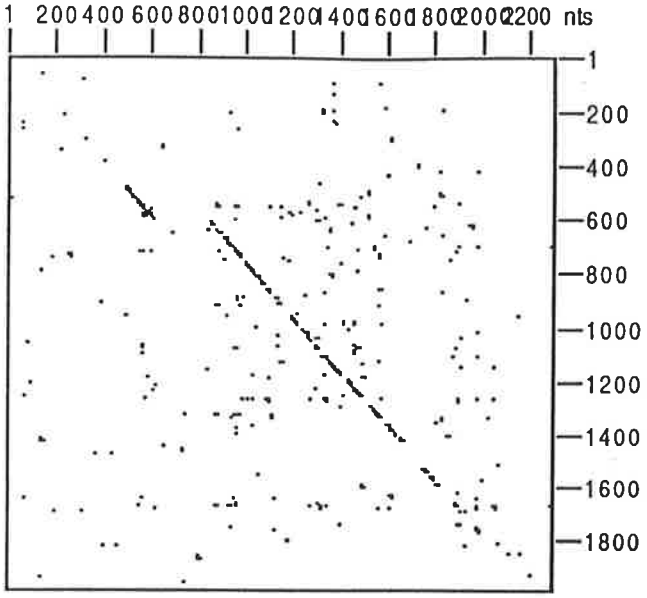
B. Comparison of barley α-amylase sequence p141 and wheat α-amylase sequence AMY3/33.

C. Comparison of barley α-amylase sequence p141 and wheat α-amylase sequence AMY2/34.

D. Comparison of barley α-amylase sequence p141 and wheat α-amylase sequence AMY2/54.

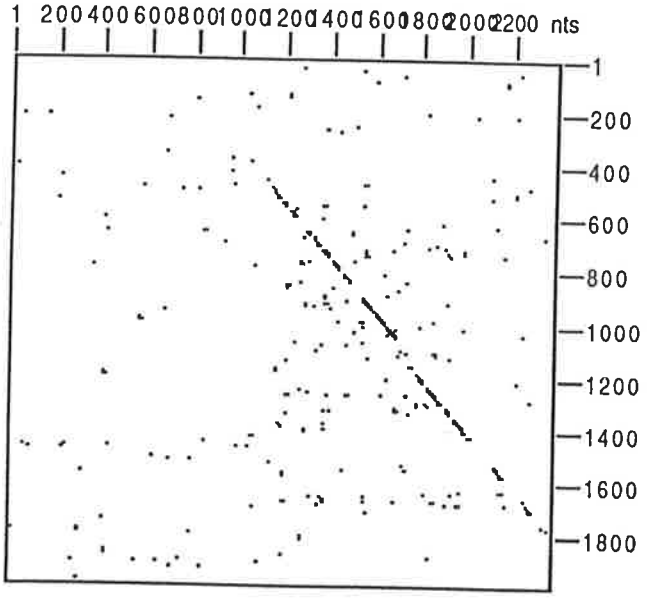
A

DNA#1 (X-axis, left ---> right): P155 (2301 nts)
DNA#2 (Y-axis, top ---> bottom): p141 (1990 nts)
matrix resolution: 200 x 200
search element length: 12
maximum # of mismatches: 2



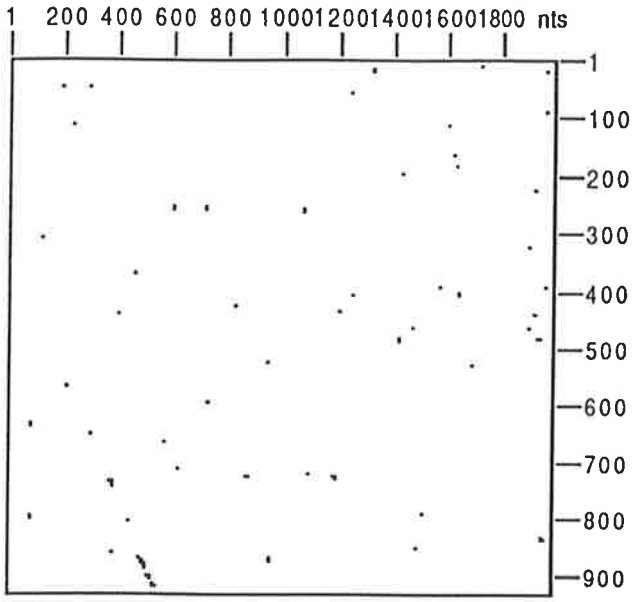
B

DNA#1 (X-axis, left ---> right): AMY3 (2397 nts)
DNA#2 (Y-axis, top ---> bottom): p141 (1990 nts)
matrix resolution: 200 x 200
search element length: 12
maximum # of mismatches: 2



C

DNA#1 (X-axis, left ---> right): p141 (1990 nts)
DNA#2 (Y-axis, top ---> bottom): amy34 (934 nts)
matrix resolution: 200 x 200
search element length: 12
maximum # of mismatches: 2



D

DNA#1 (X-axis, left ---> right): p141 (1990 nts)
DNA#2 (Y-axis, top ---> bottom): amy2/54w (847 nts
matrix resolution: 200 x 200
search element length: 12
maximum # of mismatches: 2

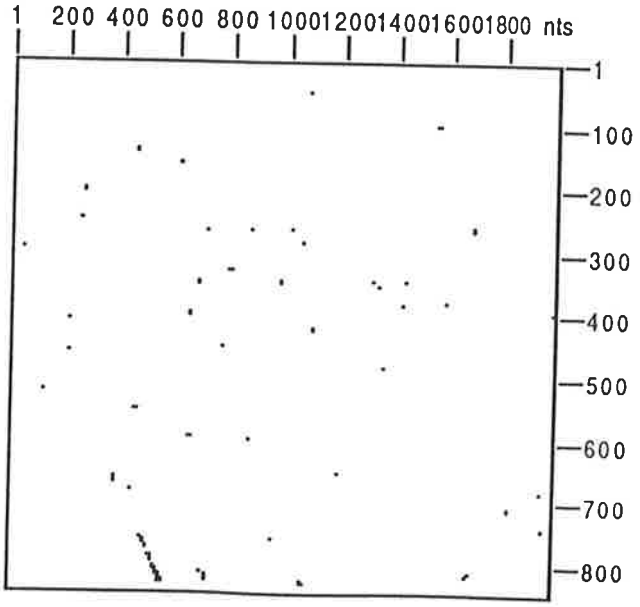


Fig. 3.2. Schematic representation of the location of the A1 and A2 primers.

The data were from Fig. 3.1. A1 and A2 are the two primers derived from conserved sequences of α-amylases.

A1: 5'GCACGCCGGCGGGTGGTAC3'.

A2: 5'GCCGGCTTGCCGTACTTGG 3'.

The shaded box indicates the sequences homologous to A1 and A2, the open box the unique sequences. The sources of the sequences are indicated in Table 3.1.

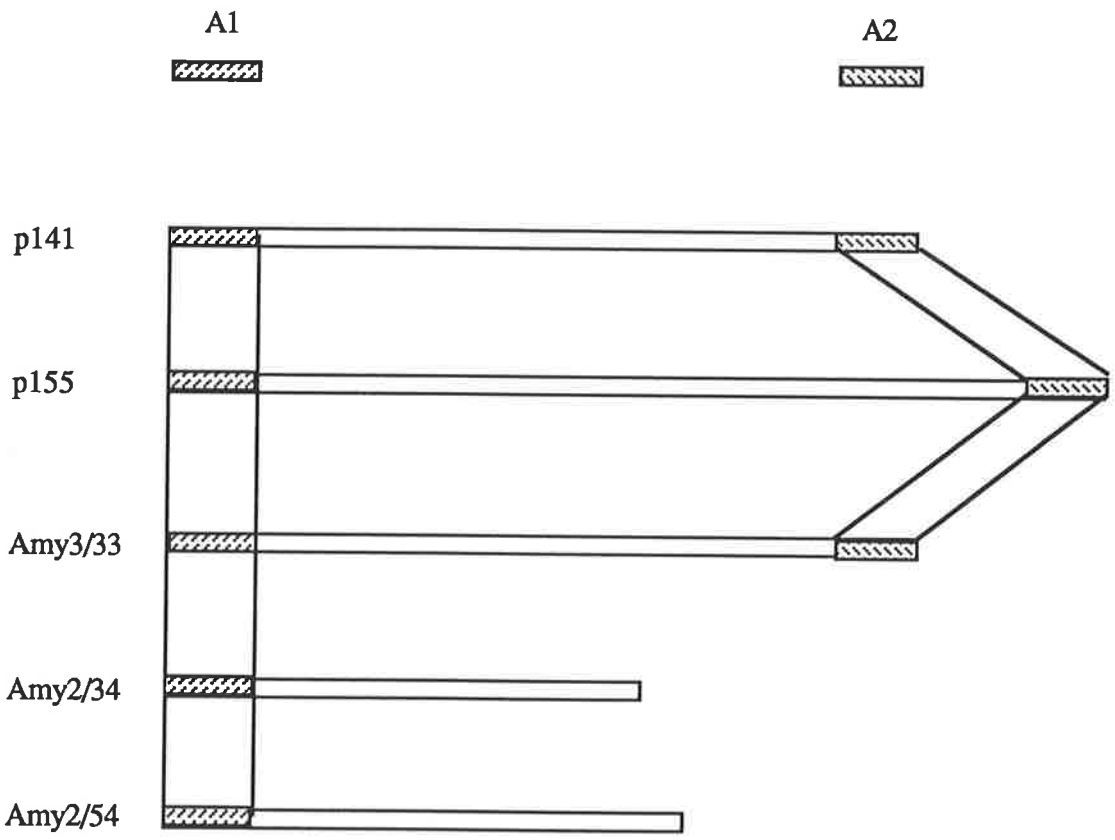
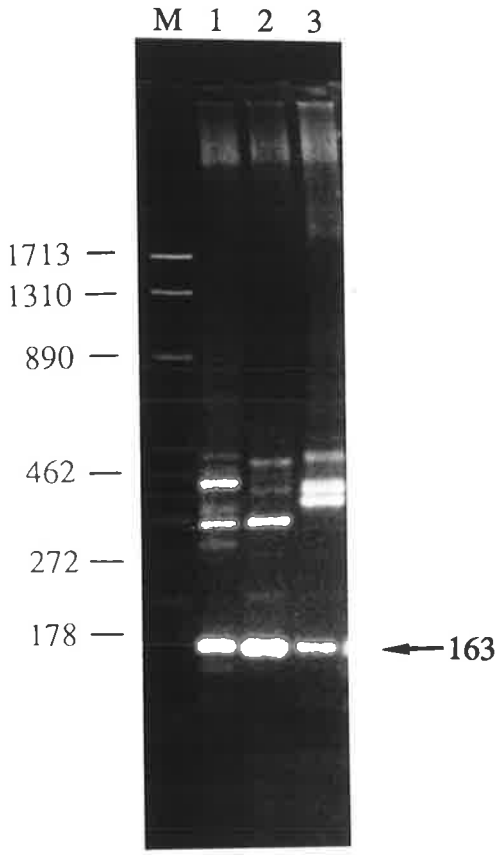

A1
A2
p141
p155
Amy3/33
Amy2/34
Amy2/54

Fig. 3.3. Amplification of cereal DNA using α-amylase primers A1 and A2.

The template DNA was from wheat (Chinese Spring) in Lane 1, rye (Imperial) in Lane 2 and barley (Betzes) in Lanes 3, respectively. Amplification was performed as described in Section 2.2.12 and 2.2.13. DNA size markers are shown in Lane M with the sizes of the marker bands given on the left in base pairs.


M 1 2 3
1713
1310
890
462
272
178
163

The results showed that mini-prep DNA from cereals (section 2.2.4) could be used to allow amplification of the α-amylase sequences with 32 cycles of PCR and using 55°C as the annealing temperature (Fig. 3.4). Even the presence of 5 ug of mini-prep DNA did not have any diminishing effect on amplification. Interestingly, the smears above 2 kb were increasing as more DNA templates were added. The 163 bp band was not produced when only 5 pg DNA present. Assuming that 1 barley cell contains 5.5 pg nucleic DNA, 50 pg or 9 copies of barley genomic DNA is the minimum of template DNA for amplification with α-amylase, using the 163 bp band as an indicator (Fig. 3.4).

As the optimal Mg^{++} concentration varies with different templates (Saiki, 1989), a range of PCR buffers with different Mg^{++} concentration was tested with α-amylase primers. The number of amplification products increased as the Mg^{++} concentration in the PCR reaction was changed from 1.5 up to 5.0 mM (Fig. 3.5). The 163 bp band was strongest with 1.5 mM Mg^{++} concentration, but it was almost invisible with 5.0 mM Mg^{++}. A 300 bp PCR product has the optimal Mg^{++} concentration at 2mM (lane 2, Fig. 3.5). This 300 bp was only present when the Mg^{++} concentration over 2mM. The data indicate that not only different primers have different magnesium optima (Saiki, 1989), different PCR products also require different concentration of Mg^{++} for successful amplification.

3.2.3. Mapping of polymorphisms

Fig. 3.6 shows the amplification products generated when the two primers, A1 and A2, were used with DNA from wheat (cv. Chinese Spring), barley (cv. Betzes) and the wheat-barley addition line series. The addition lines contain a single pair of barley chromosomes in a wheat background. Using addition lines of barley, the polymorphic bands in barley (cv. Betzes) could be mapped to

Fig. 3.4. Effect of DNA concentration on the amplification of barley DNA.

The template DNA was from barley (Betzes). Amplifications were performed as described in Fig. 3.3 except that DNA concentration was varied from 5pg to 5ug per reaction (as indicated at the top of each lane). DNA size markers are shown in lane MW with the sizes of the marker bands given on the right in base pairs.

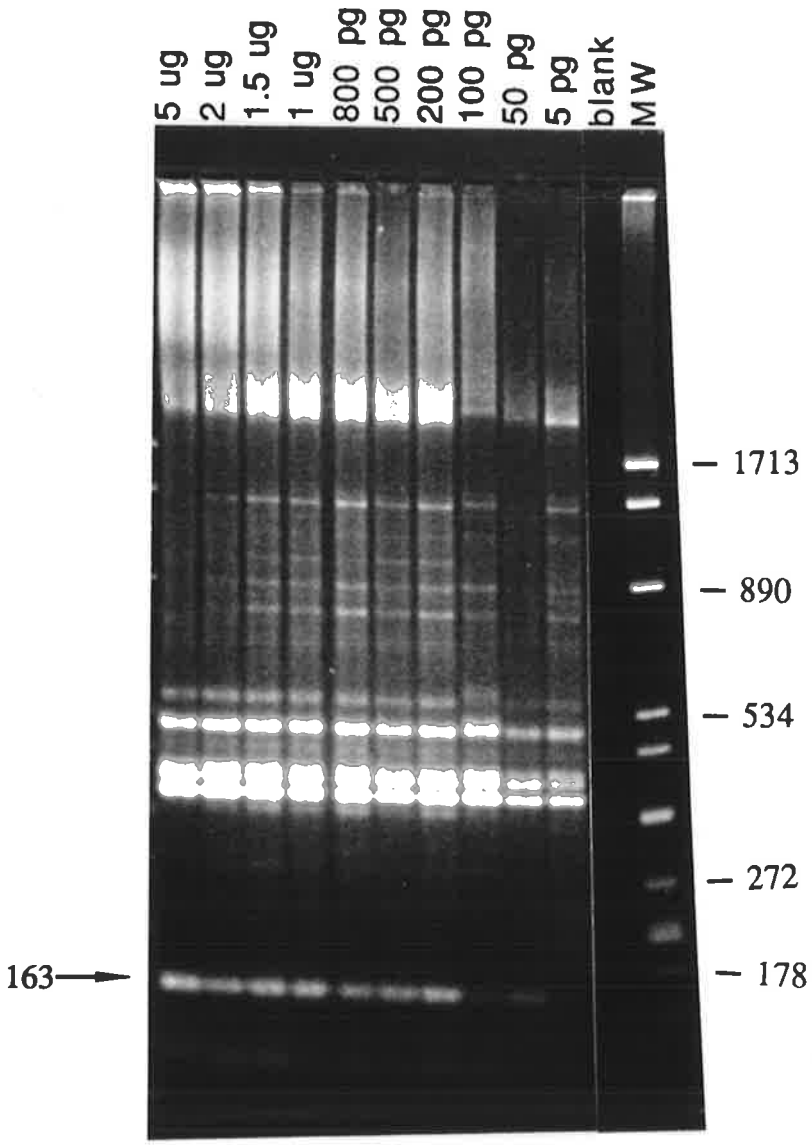

5 ug
2 ug
1.5 ug
1 ug
800 pg
500 pg
200 pg
100 pg
50 pg
5 pg
blank
MW
1713
890
534
272
178
163

Fig. 3.5. Effect of Mg^{++} concentration on the amplification of barley DNA

The template DNA was from barley (Betzes). Amplifications were performed as described in Fig. 3.3 except that $MgCl_2$ was varied from 1.5 to 5 mM: lane 1 (1.5 mM), lane 2 (2mM), lane 3 (3mM) and lane 4 (5 mM). DNA size markers are shown in Lane M with the sizes of the marker bands given on the right in base pairs.

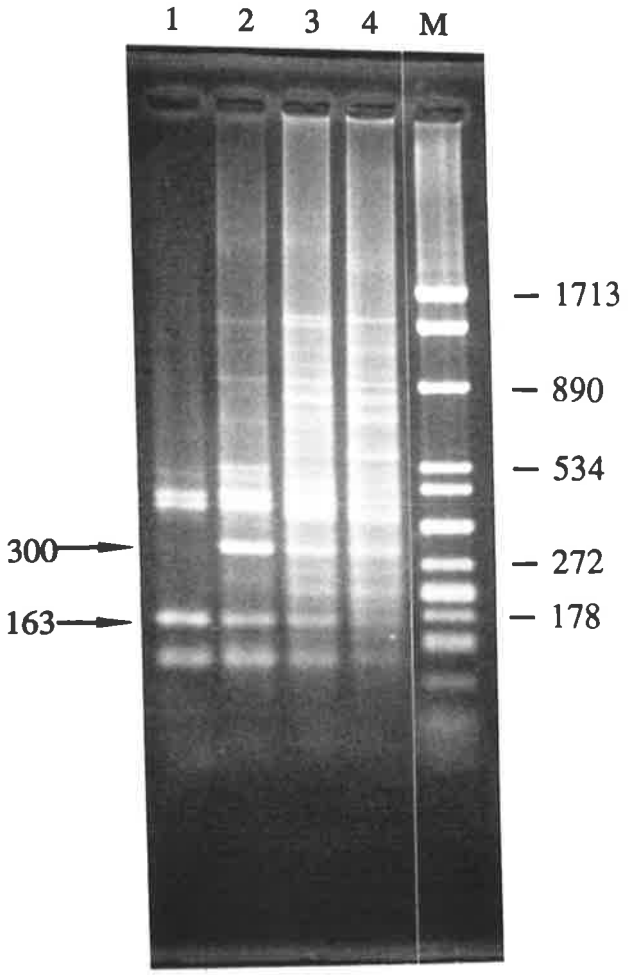
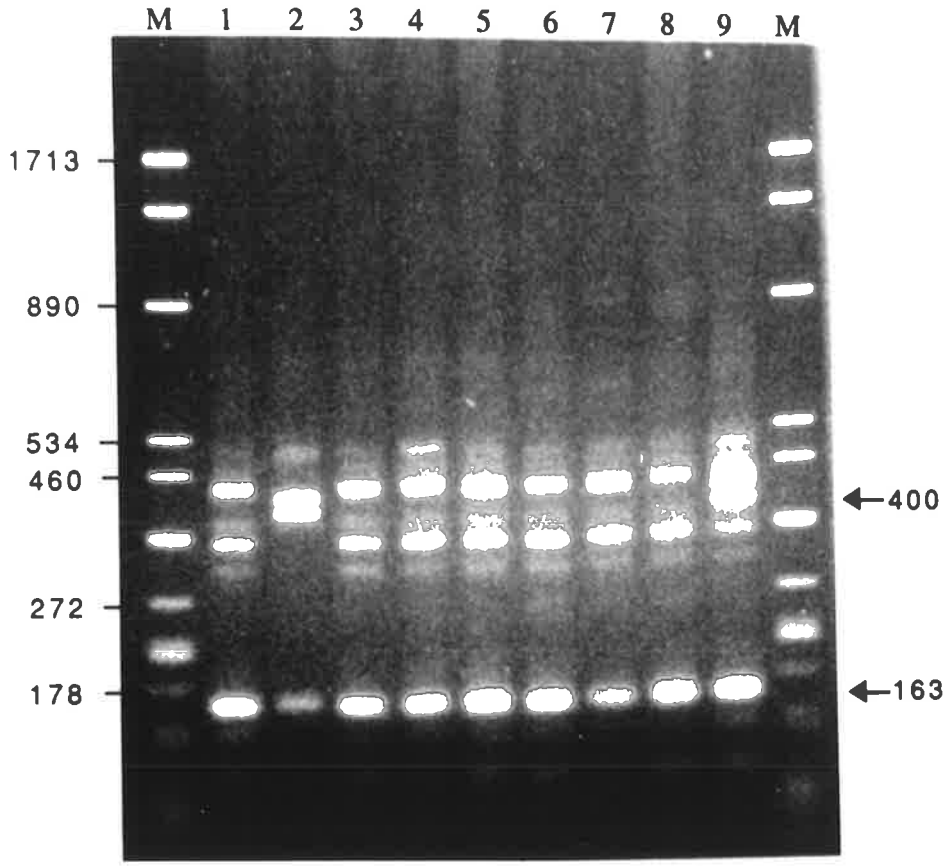

1
2
3
4
M
1713
890
534
272
178
300
163

Fig. 3.6. Mapping of PCR products with cereal DNA using α-amylase primers A1 and A2.

The template DNA was from wheat (Chinese Spring) in Lane 1, barley (Betzes) in Lane 2 and the wheat-barley addition lines 1H to 7H, in Lanes 3 to 9, respectively. DNA size markers are shown in Lane M. The barley specific band at 400 bp in Lanes 2 and 9 is indicated by an arrow. The sizes of the marker bands are given on the left in base pairs. The PCR products were derived from template DNA isolated by the small scale method (Section 2.2.4) and were fractionated on a 3% agarose gel.


M
1
2
3
4
5
6
7
8
9
M
1713
890
534
460
272
178
400
163

barley chromosome 7H (compare lanes 2 and 9, Fig. 3.6) on a 3% agarose gel. It appears that there were more than two polymorphic bands mapped to 7H.

3.3. PCR with semi-random and random primers through the polymerase chain reaction.

The results obtained with the α-amylase primers (section 3.1) demonstrated the value of the PCR technique in revealing and mapping polymorphisms in cereals. However, for this type of reaction it is necessary to have adequate information about the target sequence to allow synthesis of the primers that are likely to target regions of variable lengths. The need for this information would greatly limit the applicability of PCR as a mapping tool. An alternative strategy was developed, based on the consensus sequences for the intron splice junctions (ISJ).

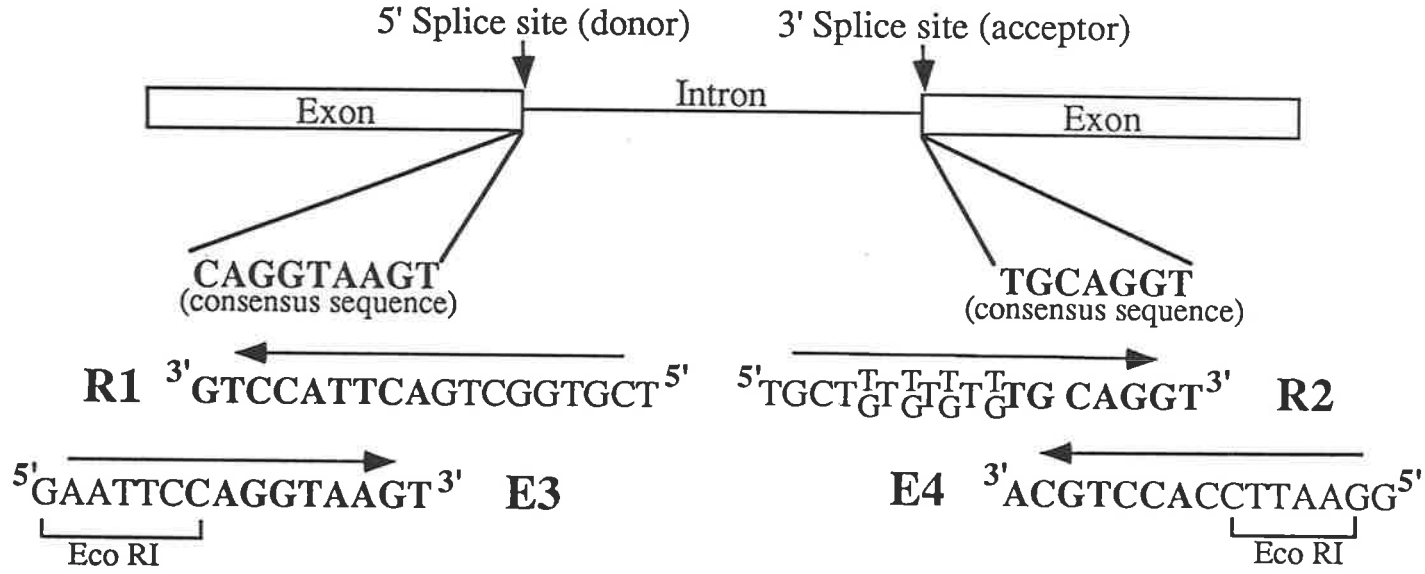
Introns have been identified in most plant genes studied (Hawkins, 1988) and the junctions to exons are highly conserved sequences. However, since the introns are generally subjected to only weak selective pressure by comparison to exons, they are usually highly variable in sequence and length. These properties would appear to make the ISJ's ideal targets for the identification of polymorphisms in PCR products.

3.3.1. The design of primers to target the intron splice junctions (ISJ)

The sequences of the ISJ primers, R1, R2, E3 and E4 (Fig. 3.7), are based on the consensus sequences of the junctions reported for plants (Brown *et al.*, 1986; Brown, 1986; Goodall and Filipowicz, 1989). The orientation of these primers at the intron junction is displayed in Fig. 3.7. The primers used were 15

Fig. 3.7. Schematic representation of the location of plant intron splice junction ISJ primers.

The sequences of R1, R2, E3 and E4 are based on the consensus sequences for the exon-intron splice junctions of plants (Brown et al. 1986; Brown 1986; Goodall and Filipowicz 1989). E3 and E4 were synthesized to contain an Eco RI site as indicated. The bases that match the consensus sequence are shown in bold type. The orientation of the primers is indicated by the arrows.


5' Splice site (donor)
3' Splice site (acceptor)
Exon
Intron
Exon
CAGGTAAGT
(consensus sequence)
TGCAGGT
(consensus sequence)
R1 3'GTCCATTCAGTCGGTGCT5'
5'TGCTTTTTTTTG CAGGT3' R2
5'GAATTCCAGGTAAGT3' E3
E4 3'ACGTCCACCTTAAGG5'
Eco RI
Eco RI

(E3 and E4) or 18 (R1 and R2) bases in length. The consensus sequence for the splice junctions is 9 bases at the 5' site and 7 bases at the 3' site. The additional bases were added at random to extend the length of the primers and provide potential sites for base pairing to the target. For the E3 and E4 primers an Eco RI site was added to facilitate cloning of the PCR products if required. The results of Sommer and Tautz (1989) suggest that only the three bases at the 3' end of the primer are critical for efficient PCR. The remainder of the primer serves to stabilise the primer-target duplex as determined by the annealing conditions of the reaction. Primers were produced to generate products from the exon regions (primers R1 and R2) or from the intron region (E3 and E4).

3.3.2. Reaction conditions for PCR with ISJ primer and random primer

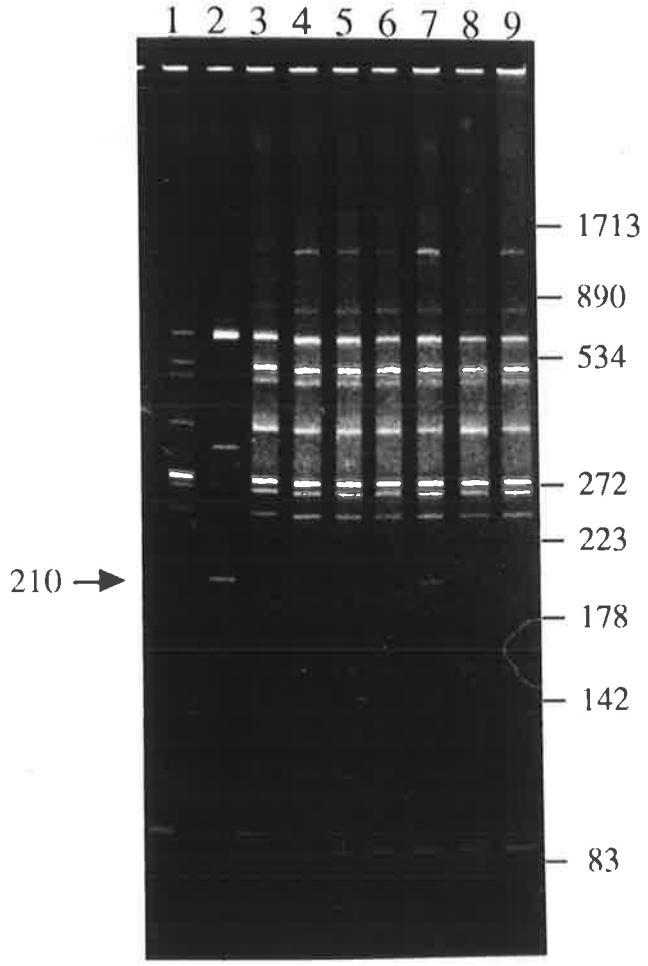
Two steps were used in the PCR reactions involving the ISJ primers: the annealing temperature during the first 6 cycles was at 40°C. This was then raised to 58°C for the final 28 cycles. A low annealing temperature was used for the initial cycles to permit amplification from targets that may only have poor homology to the primers.

