Theor Appl Genet (1999) 99: 192–198

G. J. Bryan · P. Stephenson · A. Collins
J. Kirby · J. B. Smith · M. D. Gale

Low levels of DNA sequence variation among adapted genotypes of hexaploid wheat

Received: 3 October 1998 / Accepted: 28 November 1998

Abstract PCR products from regions corresponding to sequences hybridising to wheat RFLP probes were sequenced in order to establish the level of DNA sequence variation among adapted wheat genotypes. Hexaploid bread wheat shows a very low rate of nucleotide polymorphism, approximately 1 polymorphic nucleotide per 1000 basepairs. Differences in PCR product length can be exploited to design genome-specific amplicons, which may have use in gene tagging or in diagnostic applications. Interpretation of results may be complicated by the simultaneous amplification of orthologous and paralogous sequences. These findings have significant implications for the use of STS markers in wheat and other polyploid species.

Key words PCR markers · Sequence-tagged-site · Wheat

Introduction

Restriction-fragment-length-polymorphisms (RFLPs) have been used extensively in many plant species to generate genetic maps but have been applied to a much lower extent by plant breeders. Recent developments in marker technology have allowed the partial replacement of RFLPs with polymerase chain reaction (PCR)-based markers, such as random amplified polymorphic DNAs (RAPDs), microsatellites and amplified-fragment-length-polymorphisms (AFLPs) (Williams et al. 1990; Weber and May 1989; Vos et al. 1995). The speed, sensitivity and ease of implementation of PCR-based markers will allow greater direct application in breeding programmes, increasing throughput and efficiency. However, locus-specific PCR-based markers, particularly microsatellites, despite their codominant and highly polymorphic nature, are time-consuming and expensive to develop. PCR-based markers showing a predominantly dominant mode of inheritance, such as RAPDs and AFLPs, may be inherently less "transferable" between plant genomes than RFLPs and, therefore, it is likely that RFLP and sequence-tagged-site (STS) markers will continue to play a major role.

Levels of RFLP observed vary greatly between plant species and depend on several factors, most notably the evolutionary history and mating system of the species concerned. In maize, for example, approximately 95% of low-copy sequences reveal RFLP when tested with just a few restriction enzymes (Shattuck-Eidens et al. 1990). In contrast, hexaploid bread wheat shows very low levels of RFLP and, therefore, wide crosses are required for genetic mapping (Chao et al. 1989), and in breeding programmes involving adapted material, RFLPs often show little or no polymorphism. Despite these problems, and due to the lack of a viable alternative, there has been a large investment in RFLP mapping in hexaploid wheat. The latest wheat RFLP maps contain in the order of 1000 markers. RFLPs have also found use recently in comparative mapping projects (Kurata et al. 1994).

The nature and amount of DNA sequence variation in mapped RFLPs in maize has been studied in detail (Shattuck-Eidens et al. 1990). The use of sequenced RFLPs as STS markers in barley and hexaploid wheat has been investigated (Tragoonrung et al. 1992; Talbert et al. 1994). A potential use of sequenced RFLPs is in

Communicated by G. E. Hart

G. J. Bryan[1](✉) · P. Stephenson · A. Collins[2] · J. Kirby
J. B. Smith · M. D. Gale
John Innes Centre, Norwich Research Park, Colney,
Norwich NR4 7UJ, UK
Fax: +44 1382 562426
E-mail: g.bryan@scri.sari.ac.uk

Present addresses:
[1]Crop Genetics Department, Scottish Crop Research Institute,
Invergowrie, Dundee DD2 5DA, Scotland
[2]Department of Plant Sciences, University of Oxford, South Parks
Road, Oxford, OX1 3RB, UK

the development of specific PCR "tags" that can be used to select for genes affecting desirable agronomic traits, such as bread-making quality (D'Ovidio and Anderson 1994; Van Campenhout et al. 1995) or for genes of scientific importance, such as the *Ph1* locus (Gill and Gill 1996). In hexaploid wheat this is made more difficult by polyploidy and the high levels of sequence homology between the three ancestral genomes. PCR markers developed from RFLPs are unlikely to detect the high levels of polymorphism seen, for example, with microsatellites, but the advantages of detecting loci already mapped are considerable.