3.3.3. Identification and mapping of polymorphisms

ISJ primers can be used in conjunction with specific primers based on the sequence information from defined cereal genes. Fig. 3.8 shows the results of amplification of cereal genomic DNA with the α-amylase A1 and R1 (5' splice site) primers. A clear polymorphic band can be seen between the wheat and barley patterns at 210 bp (lanes 1 and 2, Fig.3.8). This band can be localised to barley chromosome 5H (at lane 7).

Fig. 3.8. Amplification of cereal DNA using the α-amylase primer, A2, in conjunction with the ISJ primer, R1.

Lane 1 shows the products obtained with wheat DNA (Chinese Spring), Lane 2 barley DNA (Betzes) and Lanes 3-9 DNA from wheat-barley addition lines 1H to 7H, respectively. The positions of DNA size markers is indicated on the right in base pairs. The barley specific band at 210 bp in lane 2 and 7 is indicated by an arrow on the left. The PCR products were fractionated on 10% polyacrylamide gels.


1 2 3 4 5 6 7 8 9
1713
890
534
272
223
210
178
142
83

The amplification products obtained from cereal DNA with ISJ and random primers are shown in Fig. 3.10a and 3.10b. When ISJ primer R1 (5' splice site, exon targeting) was used with random primer B1, a faint smear of bands was obtained (Fig. 3.9). The PCR products were too complex to allow resolution of single bands and could not be used directly for the identification of polymorphisms. However, if the template DNA was digested with the restriction endonuclease Eco RV prior to the PCR reaction, clear banding patterns were generated. A distinctive pattern of polymorphic bands was observed between wheat and barley (compare lanes 1 and 2 in Fig. 3.10a). A barley specific band of 110bp could be mapped onto 1H* and 6H (Lanes 3 and 8 at 110bp). As 1H* is actually a double monosomic line of the barley chromosomes 1H and 6H (Islam and Shepherd, unpublished), the barley specific band is probably located on 6H alone rather than on both 1H and 6H.

The R2 primer (3' splice site, exon targeting, Fig. 3.7) coupled with random primer L1 (L1 is a primer complementary to the imm 434 portion of λgt10 template, Huynh *et al.* 1985) also revealed extensive polymorphisms between wheat and barley (compare lanes 8 and 9, Fig. 3.10b). A 450bp amplified fragment specific to barley was mapped to barley chromosome 3H using the addition lines (Lanes 5 and 8, Fig. 3.10b). Note also a band at 280bp present in wheat, barley and the addition lines (Fig. 3.10b). The intensity of this band is stronger in 1H* and 6H (Lanes 7 and 2, respectively, Fig. 3.10b). The corresponding locus may be situated on 6H in barley as discussed above for the 110 bp band in Fig. 3.10a.

Fig. 3.9. Amplification of cereal DNA using ISJ primer R1 coupled with random primer B1.

ISJ primer R1 and random primer B1 were used to amplify cereal DNA. Lane 1 shows the products obtained with wheat DNA (Chinese Spring), lane 2 barley DNA (Betzes). The reaction conditions were described in section 3.3.2. The size of the marker bands in base pairs is given at the left. The PCR products were fractionated on a 3% agarose gel.

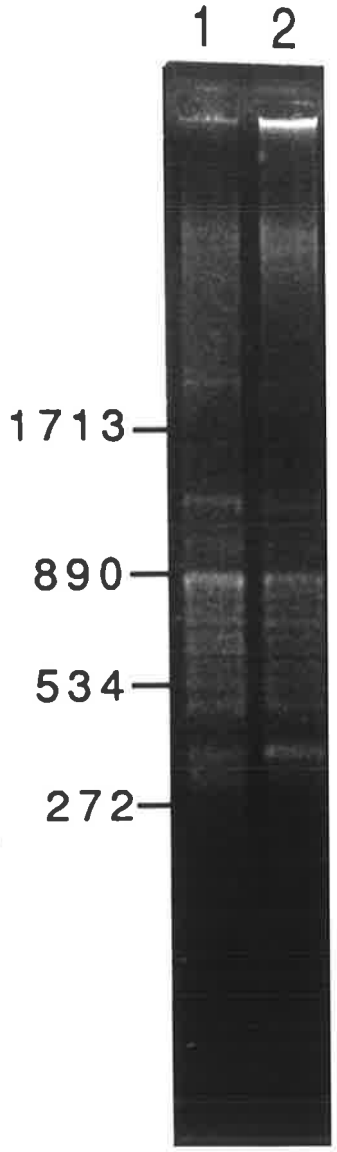

1
2
1713
890
534
272

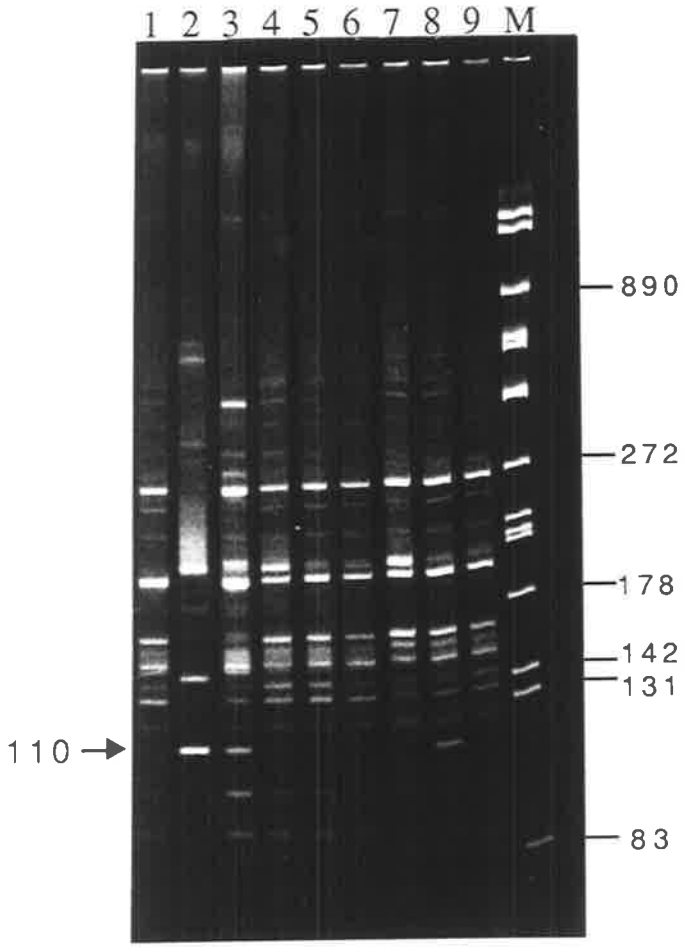
Fig. 3.10. Amplification of cereal DNA using ISJ primers coupled with random primers.

DNA size markers are shown in Lanes M with the sizes of the marker bands given on the right in base pairs. The PCR products were fractionated on a 10% polyacrylamide gel.

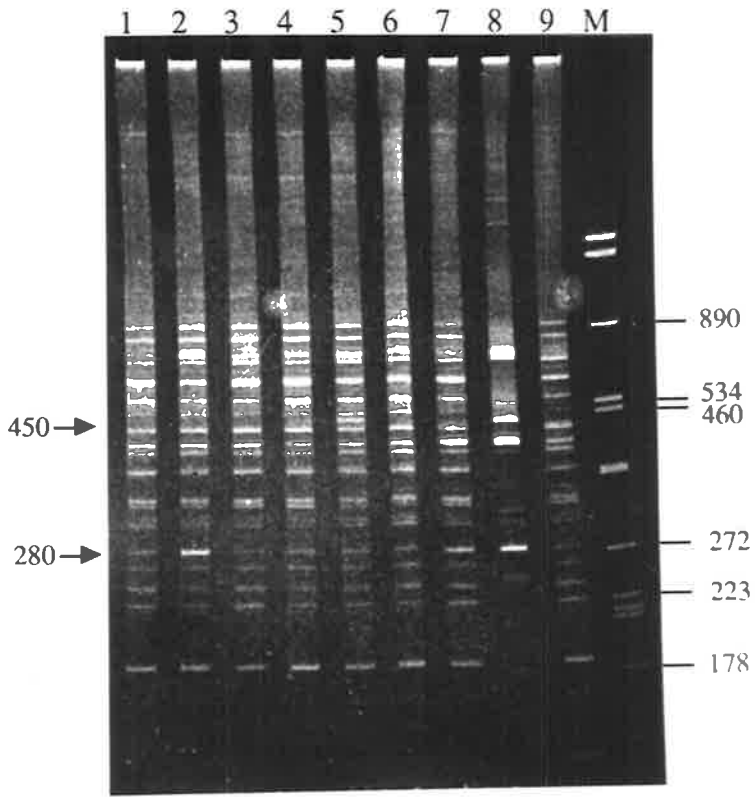
a. ISJ primer R1 and random primer B1 were used to amplify cereal DNA. The DNAs were pre-digested with Eco RV prior to amplification. Lane 1 shows the products obtained with wheat DNA (Chinese Spring), Lane 2 barley DNA (Betzes) and Lanes 3-9 DNA from wheat barley addition lines 1H to 7H, respectively. The barley specific band in lane 2, 3 and 8 at 110bp is indicated by arrows on the left.

b. ISJ primer R2 and random primer L1 were used to amplify cereal DNAs without pre-digestion. Lane 9 shows the products obtained with wheat DNA (Chinese Spring), Lane 8 barley DNA (Betzes) and Lanes 7-1 DNA from wheat barley addition lines 7H to 1H, respectively. The barley specific bands in lanes 2, 5 and 8 are indicated by an arrow on the left.

a


1 2 3 4 5 6 7 8 9 M
890
272
178
142
131
110
83

b


1 2 3 4 5 6 7 8 9 M
890
534
460
450
280
272
223
178

3.4. Discussion

In this study, the polymorphisms identified by the PCR are in the length of region amplified. One is looking for variation in the length of sequences separating the target regions with which the primers will anneal. In mammalian systems it has been demonstrated that these types of polymorphisms are abundant (Litt and Luty, 1989; Tautz, 1989; Weber and May, 1989). A major limitation for the PCR approach is the need for extensive sequence information in order to synthesize the appropriate primers. With good sequence information, the primers can be synthesized to sit on each side of regions that are likely to be of variable length. In this way specific, known sequences can be amplified out of the genomic DNA; for a simple target, such as an individual gene, one band may be generated. More complex patterns can be obtained if multi-gene families are targeted using primers that show homology to several members of the family. This type of marker is exemplified by the α-amylase gene family. The primers A1 and A2 were selected to allow the detection of multiple genes and to flank an area that was thought to be variable (Baulcombe *et al.*, 1987; Huttly *et al.*, 1988; Knox *et al.*, 1987). One of the major bands generated, at 163 bp, was as predicted. The remaining bands could not be unequivocally correlated to the published α-amylase sequences. However, the localisation of polymorphic bands to barley chromosome 7H is in agreement with the known location of a set of α-amylase genes on this chromosome (Knox *et al.*, 1987).

Although many plant sequences have been published, there are far too few available to allow the generation of extensive genetic maps. Alternative procedures must be sought to extend the number of polymorphisms that can be detected through the PCR. This can be achieved by the use of random sequences to prime the reaction (Williams *et al.*, 1990; Welsh and McClelland, 1991a) or by developing primers that target a range of sequences. For the second system to function one must identify sequences that occur frequently throughout the genome

but are likely to be evenly dispersed. It is a further advantage if the target sequences occur rarely, if at all, in heterochromatic regions since these areas are unlikely to contain genes of agronomic importance to which linkage would subsequently be sought. The criteria outlined are best met by the conserved regions of genes. Several options exist; for example, the translation start signals, the poly A addition sites, promoter and enhancer regions and the intron-exon splice junctions. Here the consensus sequence information for plant intron splice junctions (ISJ) (Brown 1986) was used.

There are four major advantages in using the ISJ sequences to generate primers. First the core of the junctions are highly conserved. In these experiments the primers were synthesized to match the 3 bases at the 3' end; this is within the highly conserved core of the consensus sequence (see Fig. 3.8). Since there is a degree of degeneracy in the consensus sequence only a sub-set of splice junctions will actually be targeted. Mismatches in the remainder of the primer-target pairing increase the number of targets and the complexity of the amplification products. The second advantage of the ISJ primers is that introns are present in most plant genes and one can, therefore, target an extremely diverse range of genes. The third and major advantage is that through the ISJ primers one avoids targeting heterochromatic regions and is likely to get a good chromosomal distribution of polymorphisms. These aspects are of particular importance in cereal mapping since repeat sequence regions are abundant, ranging from about 84% in barley to over 90% of the rye genome (Flavell *et al.*, 1974). The fourth advantage is that the cost of PCR is reduced through the use of the ISJ primers since only one addition primer is required for each reaction and the ISJ primer can be produced in large scale synthesis.

The PCR can be based on primers derived from sequences that are defined, semi-random, such as the ISJ, or purely random. The random primers used here are of about 20 bases in length. These primers were substantially longer than the

9 or 10 base primers described by Williams *et al.* (1990) and it is, therefore, not possible to make a direct comparison between the two types of PCR. If both primers are random, the probability of finding the target sequence for both primers in close vicinity to each other, is low.

It has been repoted that the optimal length for PCR primers is around 20 bases (Erlich, 1989) . For random primers longer than 9-10 bases the likelihood of success in the PCR is increased through the use of the ISJ primers in conjunction with either random or defined primers. The examples presented here demonstrate that the ISJ primers can be used singly or in conjunction with either defined or random primers if the conditions for PCR amplification are optimized.

Pre-digestion of the template DNA with a restriction endonuclease is required in some cases when the ISJ primers are used alone or in conjunction with other primers (e.g. Fig. 3.9). The major effect of pre-digestion is the reduction in complexity of bands to a point where individual bands can be resolved. The type of enzyme used appears to have little effect on the actual banding pattern although some variation has been seen (data not shown). There are two possible explanations for this effect. First, the digestion may prevent the amplification of certain regions by cleaving between the target sequences. This would lead to a reduction in band complexity and occasional variation between the patterns generated by templates digested with different endonucleases. Second, the pre-digestion may allow improved denaturation and replication of some of the template DNA fragments as a result of the reduced length and better access for the *Taq* polymerase to the DNA.

Two types of PCR can be used to detect and map polymorphisms in wheat and barley. The first uses specific primers that can target known genes or gene families. This method is rapid, cheap and does not require pre-digestion of the template with restriction endonucleases. However, sequence information is needed. PCR based on specific primers will be a valuable technique for screening

with one or few markers in a breeding program. For example, if an RFLP marker were identified that is closely linked to a gene of interest, the RFLP probe could be sequenced and PCR primers synthesized. The locus could then be monitored through the PCR rather than the more difficult RFLP. This procedure would be justified if large numbers of plants were to be screened.

A second type of PCR uses the ISJ primers. These allow the generation of a great diversity of markers without any additional sequence information. The ISJ based markers can be used for generating maps of the cereal genomes. Although pre-digestion of the template may be required in some cases, these markers are still cheaper and faster to assay than RFLP's. They are technically less demanding to use than RFLP's and, as with standard PCR, do not require radioactivity to detect.

Chapter 4

Intravarietal and intervarietal polymorphisms revealed by PCR

4.1. Introduction

The understanding of the genetics and ecology of natural populations has been greatly advanced by the application of molecular techniques, especially through the analysis of isozymes and DNA variation. Comparisons of shared and polymorphic restriction fragment lengths, RFLPs, have been utilised to infer phylogenetic relationships in different species (Havey and Muehlbauer, 1989; Saghai-Maroof *et al.*, 1984). Song *et al.* (1988 a, b; 1990) used RFLPs to estimate genome relationships within *Brassica*, as well as between *Brassica* and wild relatives. In the last few years, mini- and micro-satellite sequences have been found to detect DNA fingerprinting in many organisms (Jeffreys *et al.*, 1985; Nakamura *et al.*, 1987; Wyman and White, 1980). This type of marker has had a profound impact on forensic and legal medicine. Recently, PCR (Saiki, 1985; Mullis and Faloona, 1987) has been used with arbitrarily chosen primers for generating genetic markers and for DNA fingerprinting (Williams *et al.*, 1990; Welsh and McClelland, 1990, Welsh *et al.*, 1991; Caetano-Anolles *et al.*, 1991).

Neale *et al.* (1986) showed that the chloroplast DNA (cpDNA) of cultivated barley is less variable than its wild ancestor. Brown and Munday (1982) reported that levels of variation in landraces sampled directly from fields in Iran were at a

level between the low values of European cultivars and the high values of natural populations of wild barley in Israel. Landraces were also less variable than wild barley for cpDNA types (Clegg *et al.*, 1984). However, Jana and Pietrazak (1988) compared allozyme diversity in matched pairs of samples of wild and cultivated barley from the same site in four countries. They found that on average the diversity in cultivated landraces equalled or exceeded that in nearby wild populations. These disparate conclusions were considered to arise from the difference in the specific isozyme loci and in the ecology of the wild populations used in these studies (Brown, 1991). According to Graner *et al.* (1990), the total diversity in barley was 0.43 (mean diversity index or average chance that any two varieties differ for a random probe). Subgroups of spring and winter barley and of two- and six- row types showed less diversity.

For plant breeders, growers and the whole agricultural industry in general, cultivar identification is important both in theoretical and practical terms, especially after the introduction of plant variety rights (Simmonds, 1987). Conventional varietal identification is based on morphological descriptions (Stegemann, 1984). However, these assessments can be labour intensive, time-consuming and vulnerable to environmental factors. Biochemical markers such as isozymes and storage proteins provide significant advantages over morphological methods as they are usually rapid, cheap and unaffected by the environment. A plethora of methods is available for the identification of wheat and barley varieties (for a review see Cooke, 1988). The main drawback of these techniques is the limited degree of polymorphisms to distinguish closely related varieties and the fact that it often requires a combinations of several techniques (Ainsworth and Sharp, 1989). There has been increasing interest in the use of DNA-based procedures for varietal identification although RFLPs are costly and labour intensive (Ainsworth and Sharp, 1989). Bunce *et al.* (1986) demonstrated that RFLPs have the potential for identifying and fingerprinting

barley cultivars. RFLPs have also been used to distinguish wheat varieties with ribosomal RNA probes and a heterologous β-amylase cDNA clone (May and Appels, 1987, 1988; Sharp *et al.*, 1988). As shown in Chapters 1 and 3, PCR has emerged as the preferred method for generating genetic markers because it is faster, simpler and cheaper than RFLPs.

As shown in Chapter 3, the ISJ primers and random primers are valuable in detecting and mapping polymorphisms between wheat and barley. However, for these polymorphisms to be useful in the development of genetic maps and applying such maps, it is important that they also detect polymorphisms between varieties or, even better, isogenic lines.

This Chapter describes the identification of polymorphisms, through the use of PCR, between or within populations of some wheat varieties and breeding lines, barley varieties and breeding lines, and rye varieties. A collection of just three primers allows the identification of any of the barley varieties currently grown commercially in Australia.

4.2. Identification of polymorphisms between barley varieties and breeding lines

The barley cultivars listed here (Table 4.1) include varieties and breeding lines developed at the Waite Institute and elsewhere in Australia, varieties from Japan (e.g. Sumire Mochi), and breeding lines from America (e.g. CIMMYT 42002). Most of them are varieties commercially available in Australia. PCR amplification was performed under the same conditions as described in section 3.2.3. Only the reproducible bands in multiple runs were considered in this study. Fig. 4.1 shows the banding profiles produced by the R1 primer from 42 commercial varieties and breeding lines of barley. Nine polymorphic bands can

Table 4.1. Barley cultivars used in this study

Cultivar No.	Cultivar	Sources
F1	Clipper(1)*	Waite
F2	Galleon(1)	Waite
F3	Schooner(1)	Waite
F4	Skiff	Waite
F5	WI 2692	Waite
F6	WI 2736	Waite
F7	WI 2737(1)	Waite
F8	O'Connor	WA (Western Australia)
F9	Stirling(1)	WA
F10	Onslow	WA
F11	Moonndyne	WA
F12	Forrest	WA
F13	Windich	WA
F14	Yagan	WA
F15	Malebo	NSW (New South Wales)
F16	Ulandra	NSW
F17	Grimmett	QLD (Queensland)
F18	Franklin	TAS (Tasmania)
F19	Weeah(1)	VIC (Victoria)
F20	Waranga	VIC
B1	CIMMYT 42002	CIMMYT**
B2	CIMMYT 42026	CIMMYT
B3	CIMMYT 42107	CIMMYT
B4	CIMMYT 42119	CIMMYT
B5	CIMMYT 42120	CIMMYT
B6	unknown variety	unknown
B7	CM.72	USA
B8	Galleon(2)*	Waite
B9	Hulless Betzes WI-2768	USA
B10	M737	Denmark
B11	Nirisaki Nijo NIJO.9	Japan
B12	Sahara 3771	Algeria
B13	SB 85216 WI-2776	Canada
B14	Ingrid	Sweden
B15	Schooner(2)	Waite
B16	Skiff(2)	Waite
B17	Stirling(2)	Waite
B18	Sumire Mochi	Japan
B19	T.ANT-17	Denmark
B20	WI 2692	Waite
B21	WI 2723	Waite
B22	WI 2737 (2)	Waite
B23	WA 1087/1	WA
B24	WA 1957/2	WA
B25	WA 2649/2	WA
B26	WA 2650/1	WA
B27	WA 2870/1	WA
B28	WA 735/276	WA
B29	WA 2646/2	WA
B30	Weeah(2)	VIC
B31	Haruna Nijo (1)	Japan
B32	Haruna Nijo (2)	Japan
B34	Betzes	Europe

*: (1) and (2) indicate two different accessions of the same cultivar stored in the Waite Institute Collection.

**: CIMMYT is located in Mexico.

be seen in this example. The amplified products with the primers E2 (random primer) and E4 (ISJ primer, Fig. 3.1), respectively, from the barley varieties and breeding lines are displayed in Fig. 4.2 and Fig. 4.3. When the R1 primer was coupled with random primer S1, a different banding pattern of the PCR products was obtained (Fig. 4.4) compared with the primer R1 alone (Fig. 4.1). Extensive polymorphism among barley cultivars was revealed by each of the primers and primer combination. Some monomorphic bands could also be seen in all the primers tested. The most reliable bands are those seen in the size range below 1kb. The larger size region also shows extensive polymorphisms but these tend to be less reproducible. This may be due to the inefficiency of PCR to amplify long DNA fragments (Erlich *et al.*, 1989; Erlich, 1989).

For any given two barley cultivars, there are at least two monomorphic bands shared between them. These monomorphic bands are a good indication that the cultivars should belong to the same group: barley. Wheat cv. Chinese Spring was included in the experiment as a control (lane 23 in Fig. 4.1b). None of the monomorphic bands were shared between wheat and barley (e.g. compare lanes 23 and 24 in Fig. 4.1b). Although some of the PCR bands may happen to have the same molecular weight, barley has a completely different PCR banding pattern to that of wheat.

4.3. Barley cultivar identification by polymorphic markers

The banding patterns produced by R1 (Fig. 4.1) show at least nine polymorphisms between varieties, whereas seven and eight polymorphisms for E2 and E4 can be seen in Fig 4.2 and 4.3, respectively. Different banding profiles can be derived from the combinations of these polymorphic bands. One of cultivars was incorrectly labelled during DNA preparation and it was tagged as 'unknown' consequentely (Table 4.1). This cultivar was shown to have the

Fig. 4.1. Identification of barley cultivars using an ISJ primer R1.

The amplifications were performed as described in Section 3.3.2. DNA was amplified from a series of commercial barley varieties and some breeding lines (Table 4.1) using ISJ primer R1. The names of the barley cultivars are shown at the top of the lanes. The PCR products were fractionated on 3% agarose gels.

A. Mini-prep genomic DNAs of 20 barley cultivars were amplified. DNA size marker (Lane M) is shown on left with the sizes of the bands given in base pairs.

B. Mini-prep genomic DNAs of another 23 barley cultivars were amplified. Wheat cv. Chinese Spring is included for reference (lanes 23). DNA size marker (Lane MW) is shown on right with the sizes of the bands given in base pairs.

A
M
1 Clipper
2 Windich
3 Galleon
4 Schooner
5 Skiff
6 WI 2692
7 WI2736
8 WI2737
9 O'Connor
10 Stirling
11 Onslow
12 Moondyne
13 Forrest
14 Malebo
15 Galleon
16 Grimmet
17 Franklin
18 Weeah
19 Waraga
20 Yagan
1713
534
272
178
B
1 Haruna Nija
2 CIMMYT 42002
3 CIMMYT 42026
4 CIMMYT 42107
5 CIMMYT 42119
6 CIMMYT 42120
7 CM.72
8 Hulless Betzes W1-2768
9 M737
10 SB 85216 W1-2776
11 Sumire Mochi
12 T.ANT-17
13 WI-2723
14 WI-2737
15 WA 1087/1
16 WA 1957/2
17 WA 2649/2
18 WA 2650/1
19 WA 2870/1
20 WA 735/276
21 WA-2646/2
22 Ingrid
23 Chinese Spring
24 Betzes
25 MW
1713
890
534

Fig. 4.2. Identification of barley cultivars using a random primer E2. DNA was amplified from a series of commercial barley varieties and some breeding lines (Table 4.1) using a random primer E2. The amplifications were performed as described in Section 3.3.2. The names of the barley cultivars are shown at the top of the lanes. The PCR products were fractionated on 3% agarose gels. DNA size marker (Lane M) is shown on left with the sizes of the bands given in base pairs.

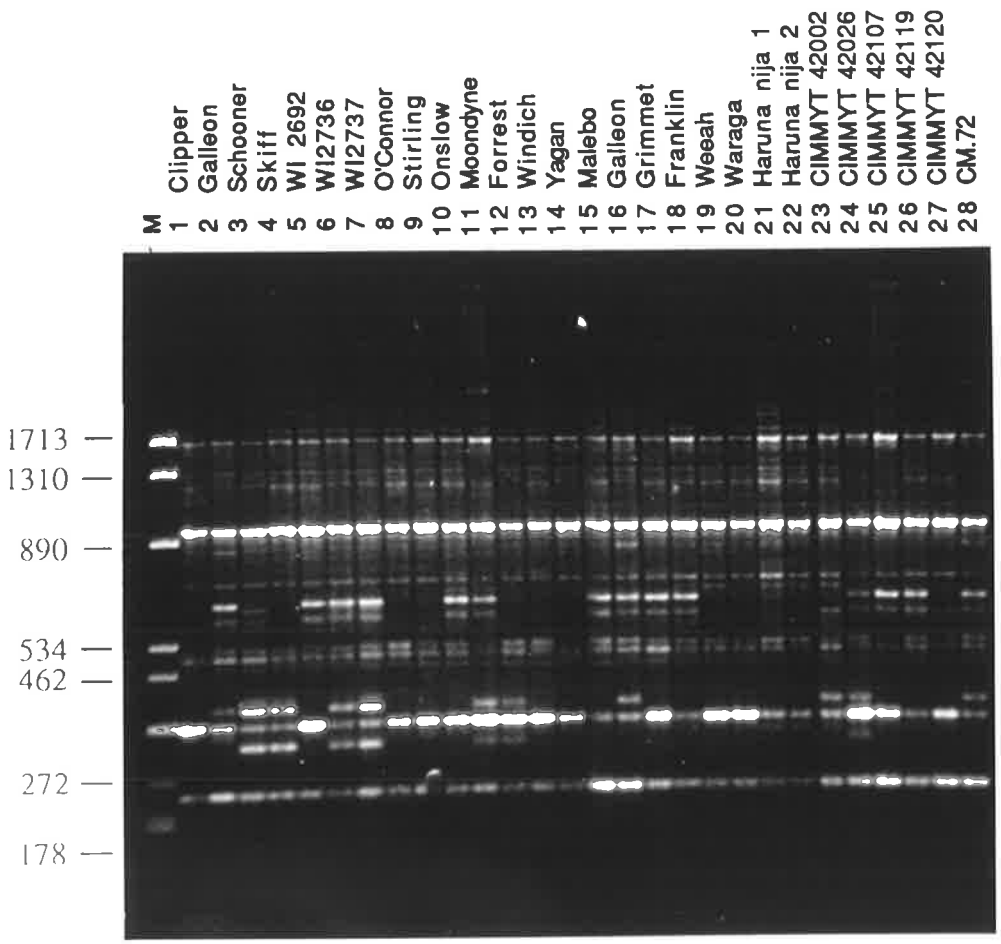

M
1 Clipper
2 Galleon
3 Schooner
4 Skiff
5 WI 2692
6 WI2736
7 WI2737
8 O'Connor
9 Stirling
10 Onslow
11 Moondyne
12 Forrest
13 Windich
14 Yagan
15 Malebo
16 Galleon
17 Grimmet
18 Franklin
19 Weeah
20 Waraga
21 Haruna nija 1
22 Haruna nija 2
23 CIMMYT 42002
24 CIMMYT 42026
25 CIMMYT 42107
26 CIMMYT 42119
27 CIMMYT 42120
28 CM.72
1713
1310
890
534
462
272
178

Fig. 4.3. Identification of barley cultivars using an ISJ primer E4.

DNA was amplified from a series of commercial barley varieties and some breeding lines (Table 4.1) using an ISJ primer E4. The amplifications were performed as described in Section 3.3.2. The names of the barley cultivars are shown at the top of the lanes. The PCR products were fractionated on 3% agarose gels. DNA size marker (Lane M) is shown on left with the sizes of the bands given in base pairs.