The objectives of the study presented here were to determine the level of DNA sequence polymorphism within regions corresponding to mapped RFLPs and to assess the feasibility of the systematic exploitation of DNA sequence polymorphism between adapted genotypes of wheat. The selected hexaploid wheat material was a small set of adapted wheat varieties and the parents of a wide mapping cross.

Materials and methods

Genetic stocks/wheat varieties

The varieties used for this analysis were chosen as representative of northern European breeding programmes as well as the parents of the mapping cross used to construct the RFLP map from this laboratory: 'Cappelle-Desprez', 'Brigadier', 'Herzog', 'Hereward', 'Soissons' and 'Chinese Spring' and a synthetic hexaploid (IPSR 1190903) hereafter referred to as 'Synthetic', (McFadden and Sears 1946; Sears 1976). The aneuploid stocks used for the determination of chromosomal origins of PCR products were the nullisomic-tetrasomic lines of 'Chinese Spring' (Sears 1954). The putative diploid ancestors *Triticum urartu*, *Aegilops speltoides* and *Aegilops squarrosa* were also used.

DNA extractions

DNA was extracted from fresh or freeze-dried plant tissue as described by Sharp et al. (1988).

RFLP probes/PCR primers

The following RFLP probe regions were selected for DNA sequencing:

PSR128 – a cDNA which detects fragments on chromosomes 5A, 5B and 5D.
PSR160 – a cDNA which detects fragments on chromosomes 4A, 7A and 7D. The 4A localisation results from the 4A/7B translocation in hexaploid wheat.
PSR551 – a genomic DNA fragment which detects a fragment on chromosome 6B.

PCR primers were made to sequenced wheat RFLP probes using the computer programme PRIMER (Genetics Computer Group, Madison, Wis., USA):

PSR128	PSP128F1: CCTCGAGAACAAGGAGAAG
	PSP128F2: CCCTCGGTTCCTAATCGGA
	PSP128R1: CGAAATCCACAAGAAGAAAG
	PSP128R2: TGCCATGTTTAGAATAACACGC
PSR160	PSP160F1: CAAATTAAATCAATATGCATGCA
	PSP160F2: CATCATCAAGAGCCTCATCTC
	PSP160R1: AACGACTTCAGCGTCAAG
	PSP160R2: ATCACCTTCAAGAACAACGC
PSR551	PSP551F1: ACCCCTCCTCTTTCTTTTCC
	PSP551R1: CTAAAACATATCCTCATGCATG

PCR amplification

Genomic sequences were amplified in 50-μl PCR reactions containing 50 ng template DNA, 10 m*M* TRIS-HCl, pH 8.3, 1.5 m*M* $MgCl_2$, 50 m*M* KCl, 100 µg/ml gelatine with 200 n*M* of each primer, 100 µ*M* of each dNTP and 1 U of *Taq* DNA polymerase. Cycling conditions were as follows: initial denaturation step of 4 min at 94°C, a followed by 25–30 cycles of a 30-s denaturation at 94°C, a 1 min annealing at 54°C and a 1-min extension at 72°C. After cycling, PCR reactions were incubated for 5 min at 72°C.

Purification and cloning of PCR products

DNA fragments were purified from agarose gels or directly from PCR reactions using Prepagene (BioRad). Purified PCR products were ligated into the vector pGEM-T (Promega) and transformed into *E. coli* strain DH5α by electroporation.

DNA sequencing and analysis

DNA sequencing was performed using the AutoRead Sequencing kit (Pharmacia) and analysed with an Automated Laser Fluorescence (ALF) sequencer (Pharmacia). Sequencing templates were produced using standard small-scale plasmid preparation techniques. For each PCR fragment, a minimum of three independent clones were sequenced. DNA sequences were compared and aligned using the GCG programmes GAP, BESTFIT and PILEUP (Genetics Computer Group, Madison, Wis., USA).

Results

PCR analysis of genomic sequence

PSR551. A single PCR product of the size predicted from the RFLP clone sequence (i.e. 784 bp) was generated for each of the seven varieties. This product was shown to originate from chromosome 6B using the aneuploid lines of 'Chinese Spring' (Fig. 1a). PCR products from the varieties 'Chinese Spring', 'Synthetic', 'Cappelle-Desprez', 'Brigadier', 'Herzog' and 'Hereward' were sequenced and compared to the cloned sequence. Two single base pair polymorphisms were observed among the seven sequences with respect to the probe sequence (A-G position 593 in var 'Brigadier'; T-A position 715 in var 'Synthetic').