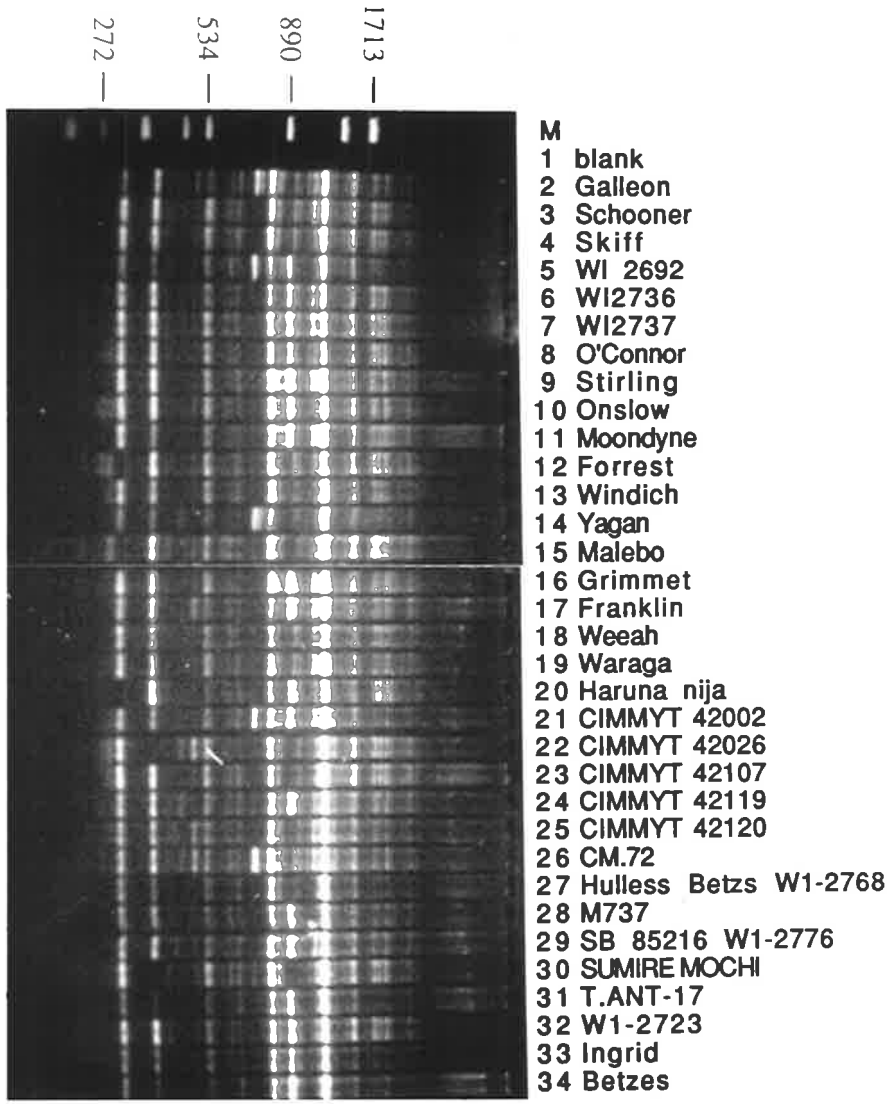
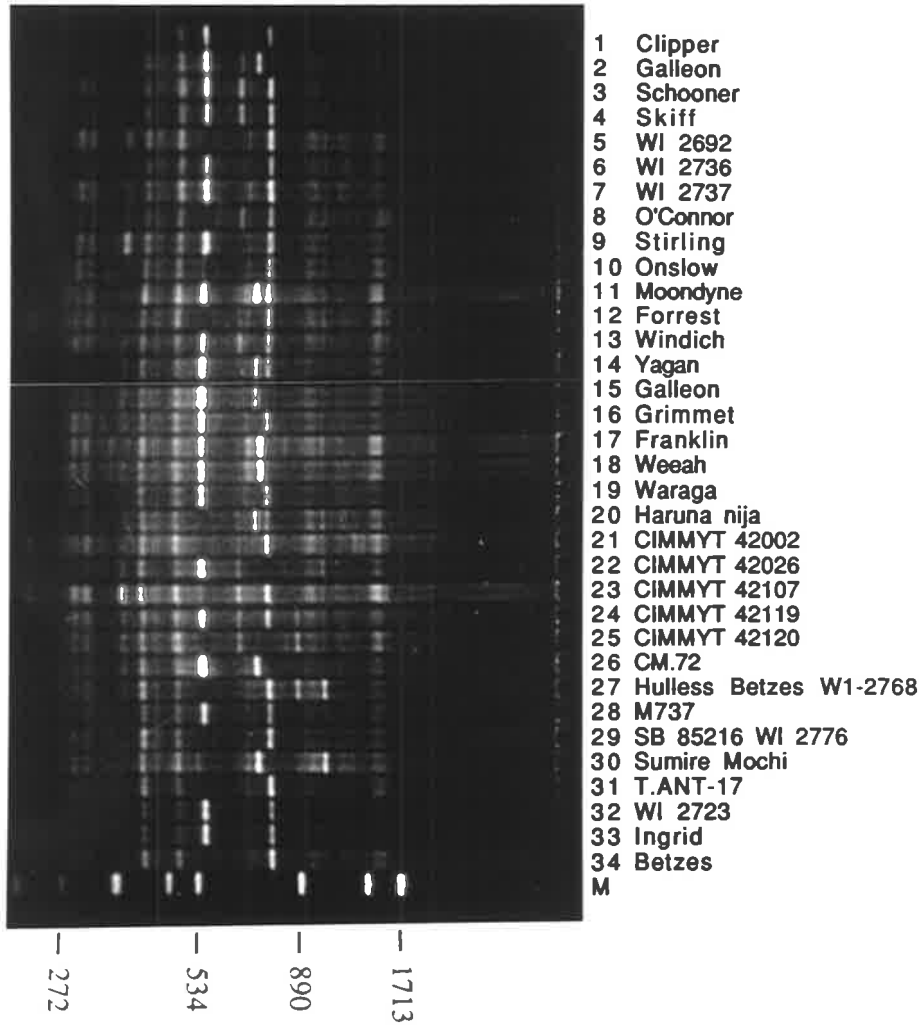

272
534
890
1713
M
1 blank
2 Galleon
3 Schooner
4 Skiff
5 WI 2692
6 WI2736
7 WI2737
8 O'Connor
9 Stirling
10 Onslow
11 Moondyne
12 Forrest
13 Windich
14 Yagan
15 Malebo
16 Grimmet
17 Franklin
18 Weeah
19 Waraga
20 Haruna nija
21 CIMMYT 42002
22 CIMMYT 42026
23 CIMMYT 42107
24 CIMMYT 42119
25 CIMMYT 42120
26 CM.72
27 Hulless Betzs W1-2768
28 M737
29 SB 85216 W1-2776
30 SUMIRE MOCHI
31 T.ANT-17
32 W1-2723
33 Ingrid
34 Betzes

Fig. 4.4. Identification of barley cultivars using an ISJ primer R1 coupled with a random primer S2.

DNA was amplified from a series of commercial barley varieties and some breeding lines (Table 4.1) using an ISJ primer R1 coupled with a random primer S2. The amplifications were performed as described in Section 3.3.2. The names of the barley cultivars are shown at the top of the lanes. The PCR products were fractionated on a 3% agarose gel. DNA size marker (Lane M) is shown on left with the sizes of the bands given in base pairs.

S2 sequence: 5' AATTTCTAGAGGTACCA 3'.


1 Clipper
2 Galleon
3 Schooner
4 Skiff
5 WI 2692
6 WI 2736
7 WI 2737
8 O'Connor
9 Stirling
10 Onslow
11 Moondyne
12 Forrest
13 Windich
14 Yagan
15 Galleon
16 Grimmet
17 Franklin
18 Weeah
19 Waraga
20 Haruna nija
21 CIMMYT 42002
22 CIMMYT 42026
23 CIMMYT 42107
24 CIMMYT 42119
25 CIMMYT 42120
26 CM.72
27 Hulless Betzes W1-2768
28 M737
29 SB 85216 WI 2776
30 Sumire Mochi
31 T.ANT-17
32 WI 2723
33 Ingrid
34 Betzes
M
272
534
890
1713

same PCR products as Clipper but not any other cultivar tested (lane 15 in Fig. 4.1a).

Most of the barley cultivars could be distinguished by specific banding patterns resulting from one primer amplification. The banding profiles of the PCR amplification products with three primers, R1, E2 and E4 respectively, allow the unequivocal identification of all the 49 varieties and breeding lines except between WI 2736 and WI 2737 (e.g. lanes 7 and 8 in Fig. 4.1a), which are two sister lines (Dr D. Sparrow, personal communication). For example, Windich and Schooner shared the same PCR banding pattern with primer R1 (lane 2 and 4 in Fig. 4.1a). Two different banding profiles were obtained between Windich and Schooner when primer E2 was used (compare lanes 3 and 13 in fig. 4.2).

Some of the PCRs with the same varieties are duplicated with different accessions chosen randomly from the Waite Institute Collection. These accessions of the same varieties are different with each other by the year of harvest or site collected. No polymorphisms were observed between accessions of a variety (e.g. lanes 21 and 22 in Fig. 4.2 for cultivar Haruna Nijo). The barley cultivars studied in this Section include all of the major varieties currently grown commercially in Australia.

4.4. Identification of polymorphisms between wheat varieties and breeding lines

The wheat cultivars investigated in this study are listed in Table 4.2. A similar approach to that described in section 4.2 was used to study the variability between wheat varieties. Fig. 4.5a shows the amplification products obtained with the R1 primer from 21 commercial varieties and breeding lines of wheat.

Table 4.2. Wheat cultivars used in this study

Cultivar No.	Cultivar
1	Tatiara
2	Oxley
3	Warigal
4	Schomburgk
5	Bindawarra
6	Festiguay
7	Yallaroi
8	Schomburk
9	Spear
10	YR 10 Schomburk
11	Molineux
12	Halberd
13	Greek 6450
14	Kenya Farmer
15	W1'MMC 1/10
16	Aus. 4784
17	(MX'Sch#3)A8/I12/8
18	Aroona
19	Smooth Sheet
20	Guillemot

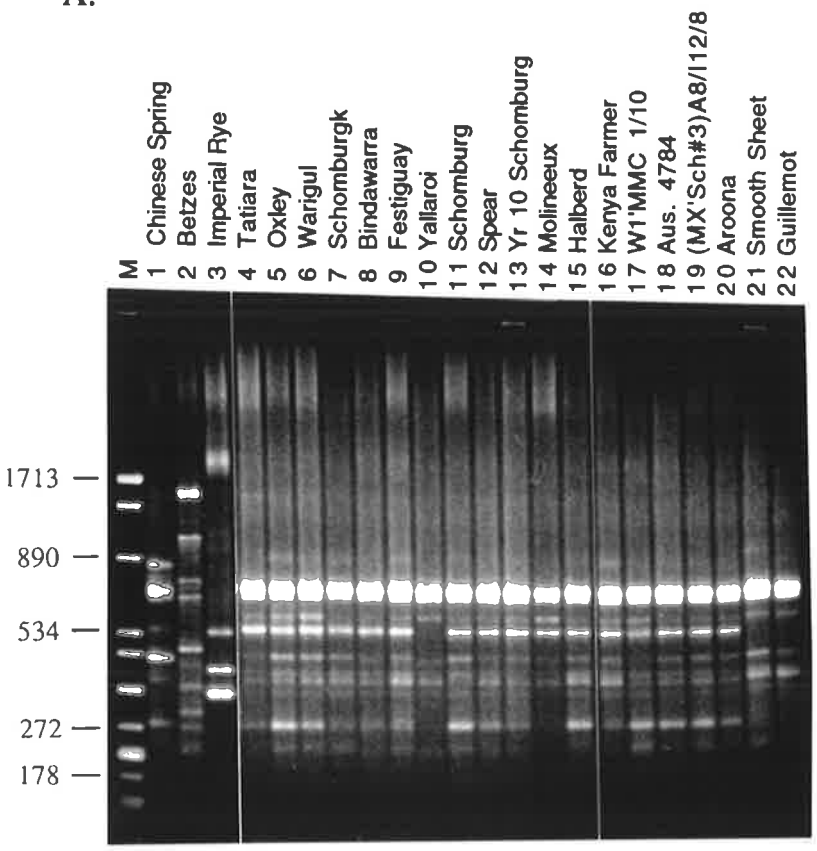
Fig. 4.5. Identification of wheat cultivars using an ISJ primer R1 and a random primer E2.

The amplifications were performed as described in Section 3.3.2. DNA size marker (Lane M) is shown on left with the sizes of the bands given in base pairs. The PCR products were fractionated on 3% agarose gels.

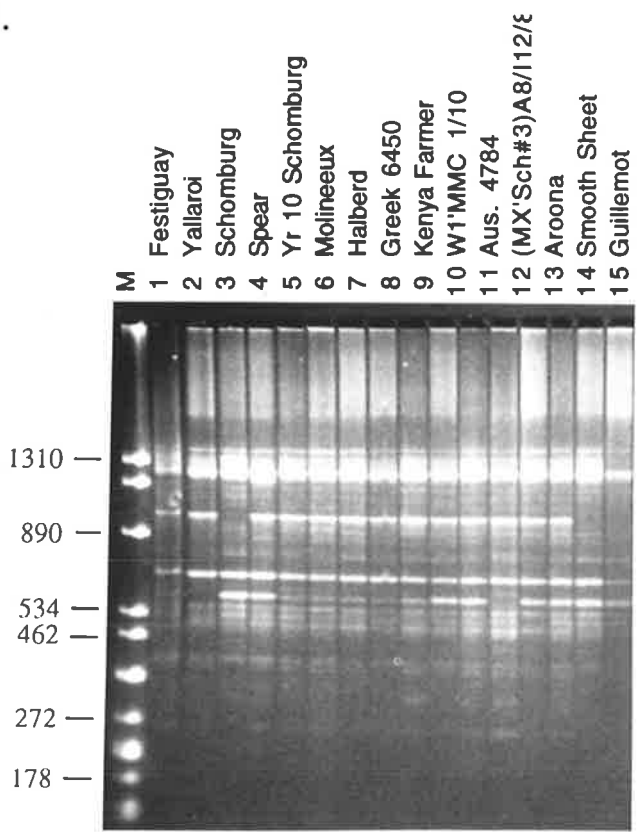
A. DNA was amplified from a series of commercial wheat varieties and some breeding lines (Table 4.2) using ISJ primer R1. The names of the wheat cultivars are shown at the top of the lanes. Barley cv. Betzes and rye cv. Imperial are included for references (lanes 2 and 3).

B. DNA was amplified from a series of commercial wheat varieties and some breeding lines (Table 4.1) using random E2. The names of the wheat cultivars are shown at the top of the lanes.

A.


M
1 Chinese Spring
2 Betzes
3 Imperial Rye
4 Tatiara
5 Oxley
6 Warigul
7 Schomburgk
8 Bindawarra
9 Festiguay
10 Yallaroi
11 Schomburg
12 Spear
13 Yr 10 Schomburg
14 Molineeux
15 Halberd
16 Kenya Farmer
17 W1'MMC 1/10
18 Aus. 4784
19 (MX'Sch#3)A8/I12/8
20 Aroona
21 Smooth Sheet
22 Guillemot
1713
890
534
272
178

B.


M
1 Festiguay
2 Yallaroi
3 Schomburg
4 Spear
5 Yr 10 Schomburg
6 Molineeux
7 Halberd
8 Greek 6450
9 Kenya Farmer
10 W1'MMC 1/10
11 Aus. 4784
12 (MX'Sch#3)A8/I12/ξ
13 Aroona
14 Smooth Sheet
15 Guillemot
1310
890
534
462
272
178

There are four reproducible polymorphic PCR bands between the wheat cultivars. Wheat cv. Chinese Spring and Barley cv. Betzes were also included as references. The banding patterns of the PCR products from the three cultivars are completely different from each other; essentially there is no common band shared between them. The PCR products with E4 are displayed in Fig. 4.5b. Three polymorphisms could be detected among the cultivars tested (Note that not all of the cultivars were tested with both of the R1 and E4 primers). It appeared that there is less variability in the length of amplified DNA fragments between wheat populations compared with that between barley populations (section 4.2). With primer E4, there were more minor PCR bands in wheat than barley (compare Fig. 4.5b and Fig. 4.3). Sometimes these minor bands appeared as a slight smear (Fig. 4.5b).

4.5. Identification of polymorphisms within populations

4.5.1. Barley

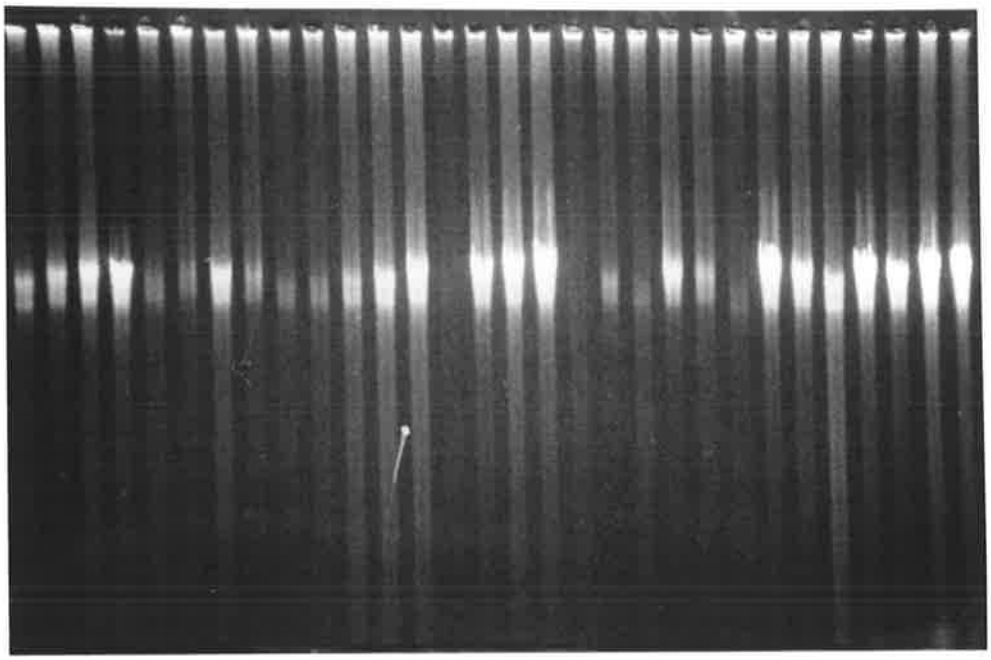
The ISJ primers R1 and E4 were used to study the variation within a population of barley. Mini-prep DNA (as described in section 2.2.4.) was made from 50 individual seedlings of barley cv. Clipper for PCR amplification with R1. The PCR results, presented in Fig. 4.6b (only 30 displayed), show that there is a consistent banding pattern between the 50 individual plants of barley cv. Clipper with one exception (indicated by an arrow). The considerable variation in the concentration and quality of the mini-prep DNA between individuals (Fig. 4.6a) seemed to have no effect on the consistency of the PCR products amplified from them (Fig. 4.6b). When the same plant DNAs were amplified with the E4, a uniform banding pattern similar to that with R1 was observed (Fig. 4.7, only 16 displayed). The individual however, which showed polymorphic PCR products with R1, did not reveal any variation with E4 (indicated by an arrow in

Fig. 4.6. Identification of intra-population polymorphism in barley cv. Clipper using an ISJ primer R1.

A. Mini-prep genomic DNA was isolated from individual Clipper plant (Section 2.2.4) and fractionated on a 1% agarose gel.

B. DNA was amplified from individual Clipper plants (Section 3.3.2). The polymorphic individual is marked with an arrow at the top of the lane. The PCR products were fractionated on a 3% agarose gel.

A.



B.

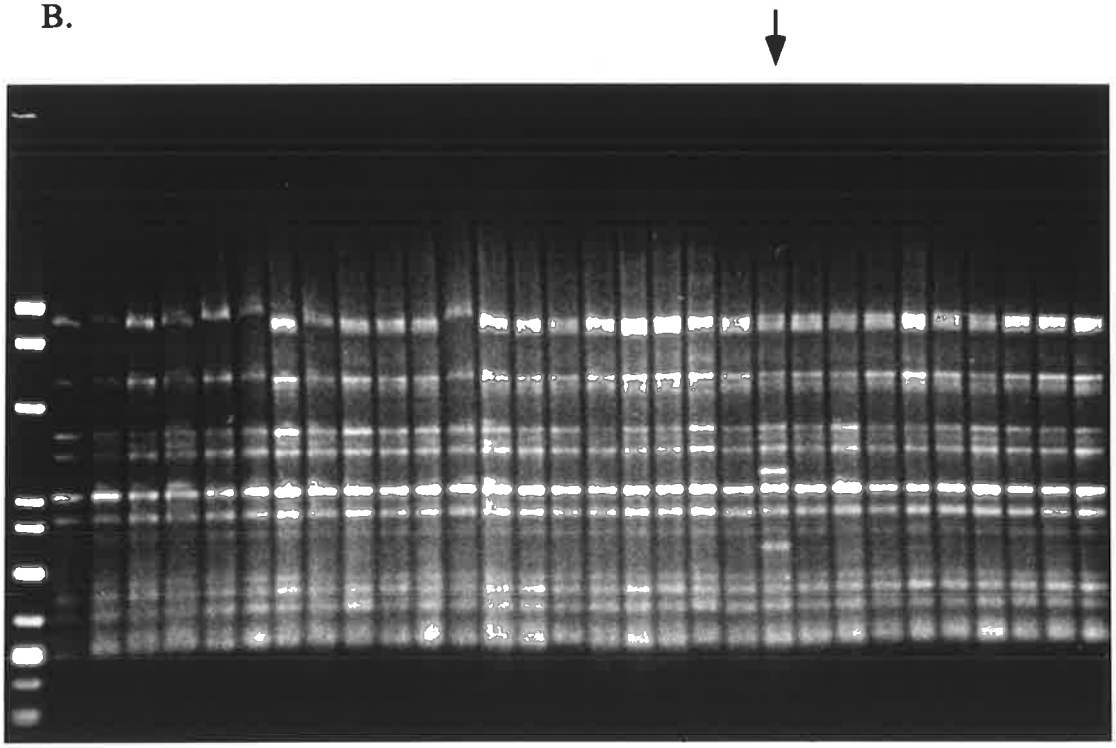


Fig. 4.7. Identification of intra-population polymorphism in barley cv. Clipper using an ISJ primer E4.

DNA was amplified from individual Clipper plants (Section 3.3.2). The polymorphic individual appeared in Fig. 4.6b is marked with an arrow at the of the lane. DNA size marker is shown on left with the sizes of the bands given in base pairs. The PCR products were fractionated on a 3% agarose gel.

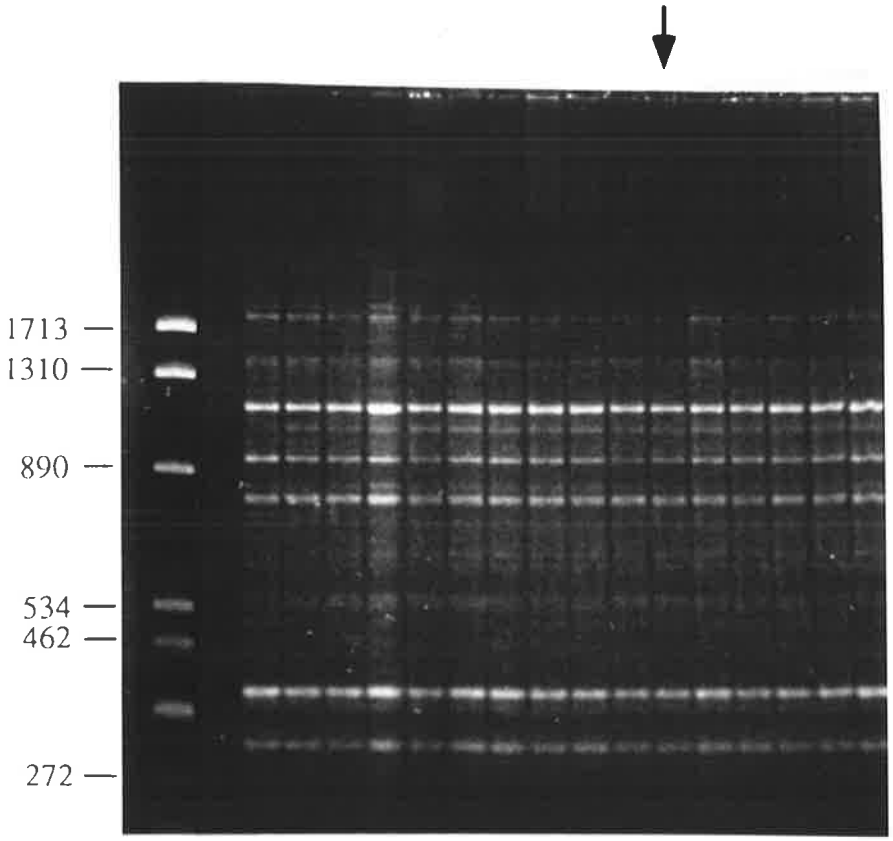

1713
1310
890
534
462
272

Fig. 4.8. Identification of intra-population polymorphism in barley ~~varietie~~ variety Haruna Nija using an ISJ primer E4.

DNA was amplified from individual Haruna Nija plants using ISJ primer E4 (Section 3.3.2). DNA size marker is shown on left with the sizes of the bands given in base pairs. The PCR products were fractionated on a 3% agarose gel.

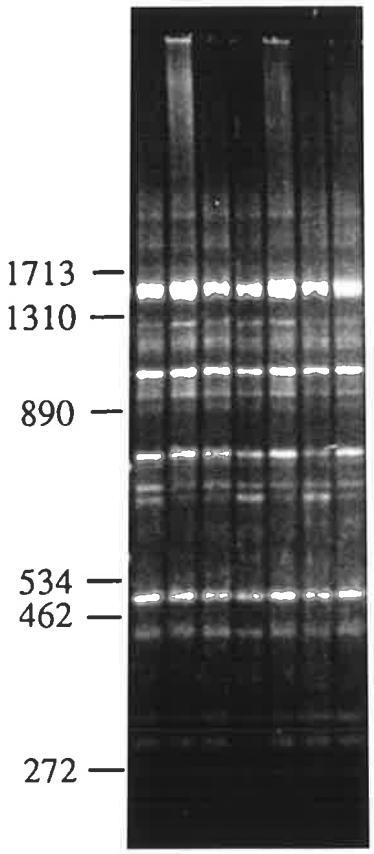

1713
1310
890
534
462
272

Fig. 4.6b). The probability of contamination during the PCR process can be ruled out because separate PCR experiments from the same DNAs resulted in the same PCR products (data not shown). Based on the monomorphic bands shared between this individual and other barley cutivars, Clipper is clearly the only one to which it should belong. As this polymorphic pattern in Clipper (indicated by an arrow in Fig. 4.6b) does not match that of any other barley, wheat or rye cultivar, it could be a true polymorphism or crosspollination. This result is consistent with a rate of 1% cross pollination typically found in barley (Dr. D Sparrow, personal communication).

Besides Clipper, some other barley varieties were also examined for variations between individuals albeit on a smaller scale. No polymorphism was observed among the individual plants of any of these varieties. An example is shown in Fig. 4.8 which displays the amplified products from 7 individual plants of Haruna Nija. All of the PCR bands appeared to be the same between these individuals of this Japanese variety.

4.5.2. Rye

The same set of primers, ISJ primers R1 and E4, was used to analyse the intrapopulation polymorphisms in rye. The rye cultivars used in this study are South Australian Rye (SA Rye) and Imperial Rye. SA Rye is a commercial variety in South Australia and Imperial Rye a variety grown on Waite Campus for research purposes. The primer R1 did not produce any polymorphic band among the individuals of Imperial Rye (Fig. 4.9a), whereas at least three polymorphic bands can be seen in SA Rye (Fig. 4.9b). When E4 was used, numerous polymorphic bands were seen among the individuals from either Imperial or SA Rye (Fig. 4.10). The levels of intrapopulation polymorphisms appeared very

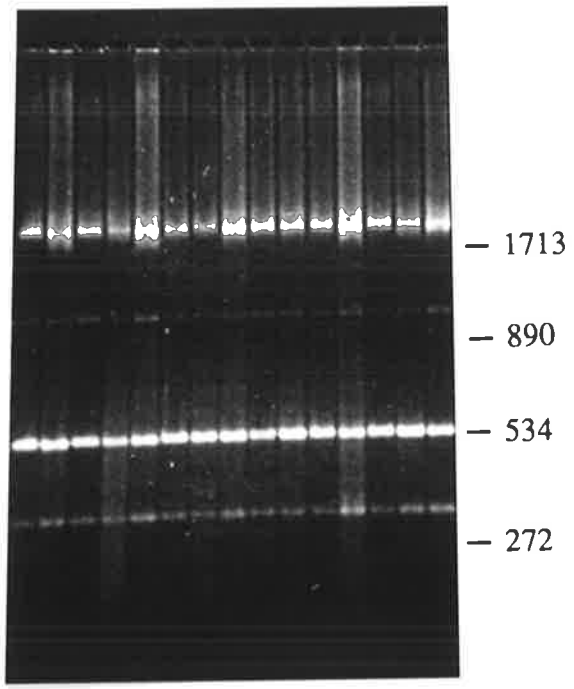
Fig. 4.9. Identification of intra-population polymorphism in rye varieties using an ISJ primer R1.

DNA size marker is shown on right with the sizes of the bands given in base pairs. The PCR products were fractionated on 3% agarose gels.

A. DNA was amplified from individual Imperial Rye plants using ISJ primer R1 (Section 3.3.2).

B. DNA was amplified from individual South Australia Rye plants using ISJ primer R1 (Section 3.3.2).

A.



B.

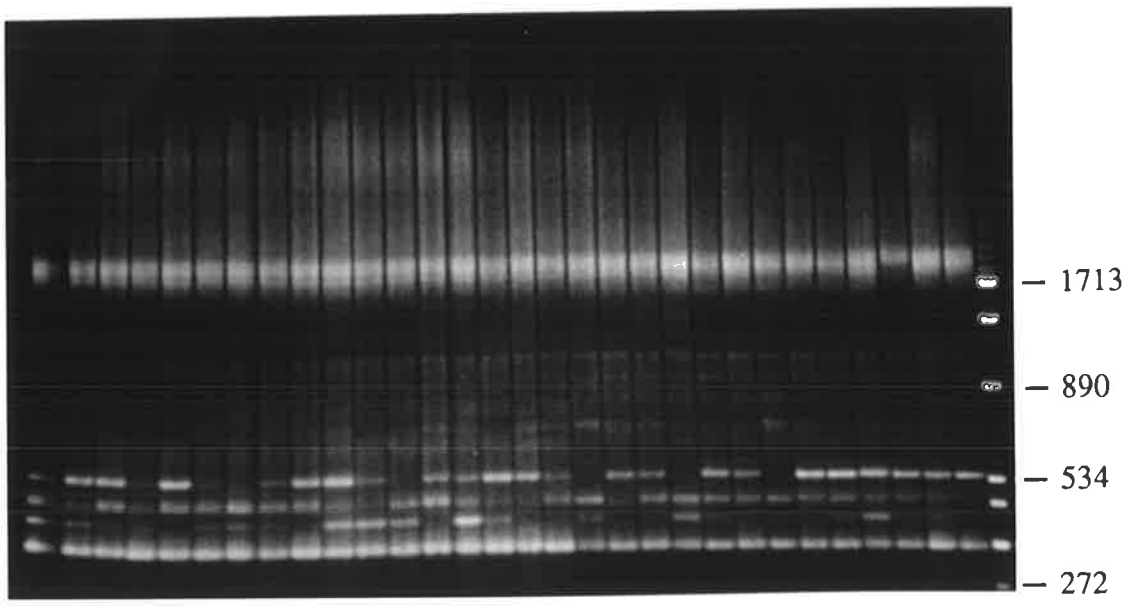
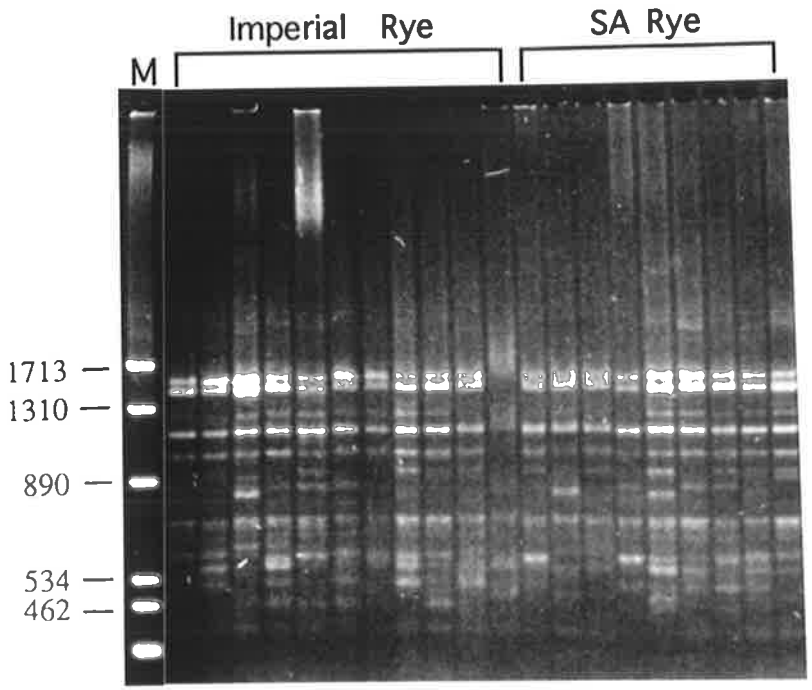


Fig. 4.10. Identification of intra-population polymorphism in rye varieties using an ISJ primer E4.

DNA was amplified from individual Imperial Rye or South Australia Rye plants using ISJ primer E4 (Section 3.3.2). DNA size marker is shown on left with the sizes of the bands given in base pairs. The PCR products were fractionated on a 3% agarose gel.


Imperial Rye
SA Rye
M
1713
1310
890
534
462

similar between these two rye populations. The intrapopulation variability among rye cultivars seemed to be higher than that of the barley cultivars.

4.6. Amplification of seed DNA by PCR

The application of molecular techniques on seeds would be very useful to researchers and especially, plant breeders. Unfortunately, a single seed does not contain enough good quality DNA for RFLP analysis (Ainsworth and Sharp, 1989). PCR has recently become the preferred tool in many areas of plant molecular biology. Generally, the templates for PCR are genomic DNAs isolated either from leaf or root tissues (Berthomieu and Meyer, 1991). The PCR amplification of genomic DNA from seeds is hampered, presumably, by unknown substances present in the DNA extraction.

Random primer E2 was tested to find the optimal conditions for seed DNA amplification. Genomic DNAs were isolated from single leaves and half seeds, respectively, of barley cv. Clipper. The genomic DNA from half seed was purified as from leaf DNA in section 2.2.4. PCR amplifications were carried out as described in section 3.2. The PCR products from leaf and seed DNAs were compared in Fig. 4.11. The results show that no PCR products could be detected in seed DNA amplification, whereas normal PCR products were obtained from leaf DNA (compare lane 1 and 2 in Fig. 4.11a). It seemed that certain substances present in seed DNA but not in leaf DNA and these inhibited the PCR reaction.

If the seed PCR was hampered by inhibitors, further dilution of the genomic DNA could reduce their concentration in the PCR reaction buffer. Higher concentration of *Taq* polymerase may increase the chance of proper amplification. Based on these assumptions, the DNA from half seed was

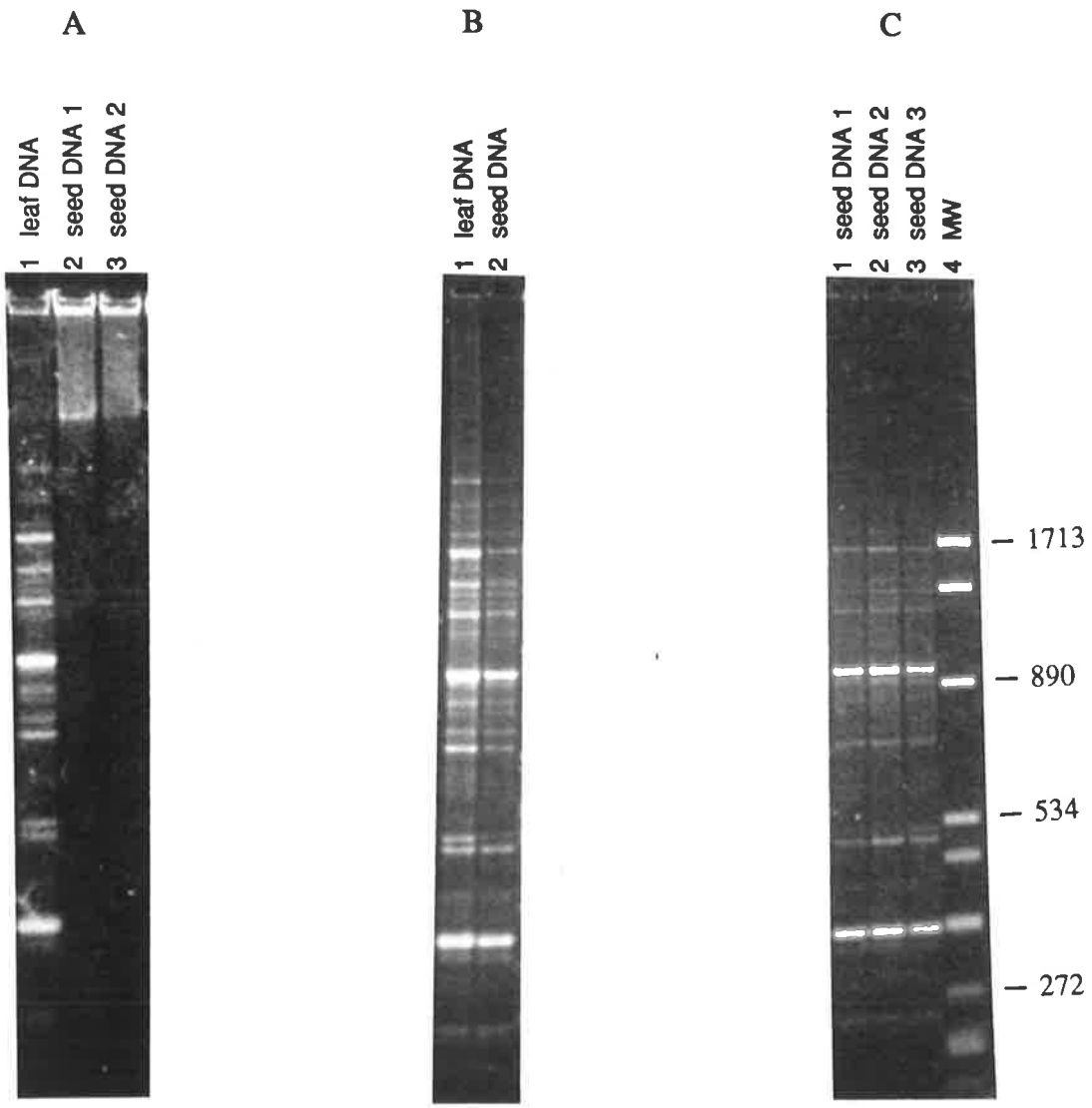
Fig. 4.11. Amplification of seed DNA.

Mini-prep genomic DNA was isolated as described in section 2.2.4. The PCR products were fractionated on 3% agarose gels.

A. DNA was amplified from half seed or leaf of barley cv. Clipper as described in Section 3.3.2.

B. DNA was amplified from half seed or leaf of barley cv. Clipper as described in Section 4.6.

C. DNA was amplified from three half seeds of barley cv. Clipper DNA as described in Section 4.6. DNA size marker is shown on right with the sizes of the bands given in base pairs.


A
1 leaf DNA
2 seed DNA 1
3 seed DNA 2
B
1 leaf DNA
2 seed DNA
C
1 seed DNA 1
2 seed DNA 2
3 seed DNA 3
4 MW
1713
890
534
272

further diluted and more *Taq* polymerase was added to each PCR reaction. The mini-prep DNA from a half seed was dissolved in 150ul instead of 50ul as the original volume (section 2.2.4). 1 ul of the seed DNA extraction was used as templates for each PCR. The concentration of *Taq* polymerase was increased from about 0.3 unit/reaction to 3 units/reaction. These modification of the PCR reaction had the desired effect. Fig. 4b showed that successful amplification could be achieved. Clearly all of the major bands with leaf DNA were also present with seed DNA (compare lane 1 and 2, Fig. 4.11b). To check the efficiency and reliability of this method, several samples from individual half seeds were used for amplification. Consistent results were achieved with each of the half-seed samples (Fig. 4.11c). However, the PCR bands with seed genomic DNA seemed to be slightly fainter compared to that from leaf DNA (compare lanes 1 and 2 in Fig. 4.11b). Apparently, this lower efficiency of amplification can be also attributed to the effect of the postulated inhibitors.

4.7. Discussion

Intra- and interpopulation variability is fundamentally important not only to the phylogenetic studies, but also to any method for identifying and fingerprinting crop cultivars. The high cost and technical complexity of RFLP has restricted its wide application: many RFLP analyses on plant phylogeny (Kochert *et al.*, 1991; Miller and Tanksley, 1990) could only afford to use one or a small number of individuals from each population instead of using populations of a reasonable size as in the case of isozymes (Brown, 1990). The procedure of PCR is as simple as conventional isozyme assays. Moreover, PCR needs only a fraction of the amount of DNA for RFLPs and it can tolerate low-quality DNA templates such as crude mini-prep DNAs. Another important feature of PCR is

that normally no radioactive material is involved. One of the drawbacks of RFLP is that a single or half seed can not provide enough good quality DNA for the analysis. The application of PCR, under the conditions described in section 4.6, solved this problem. This method provides a particularly handy tool as the cereal breeders can use half of the seed for PCR analysis while the other half with the embryo is still available for germination.

One of the outcomes from the studies in this chapter is the ready identification of any of the barley varieties currently grown commercially in Australia. Besides the Australian cultivars, the studies in this section also included a number of barley varieties and breeding lines from Asia, Europe and South America. Each of them, except two Australian sister lines, can be readily identified with the maximum of 3 primers. We can expect that it should be relatively easy to adapt this method to identify any barley cultivars from all over the world.

DNA polymorphisms detected by PCR within populations of the out crossing species rye (*Secale cereale*) cv. South Australia are higher than that of the selfing species wheat (*Triticum aestivum*) and barley (*Hordeum vulgare).* For the primer R1, cv. 'Imperial' showed no variation between individuals which is characteristic for a barley variety. However, with E4, 'Imperial' displayed a similar level of variability to that of 'South Australia'. So it seemed that the intrapopulation polymorphism of rye cv. 'Imperial' is higher than that of barley varieties but lower than rye cv. 'South Australia'. The explanation could be that rye is capable of selfing and rye cv. 'Imperial' used in this study has been grown in a confined environment at the Waite Institute for generations. As rye cv. 'South Australia' was obtained from a commercially grown field, it is not surprising that it showed higher polymorphisms than rye cv. 'Imperial'. Even from this as yet limited study, the results indicate that the geographic isolation

may have played a fundamental role in the development of intrapopulation variations in outcrosser rye. From the data presented in this section, it appears that the intrapopulation polymorphisms is higher in outcrosser than in selfer. This observation is consistent with data from isozymes and/or RFLPs for cereals and other plants (e.g. Asins and Carbonell, 1986; Baatout *et al.*, 1991). However, it should be noted that cereal breeders have put a great effort to breed pure breeding lines and varieties. Imperial rye had been inbred in the breeding program (Drs KW Shepherd and D. Sparrow, personal communication).

Compared with barley and rye cultivars, wheat varieties showed markedly fewer polymorphisms. This phenomenon could be attributed to the narrow genetic base of wheat associated with its relatively recent origin (Bell, 1987). Previous reports using RFLPs already support this statement (Chao *et al.*, 1989; Kam-Morgan *et al.*, 1989; Liu and Tsunewaki, 1990).

PCR, for its power and easy handling, will also greatly facilitate the conservation of germplasms, especially the so called "core collection" (Brown, 1989, 1991). The identification of intrapopulation polymorphism by PCR will help to provide a manageable and representative sample of the accession and to improve the evaluation and use of existing collections. The significance of the construction of genetic maps is fully discussed in section 1.1.7. The intervarietal polymorphism revealed by PCR will certainly benefit the development of such maps and the linkage studies of divergent characters, as demonstrated in Chapter 3.

To provide a quantitative assessment of relatedness of the varieties, the fingerprinting profiles revealed by PCR in this study were analysed by parsimony analysis using the PAUP and Phylip computer programs (section 2.3). However, neither of them provide consistent results. For example, the PCR banding profiles of the barley varieties were analysed by the PAUP and Phylip

programs, respectively. Family trees were produced by these programs. The trees appeared completely different if the data were keyed into the computer in a different sequence (data not shown). Similar problems were also encountered in other species with similar type of data.(Dr. J. Manner, personal communication). Further discussion on the application of PCR to phylogenetic studies will be given in Chapter 7.

The results in this Chapter demonstrate that PCR has the potential to become an invaluable tool in a diverse arsenal of techniques employed by plant breeders and systematists to unravel the mysteries surrounding the origin and evolution of plants, identifying and fingerprinting individual crop cultivars, and facilitating breeding programs and the conservation of germplasms.

Chapter 5

Attempted cloning of 6-PGD gene through the polymerase chain reaction.

5.1. Introduction

The molecular studies of individual genes is one of the most important aspects of modern molecular biology. Compared to human genetic studies, there are relatively few genes from cereals that have been cloned and sequenced. One of the traditional approaches is to use Southern blotting to test the ability of a DNA probe from one species to hybridize with the DNA from the genomes of other species. PCR has fast become an invaluable tool in this field. In normal PCR applications, the target fragment has been cloned and sequenced although frequently, the DNA sequence of a region to be analysed is not entirely known. Several new PCR-based methods have been developed to overcome this limitation (section 1.2.1.2). PCR appears to be a particularly promising technique for investigating unknown genes in cereals as many genes have been studied extensively in other organisms and their sequence data are available.

In the present study, PCR has been adopted in an attempt to isolate the cereal gene for 6-PGD (6-phosphogluconate dehydrogenase). This gene has been mapped in rye but its DNA sequences is unknown.

The 6-PGD gene is important since it is linked to a rye gene controlling resistance to the cereal cyst nematode (*Heterodera avenae* Woll) (CCN). CCN

is a serious pest of grain crops in Australia and many other countries (Meagher, 1977). A gene for CCN resistance has been found in a line of triticale T-701 (Fisher, 1982). Studies have indicated that this resistance is far more effective than any CCN resistance gene presently existing in wheat (Asiedu, 1986). The location of this character has been traced to rye chromosome 6R (Asiedu *et al.*, 1990; Dundas *et al.*, 1991).

Attempts are being made at the Waite Institute to transfer this 6R nematode resistance gene from rye to wheat (Dundas *et al.*, 1992). The principle for gene transfer was that of homoeologous recombination used successfully by Riley *et al.* (1968), Sears (1973) and Joshi and Singh (1979). Pairing between the rye 6R chromosome and 6D of wheat has been shown to occur in the presence of mutant gene *ph1b*. Recombinants have been detected by checking for the presence or absence of isozyme markers known to be located on the same arm of the rye chromosome as does the CCN resistance gene. Two of the these marker loci in routine use are GOT (glutamate oxaloacetate transaminase) and 6-PGD (6-phosphogluconate dehydrogenase) (Dundas *et al.*, 1991). The CCN resistance locus, designated as *CreR*, was further mapped on the interstitial section of the long arm of 6R between *Got-R2* and *6-Pgd r1b* (Dr I. Dundas, personal communication).

At present methods of screening for CCN resistance are difficult and time consuming, all being based on bioassay systems. The breeding program to transfer this important gene, will benefit greatly by the isolation of molecular markers closely linked to the CCN resistance gene. Therefore, cloning of 6-PGD could provide not only a useful DNA marker on chromosome 6R but may also ultimately allow "chromosome walking" to *CreR*.

An attempt to isolate a 6-PGD gene from wheat is described in this chapter.

5.2. Southern analysis cereal DNA with 6-PGD clones from *E. coli* and rat

Cross-hybridization of the 6-PGD probes, RW229/pMN1 from *E. coli* (Nasoff and Wolf, 1980) and p6PGD-1 from rat (Miksicek and Towles, 1983), to cereal genomic DNAs was attempted using a gradient of stringency of Southern hybridization. This was done by altering the concentration of formamide in the hybridization buffer. The concentration of formamide varied from 25% to 40% with 5% intervals. The remainder of the hybridization components were kept constant. No signal of cross hybridization was observed under any conditions tested (data not shown).

In another attempt, a wheat cDNA library in λgt10 was obtained from Mr LH Ji (Waite Institute). This library originated from wheat leaf and contained about 10^5 λ clones (Ji, 1992). The 6-PGD clones from *E. coli* and rat were used to screen the λ library at a plating density of about 25,000 pfu per plate. The same gradient of stringency for hybridization buffer as described above was applied for the λ screening. Five hybridization spots could be detected with RW229/pMN1 (from *E. coli*) on the autoradiograph when the formamide concentration was raised to 30% in the hybridization buffer. The corresponding regions on the plates were cored out and plated again, at a lower pfu, for second screening. No hybridization signal was detected with the rat 6PGD probe during the first or the second screening. Two of the phages displayed positive hybridization with the *E. coli* probe in the second screening. λ DNA was isolated from each of these clones (section 2.2.5). Once again, RW229/pMN1 and p6PGD-1 were used as probes for the Southern hybridization. However, neither *E. coli* or rat 6PGD probe could reveal any Southern band (data not shown). No homology was observed between these

two clones and the 6-PGD probes. The positive signals from the plate screening may have been caused by the cross hybridization between the *E. coli* and λ DNAs.

As both Southern hybridization to cereal DNA and the screening of the cDNA library had failed, a different approach was attempted for cloning of the 6-PGD gene.

5.3. Partial cloning and sequencing of the 6-PGD gene

5.3.1. Outline of the cloning procedure

The results in the previous section indicated that insufficient sequence homology exists between the cereal and *E. coli* and rat 6-PGD genes for cross hybridization. However, it has been reported that evolutionarily conserved portions of a gene can be used to construct primers for PCR that will allow the amplification of variable sequences that lie between the conserved templates for these primers (see review by Rose, 1991). It has been established that within the 6-PGD gene the amino sequences VKMVHNGIEYGDMQLI and QAQRDYFGAHTY are identical in *E. coli, Drosophila* and sheep (Dr. M. Scott, personal communication). These two conserved regions are 800 bp apart in the *E. coli* 6-PGD gene, *gnd* (Nasoff *et al.*, 1984).

The two primers for the cloning of 6-PGD, P1 (5' GTGAAGATGGTTCA CAACGGTATTGAATATGGTGATATG 3') and P2 (5' ACCAAAATAGTCACG 3'), are based on the degenerated DNA sequences from the above protein sequences. Unfortunately, the *E. coli* 6-PGD gene was the only 6-PGD gene sequence available when the present study was

in progress. The relative location of the primer P1 and P2 corresponding to these conserved regions in *E. coli* is shown schematically in Fig. 5.1.

5.3.2. Characterization and sequencing of the 6-PGD clone

When P1 and P2 primers were used to amplify cereal genomic DNA, a band of 810 bp could be observed in wheat (cv. Chinese Spring) (lane 1 in Fig. 5.2), whereas no PCR product could be detected in barley (cv. Betzes) or rye (cv. Imperial) (lanes 2 and 3 in Fig. 5.2). Using a 6-PGD probe from rat, p6PGD-1, Southern analysis showed that this band had strong homology with p6PGD-1 under standard hybridization conditions (Fig. 5.3).

The PCR band from wheat (indicated by an arrow in Fig. 5.2) was then isolated from the agarose gel by the Geneclean method (section 2.2.6). The ends of the DNA fragment were filled with Klenow fragment of DNA polymerase I and then ligated into Sma I-digested plasmid, pUC19 (section 2.2.6).

One of the recombinant clones was picked. After plasmid DNA isolation and restriction enzyme analysis, it has demonstrated that this clone contained the 810 bp band (Fig. 5.4). This clone was subsequently named 6WP. Based on the restriction enzyme analysis of 6WP (e.g. Fig. 5.4), a restriction map was constructed (Fig. 5.5).

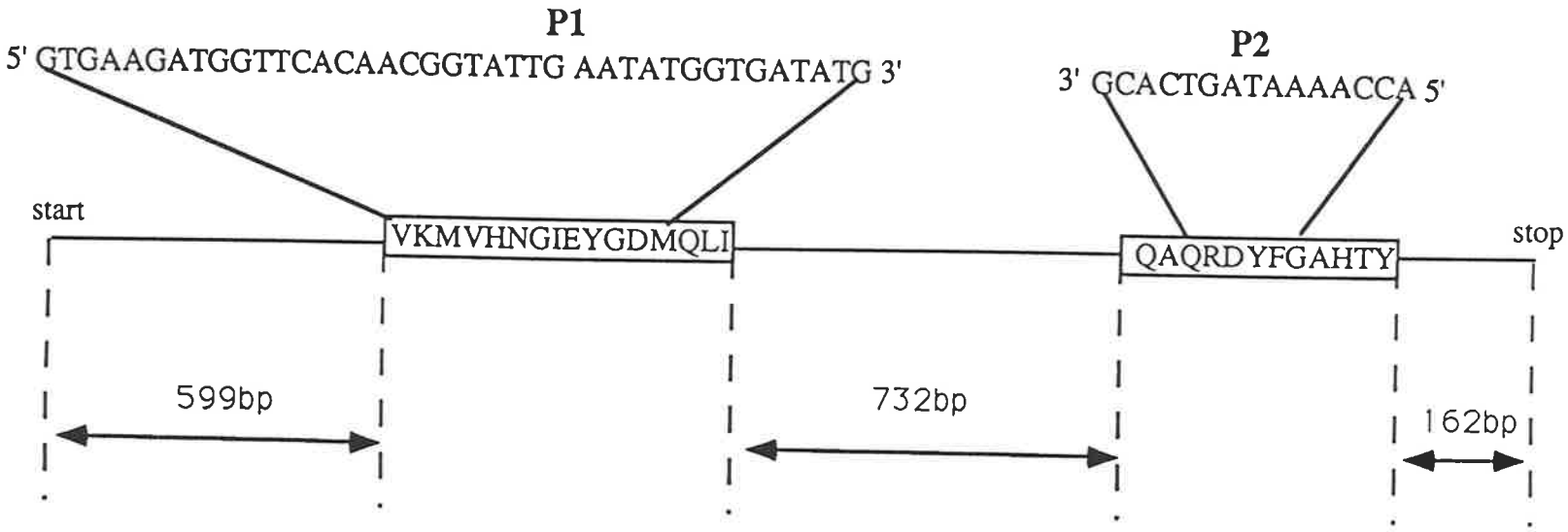


Fig. 5.1 Schematic representation of the regions homologous to the P1 and P2 primers in *E. coli*. (adapted from Nasoff *et al.*, 1984).

Fig. 5.2. Amplification of cereal DNA using 6-PGD primers P1 and P2. Amplification was performed as described in Section 2.2.12 and 2.2.13. The PCR products were fractionated on a 3% agarose gel. DNA size markers are shown in Lane MW with the sizes of the marker bands given on the right in base pairs.

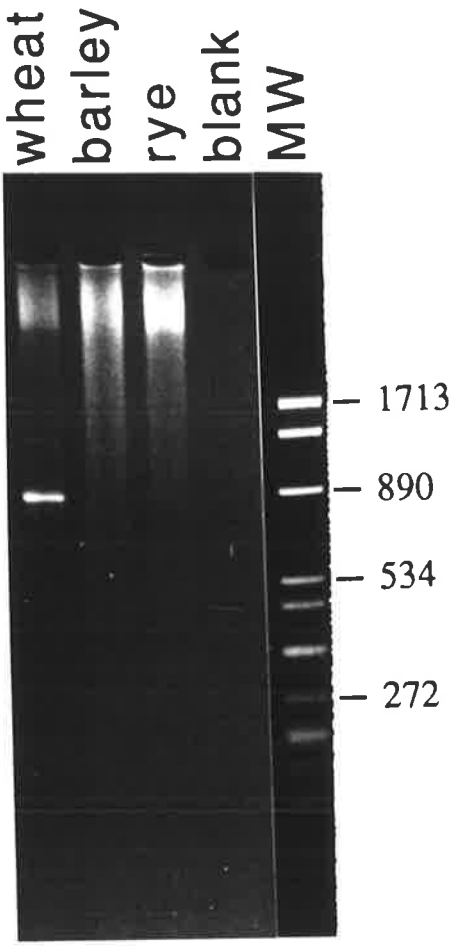
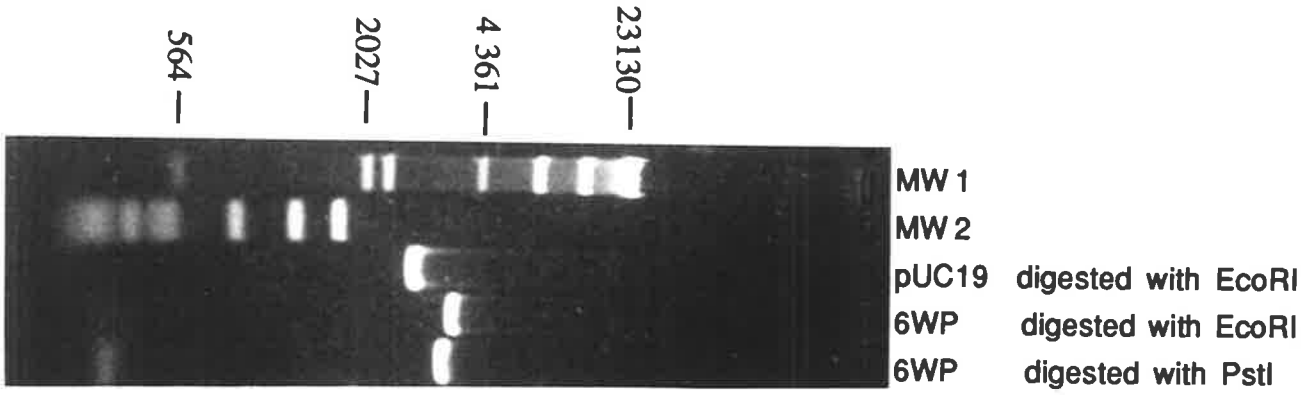

wheat
barley
rye
blank
MW
1713
890
534
272

Fig. 5.3. Autoradiograph of hybridization of 6-PGD probe p6PGD-1 to Southern filter made from gel shown in Fig. 5.2.

810
wheat
barley
rye
blank

Fig. 5.4. Restriction of clone 6WP with different endonucleases.
The size markers used for the gel were λdv1 DNA digested with Hae III (Streeck and Hobom 1975) and λ DNA digested with Hind III (lane MW1) .


564
2027
4 361
23130
MW 1
MW 2
pUC19 digested with EcoRI
6WP digested with EcoRI
6WP digested with PstI

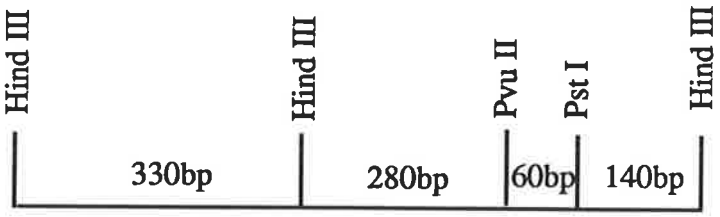


Fig. 5.5 Restriction map of 6WP

The clone 6WP was further analysed by DNA sequencing with the dideoxy chain-terminating method of Sanger (1977). The sequencing of 6WP was performed with a cycle sequencing kit in a automated DNA sequencing system 373A developed by Applied Biosystems Inc. (ABI) (section 2.2.16). One strand of the 6WP clone was sequenced with universal primer for one direction and M13 reverse primer for the other.

The nucleotide sequence obtained is displayed in Fig. 5.6. Approximately 400 bases could be determined for each sequencing reaction.