Analysis of PCR products using primers against two wheat cDNA sequences

PSR128. Cloned PCR products generated from wheat cultivars using primers PSP128F1 and PSP128R1 were of three discrete lengths: (1) 423 bp, (2) 416 bp, (3) 479–481-bp. The 423-bp and 416-bp products were absent from the nullisomic-5 A and nullisomic-5B lines, respectively, suggesting that these fragments originated from chromosomes 5A and 5B. The 479- to 481-bp product was absent from the nullisomic-5D line, suggesting that this fragment originated from chromosome 5D (Fig. 1b). The 416-bp PCR product showed a much lower level of amplification than the other two products, except in the nullisomic 5A tetrasomic 5B line. These results have been confirmed by restriction digests of the products and by their analysis on sequencing gels (Fig. 2). These data suggested that the three PCR products obtained originate from the group-5 chromosomes to which probe PSR128 hybridises.

Table 1 gives information on clones sequenced for the various products from the seven cultivars. This information should not be taken to suggest that clones "missing" from an accession imply that the relevant sequence is not detected in that accession by the primers used. Rather it reflects the sampling from clones sequenced for each cultivar following cloning of the PCR products. These PCR products were all larger than the size expected from the cDNA sequence (326 bp). This difference, as well as size differences between the products, was due to the presence of introns of different lengths in the amplified products. Sequences at the termini of the putative introns agreed with consensus nucleotide sequences for intron-exon boundaries for monocotyledonous plants (Goodall et al. 1991). A length difference among varieties in the 479- to 481-bp sequence was due to a two base pair insertion/deletion polymorphism present within in the intron. The three introns (91 bp, 97 bp and 154–156 bp) were inserted at identical positions. The two shorter introns shared approximately 75 bp of sequence homology with the 154- to 156-bp intron (Fig. 3a). Overall, there are 11 polymorphisms between the three exons, ten single base pair substitutions, and a one base pair insertion/deletion. The 416-bp and 423-bp sequences differed by virtue of two small deletions (6 bp and 1 bp) as well as several single base pair differences, the exact number being dependent on the alignment used. Overall the three sequences showed high levels of identity (PSP128-5A vs PSP128-5B – 96%, PSP128-5A vs PSP128-5D – 93%, PSP128-5B vs PSP128-5D – 93%). The differences were largely

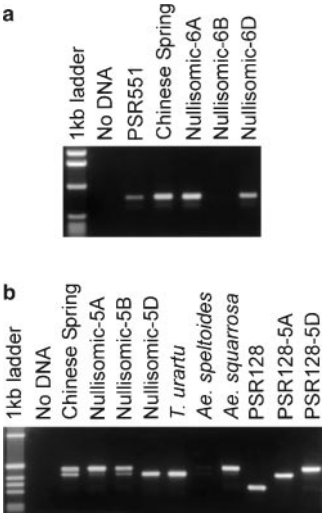

Fig. 1a, b PCR products generated using primers to RFLP probes. **a** PCR products generated with primers PSR551F1 and PSR551R1 using an annealing temperature of 54°C. Products were electrophoresed in a 1% agarose gel in TAE buffer. **b** PCR products generated with primers PSR128F1 and PSR128R1 using an annealing temperature of 54°C. Products were electrophoresed in a 3% agarose gel in TAE buffer

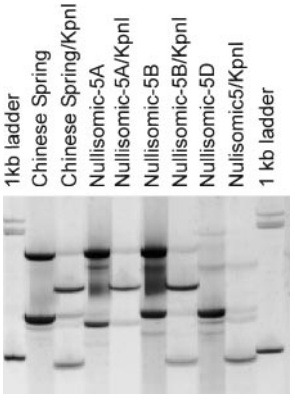

Fig. 2 PCR products of PSR128 were digested with *Kpn*I and the fragments run on a 6% sequencing gel which was then silver-stained

Table 1 DNA sequence polymorphism among hexaploid wheat varieties

RFLP probe	PCR locus	PCR product size	Number of varieties analysed[a]	Number of observed polymorphisms	Rate of observed polymorphism
PSR551	*Xpsp551-6B*	784	6 (CS, SY, CD, BR, HZ, HE)	2	4.25×10^{-4}
PSR128	*Xpsp128-5A*	423	7 (CS, SY, CD, BR, HZ, HE, SO)	3	1.01×10^{-3}
	Xpsp128-5B	416	1 (SY)	–	–
	Xpsp128-5D	479–481	4 (CS, SY, CD, SO)	4	2.09×10^{-3}
PSR160	*Xpsp160-7A*	325	4 (CS, SY, HZ, HE)	0	0
	Xpsp160-4A	346	1 (CS)	–	–
	Xpsp160-?	359	4 (CS, SY, CD, BR)	1	6.96×10^{-4}
Overall					$\sim 1.01 \times 10^{-3}$

[a] Wheat accessions/cultivars coded as follows: 'Chinese Spring', CS; 'Synthetic', SY; 'Cappelle-Desprez', CD; 'Brigadier', BR; 'Herzog,' HZ; 'Hereward', HE; 'Soissons', SO

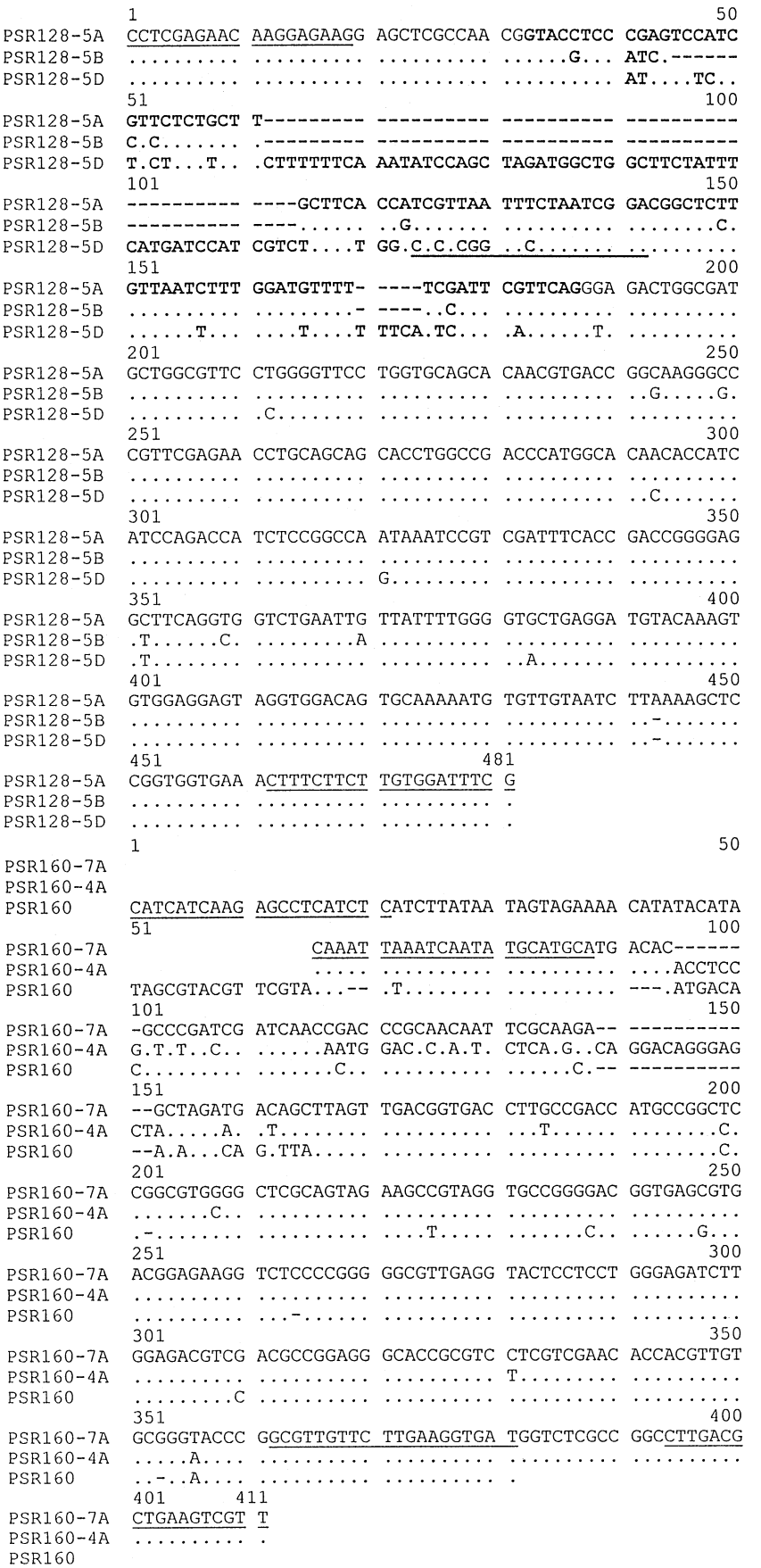

Fig. 3a, b Sequence comparisons between PCR products. Identities between sequences are indicated by (·) and gaps introduced to optimise alignments by (-). PCR primer sequences are *underlined* (see text for details). **a** Comparisons between different types of sequences obtained using PCR primers to PSR128. Intron sequence is shown in *bold type*. Note that primer PSR128F2 was designed against the lower sequence variant (i.e. PSR128-5D). Primer PSR128R2 is not shown since it lies outside the PSR128 sequence shown. **b** Comparisons between different types of sequences obtained using PCR primers to PSR160

explained by intron length variation and to sequence differences within the introns.

The sizes of PCR products obtained from *Triticum urartu* and *Aegilops squarrosa*, putative A- and D-genome donors of bread wheat, are consistent with the conclusion that the aforementioned PCR products originate from the A and D genomes (Fig 1b). Moreover, the sequence of the cloned *Aegilops squarrosa* PCR product was virtually identical to the 479- to 481-bp clones (data not shown). A fragment of this sequence was specifically amplified from wheat using primer PSP128F2 and raised against sequences unique to the larger intron, in combination with either reverse primer (Fig. 4). This product was not amplified from a nullisomic-5D line, nor from *T. urartu* and *Ae. speltoides*, possible donors of the A and B genomes of hexaploid wheat, confirming that the 479- to 481-bp product originated from chromosome 5D.