TACACCCTGG TTCACAACGG TATNGAATAT GGTGATATGC AGCTGATCGC
primer P1

AGAAGCTTAT GCTCTGNTGN ANGGCGCTCT GNGCCTGAAC AACGAAGAGC

TGGCTGAAAC CTTCACCGAGT GGAACAACGG CGAGCTGAG CAGCTACCTG

ATCGACATCA CCAGAGATAT GTTCACCAAG AAAGACGAAG AAGGTAAATA

CCTGGTTGNT GTCCTCNTGG ATGAAGCAGC AAACAAAGGT ACNGGTAAAT

GGGCCAGCCA GAGCTCACTG GATCTTNGCG AGCCTCTGTC ACTGATGNGC

GAGTCTGTTT TTGCCCGTTA CCTCTCGTCA CTGAAAACGC AGCGCGTTGC

GGCNTCGAAA GTACTNAGCG GTNCACAGGN TAAAGCTTTC ACCGGCGNCA

AAGCAGAGAA AAAAGTATCT NCCACGCTCT GTACCTTGGG NNAAAAATCG

TNNTCCNAAC TCTCAGGNCT NCTCCTCAGA CTTCGTGCCG CGCCAGACTA

GTAACAANTG GGATCTNAAC TGCGGTGAGA TCGCCAAGAT CTTACCGCCT

AGATTGCCAT CATTCGTGCA CAATTCCTGC AGAAAATCAC CGATGCATAT

GCGGAAAACG CGGGTATTGC CAACCCNCTG TGACGCCATC CGTCAAGAAG

ACTGCCGAGC AAATACCAGG AGGCGCTGCG TGTCGTTGTG GCCTACGCTG

TGCATAACAT TATTCCTGTA CCGACCGTCA CCGCAGCAGT AGCGTACTAC

GACAGCTACC NAGCTGCCGT TCTGCCTGCN AACCTGATTN AGGCTCAGCC

TGACTATTTT GAT
primer P2

Fig. 5.6. Nucleotide sequence of 6WP. The regions homologous to primers used for PCR (section 5.3.1), P1 and P2, are underlined.

6WP was also used to probe cereal genomic DNAs (wheat, barley and rye) digested with various restriction endonucleases. Standard hybridization conditions were applied. Unfortunately, no hybridization bands were detected with 6WP as a probe in the Southern analysis (data not shown).

A serious question arouses here: why did the 6WP clone not hybridize to cereal DNAs? Completion of the sequence of 6WP was then put on hold pending the clarification of the origin of 6WP.

5.3.3. Sequence comparison to database

The sequence of 6WP was compared to other sequences in the EMBL data base by FASTA program (Pearson and Lipman, 1988). The nucleotide sequences of 6WP was converted to amino acid sequences (Fig. 5.7a). The best nucleotide sequence match with 6WP was with a 6-PGD gene in *E. coli*, *gnd* (Nasoff *et al.*, 1984). 84.2% identity was found between 6WP and *gnd* over a region of 114 amino acids (Fig. 5.7b).

5.4. Discussion

From the Southern hybridization data with the 6-PGD probes, it is apparent that the cross species differences are too great for the hybridization to take place even at low stringency. Nevertheless, using conserved regions of the *E. coli* 6-PGD gene to construct PCR primers, it was possible to amplify a band from wheat in the expected size range. It was, however, surprizing that barley and rye DNAs failed to produce a PCR product.

A data base search showed that the sequence of the 6WP clone had a very high homology with published 6-PGD genes, especially with *E. coli* (over

Fig. 5.7.

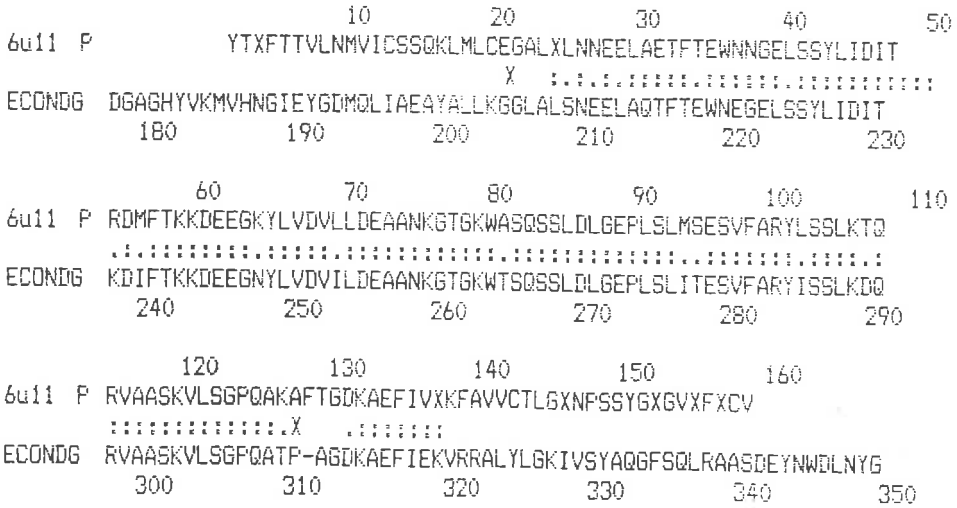
A. Amino acid sequence of 6WP deduced from the nucleotide sequence in Fig. 5.6.

B. Sequence alignment of 6WP (sequenced clone 6u11) to 6-PGD gene in *E. coli*, *gnd* (ECONDG, Nasoff *et al.*, 1984). Identities are indicated by : and conservative substitutions are indicated by ·.

A

YTXFTTVLNMVICSSQKLMLCEGALXLNNEELAETFTEWNNGELSSYLIDIT
RDMFTKKDEEGKYLVDVLLDEAANKGTGKWASQSSLDLGEPLSLMSESVF
ARYLSSLKTQRVAASKVLSGPQAKAFTGDKAEFIVXKFAVVCTLGXNPSSY
GXGVXFXCV

B



80%). On the other hand, no cross-hybridization signal was obtained from Southern analysis with cereal DNAs

It appears that the 6WP clones could be derived from the contamination of *E. coli* in the lab stock of wheat DNA. The slight difference between the sequences of 6WP clone and the *E. coli* 6-PGD gene could be attributed to the high polymorphism among *E. coli* 6-PGD genes from different strains (Prof. P.Reeves, personal communication). Another reason could be the quality of sequence data, as only one strand of 6WP clone was sequenced.

Since the advent of the PCR, the contamination of the PCR always poses a potentially serious problem for researchers. The carry-over of previously amplified sequences into new reactions is considered the most serious. There have been various methods devised to prevent carry-over (section 1.2.6). However, the presence of contamination in the template has not received enough attention.

The approach adopted in this study, using an evolutionarily conserved portion of a gene for primer construction, is particularly vulnerable to exogenous DNA contamination. Once the contamination is present in the template solution, it is impossible to destroy or separate off. It is also difficult to analyse the data for contaminating material as the target products are not clearly defined. The prevention of this type of contamination is not achievable with the presently available methods. General lab cleanliness may be the best and only remedy.

The primers used in this study were derived from a conserved protein sequence. The DNA sequence from *E. coli* also contributed to the construction of the 6-PGD primers. Hence, the primer sequence is biased toward the homologous region in the *E. coli* gene. There could be some substantial

difference with the DNA sequences between the prokaryote and the eukaryote. Unfortunately, there is still no nucleotide sequence of 6-PGD genes available from any eukaryote although partial protein sequence data from rat is available (Dr. M. Scott, personal communication).

If the sequence difference is small in respect of the size and the 3' end of the primer, the annealing temperature could be adjusted lower to tolerate the difference in the PCR. However, another danger lies in the possibility that the primers would anneal to numerous partially homologous regions and thus prevent successful PCR amplification of the specific sequence.

Various degrees of success in gene studies have been achieved with the cross hybridization of a probe from one genome to another or, "zoo blot". However, as shown in this study, the greatest obstacle is the problem of cross-species differences. Southern hybridization generally requires a high homology between the probe and the target sequences. The adaptation of PCR to clone unknown sequences has greatly facilitated the investigation of new genes. As more people turn to PCR-based amplification techniques for studying unknown sequences, greater attention to exogenous DNA contamination and careful analysis of the PCR results are needed to avoid carry-over and other types of contamination.

Chapter 6

Generation of DNA markers for specific regions on rye chromosomes

6.1. Repeat sequences in rye

Rye has a high proportion of of repetitive sequence DNA; over 90% of the rye genome (Appels *et al.*, 1986), compared to 83% of the wheat genome (Flavell and Smith, 1976; Smith and Flavell, 1977). The repetitive DNA sequences of the rye genome are usually found interspersed with each other (Rimpau *et al.*, 1978). Single or low copy sequences of the rye genome occur mainly interspersed with repetitive sequences.

Several repetitive sequence families have been described in rye: the 120, 610, and 630 bp families (Bedbrook *et al.*, 1980), the 350-480 bp family (Appels and McIntyre, 1985) and the 5.3H3 family (Appels *et al.*, 1986). Two of these families are rye specific: the 350-480 bp family and the 5.3H5 family (Appels *et al.*, 1986). The 350-480 bp family is the most abundant repeated sequences of the rye genome (about 10^6 copies: equivalent to 6.1% of the genome) and is located mainly near the telomere of each rye chromosome. The 5.3H5 family is dispersed throughout nonheterochromatic regions (Appels *et al.*, 1986). *S. cereale* and *S. montanum* both have the same four families of repeated sequences, but there is a conspicuous difference in relative proportion of 610 and 630 bp families. The 350-480 bp family is present in three species (*S. cereale*, *S. montanum* and *S. vavilouii),* but is not equally abundant in these species.

The 120 family is equally abundant in all *Secale* species analysed and is the only family in *S. silvestre* (Bedbrook *et al.*, 1980).

In this chapter, a new rye specific probe, pAW173, was identified from the pAW clones originally isolated by Dr. P. Langridge. This clone was found later to belong to a repeat DNA family which we have called R173 (Rogowsky *et al.*, 1991a). An attempt was made to isolate chromosomal region-specific RFLP probes from the flanking regions of the R173 family through the use of an inverse-PCR based method.

6.1.1. Characterization of the pAW clones

The pAW clones were kindly provided by Dr. P. Langridge. This group of clones was isolated from *Secale cereale* cv. 'Imperial' and the detailed procedure of cloning was described in Guidet *et al.* (1991). One clone of this group, pAW 161, was found to hybridize to rye but not wheat (Speck, 1986). The evidence presented here demonstrates that pAW 161 does not hybridize to barley (Fig. 6.1). Another clone from this group, pAW 174, was also found to hybridize to rye but not to wheat (Fig. 6.2). The hybridization profiles of the pAW 174 to cereal genomic DNAs, e.g. ladders in the Southerns, (lane Rye in Fig. 6.2), is similar to that of pAW 161. This finding is in agreement with previous data which suggested that this clone cross hybridized to pAW 161 (Speck, 1986).

6.1.2. Species specificity of pAW 173

pAW 173 did not hybridize to any of the other clones in the same group of clones (Speck, 1986). In this study, pAW 173 was hybridized to Eco RV, HindIII and HinfI digested DNA from wheat (cv. Chinese Spring), barley (cv. Betzes) and rye (cv. Imperial) (Fig. 6.3). Strong signals were obtained for rye (lanes 2, 5 and 8), whereas no signal could be detected in lanes 1, 3, 4, 6, 7 and

Fig. 6.1. Hybridization pattern of cereal DNA probed with pAW 161. Cereal DNAs were digested with various restriction endonucleases, fractionated in a 1% agarose gel (section 2.2.7), and probed with the insert of pAW161 (section 2.2.11). Lanes 1, 4 and 7: wheat (Chinese Spring) DNA; lanes 2, 5 and 8: rye (Imperial); lanes 3, 6 and 9: barley (Betzes). DNAs in lanes 1, 2 and 3 were digested with Eco RV; DNAs in lanes 4, 5 and 6 were digested with Hind III; DNAs in lanes 7, 8 and 9 were digested with Hinf I. The position of marker bands is given on left with the sizes of the bands in base pairs.

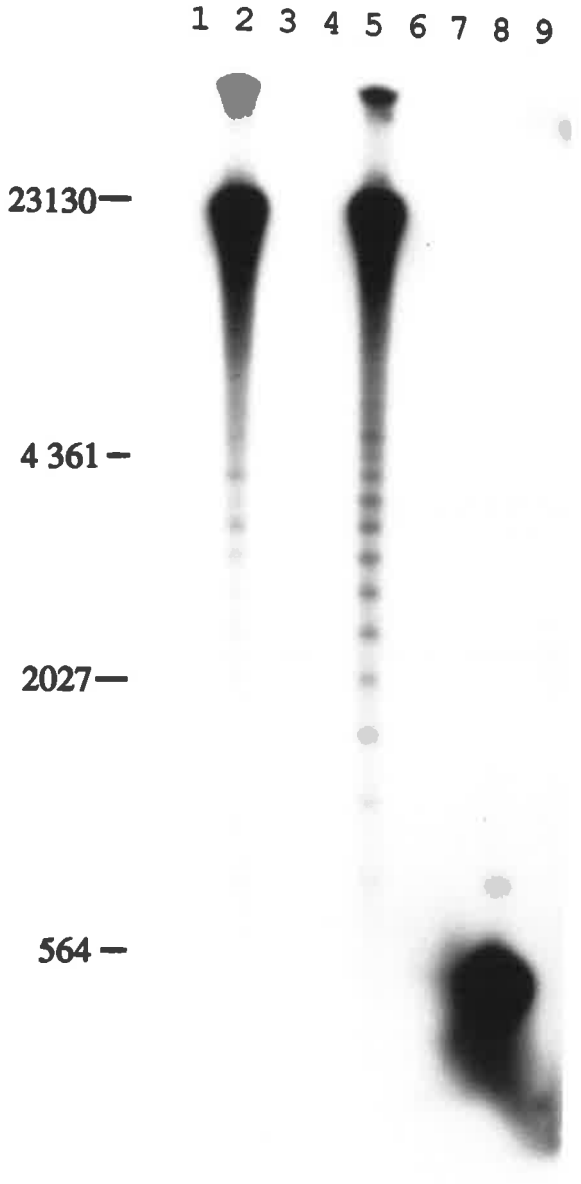

1 2 3 4 5 6 7 8 9
23130
4 361
2027
564

Fig. 6.2. Hybridization pattern of wheat, rye and wheat-rye addition lines DNAs probed with pAW 174.

Cereal DNAs were digested with Eco RI, fractionated in a 1% agarose gel (section 2.2.7), and probed with the insert of pAW 174 (section 2.2.11). The position of marker bands is given on left with the sizes of the bands in base pairs.

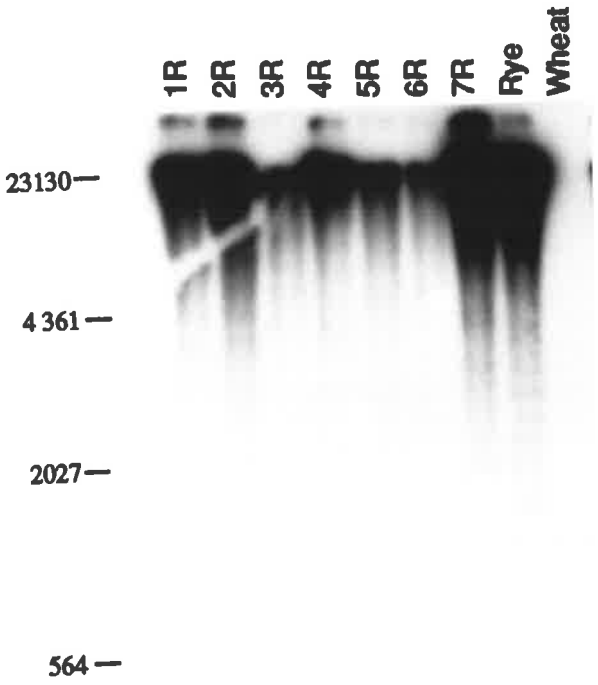
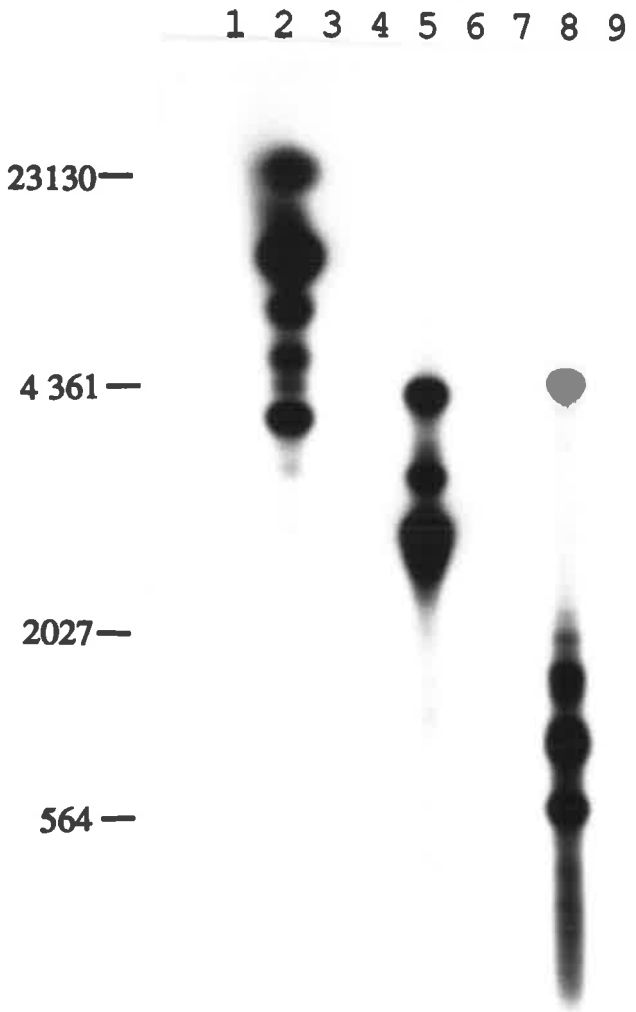

1R
2R
3R
4R
5R
6R
7R
Rye
Wheat
23130
4 361
2027
564

Fig. 6.3. Hybridization pattern of different cereals probed with pAW 173.

Cereal DNAs were digested with various restriction endonucleases, fractionated in a 1% agarose gel (section 2.2.7), and probed with the insert of pAW173 (section 2.2.11). Lanes 1, 4 and 7: wheat (Chinese Spring) DNA; lanes 2, 5 and 8: rye (Imperial); lanes 3, 6 and 9: barley (Betzes). DNAs in lanes 1, 2 and 3 were digested with Eco RV; DNAs in lanes 4, 5 and 6 were digested with Hind III; DNAs in lanes 7, 8 and 9 were digested with Hinf I. The position of marker bands is given on left with the sizes of the bands in base pairs.


1 2 3 4 5 6 7 8 9
23130
4 361
2027
564

9, corresponding to wheat or barley DNAs. Therefore, pAW 173 appears to be a rye specific probe and this sequence is absent from the wheat and barley genomes within the detection range of the present method.

6.1.3. Distribution of pAW 173 sequences in rye genome

To analyse the genomic organization of pAW 173 repetitive sequence in rye, hybridization of pAW 173 clone was carried out to 'Imperial' rye genomic DNA digested with 4-base and 6-base cutters (Fig. 6.4). Several bands or a smear could be seen for each of the 6-base cutters, e.g. HindIII revealed three major Southern bands measuring 1.0, 1.4 and 1.7 kb. The 4-base cutters generally revealed more bands than the 6-base cutters. These results imply that the full length of the repeating unit of pAW 173 is longer than 1-2 kb as sites for 6-base cutters will randomly occur only once every several thousand base pairs. The absence of 'ladders' in the genomic blots indicates that pAW 173 is probably a non-tandem repeat sequence (Sonina *et al.*, 1989).

The rye addition lines were used to shed light on the chromosomal distribution of pAW 173 in rye. When pAW 173 was used to probe the genomic DNAs of rye addition lines, there was no difference between the banding profiles of individual rye chromosomes (Fig. 6.5). It seems that the organization of the pAW 173 repeats is very much conserved on each of the rye chromosomes as no polymorphic Southern band could be detected, although there might exist some difference in copy numbers between each rye chromosomes.

6.2. Cloning and characterization of the DNA sequences flanking R173 family by 'inverse-PCR'

Rogowsky *et al.* (1991a) generated a genomic library from a wheat line with three copies of the short arm of chromosome 1 of rye, 1RS. Seventy-seven λ

Fig. 6.4. Hybridization pattern of rye DNA probed with pAW 173. Rye (Imperial) DNA was digested with various hexanucleotide or tetranucleotide restriction endonucleases, fractionated in a 1% agarose gel (section 2.2.7), and probed with the insert of pAW173 (section 2.2.11). Lane 1: Bam HI, lane 2: Bgl II; lanes 3: Eco RI; lanes 4: Hind III; lane 5: Alu I; lane 6: Hae III; lane 7: Sau 3A and lane 8 Taq I. The position of marker bands is given on left with the sizes of the bands in base pairs.

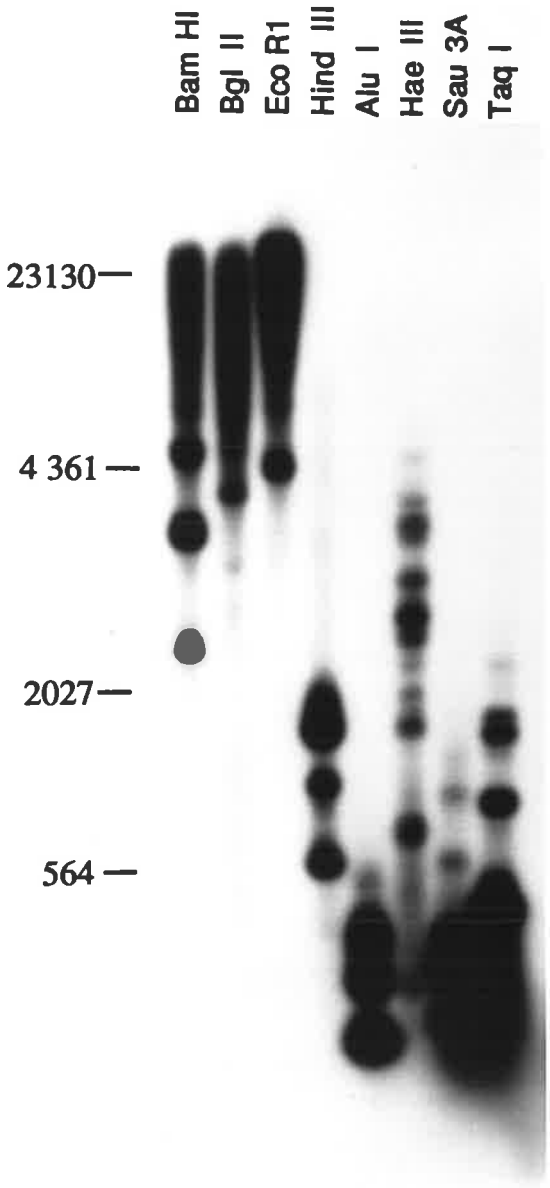
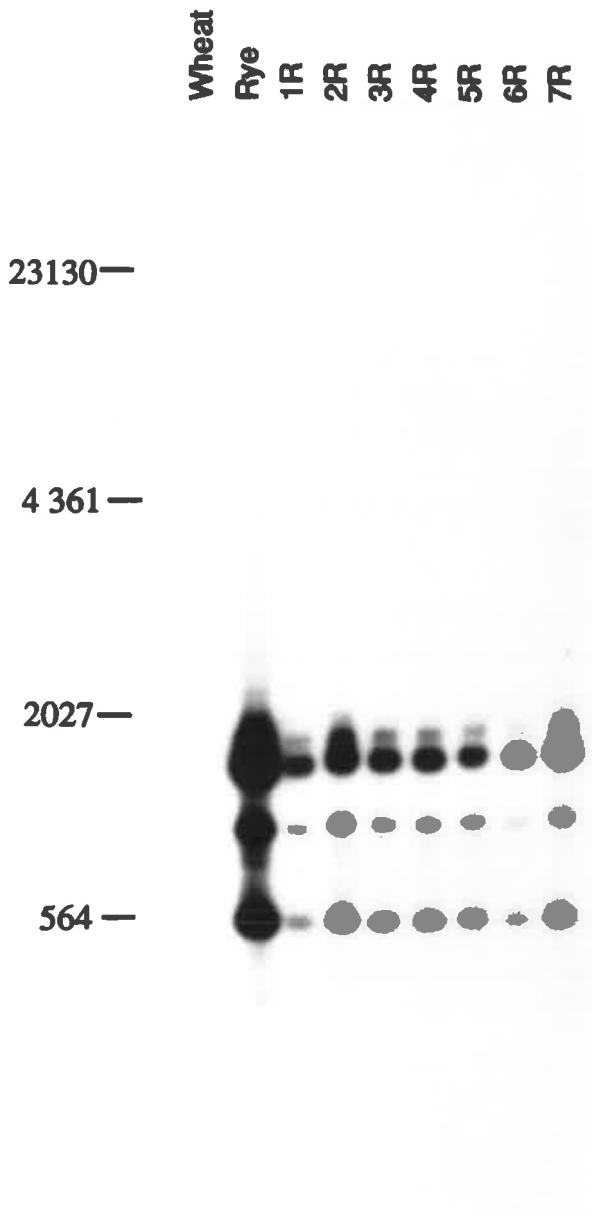

Bam HI
Bgl II
Eco R1
Hind III
Alu I
Hae III
Sau 3A
Taq I
23130
4 361
2027
564

Fig. 6.5. Hybridization pattern of wheat-rye addition lines probed with pAW 173.

Wheat-rye addition line DNA was digested with Hind III, fractionated in a 1% agarose gel (section 2.2.7), and probed with the insert of pAW173 (section 2.2.11). Wheat cv. Chinese Spring and rye cv. Imperial are included for references. The position of marker bands is given on left with the sizes of the bands in base pairs.


Wheat
Rye
1R
2R
3R
4R
5R
6R
7R
23130
4 361
2027
564

clones, representing independent members of the family, were isolated by hybridization to pAW173. They share a common region of approximately 3.5kb, which is free of large internal repeats and therefore constitutes the basic unit of the dispersed R173 family. Three of the λ clones were sequenced and the sequence analysis demonstrated a high degree of homology in conserved domains.

Repetitive sequences have been used successfully in monitoring the introgression of alien chromatin into wheat (Koebner *et al.*, 1986; Xin and Appels, 1988; Rogowsky *et al.*, 1991b). Furthermore, Appels *et al.* (1986) proposed that RFLP probes could be isolated from the flanking region of dispersed repeats present as alien material in wheat background. Rogowsky *et al.* (1991a) adopted this approach and obtained a RFLP probe, located to the short arm of rye chromosome 1, by subcloning the flanking regions of the R173 family. However, this approach yielded only two useful RFLP probes from a total of seventy-seven λ clones. Considering the time and effort involved in this type of analysis, alternative methods were investigated.

To obtain molecular information on the flanking regions of a gene or, in the case R173 family, one can clone and sequence a number of large clones which contain the segment of interest. However, large-scale cloning and sequencing are very costly and time consuming. PCR provides an alternative approach to many fundamental procedures in molecular biology. One strategy for the characterization of fragments adjacent to known sequences is "inverse-PCR" which involves PCR amplification following ligation-mediated circularization (Triglis *et al.*, 1988; Ochman *et al.*, 1988; Silver and Keerikatte, 1988).

As shown earlier, a wheat-rye recombinant line, DR-A1, contains about 100 copies of the R173 family (Rogowsky *et al.*, 1991b). The plan of this experiment is to amplify the flanking regions of the R173 family in DR-A1 by

inverse-PCR with primers V1 and V2 (Fig. 6.6) and the PCR products are to be cloned into plasmid. Region-specific (DR-A1, Fig. 6.6a) RFLP probes would be obtained by screening the inverse-PCR clones through Southern analysis. The first step is to use the clones from rye as a test system. If successful, the hybridization analysis would be extended to the clones derived DR-A1. As the recombinant line DR-A1 contains only a small rye segment, the probability of finding a specific probe homologous to this region is very low. Therefore, clones from TTL, which contains the whole short arm of chromosome 1R, would also be analysed and screened for RFLP probe.

6.2.1. Procedure of cloning the flanking region of the R173 family.through inverse-PCR

The sequence data of three R173 λ clones was kindly provided by Dr. P. Rogowsky. A schematic representation of the three clones indicating the location and extent of the R173 element is shown in Fig. 6.6. Based on the available sequences, a strategy was devised for cloning the sequences flanking the R173 family. This is shown schematically in Fig. 6.7. The region complementary to the V1 primer is located on the end of the consensus R173 sequences and the 5'-3' direction of V1 is towards the outside of R173. The region complementary to the V2 is on the 5' side of a Hind III site 270 bp away from the end and the orientation of V2 (5' to 3') is toward the centre of R173 (Fig. 6.7). Pst I sites were incorporated into both V1 and V2 primers (Table 6.1). Genomic DNAs were made from DR-A1, rye and TTL, a wheat line with three copies of the short arm of chromosome 1 of rye. Hind III digestion was then carried out to completion with these DNAs (section 2.2.15). The DNA fragments were diluted and ligated under conditions that favour the formation of monomeric circles (Collins and Weissman, 1984). The resulting circles were then used as

Fig. 6.6. Nucleotide sequence homology of three λ clones containing the R173 element (adapted from Rogowsky *et al.* 1992, unpublished). Dotted areas indicate the conserved regions. The dotted line indicates the Hind III sites. The vertical open bar indicates the regions homologous to pAW173. The relative location and orientation the V1 and V2 primers are shown in λ2.