PSR160. Sequencing of clones from the PCR products generated using primers PSP160F1 and PSP160R1 revealed small variations in length: (1) 325 bp, (2) 346 bp. These products differed by virtue of a 21-bp insertion, dispersed across a 38-bp segment of the shorter sequence, most parsimoniously explained by the presence of two insertions, of 7 bp and 14 bp, respectively (see Fig. 3b). There were five sequence differences between the two sizes of clone outside the region containing the insertion/deletion. PCR experiments using nullisomic-tetrasomic lines, and electrophoresis on sequencing gels, revealed that the 325-bp and 346-bp products originated from chromosomes 7A and 4A, respectively (data not shown). A third PCR product (PSP160-?) was obtained using a second set of primers, PSP160F2 and PSP160R2: (3) 359 bp, which resembled the 325-bp product most closely (Fig. 3b), the chromosomal origin of which was not determined. The three sequences showed high levels of sequence identity (PSR160-7A vs PSR160-4A – 90%, PSR160-7A vs PSR160-? – 91%, PSR160-4A vs PSR160-? – 87%). Table 1 gives information on the clones sequenced for the various products from the seven cultivars.

Comparison of sequences of same length from different wheat varieties and levels of polymorphism

The level of sequence polymorphism between varieties was determined for the different sequence types obtained from the various loci (Table 1). Clones containing the 416-bp PSR128-5B sequence and the 346-bp PSR160-4A sequence were only obtained from 'Synthetic' and 'Chinese Spring' respectively and, therefore, polymorphism data have not been calculated for these products. There were only seven DNA sequence polymorphisms for the other two PSR128 sequence types, four insertion/deletions and three single base pair substitutions. Insertion/deletion events affecting two or more contiguous nucleotides were assumed to represent single mutational events. Clones were obtained from four of the seven varieties for two of the PSR160 sequence types identified. The 7A products obtained from vars 'Chinese Spring', 'Synthetic', 'Herzog' and 'Hereward' were identical in sequence. The cloned 359-bp products were almost identical among four wheat varieties ('Chinese Spring', 'Synthetic', 'Cappelle-Desprez', 'Brigadier') with only a single polymorphism in var 'Brigadier'.