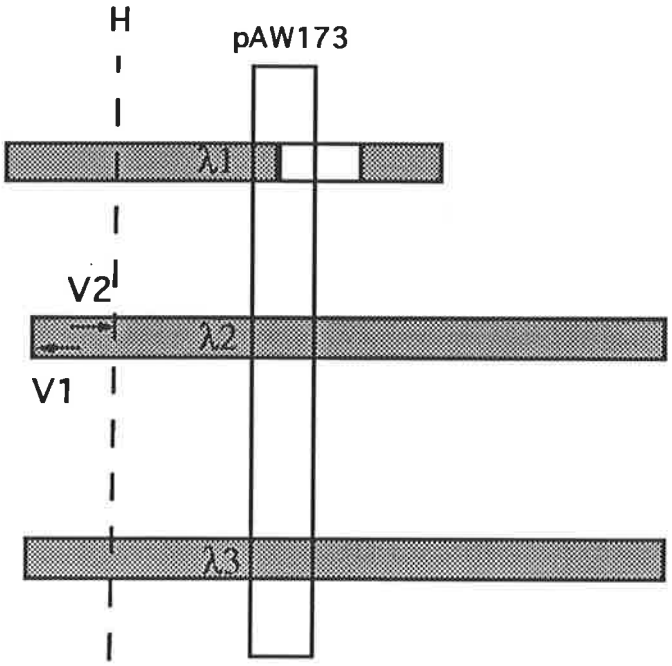
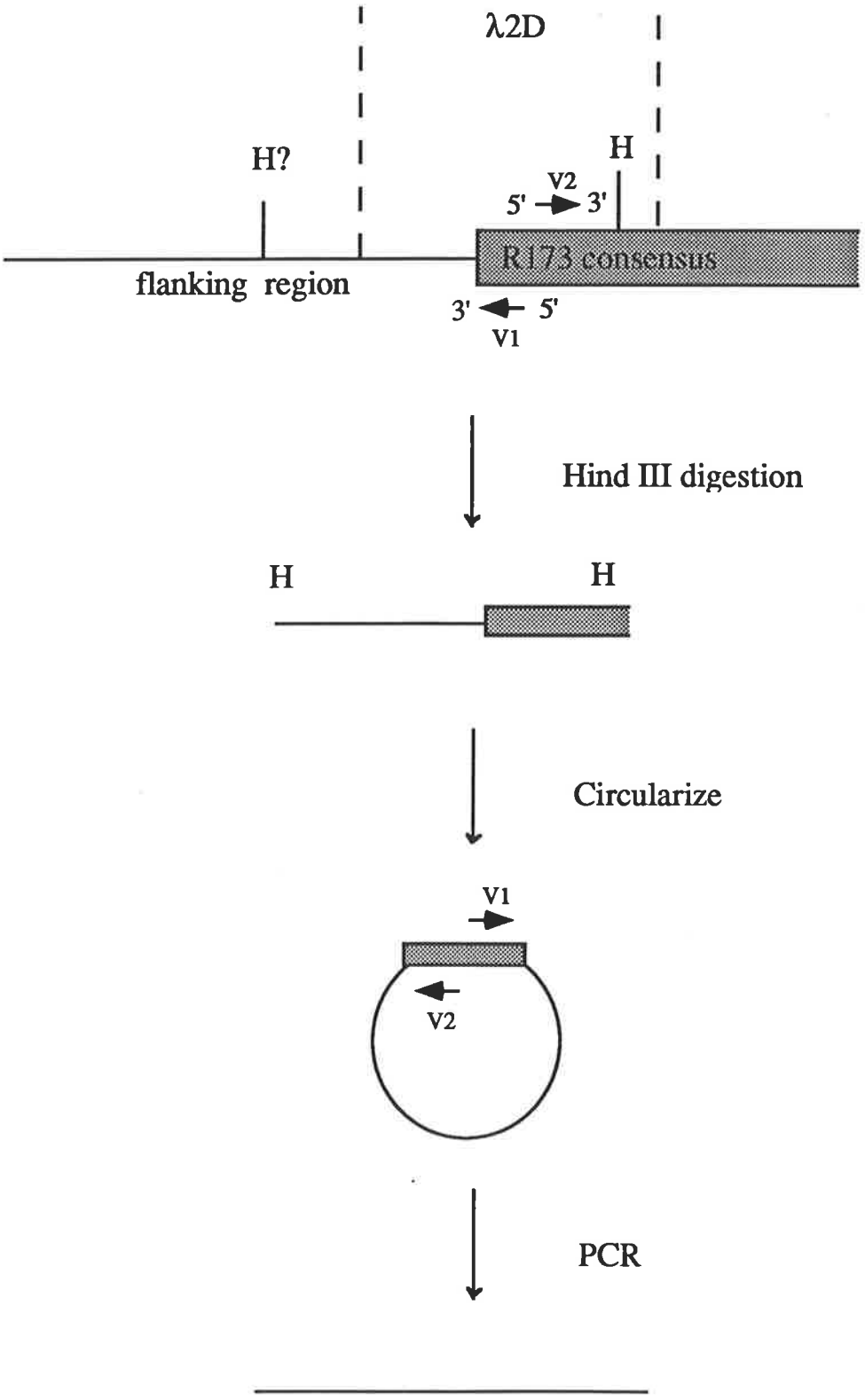

H
pAW173
λ1
V2
λ2
V1
λ3

Fig. 6.7. Schematic representation of the inverse-PCR approach. λ2D and the dotted lines indicate the location of this region. 'H' indicates the Hind III sites. Dotted areas indicate the consensus regions of the λ clones (Rogowsky *et al.* 1992, unpublished). The relative location and orientation the V1 and V2 primers are also shown here. The details of the inverse-PCR are described in Section 2.2.15.



Amplified products from the flanking region

templates for amplification by PCR using the V1 and V2 primers. The primary products of the amplification should be linear double-stranded molecules containing segments flanking one side of the R173 elements. The amplified products were then digested with Pst I and ligated into pUC 19 cleaved with Pst I.

Table 6.1 The sequences of V1 and V2 primers

Primer	Sequence	features
V1	5'AACACAC<u>CTGCAG</u>GTAGT3' Pst I	toward end of R173
V2	5'AGATT<u>CTGCAG</u>TGTCAAGC3' Pst I	toward centre of R173

6.2.2. Characterization of the inverse-PCR clones

6.2.2.1. Dot-blotting

Mini-prep DNA was made from the individual recombinant plasmids. As shown in previous section, they are derived from DR-A1, rye and TTL respectively. Ten rye-derived clones, 20 from DR-A1 and 58 from TTL were picked out randomly. A total of 88 inverse-PCR clones were screened via dot-blotting (Fig. 6.8). The location of the clones and controls on the dot blot membrane are displayed in Fig. 6.8a. The numbers in the each box are the clone numbers. Clones number 1 to 20 were from DR-A1, 21 to 78 from TTL and 79 to 88 from rye. The probes used for dot blotting were λ2D DNA, total genomic

Fig. 6.8.

A. The position of the clones and the reference DNAs in dot-blotting (6.2.2.1).

B. Hybridization pattern of the above DNAs probed with genomic DNA from wheat cv. Chinese Spring (dot-blotting was carried out as described in Section 2.2.9).

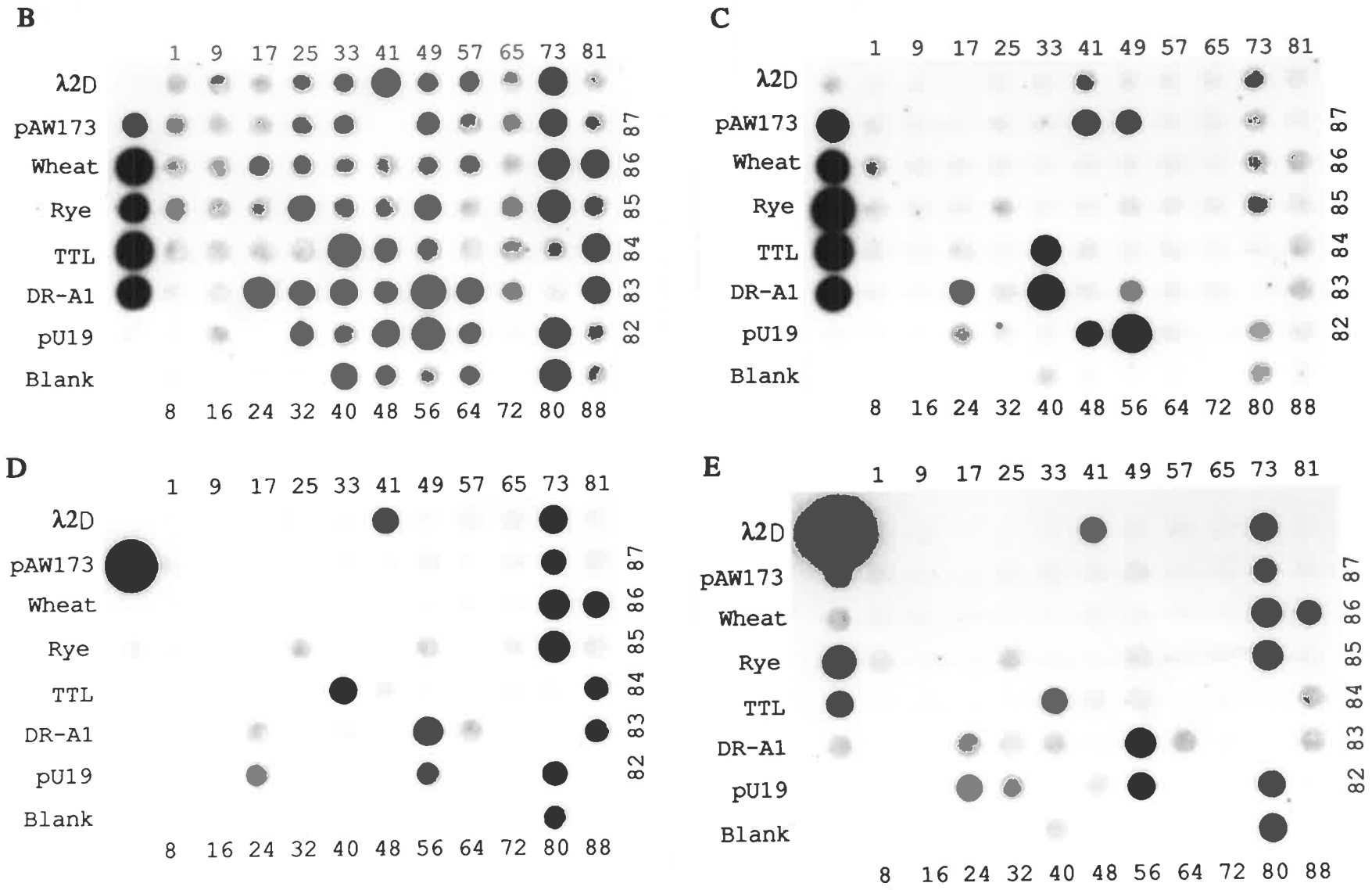
C. Hybridization pattern of the above DNAs probed with genomic DNA from rye cv. Imperial.

D. Hybridization pattern of the above DNAs probed with pAW173.

E. Hybridization pattern of the above DNAs probed with λ2D.

λ2D	1	9	17	25	33	41	49	57	65	73	81
pAW173	2	10	18	26	34	42	50	58	66	74	82
Wheat	3	11	19	27	35	43	51	59	67	75	83
Rye	4	12	20	28	36	44	52	60	68	76	84
TTL	5	13	21	29	37	45	53	61	69	77	85
DR-A1	6	14	22	30	38	46	54	62	70	78	86
pU19	7	15	23	31	39	47	55	63	71	79	87
Blank	8	16	24	32	40	48	56	64	72	80	88

A


B
1 9 17 25 33 41 49 57 65 73 81
λ2D
pAW173
Wheat
Rye
TTL
DR-A1
pU19
Blank
82 83 84 85 86 87
8 16 24 32 40 48 56 64 72 80 88
C
1 9 17 25 33 41 49 57 65 73 81
λ2D
pAW173
Wheat
Rye
TTL
DR-A1
pU19
Blank
82 83 84 85 86 87
8 16 24 32 40 48 56 64 72 80 88
D
1 9 17 25 33 41 49 57 65 73 81
λ2D
pAW173
Wheat
Rye
TTL
DR-A1
pU19
Blank
82 83 84 85 86 87
8 16 24 32 40 48 56 64 72 80 88
E
1 9 17 25 33 41 49 57 65 73 81
λ2D
pAW173
Wheat
Rye
TTL
DR-A1
pU19
Blank
82 83 84 85 86 87
8 16 24 32 40 48 56 64 72 80 88

DNAs from wheat (cv. Chinese Spring), rye (cv. Imperial) or DR-A1 and pAW 173. λ2D is the segment which contain the flanking region of R173 consensus sequence (Fig. 6.7). λ2D DNA, the above cereal genomic DNAs, and pUC19 were also used as control (Fig. 6.8a). The purpose of using λ2D DNA (the location of this probe on R173 element is shown in Fig. 6.7) is to find out if any clones are homologous to this region. The genomic DNA probes serve to identify the clones with repetitive or organellar DNA sequences and also to identify species-specific clones. The results of dot blotting are summerized in Table 6.2. According to the hybridizing signals to cereal genomic DNAs, 21 clones (26%) were likely to contain repetitive or organellar DNA sequences (e.g. clone 55, Table 6.2). The results indicate that some of the clones might have homology to the λ2D clone (e.g. clone 76 in Fig. 6.8e). As λ2D contains repetitive sequences other than pAW 173 (Dr. P. Rogowsky, personal communication), it is not certain whether these clones were derived from the flanking region of R173 element or from other part of the genome. Not surprisingly, all of the probes displayed strong homology to their own sequences (e.g. λ2D in Fig 6.8e).

When rye (cv. Imperial) was used as a probe, the amount of rye chromatin in the samples could be quantified in the order: rye >TTL>DR-A1>wheat (Fig.6.8c). There was no significant difference between the signals with pAW173 or wheat. The reason could be that the wheat genome shares considerable homology with the rye genome although pAW173 is a rye specific probe. The cereal genomic DNA probed with λ2D displayed similar intensity as seen when rye DNA was used as probe(Fig.6.8e) indicating that λ2D is essentially a rye specific probe. As expected, there is no homology between pAW173 and λ2D detected by dot blotting (Fig.6.8d and Fig.6.8e).

6.2.2.2. Southern-blotting

In the next step, the inverse-PCR clones were used as probes on Southern blots of genomic DNAs from wheat (cv. Chinese Spring), barley (cv. Betzses), rye (cv. Imperial), wheat /barley addition lines and/or wheat/rye addition lines. Rye-derived inverse PCR clones were screened first by this method. One of the clones (number 88) identified a wheat/rye polymorphism that could be mapped to 2R and 7R (Fig. 6.9). The difference in the intensity of signals between rye additional lines may be caused by uneven loading during the Southerns. Although only ten clones were tested, it appeared that this approach is feasible for obtaining useful RFLP probes. However, the efficiency seems to be too low for routine analysis.

Several other clones were also tested by Southern analysis. The results of Southern blots are summarised in Table 6.2. Most clones, which seemed to contain single or low copy sequences in dot-blotting, were found to contain repetitive sequences common to both wheat and rye in Southern blots (e.g. clone 63, Fig. 6.10). Six clones appeared to be rye specific and displayed a pattern similar to pAW 173 (e.g.clone 50, Fig. 6.11). Only seven of the clones (13.4%) appeared to be of single or low-copy nature among the clones tested in Southern blots (Table 6.2). Two could be mapped to 5H and 6H respectively (clone 26 and 52, Fig. 6.12a,b). Both of these two clones were derived from TTL. However, other clones of this type could not be mapped to any chromosomes (e.g. clone 9, Fig. 6.13) using barley and rye addition lines, as no polymorphisms were identified between wheat and rye or wheat and barley. Generally, four restriction enzymes were used to digest genomic DNAs for each clone.

Fig. 6.9. Mapping of clone 88 to rye chromosomes.
Genomic DNA was digested with Hind III, resolved on a 1% agarose gel, Southern blotted and probed with clone 88. The chromosomal locations of rye DNA bands are indicated by arrows.

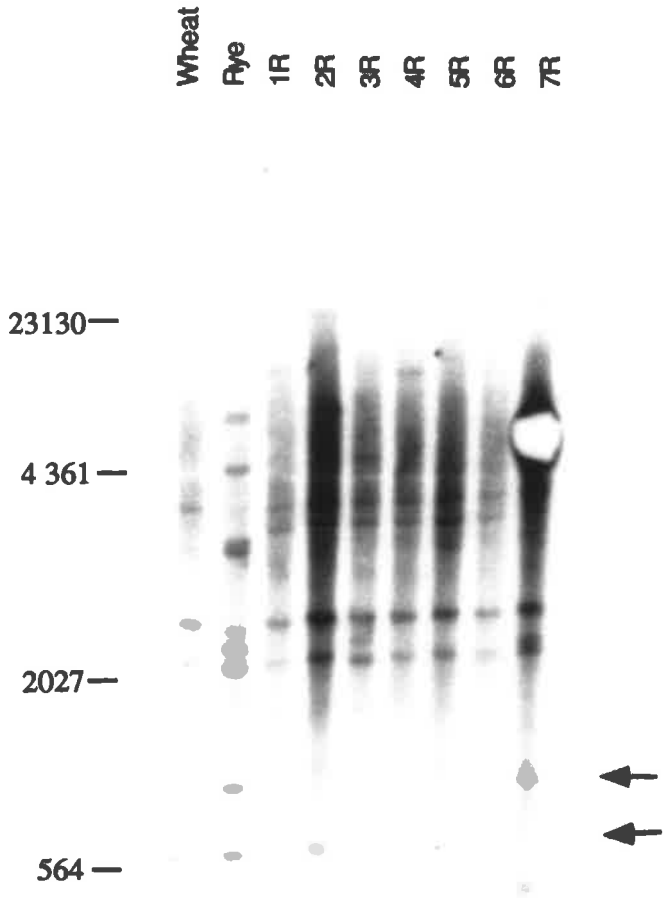

Wheat
Rye
1R
2R
3R
4R
5R
6R
7R
23130
4 361
2027
564

Fig. 6.10. Hybridization pattern of wheat, rye and wheat-rye addition lines DNAs probed with clone 63.

Genomic DNA was digested with Hind III, resolved on a 1% agarose gel, Southern blotted and probed with clone 63.

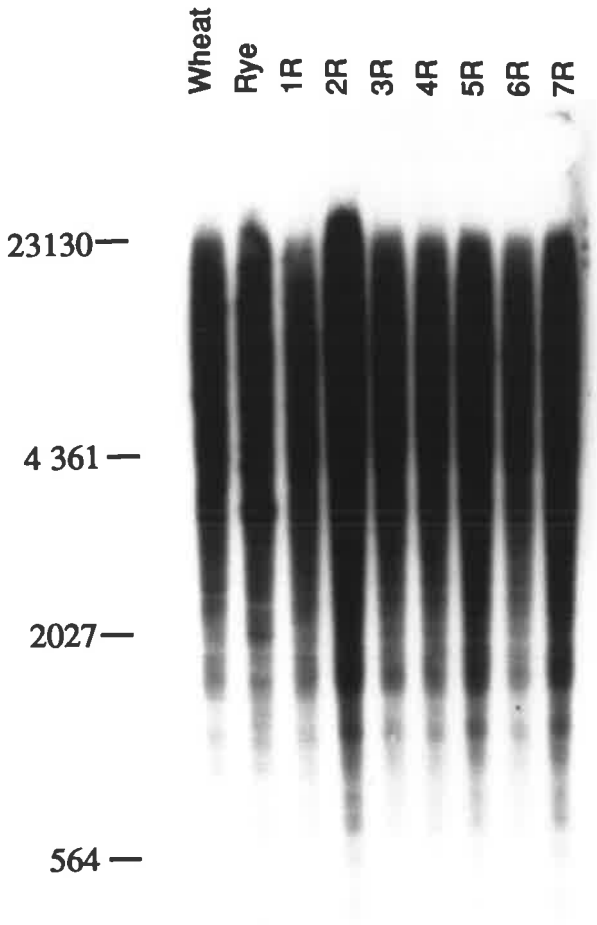

Wheat
Rye
1R
2R
3R
4R
5R
6R
7R
23130
4 361
2027
564

Fig. 6.11. Hybridization pattern of wheat, rye and wheat-rye addition lines DNAs probed with clone 50.

Genomic DNA was digested with Hind III, resolved on a 1% agarose gel, Southern blotted and probed with clone 50.

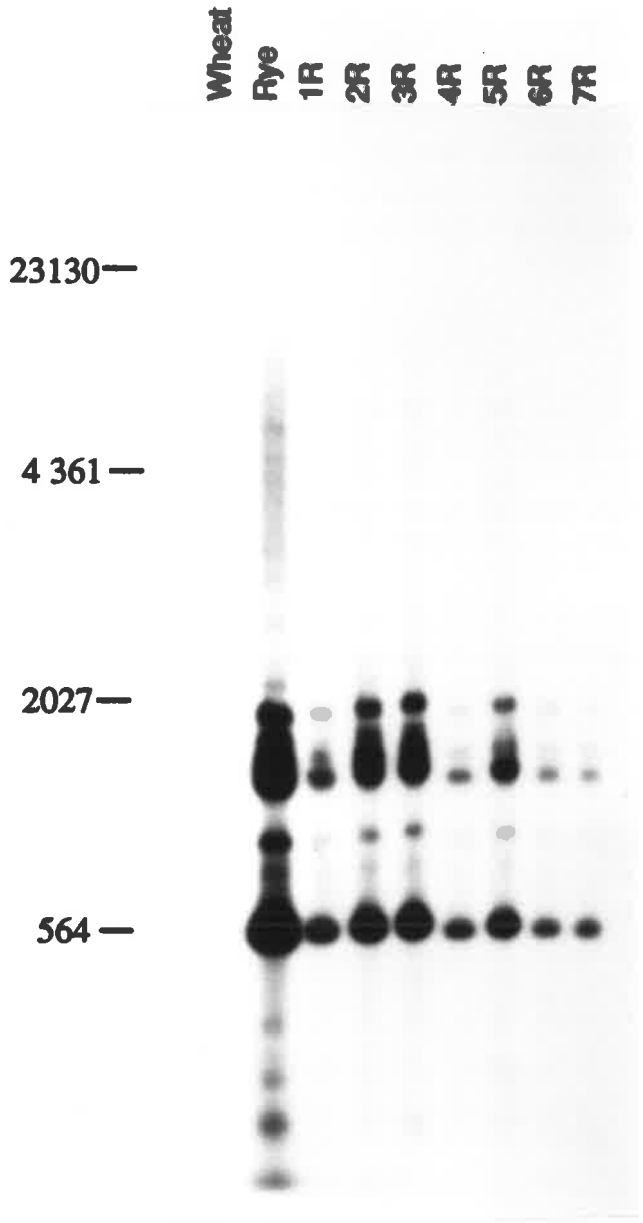

Wheat
Rye
1R
2R
3R
4R
5R
6R
7R
23130
4 361
2027
564

Fig. 6.12. Mapping of clones 26 and 52 to barley chromosomes.

A. Genomic DNA was digested with Hind III, resolved on a 1% agarose gel, Southern blotted and probed with clone 26.

B. Genomic DNA was digested with Hind III, resolved on a 1% agarose gel, Southern blotted and probed with clone 52.

The chromosomal locations of barley DNA bands are indicated by arrows.

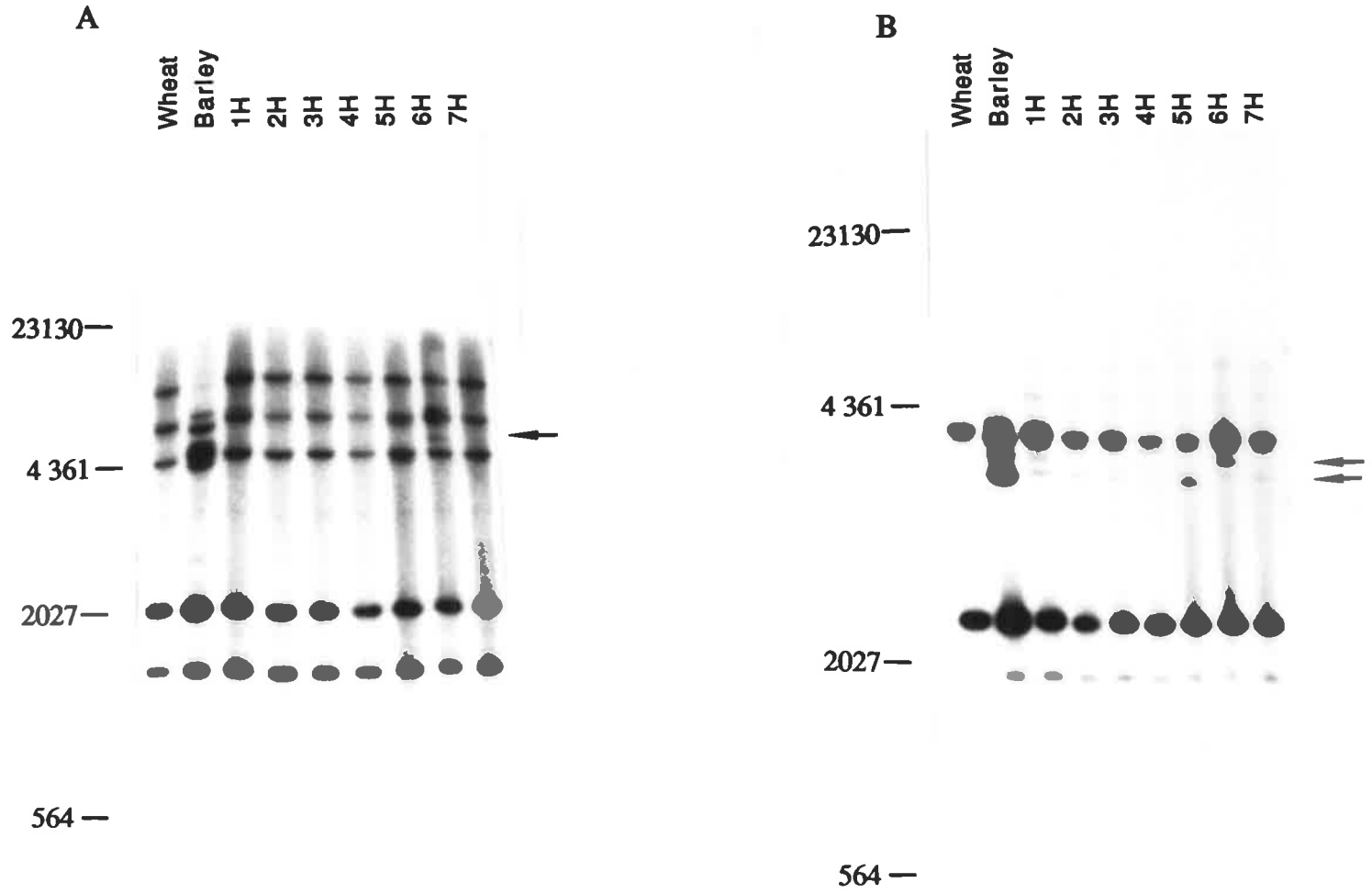

A
Wheat
Barley
1H
2H
3H
4H
5H
6H
7H
23130
4 361
2027
564
B
Wheat
Barley
1H
2H
3H
4H
5H
6H
7H
23130
4 361
2027
564

Fig. 6.13. Hybridization pattern of wheat, rye and wheat-rye addition lines DNAs probed with clone 9.

Genomic DNA was digested with Hind III, resolved on a 1% agarose gel, Southern blotted and probed with clone 9.

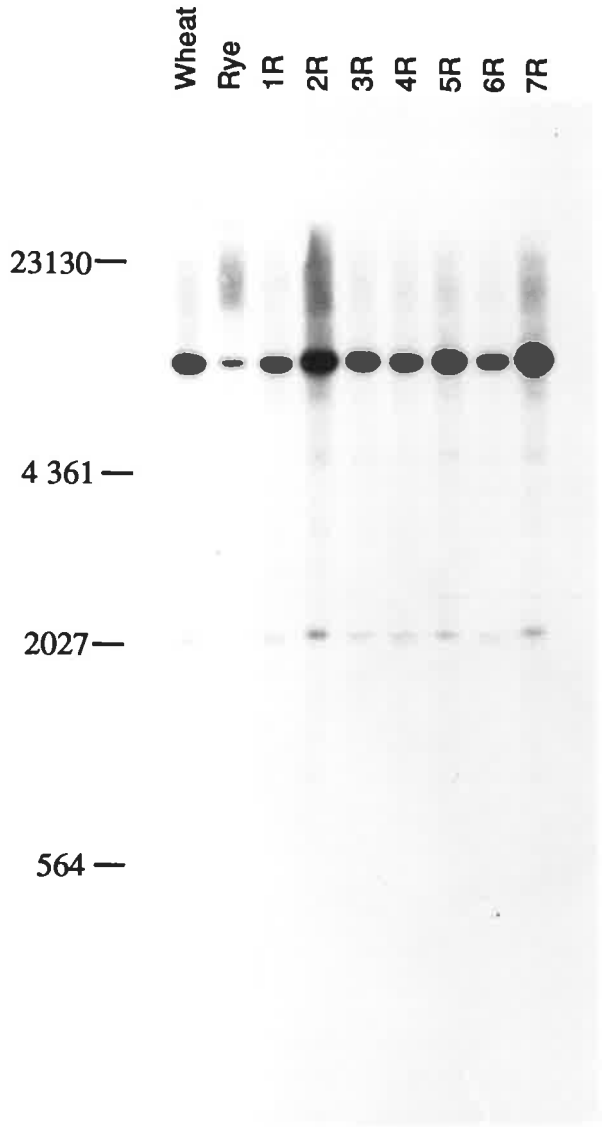

Wheat
Rye
1R
2R
3R
4R
5R
6R
7R
23130
4 361
2027
564

Table 6.2 Summary of hybridization data

Clone No	Probes used for hybridization				Copy No	Homology	173 pattern	Mapping
	CS	Imp	173	D-2				
	Dot blot data				Southern blot data			
1	-	-	-	-	Low	W/R		
2	-	-	-	-	Repeat	R	yes	
3	-	-	-	-	Repeat	W/R		
4	-	-	*	*	Repeat	R		
5	-	-	-	-	Low	W/R		
6	-	-	-	-	Repeat	W/R		
7	-	-	-	-	Repeat	W		
8	-	-	-	-	Repeat	W/R		
9	-	-	-	-	Low	W/R/B		
10	-	-	-	-	Repeat	W/R		
11	-	-	-	-	Low	W/R		
12	-	-	-	-	Repeat	W		
13	-	-	-	-				
14	-	-	-	-	Repeat	W/R		
15	-	-	-	-	Repeat	W/R		
16	-	-	-	-	Repeat	W/R		
17	-	-	-	-	Repeat	W/R		
18	-	-	-	-	Repeat	W		
19	-	-	-	-	Repeat	W/B		
20	-	-	-	-	Repeat	W/B		
21	-	-	-	-	Repeat	W/R		
22	*	*	**	*	Repeat	W/R		
23	-	-	**	*				
24	-	-	?	-	Repeat	W/R		
25	-	-	-	-	Repeat	W/R		
26	-	-	-	-	Low	W/B		6H
27	-	-	-	-				
28	*	-	*	*				
29	-	-	-	-	Repeat			
30	-	-	*	-				
31	-	-	*	*				
32	-	-	-	-				
33	-	-	-	-				
34	-	-	-	-				
35	-	-	-	-				
36	-	-	-	-				
37	*	*	***	*	Repeat	W/R		
38	*	**	*	*	Repeat	R	yes	
39	-	**	-	-				
40	*	-	**	*				
41	*	-	***	*				

Clone No	Probes used for hybridization				Copy No	Homology	173 pattern	Mapping
	CS	Imp	173	D-2				
	Dot blot data					Southern blot data		
42	-	-	**	-	Repeat	W/R		
43	-	*	-	-	Repeat	W/R		
44	-	-	-	-	Repeat	W/R		
45	-	-	-	*	Repeat	W/R		
46	-	-	-	-				
47	-	*	*	-	Repeat			
48	-	-	-	-	Repeat	W/R		
49	-	-	-	-				
50	-	*	*	-	Repeat	R	yes	
51	-	*	-	-				
52	*	-	*	-	Low	W/R/B		5H,6H
53	-	-	*	-				
54	*	-	***	**				
55	*	**	**	*			yes	
56	-	-	-	-	Repeat	W/R		
57	-	-	-	-	Repeat	W/R		
58	-	-	-	-				
59	-	-	-	-				
60	-	-	-	-	Repeat	R	yes	
61	-	-	-	-	Repeat	R	yes	
62	*	-	*	*				
63	-	-	-	-	Repeat	W/R		
64	-	-	-	-				
65	-	-	-	-				
66	-	-	-	-				
67	-	-	-	-				
68	-	-	-	-				
69	-	-	-	-				
70	-	-	-	-				
71	-	-	-	-				
72	-	-	-	-				
73	*	-	**	*				
74	*	-	**	*				
75	*	-	***	*				
76	*	-	***	*				
77	-	-	-	-	Repeat	W/R		
78	-	-	-	-				
79	*	-	***	*				
80	*	-	***	*	Repeat	W/R		

6.2.2.3. PCR analysis using inverse-PCR primers

As shown earlier, R173 represents a group of dispersed repetitive sequences. The orientation of each unit and the relationship between units is not clear. The inverse-PCR primers, V1 and V2, were used directly to amplify cereal genomic DNAs. Although they were designed facing in opposite direction, many amplified products could be seen from the cereal genomic DNAs (Fig. 6.14). One rye band (1340bp) could be mapped to chromosome 1R (lane rye and 1R, Fig. 6.14) and another to 2R (lane rye and 2R, Fig. 6.14). There are also several faint bands present , but they tend to be less reproducible (e.g. lane DR-A1, Fig. 6.14).