Adopting the rationale of comparing only sequences of identical length originating as far as it is possible to determine from the same locus, we were able to obtain estimates of the approximate level of DNA sequence polymorphism among wheat varieties. Table 1 summarises the polymorphism data, and from this we calculated an overall average level of DNA sequence polymorphism of approximately 0.001 per nucleotide.

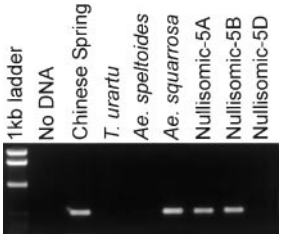

Fig. 4 Genome-specific amplification for PSR128-5D with primers PSP128F2 and PSP128R2 using an annealing temperature of 54°C. Products were electrophoresed in a 1% agarose gel in TAE buffer

Discussion

Conversion of RFLPs to PCR primers in hexaploid wheat

PCR primers generated from wheat RFLP clones can be used to generate PCR products of the predicted size from hexaploid wheat (Talbert et al. 1994). However, it is unclear whether PCR primers are predominantly amplifying from the loci to which the RFLP was mapped. Erpelding et al. (1996) reported that only 69% of PCR primer pairs amplified fragments originating from loci expected from RFLP mapping and also used Southern analysis to show that certain amplification products are dissimilar to the original RFLP probe. In contrast, our studies show that seven PCR products characterised at the sequence level show very strong sequence homology to the corresponding RFLP probe, differing mainly by the presence of introns or short insertions. We also provide evidence that at least six of seven PCR poducts originate from chromosomes expected from RFLP mapping. Our analysis of primers from other wheat cDNAs indicates that there are PCR fragments which are not absent from any nullisomic-tetrasomic lines (G. Bryan, unpublished observations), suggesting the possible amplification of the same size product from loci from the A, B and D genomes.

Use of cloned PCR products to study DNA sequence polymorphism

PCR using primers to wheat cDNAs has a tendency to give several PCR products of very similar size, rendering direct sequencing impossible. The use of cloned PCR products requires the sequencing of a relatively large number of clones to measure rates of sequence polymorphism. Some primer pairs (e.g. PSP160F2 and PSP160R2) produce PCR products which clearly differ in size, and yet, all clones obtained are of identical length, suggesting that there are certain sequences which are preferentially ligated into "T-tailed" PCR product cloning vectors, such as pGEM-T.

Levels of polymorphism among wheat varieties

Our overall estimate of approximately 1 difference per 1000 base pairs between any two wheat varieties should be viewed as a "baseline" estimate of the level of DNA sequence polymorphism for low-copy-number sequences in hexaploid wheat. The number of mutational events is very low and, therefore, PCR and sequencing errors are significant factors. The amount of DNA sequencing required to carry out a more extensive analysis would be considerable, and the benefits of such a study are questionable, given the low rate of polymorphism. The selected wheat varieties were chosen to represent a cross section of northern European varieties as well as more diverse material ('Chinese Spring' and 'Synthetic'). The low observed level of DNA sequence polymorphism raises doubts as to the utility of STS methodology for the routine detection of polymorphism in hexaploid wheat but does not preclude the conversion of particular RFLPs to specific PCR markers (Penner et al. 1995). Other authors that have used STS methodology in wheat and barley (e.g. Tragoonrung et al. 1992; Talbert et al. 1994), genomic RFLP clones, which may be inherently more polymorphic than cDNAs, and were able to survey larger sets of material, since they relied upon indirect methods (i.e. restriction endonuclease digestion of PCR products) for the detection of polymorphism.

RFLP data based upon a more diverse set of material give rise to an average pairwise difference of about 9% (Chao et al. 1989). RFLPs, unlike STS markers, assay sequence polymorphism outside the probe region. If all RFLPs were due to single base pair changes, we would expect, extrapolating from our data, an approximately 1% (12/1000) chance of observing RFLP between two varieties for any restriction enzyme with a six base pair recognition site. Since the RFLP level is generally higher, by about tenfold, this may suggest that a high proportion of RFLP in hexaploid wheat is due to insertion/deletion/rearrangement events.