6.3. Discussion

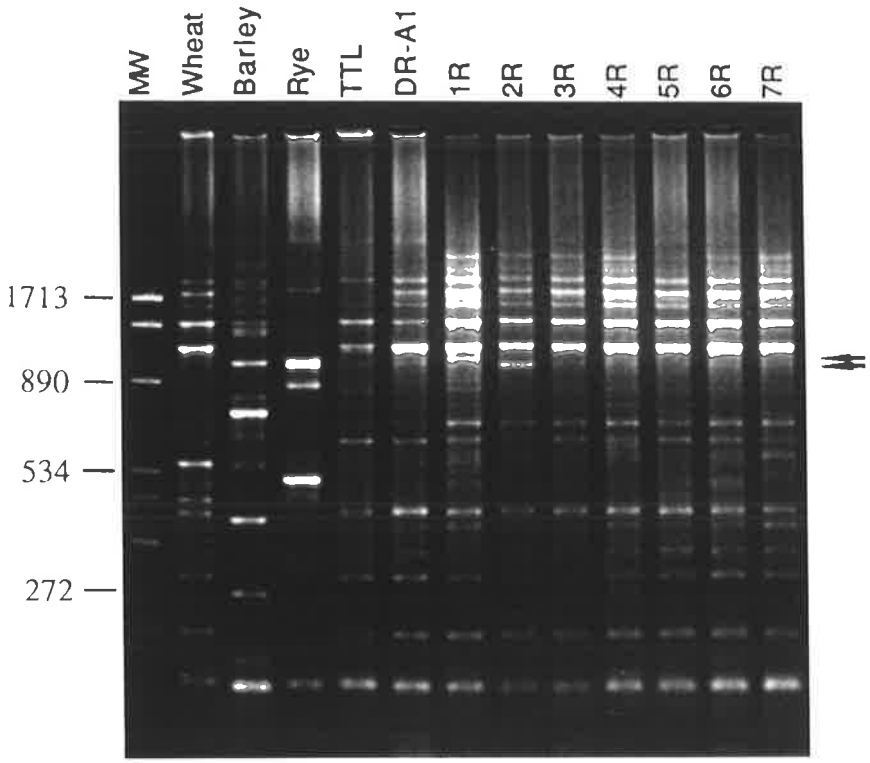
The data presented above indicate that both pAW 173 and 174 are rye specific DNA sequences. Further characterization of pAW 161 in this laboratory (Guidet *et al.*, 1991) showed that it is a member of the 350-480 bp family (Bedbrook *et al.*, 1980; Appels and McIntyre, 1985). According to Speck (1986), pAW 174 also belongs to this family.

The pAW 173 sequences have been found to be rye specific in the present study. The 173 repeats element is present on all rye chromosomes and has been shown by *in situ* hybridization to be dispersed over the entire rye genome (Guidet *et al.*, 1991).

The dispersed nature of the R173 elements poses an interesting question: what is the feature of its flanking regions? The R173 family could be flanked by conserved sequences or random ones. The sequences can be single copy or multicopy. If these regions mainly consist of single copy sequences, they could become an excellent source for RFLP probes as these markers would be

Fig. 6.14. Mapping of PCR products with cereal DNA using primers V1 and V2.

The template DNA was from wheat (Chinese Spring), barley (Betzes), rye (Imperial), the triple translocation line (TTL) and the wheat-rye addition lines 1R to 7R. Amplification was performed as described in Section 2.2.12 and 2.2.13. The PCR products were fractionated on a 3% agarose gel. DNA size markers are shown in lane MW. The sizes of the marker bands are given on the left in base pairs. The chromosomal locations of rye specific bands are indicated by arrows.


MW
Wheat
Barley
Rye
TTL
DR-A1
1R
2R
3R
4R
5R
6R
7R
1713
890
534
272

numerous and evenly distributed as they are associated with R173. On the other hand, various amount of rye has been introduced into wheat and pAW 173 has been used successfully in monitoring the introgression of rye into wheat (Rogowsky *et al.*, 1991b). The flanking region of R173 in these rye segment could contain region-specific RFLP probes. One approach was to clone and analyse the flanking regions a number of λ clones which contain the pAW173 segment (Rogowsky *et al.*, 1991b). The present study used inverse-PCR for the same purpose.

Although three RFLP probes were mapped to 2R, 7R, 5H and 6H, no 1RS or 1H specific probe was found among the clones generated by inverse-PCR. The reason could be that only 88 colonies were analysed by dot blotting and 52 screened by Southern blotting. Most of the inverse-PCR clones were found to represent repetitive sequences. This implies that the flanking regions of R173 family mainly consist of repetitive sequences, with few single or low copy sequences. The data from this study confirm the results of Southern analysis of the R173 family and the analysis of flanking regions in lamda clones (Rogowsky *et al.*, 1991a). The above observations agree well with the prediction of Rimpau *et al.* (1978) that, in the case of rye specific sequences, 84% of the flanking sequences are expected to be interspersed with repetitive sequences common to wheat and rye, whereas 16% should be single or low copy DNA. In the present study, 86.6% of the inverse-PCR clones represent repetitive sequences, whereas 13.4% single or low copy DNA. Therefore, obtaining region-specific RFLP probes through inverse-PCR via the R173 element may require the screening of a large number of clones.

On the other hand, two PCR bands was mapped on 1R and 2R, respectively, by direct PCR amplification using the inverse-PCR primers. More markers are expected to be found through this approach if more primer and/or

primer are used. These results demonstrated that primers can be constructed from interspersed repetitive sequences and used directly for mapping purposes.

It remains unclear, however, whether these bands were truly derived from the border region of R173 elements. The Southern analysis showed that the core sequences of the R173 family, pAW 173, is conserved on every rye chromosome. Nevertheless, the region near the border of the R173 element could be more polymorphic. Inversion or other changes may have occurred in these regions. The amplification event with V1 and V2 might have been the reflection of these changes. The possibility that two R173 elements are located closely to each other and one of the elements is inverted, seems remote because of the dispersed nature of the R173 family. It is more likely that inversions, which took place in the region between V1 and V2 homologous sequences, gave rise to the 1R and 2R specific bands. Nevertheless, it can not be ruled out that these bands were the amplification products of unrelated sequences annealing to the primers.

Both of data from Rogowsky *et al.* (1991a) and this study indicates that the flanking region of R173 is not a particularly good source for RFLP probes. With a fraction of the time and effort for RFLP, PCR has yielded good markers although the mechanism is not clear. As more sequence data of repeat sequences from the cereal genome become available, it is expected that the combination of primers derived from different repeat sequences will play an important role in genome mapping.

Chapter 7
General discussion and future prospects

7.1. Introduction

At the commencement of this project, RFLPs were in extensive use in human genetic studies but few useful maps were available for crop plants (Tanksley *et al.*, 1989). There was almost no report for RFLP mapping in wheat or barley. This situation has now changed dramatically through joint efforts by research groups world-wide. Several RFLP maps have been published with barley (for example, Heun *et al.*, 1991) and wheat (e.g. Liu and Tsunewaki, 1991). However, genome analysis of cereals has been difficult as these genomes are larger than those of most other eucaryotes for which maps have been generated, including tomato and maize. Furthermore, little was known about the likely levels of polymorphism and variability between cultivars of cereals. It was perceived that the level of polymorphism might be low in common wheat because of its narrow genetic base and recent origin (Bell, 1987). The picture for barley was not clear either except that cultivated barley was reported to be less variable than its wild ancestor (Neale *et al.*, 1986). The complexity of the wheat and barley genomes also caused problems for RFLP detection. The Southern hybridization technique operates close to its resolution limit when single copy genes are sought in wheat.

The cost and technical difficulty of RFLPs suggested that there may be major problems with the implementation of this technique into plant breeding programs. Reports in the last few years have confirmed that the genome complexity and the low

level of polymorphism are the main reasons why wheat genome mapping has lagged behind that in plants (for example Kam-Morgan *et al.*, 1989).

The advent of PCR (Saiki *et al.*, 1985) offered a possible solution to some of these limitations to cereal genome mapping. This new technique offered several advantages including improved sensitivity, speed and simplicity. However, sequence information for primer design was much less abundant for cereals than for humans or other animals.

7.2. Sources of primers for PCR

To adopt PCR in genome mapping, an important consideration is the source of suitable primers. At present three general types of primers can be used to reveal and map polymorphisms in cereals: those that target specific, known sequences (eg, α-amylase), those that target intron-exon splice junctions and random primers(RAPDs).

With the rapid progress in molecular biology, especially after the introduction of PCR, the amount of sequence information has increased dramatically. For example, only few sequences from wheat and barley were recorded in the EMBL database in 1988 but in 1992, 180 and 120 sequences were registered for wheat and barley, respectively . Using gene sequence information from a data bank is an important approach for primer construction as demonstrated by the PCR for the α–amylase gene described in the present study. However, there are a number of limitations to this type of analysis. The level of polymorphism is expected to be low as the available sequences and, hence the target regions, are mainly from conserved regions within genes. Furthermore, the number of wheat and barley sequences within the database is not a true reflection of the number of potential PCR markers as many of the sequences are for the same or related genes.

Another approach is to use sequence information from other organisms to develop PCR primers. This can be done for sequences that show high levels of conservation. In order to test this concept, primers were tested for 6-PGD based on rat and *E. coli* protein and DNA sequences. The result presented in Chapter 5 is disappointing as the PCR and cloned product is probably the result of contamination. Although the technical steps of PCR cloning and sequencing appeared to work well in this case, the product was useless. This technique is too risky and time-consuming for general use. Furthermore, it may be inappropriate to use sequence data from bacterial genes. However, the approach of using an evolutionarily conserved portion of a gene for primer construction is quite common in gene cloning and genetic studies of various organisms. In these cases, the risk of artefacts was taken as PCR offered the only means for cloning the genes of interest.

Attempts have also be made to use inverse-PCR to generate chromosomal segment specific RFLP probes. However, the number of clones generated and analysed was too small to fully evaluate the feasibility of this approach. The present study indicated that the flanking regions of the R173 sequence family mainly consist of other classes of repetitive sequences, with few single or low copy sequences. It appears that the efficiency of this approach to the isolation of probes is not higher than with a random clone library. The inverse-PCR approach may better suit genomes with a lower content of repetitive sequences, such as *Arabidopsis* , as the probability of single copy sequence flanking the target repeat sequence is higher. Preliminary results in this study suggested that the primers derived from repeat sequences could be useful for generating markers for specific chromosome segment .

The use of short random primers (RAPDs) as molecular markers was proposed well after this project commenced (Williams *et al.*, 1990). RAPDs have proved to be very useful for genetic map construction and molecular genetic studies when a specific cross or mapping population is used. However, RAPDs suffer from

poor reproducibility. In addition, the application of RAPDs to plant breeding is restricted in that the map produced by RAPDs is not transferable from one cross to another. This means that a marker shown to have close linkage to a character can rarely be used in tracking that character in new crosses. A further difficulty with RAPDs is that bands may be suppressed in new background.

The ISJ primers, proposed in this thesis, offer a compromise between primers that target specific genes and RAPDs. In order to generate large numbers of markers distributed evenly over the chromosomes of the organism, we need to develop primers that target a range of sequences. As discussed in Chapter 3, several options exist; such as the translation start signals, the poly A addition sites, promoter and enhancer regions and the intron-exon splice junctions. All of these sequences have the advantages of frequent occurrence throughout the genome but even dispersal, lack of association with heterochromatic regions (since such areas are unlikely to contain genes) and the potential to provide hundreds or thousands of sites for the primers to anneal. A good candidate sequence to test this approach was the consensus sequence for plant intron splice junctions (ISJ) as the core of the junctions are highly conserved and the length of the conserved section is very close to the optimal length for PCR primers (Brown, 1986).

The application of PCR with ISJ primer and other primers proved to be successful for genome mapping in wheat, rye and barley. Many PCR bands, usually more than ten, could be generated in one amplification reaction. Significantly, the PCRs with ISJ primers were reproducible over varying quality and quantity of the template DNAs. Useful levels of polymorphism both between and within species were revealed with this approach.

The results presented here imply that PCR using ISJ primers is more robust than RAPD systems. PCR conditions for RAPD must be optimised for each species in respect to primer and template concentration and the temperature ramp in the thermal

cycle profile (Klein-Lankhorst *et al.*, 1991). The reason for the difference in reproducibility between RAPD and ISJ could be due to the length of primers. The primers used here are around 20 bases which has been reported as the optimal length for PCR amplification (Erlich, 1989). Moreover, the annealing temperature is raised at the second stage and this has been shown to increase the reproducibilty of PCR (Chapter 4). For the long random primers, the likelihood of success in the PCR is increased through the use of the ISJ primers in conjunction with either random or defined primers. By comparison, the RAPD primers are only 9 or 10 bases in length and annealing temperatures around 37^{o}C must be used. Primer annealing requires only loose association under these conditions and may cause the problems of sensitivity to DNA quality and quantity and the reaction conditions.

Armed with the ISJ PCR technique, the investigation on the intra and interspecific polymorphism was carried out with varieties and breeding lines of barley, wheat and rye. This study reveals that there is extensive diversity at the DNA level among the barley varieties. With the assistance of the barley addition lines, the bands can be mapped on to individual barley chromosomes.

The investigation of the levels of polymorphism between barley varieties provides a useful technique for the identification of barley varieties grown commercially in Australia. Little variation is seen within varieties while extensive and reliable intervarietal variation was detected. This method could unequivocally distinguish all of the commercially grown Australian barley varieties available at the Waite Institute. More care is needed to apply this approach to rye variety identification as there appears to be a high degree of intravarietal polymorphism. For application to wheat variety identification, improvements are needed in the detection system.

The results presented in this study show that the variability between pure line varieties of cultivated wheat is markedly lower than that in barley and rye. These findings confirm that common wheat varieties assayed here have a narrow genetic

base. Wheat frequently showed many minor PCR bands and sometimes slight smears of bands in contrast to the fewer, sharper bands seen with barley. This phenomenon may be associated with the greater complexity of the wheat genome relative to barley. In agarose and polyacrylamide gel systems employed in this study, the slight smear may obscure the detection of minor polymorphisms. Single stranded conformation polymorphism (SSCPs) (Dean *et al.*, 1990) and denaturing gradient gel electrophoresis (DGGE) (Myer *et al.*, 1986) may prove to be more powerful and suitable tools for revealing minor sequence changes, such as base substitution, in wheat cultivars.

7.3. Implementation of PCR analysis

The ultimate test of a molecular marker system will be in the success of its implementation into crop improvement programs. The effort here has concentrated upon the development and testing of the ISJ primer system. This presents a new type of marker that can be easily and reliably used to detect and map polymorphisms. However, these primers would never provide the sole source of markers for mapping. In a mapping program a range of marker systems would be used. RFLP's will have a key role and are likely to form the basis of maps. With ISJ primers it would be relatively straight forward to generate an average of two or more markers per chromosome arm. This would be achieved by using the series of ISJ primers described here alone or in combination with two or more random primers. The ability to use seed as a DNA source for the PCR analysis would allow the initial screening to be done without the need to recover whole plants. At two markers per chromosome arm, the ISJ primers would provide a crude map for initial analysis of a population, further analysis would proceed with standard RFLP markers. In many cases it is likely that the ISJ primers could be used directly to find linkage to the trait of interest.

The emphasis in this work has been on the development of markers for cereals. The ISJ primer technique would, however, have applicability to many eucaryotes. Indeed preliminary analysis in this laboratory indicate that these primers are useful for many other organisms, including fungi. This study has not included an analysis of the applicability of ISJ primers to organisms other than wheat, rye and barley. The ISJ primers were developed to test the model of the use of conserved regions as sources of PCR primers. The other types of conserved sequences described above have not been tested. An extensive array of new molecular markers would be available to molecular geneticists and cereal breeders should these prove as successful as ISJ primers in revealing polymorphisms in cereals.

BIBLIOGRAPHY

Ainsworth CC, Sharp PJ (1989) The potential role of DNA probes in plant variety identification. Plant Var Seeds 2:27-34

Allard RW (1988) Genetic changes associated with the evolution of adaptedness in cultivated plants and their wild progenitors. J Hered 79:225-238

Appels R, McIntyre CL (1985) Cereal genome organization as revealed by molecular probes. Oxf Surv Plant Mol Cell Biol 2:235-252

Appels R, Driscoll C, Peacock WJ (1978) Heterochromatin and highly repeated DNA sequences in rye (*Secale cereale*). Chromosoma 70:67-89

Appels R, Gerlach WL, Dennis ES, Swift H, Peacock WJ (1980) Molecular and chromosomal organization of DNA sequences coding for the ribosomal RNAs in cereals. Chromosoma 78: 293-311

Appels R, Moran LB, Gustafson JP (1986) Rye heterochromatin. I. Studies on the clusters of the major repeating sequence and the identification of a new dispersed repetitive sequence element. Can J Genet Cytol 28:645-657

Asiedu R (1986) Resistance to *Heterodero avenae* Woll located in the rye genome of triticale. PhD thesis, University of Adelaide, South Australia

Asiedu R, Fisher JM, Driscoll CJ (1990) Resistance to *Heterodero avenae* in the rye genome of triticale. Theo Appl Genet 79:331-336

Asins MJ, Carbonell EA (1986) A comparative study on variability and phylogeny of Triticum species 1. Intraspecific variability. Theor Appl Genet 72:551-558

Baima S, Haegi A, Stroman P, Casadoro P (1989) Characterization of a cDNA clone for barley leaf glutamine sythetase. Carleberg Res Commun 54:1-9

Balnaves ME, Nasioulas S, Dahl HHM, Forrest S (1991) Direct PCR from CVS and blood lysates for detection of cystic fibrosis and Duchenne muscular dystrophy deletions. Nucleic Acids Res 19:1155

Baulcombe DC, Huttly AK, Martienssen RA, Barker RF, Jarvis MG (1987) A novel wheat α-amylase gene (α-Amy3). Mol Gen Genet 209:33-40

Baum M, Appels R (1991) The cytogenetics and molecular structure of chromosome 1R--one of the most widely utilized sources of alien chromatin in wheat varieties. Chromosoma 101:1-10

Beckmann JS (1988) Oligonucleotide polymorphisms: a new tool for genomic genetics. Bio/Technology 6:1061-1064

Beckmann JS, Soller M (1986) Restriction fragment length polymorphisms and genetic improvement of agricultural species. Euphytica 35:111-124

Beckmann JS, Soller M (1990) Toward a unified approach to genetic mapping of eukaryotes based on sequence tagged microsatellite sites.Bio/Technology 8:930-932

Bedbrook JR, Jones J, O'Dell M, Thompon RD, Flavell RB (1980) A molecular description of telometric heterochromatin in Secale species. Cell 19:545-560

Bell (1987) The history of wheat cultivation. In: Lupton FGH (ed) Wheat breeding: its scientific basis. Chapman and Hall, London, pp 31-49

Bennett MD (1972) Nuclear DNA content and minimum generation time in herbaceeous plants. Proc R Soc Lond B181:109-135

Bennett MD, Smith JD (1976) Nuclear DNA amounts in angiosperms. Phil Trans R Soc Lond 274:227-274

Berthomieu P, Meyer C (1991) Direct amplification of plant genomic DNA from leaf and root pieces using PCR. Plant Mol Biol 17:555-557

Blake T (1987) Strategies for restriction fragment length polymorphism analysis in barley. In: Yasuda S, Konishi T (eds) Barley genetics V. Sanyo Press, Okayama, pp 503-508

Blake T, Shin JS, Sanchez M (1989) Development of the barley RFLP map. In: Helentjaris T, Burr B (eds) Development and application of molecular markers to problems in plant genetics, Cold Spring Harb, pp 139-143

Bothmer R von (1991) The wild species of *Hordeum*: relationship and potential use

for improvement of cultivated barley. In: Shewry (ed) Barley: genetics, molecular biology and biotechnology. CAB International, (in press)

Botstein D, White RL, Skonic M, David RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314-331

Brown AHD (1989) Core collections: A practical approach to genetic resource management. Genome 31:818-824

Brown AHD (1990) The role of isozyme studies in molecular systematics. Aust Syst Bot 3:39-46

Brown AHD (1991) Genetic variation and resources in cultivated barley and wild *Hordeum*. 6th International barley genetic symposium Barley genetics 6:VolI (in press)

Brown AHD, Munday J (1982) Population-genetic structure and optimal sampling of land races of barley from Iran. Genetica 58:85-96

Brown JW (1986) A catalogue of splice junction and putative branch point sequences from plant introns. Nucleic Acids Res 14:9549-9559

Brown JW, Feix G, Frendewey D (1986) Accurate in vitro splicing of two pre-mRNA plant introns in a HeLa cell nuclear extract. EMBO J 5:2749-2758

Bruzdzinski C, Gelehrter TD (1989) Determination of exon-intron structure: a novel application of the polymerase chain reaction technique. DNA 8:691-696

Bunce NAC, Forde BG, Kreis MK, Shewry PR (1986) DNA restriction fragment polymorphism at hordein loci: application to identifying and fingerprinting barley cultivars. Seed Sci Technol 14:419-429

Burr B, Evola E, Burr FA, Beckman JS (1983) The application of restriction fragment polymorphism to plant breeding. In: Setlow JK, Hollaender A (eds) Genetic engineering principles and methods, Plenum Press, New York, pp 45-59

Caetano-Anolles G, Bassam BJ, Gresshoff PM (1991) DNA fingerprinting using very short arbitrary oligonucleotide primers. Bio/Technology 9:553-557

Cariello NF, (1991) Fidelity of Thermococcus litoralis DNA polymerase (Vent™)

in PCR determined by denaturing gradient gel electrophoresis. Nucleic Acids Res 19: 4193-4198

Caruthers MH, Barone AD, Beaucage SL, Dodds DR, Fisher EF, McBride LJ, Matteucci M, Stabinsky Z, Tang JY (1987) Chemical synthesis of deoxyoligonucleotides by the phosphoramidite method. Methods Enzymol 154:287-313

Chamberlain JS, Gibbs RA, Ranier JE, Nguyen PN, Caskey CT (1988) Deletion screening of the Duchenne muscular dystrophin locus via mutiplex DNA amplification. Nucleic Acids Res 16:11141-11156

Chandler PM, Zwar JA, Jacobsen JV, Higgins TJV, Inglis AS (1984) The effects of gibberellic acid and abscisic acid on α-amylase mRNA levels in barley aleurone layers studies using an α-amylase cDNA clone. Plant Mol Biol 3:407 -418

Chao S, Sharp PJ, Worland AJ, Warham EJ, Koebner RMD, Gale MD (1989) RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. Theo Appl Genet 78:495-504

Clegg MT, Brown AHD, Whitefield PR (1984) Chloroplast DNA diversity in wild and cultivated barley: Implications for genetic conservation. Genet Res Camb 43:339-343

Collins FS, Weissman SM (1984) Directional cloning of DNA fragments at a large distance from an initial probe: a circularization method. Proc Natl Acad Sci USA 81:6812-6816

Connell C, Fung S, Heiner C, Bridgham J, Chakerian V, Heron e, Jones B, Menchen S Mordan W, Raff M, Recknor M, Smith L, Springer J, Woo S, Hunkapiller T (1987) Automated DNA sequence analysis. Bio/Techniques 5:3428-3433

Cooke RJ (1988) The use of electrophoresis for the distinctness testing of varieties of autogamous species. Plant Var Seeds 2:3-13

Day PR, Lupton FGH (1987) Future prospects. In: Lupton FGH (ed) Wheat breeding, its scientific basis. Chapman and Hall, London, pp 539-552

Dean M, Drumm M, Stewart C, Gerrard B, Perry A, Hidaka N, Cole J, Collins F, Iannuzzi N (1990) Approaches to to localize disease genes as applied to cystic fibrosis. Nucleic Acids Res 18:345-350

Dennis ES, Gerlach WL, Peacock WJ (1980) Identical polypyrimidine-polypurine satellite DNAs in wheat and barley. Heredity 44:349-366

Dhaliwal AS, Mares DJ, Marshall DR (1987) Effect of 1B/1R translocation on milling and quality characteristics of bread wheats. Cereal Chem 64:72-76

D'Ovidio, R, Tanzarella OA, Porceddu E (1990) Rapid and efficient detection of genetic polymorphism in wheat through amplification by polymerase chain reaction. Plant Mol Biol 15:169-171

Dundas IS, Shepherd KW, Fisher JM (1988) Transfer of cereal cyst nematode resistance from rye chromosome 6R into wheat. In: McWhite, KS, Downes RW, Reid BJ (eds), Proc 9th Aust Plant Breeding conf, Wagga Wagga, PP 101-102

Dundas IS, Shepherd KW, Fisher JM (1991) Prospects for breeding wheat resistant to the cereal cyst nematode using a gene from rye chromosome 6R. In: O'Brien L, Ellison FW (eds), Proc 6th Ass wheat breeding Soc Australia, pp 147-149

Dundas IS, Shepherd KW, Fisher JM (1992) Transfer of a nematode resistance gene from rye chromosome 6R to wheat and its mapping with deletion mutants. Abst 1st Int Crop Sci Cong, Crop Sci Soc Am & Iowa State Univ, Ames, Iowa, p 58

Eckert KA, Kunkel TA (1989) High fidelity DNA synthesis by the Thermus aquaticus DNA polymerase. Nucleic Acids Res 18:3739-3744

Economou EP, Bergen A, Warren AC, Antonarakis SE (1990) The polydeoxyadenylate tract of Alu repetitive elements is polymorphic in the human genome. Proc Natl Acad Sci USA 87:2951-2954

Egelke DR, Hoener PA, Collins FS (1988) Direct sequencing of enzymatically amplified human genomic DNA. Proc Natl Acad Sci USA 85:544-548

Erlich HA (1989) PCR Technology principles and applications for DNA amplification. Stockton Press, New York

Erlich HA, Gibbs RA, Kazazian JNR HH (1989) The polymerase chain reaction. Cold Spring Harbor Press, New York

Eskin NAM (1990) Biochemistry of Foods. 2nd ed, Academic Press, pp 167-203

FAO (1982) FAO production yearbook 1982. FAO Statistics Service No 46 Vol 36

Feinberg AP, Vogelstein B(1983) A technique for radiolabelling DNA restriction endonuclease fragment to high specific activity. Anal Biochem 132:6-13

Feldman M, Sears ER (1970) The wild gene resources of wheat. Sci Am 98-109

Felsenstein (1987) Phylogenies from molecular sequences: inference and reliability. Annu Rev Genet 22:521-565

Fisher JM (1982) Towards a consistent laboratory assay for resistance to *Hetereodera avenae*. EPPO Bull 12:445-449

Flavell RB (1980) The molecular characterization and organization of plant chromosomal DNA sequences. A Rev Pl Physiol 31:569-596

Flavell RB (1983) Repeated sequences and genome architecture. In: Ceferri O, Dure L (eds) Structure and function of plant genomes, Plenum Press, New York, pp 1-14

Flavell RB (1986) Repetitive DNA and chromosome evolution in plants. Phil Trans R Soc Lon B 312:227-242

Flavell RB, O'Dell M (1976) Ribosomal RNA genes on homoeologous chromosomes of group 5 and 6 in hexaploid wheat. Heredity 37:377-385

Flavell RB, Smith DB (1976) Nucleotide sequence organization in the wheat genome. Heredity 37:231-252

Flavell RB, Bennett MD, Smith JB, Smith DB (1974) Genome size and the proportion of repeated nucleotide sequence DNA in plants. Biochem Genetics 12:257-269

Flavell RB, Rimpau J, Smith DB (1977) Repeated sequence DNA relationships in four cereal genomes. Chromosoma 63:205-222

Flavell RB, O'Dell M, Hutchinson J (1981) Nucleotide sequence organization in plant chromosomes and evidence for sequence translocation during evolution. Cold Spring Harb Symp Quant Biol 45:501-508

Flavell RB, Bennett MD, Seal AG, Hutchinson J (1987) Chromosome structure and organization . In: Lupton FGH (ed) Wheat breeding, its scientific basis. Chapman and Hall, London, pp 211-268