Potential use of primers from wheat RFLP probe sequences which detect homoeologous loci

The majority of PCR primers raised against cloned wheat sequences appear to amplify from at least two loci (G. Bryan, unpublished data). Products obtained from PSR128 show polymorphism in intron length, allowing the design of at least one locus-specific amplicon. Most primers we have tested generate fragments of more uniform length, suggesting that significant length variation between orthologous loci in wheat is not common. This suggests a lower potential for the design of locus specific amplicons because of a smaller "target", for the design of specific primers. However, such differences, however small, can be exploited to allow the design of PCR primers specific to one locus. This may prove to be useful in cases where the tagging of a specific copy of a homoeologous locus is required. If this concept could be combined with a sensitive assay for sequence polymorphism, such as single-stranded-conformation polymorphism (Orita et al. 1989), useful locus-specific PCR markers could be developed. Nonetheless, this type of approach requires the cloning and sequencing of PCR-amplified copies of a homoeologous locus and, therefore, the amount of work involved is considerable. The cost of converting an RFLP to a PCR-based specific marker needs to be considered on a case-by-case basis and may only prove worthwhile where there is a definite goal in mind, for example, the tagging of a closely linked quantitative trait locus or resistance gene.

Acknowledgements This work was undertaken as part of a programme jointly funded by the Biotechnology and Biological Sciences Research Council, Department of Trade and Industry, Zeneca plc, Plant Breeding International Ltd, Nickerson Biocem Ltd and Ciba Geigy plc.

References

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

D'Ovidio R, Anderson OD (1994) PCR analysis to distinguish between alleles of a member of a multigene family correlated with wheat bread–making quality. Theor Appl Genet 88:759–763

Erpelding JE, Blake NK, Blake TK, Talbert LE (1996) Transfer of sequence-tagged-site (STS) markers between wheat and barley. Genome 39:802–810

Gill KS, Gill BS (1996) A PCR–based screening assay of *Ph1*, the chromosome-pairing regulator gene of wheat. Crop Sci 36:719–722

Goodall GJ, Kiss T, Filipowicz W (1991) Nuclear RNA splicing and small nuclear RNAs and their genes in higher plants. Oxford Surv Plant Mol and Cell Biol 7:255–296

Kurata N, Moore G, Nagamura, Foote T, Yano M, Minobe Y, Gale M (1994) Conservation of genome structure between rice and wheat. Bio/technology 12:276–278

McFadden ES, Sears ER (1946) The origin of *Triticum spelta* and its free threshing hexaploid relatives. J Hered 37 : 81–89

Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proc Natl Acad Sci USA 86 : 2766–2770

Penner GA, Stebbing J, Legge B (1995) Conversion of an RFLP marker for the barley stem rust resistance gene *Rpg1* to a specific PCR-amplifiable polymorphism. Mol Breed 1 : 349–354

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

Sears ER (1976) A synthetic hexaploid wheat with fragile rachis. Wheat Inf Serv 41–42 : 31–32

Sharp PJ, Kreis M, Shewry P, Gale MD (1988) Location of β-amylase sequence in wheat and its relatives. Theor Appl Genet 75 : 289–290

Shattuck-Eidens DM, Bell RN, Neuhasen SL, Helentjaris T (1990) DNA sequence variation within maize and melon: observations from polymerase chain reaction amplification and direct sequencing. Genetics 126 : 207–217

Talbert LE, Blake NK, Chee PW, Blake TK, Magyar GM (1994) Evaluation of "sequence-tagged-site", PCR products as molecular markers in wheat. Theor Appl Genet 87 : 789–794

Tragoonrung S, Kanazin V, Hayes PM, Blake TK (1992) Sequence-tagged-site-facilitated PCR for barley genome mapping. Theor Appl Genet 84 : 1002–1008

Van Campenhout S, Vander Stappen J, Sagi L, Volckaert G (1995) Locus–specific primers for LMW glutenin genes on each of the group 1 chromosomes of hexaploid wheat. Theor Appl Genet 91 : 313–319

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nuclic Acids Res 23 : 4407–4414

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

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