Foulkes NS, Pandolfi de Rinaldis PP, McDonnll J, Cross NCP, Luzzato L (1988) Nucleic Acids Res 16:5687-5688

Freeman J, Marquez AJ, Wallsgrove RM, Saarelainen R, Forde BG (1990) Molecular analysis of barley mutants deficient in chloroplast glutamin synthetase. Plant Mol Biol 14:297-311

Gale MD, Hart GE (1987) Biochemical/molecular loci in hexaploid wheat (*Triticum aestivum* , 2n=42, genomes AABBDD). In: O'Brien SJ (ed) Genetic maps. Cold Spring Harb Lab, New York, p 678

Gale MD, Miller TE (1987) The introduction of alien genetic variation into wheat. In: Lupton FGH (ed) Wheat breeding, its scientific basis. Chapman and Hall, London, pp 173-210

Gelfand DH, White TJ (1990) Thermostable DNA polymerase. In: Innis M, Gelfand DH, Sninsky J, White TJ (eds) PCR Protocol. Academic, New York, pp 129-141

Gerlach WL, Bedbrook JR (1979) Cloning and characterization of ribosomal RNA genes from wheat and barley. Nucleic Acids Res 7:1869-1885

Gerlach WL, Dyer TA (1980) Sequence organization of the repeating units in the nucleus of wheat which contain 5S rRNA genes. Nucleic Acids Res 8:4851-4865

Gertz A, Wricke G (1989) Linkage between the incompatibility locus Z and a β-glucosidase locus in rye. Plant Breed 102:255-259

Gill BS, Chen PD (1987) Role of cytoplasm specific introgression in the evolution of the polyploid wheats. Proc Natl Acad Sci USA 84:6800-6804

Gill KS, Lubbers EL, Gill BS, Raupp WJ, Cox TS (1991) A genetic linkage map of *Triticum tauschii* (DD) and its relationship to the D genome of bread wheat (AABBDD). Genome 34:362-374

Good AG, Pelcher LE, Crosby WL (1989) Nucleotide sequence of a complete barley alcohol dehydrogenase-1 cDNA. Nucleic Acids Res 16: 7182

Goodall GJ, Filipowicz W (1989) The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. Cell 58: 473-483

Graner A, Siedler H, Jahoor A, Herrmann RG, Wenzel G (1990) Assessment of the degree and the type of restriction fragment length polymorphism in barley (*Hordeum vulgare*). Theo Appl Genet 80:826-832

Guidet F, Rogowsky P, Taylor C, Song W, Langridge P (1991) Cloning and characterisation of a new rye-specific repeated sequence. Genome 34:81-87

Gupta PK (1971) Homoeologous relationships between wheat and rye chromosomes-present status. Genetica 42:199-213

Gustafson JP, Butler E, McIntyre CL (1990) Physical mapping of a low copy DNA sequence in rye (*Secale cereale*). Proc Natl Acad Sci USA 87:1899-1902

Gyllensten UB (1989) PCR and DNA sequencing. Bio/Techniques. 7:700-708

Gyllensten UB, Erlich HA (1988) Generation of single stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA DQA locus. Proc Natl Acad Sci USA 85:7652-7656

Hamada H, Petrino MG, Kakunaga T (1982) A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc Acad Sci USA 79:6465-6469

Hanahan D (1983) Plasmid screening at high colony density. Methods Enzymol 100:333

Harlan JR, DeWet JMJ, Price BE (1973) Comparative evolution of cereals. Evolution 27:311-320

Hart GE (1987) Genetic and biochemical studies of enzymes. In: Heyne EG (ed) Wheat and wheat improvement, Am Soc Agr, Wisconsin, pp 199-214

Hawkins JD (1988) A survey on intron and exon length. Nucleic Acids Res 16: 9893-9905

Havey MJ, Muehlbauer FJ (1989) Linkage between restriction fragment length isozyme, and morphological markers in lentil. Theo Appl Genet 77:395-401

Heun M, Kennedy AE, Anderson JA, Lapitan NLV, Sorrells ME, Tanksley SD (1991) Construction of a restriction fragment length polymorphism map for barley (*Hordeum vulgare*). Genome 34:437-447

Hunkapiller T, Kaiser RJ, Koop BF, HOOD L (1991) Large-scale DNA sequencing. Curr Opin Biotech 2:92-101

Hutchinson JR, Flavell RB, Jones J (1981) Physical mapping of plant chromosomes by in situ hybridization. In: Setlow JK, Hollaender (eds) Genetic engineering, Vol III, Plenum Press, New york, pp 207-222

Huttly AK, Martienssen RA, Baulcombe DC (1988) Sequence heterogeneity and differential expression of the α-Amy2 gene family in wheat. Mol Gen Genet 214:232-240

Huynh TV, Young RA, Davis RW (1985) Constructing and screening cDNA libraries in λgt10 and λgt11. In: Glover DM (ed) DNA cloning: a practical approach , vol. 1. IRL press, Oxford, pp49-78

Innis M, Myambo KB, Gelfand DH, Brow MAD (1988) DNA sequencing with *Thermus aquaticu* DNA polymerase and direct sequencing of polymerase chain reaction-amplified products. Proc Natl Acad Sci USA 85:9436-9440

Isaacs ST, Tessman JW, Metchette KC, Hearst JE, Cimino GD (1991) Post-PCR sterilization: development and application to an HIV-1 diagnostic assay. Nucleic Acids Res 19: 109

Islam AKMS, Shepherd KW, Sparrow DHB (1981) Isolation and characterization of euplasmic wheat-barley chromosome addition lines. Heredity 46:161-174

Islam AKMS, Shepherd KW (1991) Alien genetic variation in wheat improvement. In: Gupta PK, Tsuchiya T (eds) Chromosome engineering in plant genetics, breeding, evolution. A, Elsevier, Amsterdam, pp 291-312

ITMI (1989) Comparative maps of triticeae based on RFLP markers.

Jahoor A, Backes G, Graner A, Herrmann RG, Fischbeck G (1991) Development of RFLP markers for barley. Plant Breed 107:73-76

Jana S, Pietrzak LN (1988) Comparative assessment of genetic diversity in wild and primitive cultivated barley in a center of diversity. Genetics 119:981-990

Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable "minisatellite" regions in human DNA. Nature 314:67-73

Ji LH (1992) Ph D thesis University of Adelaide

John B, Miklos GLG (1979) Functional aspects of satellite DNA and heterochromatin. Int Rev Cytol 58:1-114

Jones JDG, Flavell RB (1982a) The mapping of highly repeated DNA families and their relationship to C bands in chromosomes of *Secale cereale*. Chromosoma 86:595-612

Jones JDG, Flavell RB (1982b) The structure amount and chromosomal localization of defined repeated DNA sequences in species of the genus *Secale*. Chromosoma 86:613-641

Jorgenson RB (1986) Relationship in the barley genus (*Hordeum*): an electrophoretic examination of proteins. Hereditas 104:273-291

Joshi BC, Singh D (1979) Introduction of alien variation into bread wheat. In: Ramanujam S (ed) Proc 5th Intern Wheat Genet Symposium New Delhi, India, pp 342-348

Kam-Morgan LNW, Gill BS, Muthukrishnan S (1989) DNA restriction fragment length polymorphisms: a strategy for genetic mapping of D genome of wheat. Genome 32:724-732

Kan Y, Dozy A (1978) Antenatenal diagnosis of sickle-cell anaemia by DNA analysis of amniotic-fluid cells. Lacet 2:910-912

Keim P, Diers BW, Olson TC, Shoemaker RC (1990) RFLP mapping in soybean: association between marker loci and variation in quantitative traits. Genetics 126:735-742

Kesseli RV, Paran I, Michelmore R (1990) Genetic linkage map of lettuce (*Lactuca sativa*, ZN=10). In: O'Brien S (ed) Genetic maps (5th ed), Cold Spring Harb Lab Press, pp 6100-6102

Kimber G, Feldman M (1987) Wild wheat, an introduction. Special report 353, Univ Missouri, Columbia, pp 1-142

Kimber G, Sears ER (1987) Evolution in the genus *Triticum* and the origin of cultivated wheat. In: Heyne EG (ed) Wheat and wheat improvement, Am Soc Agr, Wisconsin, pp 154-164

Kinzler KW, Vogelstein B (1989) Whole genome PCR: application to the identification of sequences bound by gene regulatory proteins. Nucleic Acids Res 17:3645-3653

Klein-Lankhorst RM, Yermunt A, Weide R, Liharska T, Zabel P (1991) Isolation of molecular markers for tomato (*L. esculentum*) using random amplified polymorphic DNA (RAPD). Theor Appl Genet 83:108-114

Knott DR (1987) Transferring alien genes to wheat. In: Heyne EG (ed) Wheat and wheat improvement, Am Soc Agr, Wisconsin, pp 462-506

Knox CAP, Sonthayanon B, Chandra GR, Muthukrishnan S (1987) Structure and organization of two divergent α-amylase genes from barley. Plant Mol Biol 9:3-17

Kochert G, Halward T, Branch WD, Simpson CE (1991) RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. Theor Appl Genet 81:563-570

Koebner RMD, Shepherd KW (1986) Controlled introgression to wheat of genes from rye chromosome arm 1RS by induction of allosyndesis 1. Isolation recombinants. Theo Appl Genet 73:197-208

Lai EC, Stein JP, Catterall JF, Woo SLC, Mace ML, Means AR, O'malley BW (1979) Molecular structure and flanking nucleotide sequences of the natural chicken ovomucoid gene. Cell 18:829-842

Langudah ES, Appels R, Brown AHD (1991) The molecular-genetic analysis of *Triticum tauschii,* the D-genome donor to hexaploid wheat. Genome 34:375-386

Lawrence GJ, Appels R (1986) Mapping of the nucleolus organizer region, seed protein loci and isozyme loci on chromosome 1R in rye. Theo Appl Genet 71:742-749

Lazarus CM, Baucombe DC, Martienssen RA (1985) α-amylase genes of wheat are two multigene families which are differentially expressed. Plant Mol Biol 5:13-24

Ledbetter SA, Nelson DL, Warren ST, Ledbetter DH (1990) Rapid isolation of DNA probes within specific chromosome regions by interspersed repetititve sequence polymease chain reaction. Genomics 6:475-481

Li H, Cui X, Arnheim N (1990) Direct electrophoretic detection of the Allelic state of single DNA molecules in human sperm by using the polymerase chain reaction. Proc Natl Acad Sci USA 87:4580-4584

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene, Am J Hum Genet. 44:397-401

Liu YG, Mori N, Tsunewaki K (1990) Restriction fragment length polymorphism (RFLP) analysis in wheat. I. Genomic DNA library construction and RFLP analysis in common wheat. Jpn J Genet 65:367-380

Liu YG, Tsunewaki K (1991) Restriction fragment length polymorphism (RFLP) analysis in wheat. II. Linkage maps of the RFLP sites in common wheat. Jpn J Genet 66:617-633

Loh EY, Elliot JF, Cwerla S, Lanier LL, Davis MM (1989) Polymerase chain reaction with single sided specificity:analysis of T-cell receptor delta chain. Science 243:217-220

Mack DH, Sninsky JJ (1988) A sensitive method for the identification of uncharacterized viruses related to known virus groups: Hepadnvirus model system. Proc Natl Acad Sci USA 85:6977-6981

Maniatis T, Fritsch EF, Sambrook, J (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harb, NY

Martienssen RA (1986) The molecular genetics of α-amylase families in wheat (*Triticum aestivum*). PhD thesis, Cambridge University, UK

Martienssen RA, Baulcombe DC (1989) An unusual wheat insertion sequence (WIS1) lies upstrean of an α-amylase gene in hexaploid wheat and carried a "minisatellite" array. Mol Gen Genet 217:401-410

Martin GB, Williams JGK, Tanksley SD (1991) Rapid identification 0f markers linked to *Pseudomonas* resistance gene in tomato by using random primers and near-isogeneic lines. Proc Natl Acad Sci USA 88:2336-2340

May CE, Appels R (1987) The molecular genetics of wheat: Toward an understanding of 16 billion base pairs of DNA. In: Heyne EG (ed) Wheat and wheat improvement, Am Soc Agr, Wisconsin, pp 165-198

May CE, Appels R (1988) Allelism of the nucleolus organizer regions of hexaploid wheat. Proceedings of the 7th International Wheat Genet Symp pp 577-584

Meagher JW (1977) World dissemination of the cereal-cyst nematode (*Heterodera avenae*) and its potential as a pathogen of wheat. J Nematol 9:9-15

Meyerowitz EM, Bowman JL, Chang C, Kempin S (1990) RFLP map of *Arabidopsis Thaliana* ZN=10. In: O'Brien S (ed) Genetic maps (5th ed), Cold Spring Harb Lab Press, pp 698-699

Miksicek RJ, Towles HC (1983) Use of a cloned cDNA sequence to measure changes in 6-phophogluconate dehydrogenase mRNA levels caused by thyroid hormone and dietary carbohydrate. J Biol Chem 258:9575-9579

Miller JC, Tanksley SD (1990) RFLP analysis of phylogenetic relationship and genetic variation in the genus *Lycopersicon*. Theo Appl Genet 80:437-448

Mitchell LG, Merril CR (1989) Affinity generation of single-stranded DNA for dideoxy sequencing following the polymerase chain reaction. Anal Biochem 178:239-242

Molnar SJ, Gupta PK, Fedak G, Wheatcroft R (1989) Ribosomal DNA repeat unit polymorphism in 25 *Hordeum* species. Theo Appl Genet 78:387-392

Morris R, Sears ER (1967) The cytogenetics of wheat and its relatives. In: Quisenberry KS, Reitz LP (eds) Wheat and wheat improvement, Agon:13, pp 19-87

Mullis K B, Faloona F (1987) Specific synthesis of DNA in vitro via a polymerase catalysed chain reaction. Methods Enzymol 155:335-350

Murai K, Naiyu X, Tsunewaki K (1989) Studies on the origin of crop species by restriction endonuclease analysis of organellar DNA. III. Chloroplast DNA variation and interspecific relationships in the genus Secale. Jap J Genet 64:35-47

Muthukrishnant S, Gill BS, Swegle M, Chandra GR (1984) Structural genes for α-amylase are located on barley chromosomes 1 and 6. J Biol Chem 259:13637-13639

Myer RM, Maniatis T, Lerman LS (1986) Detection and localization of single base

changes by denaturing gradient gel electrophoresis. Methods Enzymol 155:501-527

Nakamura Y, Leppert M, O'Connell P, Wolf R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Kumlin E, White R (1987) Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 235:1616-1622

Nasoff MS, Wolf RE (1980) Molecular cloning , correlation of genetic and restriction maps, and determinationn of the direction of transcription of gnd of *Escheichia coli*. J Bact 143:731-741

Nasoff MS, Baker HV, Wolf RE (1984) DNA sequence of the *Escheichia coli*.gene, gnd, for 6-phosphogluconate dehydrogenase. Gene 27:253-264

Nath J, McNay JW, Paroda CM, Gulati SC (1983) Implication of *Triticum searsii* as the B genome donor to wheat using DNA hybridizations. Biochem Genet 21:745-759

Nath J, Thompson JP, Gulati SC (1985) Identification of the G-genome donor to *Triticum timopheevii* by DNA hybridization. Biochem Genet 23:125-137

Neale DB, Saghai-Maroof MA, Allard RW, Zhang Q, Jorgensen RA (1986) Chloroplast DNA diversity in populations of wild and cultivated barley. Genetics 120:1105-1110

Nelson DV, Ledbetter SA, Corbo L, Victoria MF, Ramiro RS, Webster TD, Ledbetter DH, Caskey CT (1989) Alu polymerase chain reaction: A method for rapid isolation of human-specific sequences from complex DNA sources. Proc Natl Acad Sci USA 86:6686-6690

Nilan RA (1964) The cytology and genetics of barley, 1957-1962. Research studies, Washington State University. 32

Nilan RA (1974) Barley (Hordeum vulgare). In King RC (ed.) Handbook of genetics, Plenum Press, New York, pp 93-110

O'Brien (1990) Genetic Maps. Cold Spring Harb Univ Press, Cold Spring Harb

Ochman H, Gerber AS, Hartl DL (1988) Genetic applications of an inverse polymerase chain reaction. Genetics 120:621-623

Olsen M, Hood L, Cantor C, Botstein D (1989) A common language for physical mapping of the human genome. Science 245:1434-1435

Ott J (1985) Analysis of human genetic linkage. Johns Hopkins Univ Press, Baltimore

Panaccio M, Lew A. (1991) PCR based diagnosis in the presence of 8% (v/v) blood. Nucleic Acids Res 19:1151

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721-726

Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci 85:2444-2448

Peddy P, Appels R (1989) A second locus for the 5S multigene family in Secale L.: sequence divergence in two lineages of the family. Genome 32:465-467

Rajaram S, Mann CE, Ortiz-Ferrara G, Mujeeb-Kazi A (1983) Adaptation, stability and high yield potential of certain 1B/1R CIMMYT wheats. In: Sakamoto S (ed) Proc 6th Int wheat Genet Symp, Kyoto, pp 613-621

Ranjekar PK, Palotta D, Lafontaine JG (1978) Analysis of plant genomes III. Denaturation and reassociation properties of cryptic satellite DNAs in barley (*Hordeum vulgare*) and wheat (*Triticum aestivum*). Biochem Biophys Acta 520:103-110

Reeves P, Stevenson G (1989) Cloning and nulceotide sequence of the *Salmoneella typhimurium* LT2 gnd. Mol Gen Gene 217:182-184

Rigby PWJ, Dieckmann M, Rhodes C, Berg P (1977) Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J Mol Biol 113:237

Riley R, Chapman V (1958) Genetic control of the cytologically diploid behaviour of hexaploid wheat. Nature 182:713-715

Riley R, Chapman V (1968) Introduction of yellow rust resistance of Aegilops comosa into wheat by genetically induced homoeologous recomnination. Nature 217:383-384

Rimpau J, Smith DB, Flavell RB (1978) Sequence organization analysis of the wheat and rye genomes by interspecies DNA/DNA hybridization. J Mol Biol 123:327-359

Rimpau J, Smith DB, Flavell RB (1980) Sequence organization in barley and oats chromosomes revealed by interspecies DNA/DNA hybridization. Heredity 44:131-149

Robertson DW (1939) Genetics of barley. J Am Soc Agro 31:273-283

Robyt JF, Whelan WJ (1968) In: Radley TA (ed) Starch and its derivatives, Chapman and Hall, London, p 423

Rogowsky PM, Manning S, Liu J, Langridge P (1991a) The R173 family of rye-specific repetitive DNA sequences: a structural analysis. genome 34:88-95

Rogowsky PM, Guidet FLY, Langridge P, Shepherd KW, Koebner RMD (1991b) Isolation and characterization of wheat-rye recombinabts involving chromosome arm 1DS of wheat. Theo Appl Genet 82:537-544

Rose EA (1991) Application of the polymerase chain reaction to genome analysis. FASEB J 5:46-54

Rothstein R. (1985) In Glover,D.M. (ed.), DNA Cloning-A practical Approach. IRL Press, Oxford, Vol.II, pp 45-64.

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81:8014-8018

Saiki RK (1989) The design and optimization of the PCR. In: Erlich HA (1989) PCR Technology: principles and applications for DNA amplification. Stockton Press, New York, pp 7-15

Saiki RK, Scarf S, Faloona F, Mulliks KB, Horn GT, Erlich HA, Arnheim N (1985) Ezymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230:1350-1354

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mulliks KB, Erlich HA (1988) Primer-directed enzymatic amplification of DNA with a

thermostable DNA polymerase. Science 239:487-491

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain -terminating inhibitors. Proc Natl Acad Sci USA 74:5463-5467

Sarkar G, Sommer SS (1990) Shedding light on PCR contamination. Nature 343:27

Schlegel R, Melz G, Mettin D (1986) Rye cytology, cytogenetics, and genetics-current status. Theo Appl Genet 72:721-734

Scoles GJ, Gill BS, Xin ZY, Clarke BC, McIntyre CL, Chapman C, Appels R (1988) Evolutionary change in grasses from the Triticeae. V. Frequent duplication and deletion events in the 5S RNA genes and the associated spacer regions. Plant Sys Evol 160:105-122

Sears ER (1952) Homoeologous chromosomes in *Triticum aestivum*. Genetics 37:624-631

Sears ER (1954) The aneuploids of common wheat. Mo Agric Exp Stn Res Bull 572:1-59

Sears ER (1966) Nullisomic-tetrasomic combinations in hexaploid wheat. In: Riley R, Lewis KR (eds) Chromosome manipulations and plant genetics, Oliver Boyd, London, pp 29-45

Sears ER (1973) Agropyron-wheat transfers induced by homoeologous pairing. In: Sears ER, Sears LMS (eds) Proc 4th Intern Wheat Genetics Symposium, Univ Miss, USA

Sears ER (1977) An induced mutation with homoeologous pairing in wheat. Can J Genet Cytol 19:585-593

Sears ER, Okamoto M (1958) Intergenomic chromosome relationships in hexaploid wheat. In: Proc 10th Int Congress Genet , Montreal, 2:258-259

Sharma HC, Gill BS (1983) Current status of wide hybridization in wheat. Euphytica 32:17-31

Sharp PJ, Soltes-Rak E (1988) Homoeologous relationships between wheat group 2 chromosome arm as determined by RFLP analysis. In Miller TE, Koebner RMD (eds) Proc 7th Int Wheat Genet Symp, IPSR, Cambridge, pp 635-638

Sharp PJ, Desai S, Gale MD (1988) Isozyme variation and RFLPs at the β-amylase loci in wheat. Theo Appl Genet 76:691-699

Sharp PJ, Chao S, Desai S, Gale MD (1989) The isolation, characterization and application in the Triticeae of a set of wheat RFLP probes identifying each Homoeologous chromosome arm. Theo Appl Genet 78:342-348

Sheffield VC, Cox DR, Lerman LS, Myers RM (1989) Attatchment of a 40-base-pair G+C-rich sequence (GC-clamp) to genomic fragments by the polymerase chain reaction results in improved detection of single-base changes. Proc Natl Acad Sci USA 86:232-236

Shepherd KW (1973) Homoeology of wheat and alien chromosomes controlling endosperm protein phenotypes. In: Sears ER, Sears LMS (eds) Proc 4th Int Wheat Genet Symp, University of Missouri, Columbia, pp 745-760

Shewry PR, Pratt HM, Miflin BJ (1978) Varietal identification of single seeds of barley by analysis of hordein polypeptides. J. Sci Food Agric 29:587-596

Shin JS, Chao S, Corpuz L, Blake T (1990) A partial map of the barley genome incorporating restriction fragment length polymorphism, polymerase chain reaction, isozyme, and morphological marker loci. Genome 33:803-810

Silver J, Keerikatte V (1988) "Inside-out" PCR to amplify cellular DNA flanking integrated proviruses. J Cell Biochem Suppl 13E WH239

Simmonds NW (1987) Trials and multiplication. In: Principles of crop improvement Longman Scientific Technical, UK

Singh NK, Shepherd KW (1988) Linkage mapping of the genes controlling endosperm proteins in wheat II. Genes on the long arms of group 1 chromosomes. Theo Appl Genet 75:642-650

Skolnick MH, Wallace RB (1988) Simultaneous analysis of multiple polymorphic loci using amplified sequence polymorphism (ASPs). Genomics 2:273-279

Slocum MK, Figdore SS, Kennard WC, Suzuki JY, Osborn TC (1990) Linkage arrangement of restriction fragment length polymorphism loci in *Brassica oleracea*. Theo Appl Genet 80:57-64

Smith GP (1976) Evolution of repeated DNA sequences by unequal crossover.

Science 191:528-535

Smith DB, Flavell RB (1974) The relatedness and evolution of repeated nucleotide sequences in the genome of some Gramineae species. Biochem Genet 12:243-256

Smith DB, Flavell RB (1977) Nucleotide sequence organization in the rye genome. Biochem Biophys Acta 474:82-97

Smyth DR (1991) Dispersed repeats in plant genomes. Chromosoma 100:355-359

Sogaard B, Wettstein-Knowles P von (1987) Barley: genes and chromosomes. Carlsberg Res Commun 52:123-196

Sommer R, Tautz D (1989) Minimal homology requirement for PCR. Nucleic Acids Res 17:6749

Sonina NV, Lushnikova AA, Tihonov AP, Ananiev EV (1989) Dialect-I, species-specific repeated DNA sequence from barley, *Hordeum vulgare*. Theo Appl Genet 78:589-593

Song KM, Osborn TC, Williams PH (1988) *Brassica* taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs) 2. Preliminary analysis of subspecies within *B. rapa* (*syn. campestris*) and *B. oleracea.*. Theor Appl Genet 76:593-600

Song KM, Osborn TC, Williams PH (1990) *Brassica* taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs) 3. Genome relationships in and related genera and the origin of *B. oleracea.* and *B. rapa* (*syn. campestris*). Theor Appl Genet 79:497-506

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503-517

Speck PG (1986) Mapping repeat DNA in rye. Honours thesis, University of Adelaide

Stallings RL, Torney DC, Hildebrand CE, Longmire JL, Deaven LL, Jett JH, Doggett NA, Moyzis RK (1990) Physical mapping of human chromosomes by repetitive sequence fingerprinting. Proc Natl Acad Sci USA 87:6218-6222

Stegemann H (1984) Retrospect on 25 years of cultivar identification by protein pattern and prospects for the future. In: Draper SR, Cooke RJ (eds) Biochemical tests for cultivar identification, Internat Seed Testing Assoc, pp 20-31

Stephens JC, Cavanaugh ML, Gradie MI, Mador ML, Kidd KK (1990) Mapping the human genome: current status. Science 250:237-250

Streeck RE, Hobom G (1975) Mapping of cleavage sites for restriction endonucleases in λdv plasmid. Eur J Biochem 57: 596-606

Studer R, Epplen JT (1990) On the organization of animal genomes: ubiquitous interspersion of repetitive DNA sequences. In: Geldemam H, Ellendorff F (eds) Genome analysis in domestic animals, VCH, Weinheim, pp 17-33

Tanksley SD, Orton TJ (1983) Isozymes in plant genetics and breeding. Elsvier, Amsterdam

Tanksley SD, Mutschler HA (1990) Linkage map of tomato (*Lycopersicon esculentum,* ZN=24). In: O'Brien S (ed) Genetic maps (5th ed), Cold Spring Harb Lab Press, pp 6.3-6.15

Tanksley SD, Young ND, Paterson AH, Bonierbale MW (1989) RFLP mapping in plant breeding: New tools for an old science. Bio/Technology 7:258-264

Tautz D (1989) Hypervariability of simple sequences as a general sources for polymorphic DNA markers. Nucleic Acids Res 17:6463-6471

Thompson JP, Nath J (1986) Elucidation of the B genome donor to *Triticum turgidum* by unique and repeated sequence DNA hybridization. Biochem Genet 24:39-50

Trick M, Dennis ES, Edward KJ, Peacock WJ (1988) Molecular analysis of the alcohol dehydrogenase gene family of barley. Plant Mol Biol 11:147-160

Triglia T, Peterson MG, Kemp DJ (1988) A procedure for in vitro amplification of DNA segment that lie outside the boundaries for known sequence. Nucleic Acids Res 16:8186

Wang ML, Atkinson MD, Chinoy CN, Devos KM, Liu C, Rogers J, Gale M (1991) RFLP-based map of rye (*Secale cereale*) chromosome 1R. Theo Appl Genet 82:174-178

Watkins PC (1988) Restriction fragment length polymorphism (RFLP): applications in human chromosome mapping and genetic disease research. Bio/Technology 6:310-320

Watson JD, Cook-Deegan RM (1991) Origins of the human genome projects. FASEB J 5:8-11

Weber JL (1990) Human DNA polymorphisms and methods of analysis. Curr Opin Biotech 1:166-171

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388-396

Wehling P (1985) Electrophoretic analysis of 10 enzyme systems in rye: linkage relationships and chromosomal location of isozyme loci. In: Proc EUCARPIA Meeting of the cereal section on rye, Svalof, Sweden, pp 101-124

Wehling P, Uphoff H (1990) Untersuchungen zur Markierung von Identifizierung Inkompatibilitatsloci beim Roggen. Vortr Pflanzenzucht 18:45-69

Weining S, Langridge P (1991) Identification and mapping of polymorphisms in cereals based on the polymerase chain reaction. Theo Appl Genet 82:209-216

Welsh J, McClelland C (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res 18: 7213-7218

Welsh J, McClelland C (1991) Genomic fingerprinting using arbitrarily primed PCR and a marix of pairwise combination of primers. Nucleic Acids Res 19: 5275-5279

Welsh J, Peterson, McClelland C (1991) Polymorphisms generated by arbitrarily primed PCR in the mouse: application to strain identification and genetic mapping. Nucleic Acids Res 19: 303-306

Wettstein-Knowles P von (1991) Cloned and mapped genes: current status. In: Shewry (ed) Barley: genetics, molecular biology and biotechnology. CAB International, (in press)

White TJ, Arnheim N, Erlich NA (1989) The polymerase chain reaction. Trend Genet 5:185-189

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531-6535

Williams MNV, Pande N, Nair S, Mohan M, Bennett J (1991) Restriction fragment length polymorphism analysis of polymerase chain reaction products amplified from mapped loci of rice (*Oryza sativa* L.) genomic DNA. Theor Appl Genet 82:489-498

Wricke G, Wehling P (1985) Linkage between an incompatibility locus and a peroxidase isozyme locus (Prx7) in rye. Theo Appl Genet 71:289-291

Wyman A, White R (1980) A Highly polymorphic locus in human DNA. Proc Natl Acad Sci USA 77:6754-6758

Xin ZY, Appels R(1988) Occurrence of rye (*Secale cereale*) 350-family DNA sequences in *Agropyron* and other Triticeae. Plant Syst Evol 160:65-76

Young ND ,Tanksley SD (1989) Restriction fragment length polymorphism maps and the concept of graphical genotypes. Theo Appl Genet 77:95-101

Zohary D (1969) The progenitors of wheat and barley in relation to domestication and agricultural dispersal in the Old World. In: Uco PJ, Dimbleby GW (eds) The domestication and exploitation of plants and animals, Duckworth, London